

Perspectives in Pragmatics, Philosophy & Psychology 1

Alessandro Capone
Franco Lo Piparo
Marco Carapezza *Editors*

Perspectives on Pragmatics and Philosophy

 Springer

Perspectives in Pragmatics, Philosophy & Psychology

Volume 1

Editor-in-Chief

Alessandro Capone, Barcellona, Italy

Consulting editors

Wayne Davis, Washington, USA

Igor Douven, Groningen, Netherlands

Franco Lo Piparo, Palermo, Italy

Louise Cummings, Nottingham, UK

For further volumes:

<http://www.springer.com/series/11797>

Alessandro Capone · Franco Lo Piparo
Marco Carapezza
Editors

Perspectives on Pragmatics and Philosophy

 Springer

Editors

Alessandro Capone
University of Messina/Palermo
Barcellona Pozzo di Gotto
Italy

Franco Lo Piparo
Marco Carapezza
University of Palermo
Palermo
Italy

ISSN 2214-3807

ISBN 978-3-319-01010-6

DOI 10.1007/978-3-319-01011-3

Springer Cham Heidelberg New York Dordrecht London

ISSN 2214-3815 (electronic)

ISBN 978-3-319-01011-3 (eBook)

Library of Congress Control Number: 2013951326

© Springer International Publishing Switzerland 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law. The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgments

I would like to give thanks to a number of people who made this project possible. Of course, the Authors of this volume need special thanks and praise for writing the papers within strict limitations of time and, also, of subject-matter. My friends Igor Douven, Wayne Davis, Istvan Kecskes, Jacob L. Mey, Yan Huang, Katarzyna Jaszczolt, Michel Seymour, Neil Feit, and Keith Allan have tangibly helped the cause of this book, directly or indirectly, by providing feedback. Words do not suffice to express my gratitude. Ties Nijssen (Springer) has stimulated this project by providing us editorial guidance and also by persuading the publisher to accept the manuscript. I would like to thank him warmly. I would like to give thanks to Francesca Piazza, Pietro Perconti, Marco Mazzone, and Francesco La Mantia, whose reassuring attitudes have played an important role in launching this project. I would like to thank Eleni Gregoromichelaki for her continuous support, help, and encouragement.

Professor Tullio De Mauro stimulated my research activity on pragmatics and philosophy. I owe my various papers on this topic to his benevolent and stimulating attitude to me. I am eternally grateful.

I would like to give warm thanks to my coeditors Franco Lo Piparo and Marco Carapezza for their fruitful and tireless collaboration.

But chiefly I would like to dedicate this volume to my father and my mother, Giuseppe Capone and Venerina Scilipoti, who have never stopped providing me with books and paper.

We would like to give thanks to Routledge, John Benjamins, Elsevier, and OUP for granting permissions to republish the following papers.

Dascal, Marcelo. 2005. Debating with myself and debating with others. In *Controversies and subjectivity*, eds. Pierluigi Barrotta and Marcelo Dascal, 33–73. Amsterdam: John Benjamins.

Saka, Paul. 1998. Quotation and the use-mention distinction. *Mind* 107/425: 113–135.

Michel, Seymour. 2010. Speech act pluralism, minimal content and pragmemes, *Journal of Pragmatics* (Elsevier, Oxford), In *Pragmemes*. Special issue of the journal of Pragmatics 42, ed. A. Capone, 2870–288. Oxford: Elsevier. ISSN 0378–2166

Steven Gross. What is a context? This paper originally appeared as Chapter II of Gross 1998 Harvard Ph.D. Dissertation, which was subsequently published by Routledge as Gross 1998/2001. *Essays on Linguistic Context-Sensitivity and its Philosophical Significance*.

Alessandro Capone

Contents

Conversation and Collective Belief	1
Margaret Gilbert and Maura Priest	
The Single Norm of Assertion	35
Martin Montminy	
From Thought Experiments to Real Experiments in Pragmatics	53
András Kertész and Ferenc Kiefer	
What Makes a Property “Semantic”?	87
Michael Devitt	
What is a Context?	113
Steven Gross	
Implicature, Inference and Cancellability	133
Michael Haugh	
Grice, Conversational Implicature and Philosophy	153
Siobhan Chapman	
Writing Letters in the Age of Grice	189
Claudia Bianchi	
Implicatures as Forms of Argument	203
Fabrizio Macagno and Douglas Walton	
Some Remarks About Speech Act Pluralism	227
Marina Sbisà	
Speech Act Pluralism, Minimal Content and Pragmemes	245
Michel Seymour	

Language Adds to Context.	277
Paolo Leonardi	
Squaring the Circle.	291
Kepa Korta and John Perry	
Irregular Negations: Pragmatic Explicature Theories.	303
Wayne A. Davis	
The (in)Significance of the Referential-Attributive Distinction.	351
Anne Bezuidenhout	
Quotation and the Use-Mention Distinction.	367
Paul Saka	
Indirect Reports and Pragmatics.	389
Nellie Wieland	
Immunity to Error Through Misidentification (IEM), ‘de se’ and Pragmatic Intrusion: A Linguistic Treatment.	413
Alessandro Capone	
Further Reflections on Semantic Minimalism: Reply to Wedgwood. . .	437
Alessandro Capone	
Putting the Pragmatics of Belief to Work.	475
Igor Douven	
Contexts, Fiction, and Truth.	489
Alberto Voltolini	
Pragmatics and Philosophy: Three Notes in Search of a Footing.	501
Alec McHoul	
A Brief Essay on Slurs.	507
Luvell Anderson and Ernie Lepore	
Viewing the Study of Argumentation as Normative Pragmatics.	515
Frans H. van Eemeren and Bart Garssen	
Rhetoric and Pragmatics: Suggestions for a Fruitful Dialogue.	537
Francesca Piazza	

Debating with Myself: Towards the Psycho-Pragmatics and Onto-Pragmatics of the Dialectical Self	557
Marcelo Dascal	
Truth, Negation and Meaning	593
Franco Lo Piparo	
Habermas' <i>Universal Pragmatics</i>: Theory of Language and Social Theory	609
Alessandra Pandolfo	
For in Psychology There are Experimental Methods and <i>Conceptual Confusion</i>: From Embodied Cognition to Wittgenstein on Language and Mind	637
Felice Cimatti	

Introduction

We would like to give thanks to Ties Nijssen who encouraged the idea of this book in a number of ways, giving us feedback and advice.

Pragmatics (or the Pragmatics of Language) is one of those areas of human knowledge that are rapidly expanding, transcending the boundaries of linguistics proper (the most recent applications outside linguistics concern anthropology, law studies, literary theory, etc.). No one really denies the importance of pragmatic theory, but very often the wars at the boundaries (e.g., the boundaries between semantics and pragmatics or between syntax and pragmatics) are fierce, and often produce no definitive results. So we are clearly faced with a discipline which on the one hand is evolving rapidly, on the other hand is constantly under attack and runs the risk of regression. In this volume, we only confine ourselves to results which appear to be solid enough to allow us to draw the boundaries and give shape to what promises to be a stable picture of pragmatics and philosophy. We pass on to students and scholars the territory of this subject confident enough that the theoretical questions presented here will produce further theoretical questions and discussions as offspring.

We are pleased to introduce a volume that is nonstandard in its coverage of topics pertaining to pragmatics and philosophy. The idea of assembling this volume came to us during numerous conversations with philosophers like Igor Douven and Wayne Davis, but we should say that other people encouraged the idea, like Istvan Kecskes, Yan Huang, Jacob L. Mey, Keith Allan, Ferenc Kiefer, and many others whose names figure in the Editorial Board of the series in which this volume appears.

Pragmatics deals with inferential processes allowing hearers to recover the speaker's intention(s) as evinced through her speech with the aid of numerous cues and clues. A genuine pragmatics of language must be able to situate the message in the context from which it draws its meaningfulness (see Gross, in this volume) and thus also requires a sensitivity to socio-linguistic considerations (Mey 2001). Initially, pragmatics started in the philosophers' publications, but soon it was divulged and spread to other fields of inquiry, such as linguistics. Within linguistics, during the twentieth century we have seen the most important pragmatic theories thrive and persuade numerous practitioners to follow this path of inquiry, being finally integrated in linguistic methodology going beyond the idea of the

Saussurian sign with its strict relationship between the signifier and the signified. The most spectacular achievement of this period is the application of pragmatics to linguistic anaphora, ellipsis, etc. In the new millennium, we witness a trend that is the reverse of what happened in the twentieth century. Pragmatics, as we have it today, deeply influenced by linguistic ideas, can be put to use in philosophical enterprises, following the model of Paul Grice's original ideas. We cannot easily forget, in fact, that pragmatics started as a way to rescue philosophy from skepticism (see Grice's essay on perception; see also Chapman, in this volume). Recently, Igor Douven has proven that pragmatics can play a role in the dissolution of Gettier's problem and utilizes it constructively to provide an adequate pragmatics of belief. In other words, Douven uses pragmatics to solve epistemological problems (see also Douven, in this volume). In essentials Douven develops an Epistemic Hygienics:

Epistemic Hygienics (EH): Do not accept sentences that could mislead your future selves.

This principle is clearly involved when we fail to memorize sentences or thoughts in a form which might mislead our future selves and clearly has an effect on what we believe. I think that one should find a way to derive the Epistemic Hygienics, developed by Igor Douven, from general cognitive principles and, in particular the Principle of Relevance by Sperber and Wilson (1986); but this is an issue which is beyond the scope of this Introduction. I have reasons to believe that Douven's Epistemic Hygienics can be applied to (philosophical) issues such as knowing how, quotation, and indirect reports.

Other people did what Igor Douven is doing for the pragmatics of belief; thus a number of scholars have grappled with the pragmatics of belief reports showing how Millian ideas can be reconciled with the view that substitutions of co-referential NPs in that-clauses of belief reports appear to be illicit. Consider the following example. Mary meets Elisabeth and comes to know a number of things about her. However, she ignores (as that information has not been given to her) that Elisabeth is the Queen of England. Then one can very well describe one of Mary's beliefs as in the statement "Mary believes that Elisabeth is playing tennis" but attributing her the belief that the Queen of England is playing tennis will not do, essentially, because we have no idea whether she also believes that Elisabeth is the Queen of England or we positively know that she believes that Elisabeth is not (likely to be) the Queen of England. A statement such as "Mary believes that the Queen of England is playing tennis" pragmatically communicates that Mary has that belief under the mode of presentation "Queen of England" while, in fact, the mode of presentation of the reference that is operative in her belief is "Elisabeth". Presumably the same kind of story can be provided for the analysis of indirect reports, which pragmatically mirror belief reports.

Pragmatics does not only deal with defeasible or, in any case, nonconventional inferences, but also with rules of discourse (or rules of usage). The papers by Montminy and Anderson and Lepore discuss discourse rules correlating with assertion and the prohibition of slurs. These papers are of interest because they show that assertions are rooted in discourse practices which have been consolidated

through the sedimentation of collective beliefs (see also Gilbert and Priest, this volume).

Now, if we have managed to extend somehow the number of topics a pragmatic theory must deal with, it is clear that in a volume with a specific focus on pragmatics and philosophy there is greater hope to extend still further the number of topics that can be treated with the aid of a pragmatic theory. Thus, we asked for the collaboration of numerous colleagues, experienced authors who have dealt with the topics of interest to us, and we found that our proposal has had a deep resonance and impact on their imaginative powers so much so that in about a year we have been ready to offer a thick volume on the topics proposed to readers. We also asked our colleagues to write in a way that could be easily grasped by those outside philosophy, say anthropologists, literary theorists, etc. So an ideal audience is in front of us: all those who would like to familiarize with pragmatics and would like to see in what ways pragmatics can deal with philosophical problems. We invited philosophers to write papers, with an aim to let philosophy and pragmatology interact in a fruitful way.

Retrospectively, we should say that we were very glad to have persuaded world renowned academics to write on areas in which they are specialists. So, this volume works well both as a stimulus to further research, as our authors provide guidelines for future research by asking important theoretical questions, and as a guide to less experienced researchers and students who would like to know more about this vast, complex, and difficult field of inquiry. We have both asked theoretical questions which are in need of important replies and provided a list of topics that is ambitious. This is not to say that this list of topics is exhaustive, but I am sure that it will allow students to find out what they are interested in. To use the words of Kasia Jaszczolt (personal communication), one should also find the topics that one tackles stimulating and intriguing ('enjoyable') from the point of view of the researcher. Furthermore, this volume will allow readers to find out more about their inner intellectual resources, as research is often a matter of "bringing out" what one has in one's mind. The maieutical role of pedagogy is embodied in the principles which allowed the editors to assemble the articles.

In this book, we largely have discussions about conversational implicatures and (conversational) explicatures. Explicatures figure prominently in what is said—as in many cases it is not possible to recover the proposition intended by the speaker without unpacking the implicit parts of the message (by using resources such as cues and clues, background knowledge, etc.). While ambiguity reduction and reference fixing (saturation of deictic elements) are surely part of explicatures, recent discussions in pragmatics have stressed the role played by free enrichment, that is the provision of constituents of thoughts that are combined with the ones that belong to the explicit logical form (see Carston 2002). Conversational implicatures are generally meaning augmentations on top of explicatures—even if the current views are that explicatures and implicatures are worked out in parallel (through mutual adjustments). There is a tension within current theories between those authors who adhere to a more classical view of conversational implicatures, and those who want to analyse some of these implicatures as explicatures. What

characterizes both explicatures and conversational implicatures is that they are calculable and nonconventional (that is, they are not (directly) tied to conventional meaning. Presumably nondetachability also correlates with explicatures and conversational implicatures (even if there are some controversial cases).

The volume is largely about Gricean pragmatics, which mainly deals with a speaker's communicative intentions as manifested through her speech (or animating her speech). Intentions, of course, are in the speaker's head; however, if they were solely in the speaker's head, they could not come out, they could not emerge from her mind and be shared with the co-participants. This sharing of intentions happens when a speaker manages to get her message across. Now, clearly this practice is not like putting a message into a bottle and putting it into the sea in the hope that someone, if anyone at all picks up the bottle, will be induced to open it, read the message and then carry out an action. Sending a message in daily interaction amounts to shaping the message in such a way that the speaker's intention(s) can be recognized by the hearers. Recognition can be aided through contextual cues and clues. Cues may lead a hearer to detect an interpretation problem (e.g., a mismatch between the literal and the nonliteral level of meaning) and clues can serve to fill the gaps left in the text (Dascal and Weizman 1987; see also Sbisà in this volume). But now it is clear that the pragmatic enterprise is not like bottle throwing in the hope that the message in the bottle will be recovered by someone; the message must be shaped by taking into account the concrete situation of the utterance and the predispositions of the hearers to pick up the cues and clues that direct the communication process. We have a hearer whose history we know, we have a person in front of us, and we use part of the history we share with her and the background assumptions that normally have a bearing on the understanding of a message to anticipate the direction which the interpretation work will take. We furthermore use principles of language use (whether we accept Gricean pragmatics (see Chapman, this volume) or a theory à la Sperber and Wilson to anticipate the direction which the interpretation process will take. So, a theory of interpretation must also be one of the speaker's intentions. Intentions and recovered messages are both the objects of pragmatics, since speakers and hearers both shape their codification and interpretation processes by taking into account principles of language use (whether the Gricean maxims or a single cognitive/communicative principle). This topic also intersects with that of collective beliefs (Gilbert and Priest in this volume), since while a conversation unfolds, both the speaker and the hearer construct shared files which are mutually recognized as being shared. Intentions are not only communicated but recognized and this process of communication/recognition is validated by shared files that are part of officially recognized resources.

Can intentions communicated through pragmatics be canceled? (And are they different, in this respect, from intentions projected through semantics (if there are such things)). Cancellability is the focus of an acrimonious controversy. There are some authors who take the nonstandard position that once intentions are communicated (even if through pragmatics means) they cannot be unimplicated, as numerous clues and cues have been mobilized to make the speaker and the hearer converge toward the intended message. Some Authors also make a distinction

between potential and actual explicatures—potential explicatures being cancellable. Some suggest that there are cases of explicatures that cannot be canceled, even if explicatures in general can be canceled. Presumably this intermediate position is likely to be consolidated in the future. This is an intriguing topic, which has been discussed recently by Burton-Roberts (2005), as well as myself. Scholars are split on this issue. While Capone (2009), drawing on ideas by Burton-Roberts (2005), develops the idea that in many important cases explicatures are not cancellable, Seymour (this volume) and Michal Haugh (this volume) take the contrary stance. Of course, only through future work can we arrive at a clearer view of the way things stand. But it is important to disseminate the seeds of the controversy in this book.

Is it true that semantics only provides schematic information which is then expanded into propositional forms through pragmatic processes? This is clearly one of the problems tackled in this book, and one promising to receive different answers. Our approach is not dogmatic, but we invite readers to be open to the multifarious proposals in this area of the semantics/pragmatics debate. What is clear is that, whatever our approach to pragmatic intrusion (whether we accept a minimal semantics which is fully propositional or a semantic theory which only provides interpretative schemas) pragmatic intrusion can be put to use in the understanding of philosophical topics. Which are the topics that can benefit most from pragmatic intrusion? These are clearly those of indirect reports (see Wieland in this volume), quotation, mixed quotations (see Saka in this volume), belief reports, knowing how, “de se” attitudes, Immunity to Error through misidentification (see Capone in this volume), the attributive/referential distinction (see Bezuidenhout, in this volume), fictional names (see Voltolini in this volume), negation (see Davis in this volume), and argumentation theory (see van Eemeren and Garssen), etc. So we are sure that by providing multiple connections between linguistic and philosophical topics, we can find out how to fuel the philosophical topics and how to construe them in such a way that we can provide constructive answers.

Is there any hope that we can find out further topics where pragmatics and philosophy interconnect? Our answer is positive. At this point we dare say that the issue of how to extend these topics is one of method and not of content. Whenever we find that a philosophical question can be reduced to a linguistic one (even if partially) and if the linguistic question reduces in one way or another to a pragmatic one, then we are sure that pragmatics can be applied to the resolution or dissolution of a given philosophical problem (see also Chapman on Grice and philosophy, in this volume). But how can a philosophical question reduce to a pragma-linguistic question? If a metaphysical or epistemological issue hinges on a logical problem and the logical problem has a linguistic structure using connectives or words admitting pragmatic intrusion, then the metaphysical or epistemological issue promises to be reducible to a pragmatic issue (See Capone on Immunity to Error through Misidentification and pragmatic intrusion, in this volume).

In general, if pragmatics makes progress and demonstrates that a linguistic phenomenon evinces partial or drastic pragmatic intrusion and if it turns out that an

epistemological issue uses that linguistic structure as part of the discussion and the essential part of the discussion hinges on that linguistic structure (or part of the argumentation is based on that linguistic structure), then we know in advance that pragmatic intrusion can be applied in a fruitful way to that epistemological problem (see Douven, in this volume). The question, essentially, is whether a certain epistemological problem is also partly a semantic problem and whether that semantic problem turns out to be a semantico/pragmatic problem (one allowing or requiring pragmatic intrusion). Of course, it is not easy to assess whether an epistemological problem is also partially a semantic problem and every attempt must/can be made to disentangle the epistemological from the semantic side of the problem. However, if, despite all attempts, the epistemological problem hinges on some semantic categories (say pronominals or indexicals or some other linguistic structure), and such categories are amenable to pragmatic analysis, then it is clear that applicability of pragmatic categories to epistemology must be granted. Is not this a good thing for pragmatics? Is not this a good thing for the view that much of our philosophy hinges on linguistic structures? We assume positive answers to these questions. This is not surprising since much of our theorizing is done through arguments which are crucially advanced in virtue of the linguistic resources of our theoretical arsenal.

Alessandro Capone
 Franco Lo Piparo
 Marco Carapezza

References

- Burton-Roberts, Noel. 2005. Robyn Carston on semantics, pragmatics and encoding. (Review article of Carston 2002). *Journal of Linguistics* 41: 389–407.
- Capone, Alessandro. 2006. On Grice's circle (further considerations on the semantics/pragmatics debate). *Journal of Pragmatics* 38: 645–669.
- Capone, Alessandro. 2008. Belief reports and pragmatic intrusion (the case of null appositives). *Journal of Pragmatics* 40: 1019–1040.
- Capone, Alessandro. 2009. Are explicatures cancellable? Towards a theory of the speaker's intentionality. *Intercultural Pragmatics* 6/1: 55–83.
- Capone, Alessandro. 2010a. On the social practice of indirect reports (further advances in the theory of pragmemes). *Journal of Pragmatics* 42: 377–391.
- Capone, Alessandro. 2010b. What can modularity of mind tell us about the semantics/pragmatics debate. *Australian Journal of Linguistics* 30/4: 497–520.
- Capone, Alessandro. 2010c. The attributive/referential distinction, pragmatics, modularity of mind and modularization. *Australian JL* 31/2: 153–186.
- Capone, Alessandro. 2010d. "Between Scylla and Charybdis": The semantics and pragmatics of attitudes 'de se'. *Journal of Intercultural Pragmatics* 7/3: 471–503.
- Capone, Alessandro. 2013. The pragmatic of quotation, explicatures and modularity of mind. *Pragmatics and Society* 4/3: 259–284..
- Carston, Robyn. 2002. *Thoughts and utterances. The pragmatics of explicit communication*. Oxford: Blackwell.

- Dascal, Marcelo, and Weizman, Elda. 1987. Contextual exploitation of interpretation clues in text understanding: An integrated model. In *The Pragmatic perspective*, eds. Verschueren, J and Bertuccelli-Papi, M, 31–46
- Dascal, Marcelo, Hintikka, J., and Lorenz, K. 1996. Games in language. In *Sprachphilosophie* ed. Dascal, M. Gerhardus, D., and Lorenz, K. 1996, vol II,(1371–1392). Berlin: Mouton De Gruyter.
- Douven, Igor. 2010. The pragmatics of belief. *Journal of Pragmatics* 42: 35–47.
- Grice, H. Paul. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.
- Huang, Yan. 2000. *Anaphora. A cross-linguistic study*. Oxford: Oxford University Press.
- Lo Piparo, Franco. 2003. *Aristotele e il linguaggio. Cosa fa di una lingua una lingua*. Roma-Bari: Editori Laterza.
- Mey, Jacob, L. 2001. *Pragmatics. An Introduction, second edition*. Oxford: Blackwell.
- Perry, John. 2001. *Reference and reflexivity*. Stanford: CSLI Publications
- Sperber, D. and Wilson, D. 1986. *Relevance theory* (Reprinted with postface in 1995). Oxford: Blackwell.
- Van Eemeren, F. and Grootendorst, R. 1982. The speech acts of arguing and convincing in externalized discussions. *Journal of Pragmatics* 6: 1–24.
- Wierzbicka, Anna. 2006. *English. Meaning and Culture*. Oxford: Oxford University Press.
- Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Oxford: Blackwell.
- Wong, Jock. 2010. The triple articulation of language. In *Pragmemes*, ed. Capone, A. *Journal of Pragmatics*.

Pragmatics and Philosophy

Margaret Gilbert and Maura Priest

This paper introduces some ideas developed over the past 20 years concerning the phenomenon referred to here as *collective belief*, and relates those ideas to the topic of conversation. Generally it suggests the benefits of a less individualistic approach to the social interaction constituted by a conversation than is standard. It proposes in particular that paradigmatic conversations involve the development of a *collective cognitive profile* of the parties. This occurs through the negotiation of a series of collective beliefs. These are constituted by commitments that are *joint* in a sense that is explained. The parties to any joint commitment have obligations towards each other to conform to the commitment. This helps to entrench a given collective belief once established. Even when interlocutors do not manage to negotiate a collective belief whose content has explicitly been specified, they are likely to establish one or more associated *implicit* collective beliefs. This idea is briefly related to some of the existing literature on conversation, including classic articles by Stalnaker and Lewis on presupposition and conversational score.

Martin Montminy

Timothy Williamson holds that the knowledge rule of assertion, according to which one should assert only what one knows, is the single norm of assertion. Montminy explains and defends Williamson's thesis. He identifies three key features associated with the thesis: the single norm of assertion should be *constitutive*, *individuating* and *basic*. Roughly, a constitutive rule of a speech act governs every possible performance of the act. A rule of assertion is individuating if it differentiates assertion from every other speech act. Finally, a rule of assertion is basic if it does not derive from other rules governing assertion, or other speech acts. He thus construes Williamson's thesis as the claim that the knowledge rule is the only individuating, constitutive and basic rule of assertion. This thesis is compatible with the idea that assertion is governed by nonepistemic constitutive rules such as moral and prudential rules. Since these rules are not individuating, there is no need to understand Williamson's thesis as bearing solely on the

epistemic rule of assertion. The Author explains how the existing arguments in favor of the knowledge rule can serve to support the thesis that this rule is also the single norm of assertion. He also responds to some objections against the knowledge rule, and criticizes alternative rules.

András Kertesz and Ferenc Kiefer

The puzzle of thought experiments is a hot topic in the philosophy of science. The paper raises the puzzle with respect to pragmatics as follows: How is it possible that thought experiments in pragmatics yield new experiential information about communication, although they are carried out entirely in one's head? The paper shows, first, that the structure of thought experiments in pragmatics consists of a series of plausible inferences. Second, the function of thought experiments is to serve as the initial step in the process of plausible argumentation as well as to test the plausibility of rival hypotheses. Third, while on the one hand, thought experiments and real experiments may be continuous, on the other hand, the former may be also indispensable components of the latter. Fourth, these properties provide a solution to the puzzle of thought experiments in pragmatics. The key idea of the solution is that thought experiments in pragmatics cannot generate new experiential information; rather, during the process of plausible argumentation they contribute to the retrospective re-evaluation of experiential information already given.

Michael Devitt

It is common to distinguish the “semantic” properties of an utterance from its “pragmatic” properties, and what is “said” from what is “meant”. What is the basis for putting something on one side rather than the other of these distinctions? Such questions are usually settled largely by appeals to intuitions. The paper rejects this approach arguing that we need a theoretical basis for these distinctions. This is to be found by noting that languages are representational systems that scientists attribute to species to explain their communicative behaviors. We then have a powerful theoretical interest in distinguishing, (a) the representational properties of an utterance that arise simply from the speaker's exploitation of a linguistic system from, (b) any other properties that may constitute the speaker's “message”. Devitt calls the former properties “semantic”, the latter, “pragmatic”. The semantic ones are constituted by properties arising from linguistic conventions, disambiguations, and reference fixings. Devitt foreshadows an argument in a forthcoming book, *Overlooking conventions: the trouble with linguistic pragmatism*, that many of the striking examples produced by linguistic pragmatists exemplify semantic rather than pragmatic properties. This argument counts against the popular pragmatist theses of “semantic underdetermination” and “truth-conditional pragmatics”. It is very much in the spirit of the tradition that pragmatists reject.

Steven Gross

Gross examines a variety of criteria for context individuation, with an eye towards identifying and distinguishing various theoretical projects for which they would be appropriate in semantics and pragmatics. The discussion is organized around six possible desiderata for contexts: (1) that they metaphysically determine proposition expressed, (2) that they be nontrivial (in a sense explained in the text), (3) that they provide epistemic illumination of how hearers understand what proposition was expressed, (4) that they be (in various senses) finite, (5) that they be nonintentionally characterized, and (6) that they be context-insensitively characterized. Particular attention is directed to the conceptions of context developed by Lewis, Stalnaker, and Davidson(ians).

Michael Haugh

The standard position in pragmatics to date has been that cancellability is useful way of differentiating implicatures from logical implications, semantic entailments and the like. In recent years, however, there has been considerable debate as to whether implicatures are in fact always cancellable, or indeed whether they are cancellable at all, amongst linguistic pragmaticians and language philosophers. In this chapter, it is suggested that cancellability encompasses a range of actions that play out in different ways depending on whether we are analyzing inferences that can lead to implicatures or the implicatures themselves. In this way, we can see how analysts have often underplayed the contingency of inferences as well as the inherent indeterminacy of implicatures in such debates. Haugh proposes that cancellability should thus be the subject of further empirically-driven analyses in order to provide a solid foundation for the theorization of implicature.

Siobhan Chapman

This chapter addresses the contrast between Grice's philosophical motivations in developing his account of conversational implicature, and the linguistic framework in which it has subsequently generally been discussed. It begins with a review of the dichotomy in twentieth century analytic philosophy between 'ideal language' and 'ordinary language' philosophy, and discusses Grice's work as an attempt to demonstrate some fundamental misconceptions in both positions. It offers an exegesis of Grice's conception of conversational implicature and then considers his application of it to a range of established philosophical problems, including the viability of his own earlier work on "nonnatural meaning", the contrasting claims of realism and skepticism, apparent differences between logic and natural language and the debate over Russell's logical account of definite descriptions. This chapter concludes with an assessment of the significance of Grice's delineation of "what is said" as a defining opposite of "what is implicated". For Grice himself, although it was central to his original philosophical motivations, this remained one of the most troublesome and least successful

aspects of his philosophy of language. For pragmaticists it has proved one of the most enduring, challenging and intriguing topics of debate.

Claudia Bianchi

This article aims to investigate the notion of implicature and its connections with speaker's intentions, communicative responsibility and normativity. Some scholars stress the normative character of conversational implicatures more than their psychological dimension. In a normative perspective, conversational implicatures don't correspond to what the speaker intends to implicate, but should be interpreted as enriching or correcting inferences licensed by the text. Bianchi's paper aims to show that the idea of an implicature that the speaker does not intend to convey is not persuasive. In Grice's theory conversational implicatures are speaker-meant: this means that inferences derived by the addressee but not intended by the speaker should not count as conversational implicatures. On the contrary, Bianchi claims that propositions intended by the speaker and not recognized by the addressee should count as implicatures, if the speaker has made her communicative intention available to her audience.

Douglas Walton and Fabrizio Macagno

In this paper, the Authors use concepts, structure and tools from argumentation theory to show how conversational implicatures are triggered by conflicts of presumptions. Speech acts are conceived as grounded on different types of presumptions, which can be pragmatic, factual or epistemic in kind. Ordinary conversation is based on generic principles valid most of the time, such as the generalization that an interrogative statement is presumed to request information, or that people can usually perform ordinary actions. However, such presumptions can conflict with each other. For instance, the two presumptions above are the conflicting ground on which the question whether the hearer can pass the salt or not is grounded. This apparent contradiction needs to be explained or solved. Implicatures are shown to be forms of explanation, based on defeasible forms of inference used in conditions of lack of knowledge, such as abductive reasoning, argument from sign, pragmatic or causal argument.

Marina Sbisà

This paper moves from a well-known pragmatic theme in the philosophy of language—speech acts—to the philosophy of action. It focuses on the problem whether an utterance token can carry out more than one speech act, to which the attention of philosophers of language was drawn by Cappelen and Lepore's (2005) defense of speech act pluralism. It is argued that Austin's notions of the locutionary, illocutionary and perlocutionary acts already presuppose speech act pluralism. In particular, it is shown that the notion of perlocutionary act, as defined by Austin, makes sense only within a pluralist framework, where no additional

behavior on the part of the speaker is needed in order for her performance of a perlocutionary act to take place, besides those of the locutionary and illocutionary acts. The central section of the paper discusses speech act pluralism as it applies to illocutionary acts, in the light of an Austin-inspired view of illocution developed by Sbisà herself and of the analysis of the conversational practices of multiact ascription recently outlined by Levinson. As to the locutionary act, it is argued that plural ascriptions are by far more limited than is claimed by Cappelen and Lepore, if they exist at all. In conclusion, it is shown that all the views admitting of speech act pluralism that are examined in the paper (those of Cappelen and Lepore, Levinson, and Sbisà) require an ascription-centered conception of action, such as that outlined by Austin in his papers in the philosophy of action.

Michel Seymour

Are there pragmemes determining primary sentence meanings that are not prescribed by semantic features and that are not cancelable? In this paper, Seymour argues that there are no such examples. Pragmemes may contribute to the determination of the content of certain assertions, but they do not contribute to the determination of minimal content of the sentence-types used in these utterances. Seymour concludes that a proper appreciation of the role of pragmemes forces us to accept speech act pluralism and bifurcationism, the idea that there are two levels of content: minimal and maximal. That is, different pragmemes produce different inferential augmentations of a minimal level of linguistic meaning. But this is precisely what semantic minimalism is all about.

Paolo Leonardi

Contextualism deems that the meaning of most utterances, if not any, varies from context to context. Here, Leonardi argues, instead, that utterances added to a context change the context by shifting salience. Considering truth-conditions together with appropriateness, he tries to show that utterances claimed to make the same assertion—such as “Rain is over”, “Here, rain is over”, “In Florence, rain is over” and “Here in Florence, rain is over”—differently affect salience, and hence in an important sense do *not* assert the same. The idea is rather that “Rain is over”, for instance, says the same in any context, but that it differently shifts salience in different contexts, and the various assertions listed may get at the same salience only in different contexts.

Kepa Korta and John Perry

Making the distinction between semantics and pragmatics has proven to be a tricky task, leading to several problems that look like Gordian knots, or worse; perhaps semantics and pragmatics are so tangled that separating them is impossible, like squaring the circle. A widespread, plausible, Grice-inspired view of the distinction is threatened by what (Levinson 2000) called “Grice’s circle.”

Gricean inferences to derive the pragmatic content of the utterance (such as conversational implicatures) require the determination of what is said (also known as the “semantic content” or the “literal truth-conditions”); but determining what is said (by processes of disambiguation, precisification, reference fixing, etc.) requires pragmatic inference. In a nutshell, pragmatic inference both requires and is required by the determination of what is said. Thus, there is no way to unravel semantics and pragmatics. In this paper, Kępa Korta and John Perry will show how to square Grice’s circle. They untie the semantics/pragmatics knot, without using any of Alexander’s methods: slicing it with a sword or removing the (semantic) pin around which it was bound. The approach consists in assuming a minimal but truth-conditionally complete notion of semantic content (Perry 2001), which doesn’t constitute what is said by the utterance, but does provide the required input for pragmatic reasoning.

Wayne Davis

Davis will examine negations that are “*irregular*” in that they are not used in accordance with the standard rules of logic, including scalar, metalinguistic, and evaluative implicature denials; presupposition denials; and contrary affirmations. The principal question is how their irregular interpretation is related to their regular interpretation. Davis focuses here on theories maintaining that one of the interpretations is an “*explicature*” (Carston) or “*implicature*” (Bach) rather than an implicature (Grice, Horn, Burton-Roberts). Three types can be distinguished, depending on whether the theory claims that the regular interpretation can be “*pragmatically derived*” from the irregular, the irregular from the regular, or both the regular and the irregular from the meaning of the negation. Davis argues that no compelling reason has been provided for distinguishing explicatures or implicatures from implicatures, and that claims that one interpretation can be pragmatically derived using either Gricean or relevance theory are completely unfounded. With one class of exceptions, Davis argues for a semantic ambiguity thesis maintaining that irregular interpretations are idioms that plausibly evolved from generalized conversational implicatures. The exceptions are evaluative implicature denials, which are still live implicatures.

Anne Bezuidenhout

This paper tackles the question as to whether or not the referential-attributive (RA) distinction has any information-structural significance. That is, does this distinction mark a contrast between different ways a speaker might package the informational content of utterances that use definite descriptions? Furthermore, are there any overt markers of this distinction? In this paper Bezuidenhout focuses on referential and attributive uses of definite descriptions, and leaves the issue of indefinites for another occasion. She answers both of the above questions negatively. While definite descriptions have an information-structural role to play,

the RA distinction does not. Moreover, after examining several potential candidates for markers of the RA distinction, she concludes that there are no such overt markers. This is in fact to be expected if one accepts my proposal to see referential and attributive uses of definite descriptions as having the same general function, namely to single out something as a center of interest. The difference is simply that referential uses focus on role bearers, whereas attributive uses focus on role properties.

Paul Saka

In “Quotation and the Use-Mention Distinction,” Paul Saka argues that the ability to quote an expression ultimately rests upon the ability to mention it. Simplifying a little, Saka says that while speaker *S* *uses* an expression *X* if and only if *S* intends to direct the thoughts of the audience to the customary extension of *X*, *S* *mentions* an expression *X* if and only if *S* intends to direct the thoughts of the audience to some item associated with *X* other than its customary extension. (In Saka 2011, this gets revised to: *S* mentions an expression *X* if and only if *S* intends to direct the thoughts of the audience to some *linguistic* item associated with *X* other than its customary extension, such as *X*'s pronunciation, *X*'s meaning, or *X* itself.) Thus, for Saka mentioning is a pragmatic phenomenon, being something that speakers can do with their languages. It is not a part of the language itself, as a speaker can choose to refer to his or her very own words regardless of what language the speaker is using, be it natural or artificial.

This natural pragmatic ability is harnessed by quotation, a language-variable convention. In English, the convention is to flank a mentioned expression with inverted commas. Thus, when a speaker *S* utters a sentence consisting of elements $\langle a \text{ “}b\text{” } c \rangle$, *S* semantically signals that *S* uses *a* and then mentions *b* and then uses *c* (Notice that even when Saka moves from discussing the pragmatics of mention to discussing the “semantics” of quotation, his framework remains oriented toward speech acts, and in that sense is thoroughly pragmatic.).

Nellie Wieland

An indirect report typically takes the form of a speaker using the locution “said that” to report on an earlier utterance. In what follows, Wieland introduces the principal philosophical and pragmatic points of interest in the study of indirect reports, including the extent to which context sensitivity affects the content of an indirect report, the constraints on the substitution of co-referential terms in reports, the extent of felicitous paraphrase and translation, the way in which indirect reports are opaque, and the use of indirect reports as pragmatic vehicles for other speech acts such as humor, insult, or irony. Throughout she develops several positions: (i) that a semantic analysis of indirect reports is insufficient, (ii) that the distinction between direct and indirect reports is not clear and that indirect reports are the predominate way of reporting while direct reports may be a para-linguistic variation on them, (iii) that most questions about the semantics and pragmatics of

indirect reports will rely on a full understanding of the nature of what is reported and how it gets reported, (iv) that an analysis of reporting requires the pragmatic tools of metarepresentation and a social, inter-personal understanding of relevance and shared knowledge.

Alessandro Capone (I.E.M.)

In this paper, Capone defends the idea that pragmatic intrusion is involved in “de se” constructions: the ego-concept being a component of the “de se” thought. He defends this idea from a number of objections. He explores the related notion of immunity to error through misidentification and claims that this too depends on pragmatic intrusion. He defends this view from obvious objections. He takes immunity to error through misidentification in “de se” thoughts to depend on the fact that the thinking subject makes an implicit use of the first-person pronominal and there is no question of attributing a referent to the pronominal, since the referent is given in the subject of the thought. In third-person “de se” attributions, some form of simulation can be used to reconstruct the thinking subject as using a form of the first-person pronominal. Immunity to error through misidentification is attributed to the thinking subject through simulation.

Alessandro Capone (Minimalism)

In this paper the author discusses a paper by Wedgwood in which he considers the possibility that Relevance Theory and Semantic Minimalism share at least some common resources. Capone maintains that the two theories have different aims and different orientations and that it might be fruitful to understand why Cappelen and Lepore stick to Semantic Minimalism despite the various objections leveled to their theory. He explores certain minimalist solutions along the lines of considerations by Michel Seymour, adopting Jaszczolt’s considerations on parsimony of levels of interpretation. He assumes that in some cases which are alleged to be cases of (radical) pragmatic intrusion logical forms contain certain variables which can be filled (or saturated) in context. As a final proposal, he broaches the idea that pragmatic enrichment at the level of the predication can be avoided by resorting to a more complex enrichment at the level of the subject. He resorts to ideas by Jaszczolt (specifically POL), to argue that parsimony considerations require that enrichments be made at the level of subjects, if possible, thus avoiding a less parsimonious view according to which both subject and predicates should be enriched (in the same utterance).

Igor Douven

In earlier work, Douven argued that pragmatic considerations similar to those that Grice has shown to pertain to assertability pertain also to the epistemic notion of acceptability; in addition to the pragmatics of assertion, there is what he called

“the pragmatics of belief”. Specifically, just as a patently true sentence may be unassertable on the grounds that by asserting it one might easily mislead one’s audience, a patently true sentence may be unacceptable on the grounds that by accepting it one might easily mislead a future self who retrieves the accepted sentence from retentive memory. In this paper, the pragmatics of belief is applied to a puzzle related to the so-called Lottery Paradox. According to virtually all popular solutions to the Lottery Paradox, so-called lottery propositions—propositions to the effect that a ticket in a given lottery will lose—are unacceptable. The puzzle then arises from the fact that, according to all standard conceptions of what our epistemic goal is, we ought to aim at holding a body of beliefs with a favorable truth-falsity ratio. For, accepting of all minus one tickets in a given lottery that they will lose seems a safe way to improve the truth-falsity ratio of one’s body of belief (while maintaining a consistent belief set). The present paper solves this puzzle by arguing that the standard conceptions of our epistemic goal should be reformulated in light of the pragmatics of belief and by further arguing that accepting a lottery proposition is wrong on account of the broadly Gricean considerations that were said to pertain to acceptability.

Alberto Voltolini

In this paper the author wants to hold that contextualism—the position according to which wide context, i.e., the concrete situation of discourse, may well have the semantic role of assigning truth-conditions to sentences—may well accommodate (along with some nowadays established theses about the semantics of proper names) three data about fiction, namely, the facts that as far as discourse involving fiction is concerned, (i) sentences about nothing are meaningful (ii) they may be true in fiction (iii) yet they may also be true in reality. Moreover, Voltolini also wants to hold that such a contextualist accommodation does not depend on endorsing a realist stance about fictional entities, although such a stance gives the best truth-conditional explanation for such data, in particular (iii).

Alec McHoul

In “Pragmatics and philosophy: Three notes in search of a footing”, Alec McHoul turns to a seminal essay by Gilbert Ryle. Here Ryle explores the possible distinctions between ordinary everyday knowledge and recondite philosophical knowledge. He uses a number of metaphors in order to achieve this including those of “morning” versus “afternoon” types of questions, and the distinction between the villager’s and the cartographer’s knowledge of the locale. The present chapter extends these metaphors by drawing on Thomas Kasulis’s distinction between the potter’s and the geologist’s knowledge of clay and uses this to reflect on lay as against professional knowledges of language. The idea of a “socio-logic” of ordinary talk, from Harvey Sacks, is alluded to as a way of seeing how such ordinary talk is grounded.

Luvell Anderson and Ernie Lepore

Slurs target race, nationality, religion, gender, sexual orientation, politics, immigrant status, line of work and many other demographics. They offend their targets—some more than others. Children who blurt out slurs are scolded regardless of their ignorance; but not uncommonly, targets appropriate slurs for themselves. The main questions of this paper are by which mechanism do slurs derogate, why some more forcefully than others, and how do targeted members succeed in mollifying some slurs?

A lacuna in the literature on slurs is the rarity with which their offenses are specified. We are told little more than that they derogate, belittle, disparage, or diminish, but never how. Explanatory options here are, however, limited. We know a lot about how words achieve efficacy; with that in mind, in what follows we will canvass alternatives in the hope of illuminating the offensive nature of slurs.

Anderson and Lepore's view, in brief, is that slurs are banished words, not because of any linguistic feature they exhibit or any content they carry, but rather because they are on a list of prohibited words—the higher up the list the more offensive their uses. This leaves a slew of questions: what determines whether a word is on or off the list, its position on the list; and why is it sometimes appropriate to flout its prohibition? The backdrop for this paper is the role of linguistics and philosophy of language in the investigation of slurs. Anderson and Lepore's answer is none.

Frans van Eemeren, Bart Garssen

In "Viewing the study of argumentation as normative pragmatics" van Eemeren and Garssen explain that the pragma-dialectical approach to argumentation involves at the same time a pragmatic and a critical treatment of argumentative discourse. Starting from the speech act dimension of pragmatics, they indicate for each of the five components of the pragma-dialectical research program what a normative pragmatic approach amounts to. First they discuss the critical conception of reasonableness developed in the philosophical component of the research program, second the model of a critical discussion designed in the theoretical component, third the resolution-oriented analysis of argumentative discourse carried out in the reconstruction component, fourth the qualitative and quantitative research of the pursuit of reasonableness and effectiveness in argumentative discourse conducted in the empirical component, and fifth the development of adequate instruments for the analysis, evaluation and production of argumentative discourse in all sorts of argumentative practices undertaken in the practical component.

Francesca Piazza

The aim of this paper is to show that classical rhetoric can provide valuable insights in the contemporary debate in pragmatics. This is especially true for Aristotelian rhetoric, due to its philosophical approach. In the first part of the paper, the author discusses the conditions under which ancient rhetoric can be a real partner of current pragmatics: (1) rhetoric must be understood as a type of knowledge (a *techne*) and not as a “jumbles of techniques”; (2) we need to consider persuasion as an anthropological feature and not only as a specific case of communication; (3) we should not exclude truth from the rhetorical field. The second part of the paper focuses on what can be considered the basic insight of classical rhetoric (stated in a well-known passage of Aristotle’s *Rhetoric*, 1358a 37-b1): *speakers and listeners are inside and not outside discourse*. If adequately interpreted, this statement has important consequences that can be interesting for current pragmatics. These consequences can be schematically summarized as follows: (1) a broader conception of persuasion, useful to overcome the opposition informative/persuasive; (2) the consideration of speaker and listener as internal component of discourse (with the latter in a key position); (3) the consideration of the stylistic elements of a discourse from a cognitive point of view; (4) the overcoming of the sharp opposition cognitive/emotional due to the consideration of *ethos* and *pathos* as discursive elements.

Marcelo Dascal

Although we tend to consider controversy, discussion, dispute, and other forms of debating as typically communicative exchanges in which we confront others, we often engage in similar dialectical interactions within ourselves. As pointed out by Michel Foucault, while “we all struggle against each other, there is always something in us which struggles against something else in us”. What this paper inquires is to what extent, if at all, these two kinds of struggle obey the same pragmatic rules and strategies. When deliberating whether to complete a task we are committed to end today or to watch a football match, does our self somehow break up into parts that argue in order to persuade each other to adopt the opposite part’s position? And if we have decided for the first option but end up performing the second, is such a flagrant violation of our own decision not only a case of uncontrollable weakness of the will, but also of the irrationality of a self whose integrity is supposed to preserve consistency? The investigation of questions such as these from a pragmatic perspective reveals that, in addition to the habitual use of language in communicative debating, we must take into account quite different dialectic practices involved in our self-debates, as well as the ontological and epistemological consequences of such differences. Philosophically, the acknowledgement of the distinctions between what I have suggested to call, respectively,

“socio-pragmatics”, “psycho-pragmatics”, and “onto-pragmatics”, along with the identification of different kinds of rationality, provides new insights into some of the most puzzling problems of the philosophy of mind, such as weakness of the will and alleged irrationality.

Franco Lo Piparo

True’ and ‘False’ are defined through a linguistic rule requiring the negation operator. This is the elaboration of an idea proffered for the first time by the Stoics on the basis of some remarks by Aristotle and then in modern times by Frege and Wittgenstein. Another thesis of this essay is the following: the true/false rule is a sort of UR-Regel underlying all linguistic practices (including prayers and commands) and all human cultures. Reinterpreting the notion of Spielraum put forward by Wittgenstein in 4.463 of the *Tractatus*, Lo Piparo will present an implicational pragmatic theory of a true proposition. Jokes and *reductio ad absurdum* are explained as examples of Spielraum.

Alessandra Pandolfo

Although Habermas’ universal pragmatics has played a marginal role in studies on pragmatics, it can still make an important and meaningful contribution, precisely because it highlights the system of validity claims that lie in speech acts. This type of analysis allows one to consider the dialogic dynamics that engage speakers in the activity of reciprocal giving and asking for reasons for saying and doing things. The reasons why a speaker knows he or she can say what he or she says in the presence of other speakers constitute an essential element of the production of meaning; and reciprocally identification of the speaker’s reasons by the listener is an indispensable condition of the activity of understanding of meaning. These factors allow one to render explicit the dialog dynamics on which the quality of the utterance and understanding, like the quality of social relationships, depend.

Felice Cimatti

Embodied Cognition (EC) is a new psychological version of an old philosophical idea: human cognition is grounded in sensorimotor experience. According to EC there is not such an entity as abstract and disembodied knowledge, that is, the root of every form of human knowledge is an acting body in the world. In this paper Cimatti will try to show that existing extensions of EC to language partly miss the point because do not fully account for the social and performative nature of language. Therefore a thorough embodied theory of language requires to consider the Wittgenstein legacy, which stresses at least two

main points: (a) a coherent theory of language is not possible if not embedded in a more comprehensive description of human way of living; (b) the meaning of a word is not an internal and psychological entity but its social use, it is the *action* we do using that word/tool. In this paper Cimatti will analyze EC literature showing that it needs to be complemented with Wittgenstein ideas on language and mind.

Conversation and Collective Belief

Margaret Gilbert and Maura Priest

Abstract This article proposes that paradigmatic conversations involve the development of a *collective cognitive profile* of the parties. This occurs through the negotiation of a series of *collective beliefs*. Collective beliefs are constituted by commitments that are *joint* in a sense that is explained. The parties to any joint commitment have associated rights and obligations. This helps to entrench a given collective belief once established. Even when interlocutors do not manage to negotiate a collective belief whose content has explicitly been specified, they are likely to establish one or more associated *implicit* collective beliefs. This supports the idea of a conversation as a collective activity whose stages are marked by the development of a relatively stable collective cognitive profile of the parties. This idea is briefly related to some of the existing literature on conversation, including classic articles by Stalnaker and Lewis on presupposition and conversational score.

Keywords Collective belief · Conversation · Conversational score · Common ground · Groups · Individualism · Negotiation · Joint commitment · Presupposition · Obligation · Rebuke

This article is dedicated to the memory of David Lewis, who suggested the connection between the idea of collective belief elaborated in Section I and his own and Robert Stalnaker's work on conversation. See section III in the text below. The authors thank the following people for reading and commenting on the penultimate draft at short notice: Alessandro Capone, Antonella Carassa, Marco Colombetti, Daniel Pilchman, Frederick Schmitt, Martin Schwab, Richmond Thomason and Philip Walsh. Responsibility for the ideas expressed here is ours alone.

M. Gilbert (✉) · M. Priest
University of California, Irvine, CA 92697–4555, USA
e-mail: margaret.gilbert@uci.edu

M. Priest
e-mail: mpriest@uci.edu

1 Introduction

This paper argues that *collective beliefs* as understood here play a central role in conversation. In particular, whatever else is going in the course of a paradigmatic conversation, the parties are negotiating the establishment of one or more collective beliefs on the basis of proposals made by one or another participant. In so arguing the paper draws on and develops in various ways an idea proposed by Margaret Gilbert some years ago.¹

The first part of the paper introduces the relevant idea of collective belief, associating it with *collective belief ascriptions*—a familiar type of everyday utterance. The second part discusses the thesis about conversation just described. The third part relates this thesis to several other ideas about conversation, including influential positions expressed in classic articles by Robert Stalnaker and David Lewis.

The discussion of collective belief in the first part brings together points on the topic from a body of previously published philosophical work by Margaret Gilbert, and is intended primarily for those readers who are unfamiliar with that work.² It not only articulates the particular concept of collective belief at issue here, but argues that there are good reasons for thinking that this is the concept expressed when everyday collective belief ascriptions are made. In other words, it is not a technical concept.

2 Collective Belief

2.1 How are Collective Belief Ascriptions to be Understood?

In everyday language, we speak about the beliefs of individuals and also about the beliefs of groups. For example, we say things like “Professor Smith believes that analytic philosophy is overly technical.” We also say things like “The Philosophy Department believes analytic philosophy is fine as it is.” The groups in question may be of a wide variety of kinds. Thus they may be informal, transient, and have no specific name or label, like the group constituted by a couple of strangers who have just begun a conversation.³ Participants in such cases, like those in the case of

¹ See Gilbert (1989: esp. 294–298).

² The initial presentation of these ideas in Gilbert (1987) and—at greater length—Gilbert (1989 ch. 5) has been followed by amplifications and clarifications in such articles and chapters as Gilbert (1994), (1996: Introduction), (2002), (2004), (2010a) (2010b).

³ Does such a conversational pair constitute a group as opposed to a mere plurality of persons? Perhaps not prior to the conversation; once the conversation has started, it would seem that the answer is affirmative. See Simmel (1971: 24). For lengthy exploration of the relatively substantial idea of a group that is in play here see Gilbert (1989: ch. 4; 2006: ch. 8). Groups of this kind differ among themselves, of course, in important ways. One intuitive starting point that will be assumed in what follows is Rousseau’s idea that a group in the sense in question is an *association* of some kind as opposed to a mere *aggregate* of persons.

more formally constituted, long-standing, named groups such as “the Philosophy Department,” may all speak of what “we” [sc. collectively, or as a group] believe.

Sometimes what looks like the ascription of a belief to a group or, as we shall put it, *a collective belief ascription*, may not be so intended at all. Thus someone who says on some occasion, of himself (or herself) and another person “We believe that Jane’s going to be late” might go on to clarify his meaning in some such way as this: “What I mean is, we all believe Jane’s going to be late” or “What I mean is, we both believe Jane’s going to be late.” In that case, the speaker disclaims any intent to ascribe a belief to “us” as a group. In more technical terms, he explicitly gives his utterance a distributive reading. From here on out, we set such cases aside.

Collective belief ascriptions may appear problematic. They apparently ascribe one or more beliefs *to a group*. But can groups have beliefs? One may be happy with the idea that individual members of a group have beliefs, but not with the idea that the group itself has beliefs. Nonetheless we often make such statements. What provokes them? That is, what do they refer to? In short, what are *collective beliefs*?

Before the account of collective belief that will be central to this paper is introduced, some of the considerations that support it will be explained.

2.2 Summative Accounts of Collective Beliefs

What may at first seem to be an attractive account of collective belief ascriptions is expressed by Anthony Quinton who writes:

We do, of course, speak freely of the mental properties and acts of a group in the way we do of individual people. Groups are said to have beliefs, emotions, and attitudes and to take decisions and make promises. But these ways of speaking are plainly metaphorical. To ascribe mental predicates to a group is always an indirect way of ascribing such predicates to its members. With such mental states as beliefs and attitudes the ascriptions are of what I have called a summative kind. To say that the industrial working class is determined to resist anti-trade union laws is to say that all or most industrial workers are so minded.⁴

The above account will be referred to as “summative” here, but the term will be used in a broader sense than Quinton’s: A *summative account* of collective beliefs is one according to which, for a group G to believe that p it is *logically necessary* that all or most members of G believe that p. The condition that *all or most members of G believe that p* will be referred to as *the summative condition*.

⁴ Quinton (1975-6: 17, see also p. 9).

Is the correct account of collective belief a summative one? Emile Durkheim, one of the founders of sociology, can be read as rejecting a summative account of collective beliefs. In *The Rules of Sociological Method* Durkheim expressed the view that anything properly called a collective belief will be “external to individual consciousnesses.” How should this phrase be interpreted? Textual evidence suggests that a collective belief is “external to individual consciousnesses” insofar as it is not necessary for *any* individual member of a group to believe that *p* in order for the group to believe that *p*.⁵

Could that possibly be correct? While this paper will not attempt further to discuss Durkheim, it will present considerations that support a positive answer. First, it argues *that fulfillment of the summative condition with respect to some proposition p is neither necessary nor sufficient for the collective belief that p*. It starts with the issue of *sufficiency*.

2.3 *The Insufficiency of the Summative Condition*

According to the type of summative account Quinton envisages:

A group *G* believes that *p*, logically speaking, if and only if all or most members of *G* believe that *p*.

This is perhaps the first account of collective belief that is likely to come to mind. Call it *the simple summative account*.

In connection with this account, suppose that every member of the marching band believes that driving at speeds of over 50 mph is dangerous. But no member has ever discussed this belief of his with any other member. On the contrary, in conversation with others most are inclined to deny they have it for fear of mockery. In this case, despite every group member holding this belief we would not be inclined to ascribe it to the marching band.

It seems, then, that fulfillment of the summative condition is not sufficient for collective belief. The simple summative account of collective belief must, then, be rejected, for it entails such sufficiency.

Perhaps something very close to the summative condition is sufficient for collective belief. The marching band case illustrates the point that individual beliefs may be kept private. Do we find that the *marching band* does not believe that driving at speeds over 50 mph is dangerous *only* because the band members’ beliefs are private? It seems not: even if every member of the band knows about the belief of every other member this does not seem enough: each member may think he is the only one with this piece of knowledge.

⁵ See Gilbert (1989: ch. 5) on this point and for further discussion of Durkheim (1895) on collective belief.

Perhaps the summative account can be saved by adding a condition of *common knowledge* of the kind familiar from the work of David Lewis (1969), Stephen Schiffer (1972), and others. According to one popular account of common knowledge: it is common knowledge in G that p if and only if (a) p; (b) everyone in G knows that p; (c) everyone in G knows that (b), and so on, *ad infinitum*. There is no need to consider whether or not this is the best account on offer for present purposes. Certainly the “ad infinitum” clause in this account has given some pause.⁶ The reader may substitute his or her preferred account in considering the following *complex summative account* of group belief:

A group G believes that p, logically speaking, if and only if (1) all most of the members of G believe that p, and (2) it is common knowledge in G that (1).

Though it may appear more promising than the simple summative account, the complex account faces a difficulty it shares with the former account. At its core lies a fact about what all or most individual members of G *personally believe*. That is what the common knowledge, in this case, concerns. A collective belief, however, is supposedly a belief *of the group*. Intuitively, that each member of the group personally believes p does not make the belief that p a belief of the group, and the addition of common knowledge as in the complex account does not change this. The group plays no essential role in such a set-up.

An example may help to back this up. Suppose—as seems likely—that that every member of a given philosophy department believes that ants are smaller than elephants. Further suppose that it is common knowledge in the department that each member believes this. Despite each member of the department so believing, and despite the common knowledge of this belief amongst all members, we would not be inclined to say “The Philosophy Department believes that ants are smaller than elephants”. At least, we would not be so inclined without more information about the case.

The Philosophy Department is already a group. One might think however that were certain parties collectively to believe something, that would make them into a group. If this is right, it too casts doubt on both of the summative accounts so far considered. Suppose—as seems likely—that all females 59.7 inches high believe that ants are smaller than elephants. Intuitively, this does not raise the population consisting of such females to the level of a group. Nor does the supposition that it is common knowledge among them that they all have the belief in question have this effect.

To help sharpen the point that the conditions posited in neither of the summative accounts considered are sufficient for collective belief, consider the following example. Imagine that the same fourteen individuals make up both the philosophy club and the humanities soft ball team. Most of these individuals believe it is always best to run on a fly ball; this belief is common knowledge among them. Let us assume that the humanities softball team believes that it is

⁶ See Vandershraaf and Sillari (2009) for an overview and comparison of a variety of accounts, some highly technical. The account in Gilbert (1989: ch. 4), introduced in the course of discussion of the nature of acting together, is one of the accounts discussed.

always best to run on a fly ball, as well it might. There is no contradiction in also saying that the *philosophy club* holds no opinion whatsoever on the matter. Indeed, this is very likely true.

It seems, then, that for a group to believe that p it is not logically sufficient either that all or most group members believe that p , or that there is common knowledge within the group that all or most members believe that p . A further argument for this conclusion will be offered shortly.⁷

2.4 *The Non-Necessity of the Summative Condition*

Imagine that the University Event committee is meeting to discuss the celebration of the University's 50th anniversary. Phyllis, a member of the committee, suggests, "J.K. Rowling would make a great keynote speaker." Some discussion ensues, with positive comments outnumbering negative ones. After a while someone asks, "So we agree that Rowling would make great speaker?" This time the members of the group either voice their agreement or remain silent. The matter is settled: Rowling would make a great speaker.

At this point it would be natural and appropriate for any member of the committee to make any of the following collective belief ascriptions:

"We agree that Rowling would be a great speaker"

"We think that..."

"In our opinion..."

"We decided that.."

"Our view at this point is that..."

Indeed, it seems that all of these statements are true.

It has already been argued that the summative condition is not sufficient for collective belief. It will now be argued that it is not necessary.

Though the Event committee now believes that Rowling would be a great speaker, it is possible that most members of the committee do not personally believe what the group believes. There are various possibilities here, including the following.

Perhaps most of the members of the Event Committee thought at the time of the discussion that Rowling would be a poor speaker. Perhaps they did not speak up during the discussion out of fear of disapproval from certain other members. Thus they let the belief that Rowling would be a great speaker become established as the

⁷ Our argument so far leave it open whether a summative account of collective belief other than those considered here may offer conditions *sufficient* for collective belief. That said, an account that offers both necessary and sufficient conditions is preferable. This, as we argue in the text below, will have to be a non-summative account. See Gilbert (1989: ch. 5) for discussion of a summative account that is more complex than those considered here. Gilbert argues that that account also fails to offer conditions sufficient for collective belief.

group's belief. Having done so, they did not then change their own minds on the topic.

Or perhaps most of the members had no opinion on the matter. They may never have heard of Rowling or did not feel they knew enough about her to form any *personal opinion* about her speaking skills. This does not prevent the Event Committee *as a group* from forming the pertinent belief about Rowling's speaking skills.

Given that these points are accepted, we must infer that the summative condition is not a necessary condition for the existence of a collective belief. Thus no form of summative account of collective belief—all forms of which, by definition, include the summative condition—can be correct as a general account. By a “general account”, here, we mean an account that purports to provide a set of conditions that are both individually necessary and jointly sufficient for collective belief.

One can easily go further here, and in so doing back up a point that was suggested earlier to be Durkheim's. That is, one can argue that not only is it not necessary to the existence of a collective belief that p that all or most members of the group in question believe that p , *it is not necessary that any members of the group personally believe that p in order that the group believes that p .*

To make this more concrete, consider again the Event committee. One might think Phyllis, who initially said that Rowling was a great speaker, must have believed this. But of course she need not have. Perhaps Phyllis believed Rowling was a terrible speaker and, having some grudge against her, wanted to see her make a fool of herself in front of the group. Evidently, in principle at least, the whole Event committee scenario could have transpired without a single member of the committee personally believing what the committee came to believe about Rowling. Nor need anyone change his or her personal position after the collective belief was formed.

It is perhaps worth emphasizing that the present point is not that such a radical disjunction between collective belief and the personal beliefs of members is *common*. Whether or not that is so is ultimately an empirical matter. The point is that such a radical disjunction is *possible*.

2.5 Collective Belief and Corresponding Rights and Obligations

Now consider a possible continuation of the Event committee example. The committee moves on to other matters. Eventually, the meeting begins to wind down. The group is reviewing their accomplishments of the day when one member unexpectedly speaks up, “J.K. Rowling will be an awful speaker.” In a tone of rebuke, another member objects: “But we thought she would be great!”

This brings to light an important aspect of collective belief: when one group member speaks against an established group belief, he is liable to be met with a negative response from the other members, such as “But...!” in the above dialogue. This response is directed specifically at the speaker’s contradiction of the collective belief as such. Thus the person who responds justifies his “But...” by saying “We thought”.⁸ As is also indicated in the above dialogue, the response in such situations may take the form of a *rebuke*. What is it to rebuke someone? Though it implies that person has done something wrong, it does more than that. One person may say to another, “You just did something wrong. Thank goodness you did! You were beginning to seem too perfect.” In this case the speaker tells someone he did something wrong without rebuking him. What, then, distinguishes telling someone of his error from rebuking him or, in other terms, *telling him off*?

A rebuke is a verbal form of punitive behavior.⁹ One needs a special standing or authority to punish a person—or to rebuke him.¹⁰ In what does that standing consist?

Pursuing this question, let us return to the example. It seems that the dissenting group member may appropriately respond with an apology, perhaps accompanied by an excuse. He might reply, “Oh, of course, you’re right. I’m sorry. We agreed that Rowling would be a great speaker. I was thinking of someone else.” His responding apologetically suggests that his comment was more than a simple error; it was an *offense* against his interlocutor, qua group member. His apology also suggests the aptness of a group member’s *demanding* that he not speak against the group belief, should it seem that he was about to do so. As with a rebuke, a demand of the kind in question requires a special standing.

All of this suggests, again, that the group members as such were *entitled* to the absence of the action that caused the offense—that a *right* of theirs been violated by the offender. The offender had a correlative obligation *to them* not to perform the offending action. This is what puts the other members in a position to *rebuke* him for performing it.¹¹ Because a right of theirs was violated, the group members, as such, may engage in this form of behavior. They will thereby enforce their right, call the offender to account, and put him on notice that any similar behavior may provoke a similarly punitive reaction.

What is it about a collective belief that gives the parties the standing to rebuke one another for gainsaying the proposition in question? How does a collective belief generate rights that, when violated, justify verbally punitive measures?

⁸ For further discussion on this point see Gilbert (1987).

⁹ Cf. Hart (1961).

¹⁰ Such “authority-presupposing” terms often have broader senses, which can lead to confusion. See Gilbert (2006: ch. 1).

¹¹ Cf. Hart (1961). Note that the obligations here are obligations *to* someone to do a certain thing. Such obligations are generally referred to as “directed” obligations. It is this type of obligation that correlates with rights of the type at issue here, referred to as “claim-rights” in rights theory.

A positive answer is proposed in the next sub-section. First we revisit the complex summative account of belief with this question in mind.

Consider a situation in which the conditions of the complex summative account are satisfied: most members of a breakfast club believe that getting up at 4:30 am is unpleasant, and this is common knowledge amongst them, qua members of the club. One day a club member, Sue, says to another member, Brian, “My new job requires me to get up at 4:30 am every day. That’s fine with me. I love being an early riser.” Is this enough to justify an offended rebuke on Brian’s part, all else being equal? It does not seem so. Brian seems not to be in a position to demand that she say anything else, or to rebuke her for saying what she says. Of course, Brian may be *surprised* by Sue’s statement—but to be surprised is not in itself to have been offended against. In expressing a minority view Sue is doing something *unusual* but this does not mean she violates anyone’s right.

This is another argument against the sufficiency of the conditions of the complex summative account of collective belief (and, a fortiori, against the sufficiency of the summative condition). As was observed earlier, the simple fact that *we believe* certain things appears to put any one member of the pertinent group in a position to rebuke a group member who says something contrary to what we believe. The Sue and Brian example shows that even when the conditions of the complex summative account are fulfilled that does not suffice to put group members in that position.

There is a non-summative account of collective belief that better accords with all the observations made so far concerning everyday collective belief ascriptions. To invoke a technical term that will be explained shortly, it is a *plural subject* account of collective belief.

2.6 A Plural Subject Account of Collective Belief

In everyday speech one hears frequent references not only to what we *believe*, but also to what we are doing, what we like, what we prefer, what we intend, and so on. In earlier work Margaret Gilbert has argued that all of these expressions refer to those who are connected in a specified manner. To use a phrase that will be explained shortly, they must be *jointly committed* in a certain way. The following discussion focuses on the *basic case* of joint commitment in which no prior authority relations are involved. This is the case most apposite for present purposes.¹²

What does it take for there to be a joint commitment? Roughly: in conditions of common knowledge, each party must have openly expressed his or her willingness

¹² On basic versus non-basic cases see e.g. Gilbert (2006: ch. 7).

jointly to commit them all to believe some proposition, accept some goal, endorse some value, and so on, as a body, thus effecting such a *joint commitment* of them all.

Gilbert has used her technical phrase “plural subject” to refer to those who are jointly committed with one another in some way: these people constitute a plural subject in her sense. The key to plural subject-hood, then, is joint commitment.

According to *the plural subject account* of collective belief, for persons A and B. A and B collectively believe that p if and only if A and B form a plural subject of the belief that p.

Alternatively and equivalently (and more perspicuously):

A and B collectively believe that p if and only if A and B are jointly committed to believe that p as a body.

Though for present purposes this type of two person case will be our focus, there would seem to be no limit in principle on how many persons can be involved.

Some further amplificatory remarks are in order.¹³ Those who are jointly committed need not know each other personally or know of each other. Thus A and B may collectively believe that p *qua persons who share a certain feature*, such as “being part of the crowd in the square” or “being members of Indiana University”.

The notion of *commitment* at issue here is a *normative* one: to be committed in the sense at issue here is to be subject to a type of normative constraint as opposed to being in a particular psychological state.¹⁴ A joint commitment falls into a class of commitments that may plausibly be thought of as *commitments of the will*.

A joint commitment as understood here is not an aggregate of independent personal commitments such as one accrues by virtue of making a personal decision. By definition here, a *personal commitment* is created by the person in question and unilaterally rescindable by him. If Sarah and Nick are jointly committed in some way they have *together committed the two of them*.

As a result each of them is, of course, *committed*, but these commitments are not personal commitments as just defined. What we may call the *associated individual commitments* of Sarah and Nick are such that neither one holds sway over his or her commitment as one does over the personal commitment engendered by one’s personal decision: if I want to change a personal decision of mine it is as simple as changing my own mind. But a joint commitment to which I am subject is not mine to change. It is *ours*, rather, and only *we* together can rescind it. Hence each is stuck with his (or her) individual commitment till the joint commitment is rescinded by both.

¹³ For extended discussion see Gilbert (2006: ch. 7); also Gilbert (2003).

¹⁴ See Gilbert (2013).

What is it for people jointly to commit one another to believe some proposition *as a body*—the particular form of joint commitment at issue in the plural subject account of collective belief? One way of explaining this is roughly as follows: the parties will fulfill their joint commitment if—at least in the context of the present interaction, and to the extent that this is possible, they emulate, by virtue of the several actions and utterances of all combined, a single body that believes that *p*.¹⁵ In more familiar terms, they are to act as if they are “of one mind” with respect to *p*.

In this connection it is important to emphasize that in order that the parties conform to their joint commitment to believe that *p* as a body *it is not necessary for each individual personally to believe that p*. Thus no one is committed through the joint commitment to himself personally believing that *p*. That is perhaps just as well since many philosophers have doubted whether one can believe something at will.

What difference does it make to a person, normatively speaking, that he is subject to a joint commitment? On the one hand, he is subject to a commitment of the general type a personal decision subjects him to. Though this matter is controversial, it can be argued that one who is subject to any such a commitment *has reason* to act accordingly, where that means something like this: all else being equal, he must act accordingly if he is to respond appropriately to the considerations bearing on his situation.¹⁶ Indeed, he must ignore his own inclinations, at least, in face of an unrescinded commitment of this type—a type that may be referred to as a *commitment of the will*.¹⁷

A *joint* commitment, however, does more than commit the individual parties in corresponding ways; *it establishes an important, many-faceted relationship between the parties*. Discussion of this here must be relatively brief.¹⁸

Consider the situation in which one party fails to conform to a joint commitment. It seems that, intuitively, the other party or parties have the standing to rebuke him for this failure. Again, if one party threatens not to conform, the other party is in a position to demand that he conform after all. As indicated in our previous discussion of the continuation of the Event Committee case, this suggests that *any one party to a joint commitment has a right to any other's conformity to the commitment, and that each party is correspondingly obligated to each party to conform*. That the parties to a joint commitment have these rights and obligations will be assumed in what follows.

¹⁵ Note that the idea is not that they are to emulate a human collectivity with the belief in question. To suppose that it is is to see the joint commitment account of collective belief as circular. “Single person” is an alternative that would presumably avoid any such thought. See also the alternative offered in the text.

¹⁶ As understood here, to *have reason* to do something is not necessarily to have *a reason* to do it, where having a reason implies that some good will come of one's doing it. See Gilbert (2006: ch 2).

¹⁷ For further discussion see e.g. Gilbert (2006: ch. 2).

¹⁸ See e.g. Gilbert (2006: ch. 7) for further discussion.

Let us go back now to the plural subject account of collective belief. Given this account, if Max and Tom form a plural subject of believing that *p*, *each of them is obligated to the other* to do his or part in believing that *p* as a body. Should Max or Tom not act accordingly, they default on an obligation to the other, and the other's corresponding right to his performance has been violated. The other is then in a position to pressure him in the authoritative manner of a rebuke. Both Max and Tom will understand this. Forming a plural subject of believing that *p* gives Tom and Max *a basis for rebuking one another should the appropriate behavior not occur*.

We can now offer the following interpretation of what went on in the Event Committee. That committee forms a plural subject that is jointly committed to believe as a body that Rowling will be a great speaker. This obligates all members of such plural subject to act in certain ways. In particular, it obligates them to act as though each is the mouthpiece of a single believer of the proposition that J.K. Rowling would be a great speaker. When one group member violates this obligation, the other members may appropriately rebuke him. The group members had rights one against the other to his acting as though Rowling is a great speaker and the dissenting member violated that right. In this way he offended against them.

2.7 Concluding Remarks on Collective Belief

In this section of the paper we have outlined an account of collective belief in terms of a joint commitment to believe that such-and-such as a body. This plural subject account appears to fit quite well the phenomenon referred to when people say things like “The Philosophy department believes that it is too small to be viable” or “We [sc. collectively] think that she will come tonight”. Among other things, it does not entail that any of the people in question personally believe what the group believes. This accords with observations made in discussion of the Event committee.

Given what has been argued so far in this paper, the plural subject account may well be correct as a general account of the phenomenon associated with everyday collective belief ascriptions.¹⁹ Rather than pursuing that question further here, we discuss a thesis that links the dynamics of everyday conversation to the phenomenon of collective belief *according to the plural subject account*. If that thesis is correct, then whatever its relation to locutions common in everyday speech, the plural subject account of collective belief describes a phenomenon that cannot be ignored by theorists of conversation.

¹⁹ Since Gilbert proposed this account of collective belief, a literature has grown up concerning whether collective belief in her sense is *belief* or rather some other cognitive state. There is no need to go into this discussion for present purposes. For references to this literature and for discussion of the issue, see Gilbert (2002).

3 The Negotiation of Collective Belief Thesis

3.1 Preliminary Remarks

The Event Committee case strongly suggests a particular thesis relating the dynamics of everyday conversation to collective belief. This section aims to introduce that thesis and explore some of its ramifications. It does not aim to present anything like a full treatment.

We shall not attempt to characterize conversations as such. We operate, rather, with an intuitive notion. We focus on verbal interchanges between two or more people that appear otherwise to have the character of conversations and that contain, *at a minimum*, an assertion made by one of the parties. The thesis at issue here can most easily be expounded in relation to assertions. The class of verbal interchanges just described may contain all *paradigmatic* conversations. To say this is not to say that the thesis we shall discuss has no relevance to other classes of conversation as well. On the contrary, it surely has such relevance, though we shall not explore that here.²⁰

Evidently, people engage in conversation for a variety of collective or individual purposes, such as becoming better informed on some topic, making plans, shooting the breeze, or increasing mutual respect. We shall not pronounce on whether or not there is some purpose, collective or individual, that characterizes all conversationalists as such. We do not see the thesis we discuss as dependent on any particular assumption in this regard. We do suppose that the conversations at issue are genuine, as opposed to pretend conversations such as one observes on the stage at the theater.

We shall not discuss whether conversationalists as such must collectively believe that they are having a conversation, though we take that to be a plausible idea. We are concerned with collective beliefs that emerge in the context of any such “framework” collective beliefs.

Most likely there is more than one class of collective belief that so emerges. Our focus is one of these classes of collective belief. To give a rough advance explanation of what that class is: we are concerned with collective beliefs such that the proposition in question—the proposition that is collectively believed—has either been verbally expressed by one of the parties or implied by such verbal expression in the course of the conversation.²¹

²⁰ Gilbert (1989: 296) briefly relates questions and orders to the thesis.

²¹ Drawing on Gilbert’s account of collective belief and her notion of joint commitment, Carassa and Colombetti (2009a) propose that, centrally, conversations involve a negotiation of *joint meaning*. Joint meaning as they understand this is a collective belief to the effect that, roughly, the speaker intended to perform a particular speech act. For instance, he was simply making a statement, or he was both doing this and issuing an invitation. We take their discussion to be complementary to that in Gilbert (1989) and the present paper. See Carassa and Colombetti (2009a), (2009b) and references therein for related discussions in which the authors characterize

3.2 *The NCB Thesis*

In the Event committee case a collective belief came about through a discussion among committee members of the merits of a particular proposal. Because of its relatively formal context, this discussion may not count as a paradigmatic conversation. That said, it strongly suggests the thesis about such conversations that we discuss in this section. This is the thesis that the process of everyday conversation is structured around collective belief formation.²² More specifically, in a paradigmatic conversation, whatever else is going on, *the parties are negotiating the establishment of one or another collective belief on the basis of proposals put forward by one or another interlocutor*. We shall call this the *negotiation of collective belief* thesis—more briefly the *NCB thesis*.²³

In the rest of this section we give some examples of how the NCB thesis construes a variety of example conversations and conversational moves, to some extent elaborating the thesis further while doing so. In all of these cases we assume that there are no special background collective beliefs or other factors that might skew one's interpretation of what is going on.

Consider two individuals, call them Jenny and Tom, waiting together at a bus stop on a sunny day. The two fall into conversation. "Sure is a beautiful day!" says Tom cheerfully. Jenny, whose actual preference is cool cloudy weather, replies, "I suppose it is." She speaks pleasantly, but without great conviction. According to the NCB thesis what happened, among other things, was this.

First, Tom proposed for collective belief the proposition that the day in question is a beautiful day. In other terms, he made a particular *collective belief proposal*. More precisely he makes what we shall refer to as an *explicit* collective belief proposal. That is, he proposes that the parties collectively believe that p by uttering a sentence that expresses proposition p.

(Footnote 21 continued)

the conversational process in terms of Gilbert's idea of joint commitment. Another important discussion that suggests an "emerging" collective belief of a particular type (though it is not couched in these terms) is Brennan and Clark (1996) on what the authors refer to as "conceptual pacts". Here the envisaged pact or agreement [which we would construe as a matter of joint commitment as in Gilbert (2006: ch. 10)] relates to *the way in which each party is to conceptualize something*. In fact one might argue that insofar as conversationalists come to make various conceptual pacts these relate to how we are *collectively* to understand certain terms, and so on, as in "What's a vixen?" "It's a female fox" "Ah". The parties now collectively believe that a vixen is a female fox (whatever they personally believe). Fixing their collective belief about what a vixen is will be the primary need of the conversational process—given that vixens are the topic of conversation. We regret that there is no space in the present article for further consideration of these and related topics and publications, though see the note on Clark (1996) in the text and notes below.

²² Here and in other statements of the thesis we refer to collective beliefs in the class roughly indicated earlier in this section.

²³ See Gilbert (1989: 294–298) and the rest of Chap. 5 of that work for further pertinent discussion.

Second, Jenny considered Tom's proposal and in spite of her personal belief to the contrary decided to go along with Tom and believe the proposition in question collectively with him, indicating that she was ready to do this by her reply. In other terms, *she accepted his collective belief proposal*.

Thus, through their conversation, Jenny and Tom together established a collective belief that would persist at least through their current interaction, all else being equal.²⁴ We shall refer to collective beliefs that are established through the making and accepting of an explicit collective belief proposal as *explicit collective beliefs*.

In this example, Jenny does not personally believe what the two collectively believe, but this does not matter. Our collectively believing that p does not entail that I personally believe that p nor does it entail that I do not personally disbelieve it. The same goes for the other person or people involved.²⁵

One can imagine all kinds of variations on the first example conversation. For instance:

Tom: "Sure is a beautiful day!"

Jennie (in a tone of doubt): "Mmm"

Tom: "But...it is so bright and sunny!"

Jennie: "You're right: it really is a beautiful day."

Tom has to work harder here than in the previous example to establish the collective belief that it is a beautiful day.

What Tom does in his second move in the negotiation is *give a reason* for believing the proposition originally proposed for collective belief. Many negotiations over a given collective belief may be expected to involve such reason-giving. This may well lead to further collective beliefs, as well as the one originally proposed, to the effect that such-and-such a proposition is true *and that its truth is a reason for believing the original proposition*.

One might be inclined think that this has happened once Jennie responds as she does in the above dialogue. That it has happened, however, would be clearer had Jennie responded, as she might very well have responded, in some such way as this.

Jennie: "True! I guess it has to be a beautiful day, then!"

Here she indicates by "True!" and "..., then!" that she accepts the following propositions that Tom has proposed for collective belief. First: it is very bright and sunny. Second: If it is very bright and sunny, then it is a beautiful day" It is true that Tom does not spell out the second proposition, as he does the first. That he is indeed proposing it for collective belief, however, this is clearly indicated by his beginning with "But..."

²⁴ In our discussion in this paper we shall not address the ways in which established collective beliefs can be amended. Suffice it to say that such amendment requires the concurrence of the parties.

²⁵ See section I above.

As the first example conversation shows, “negotiating” a collective belief can be as simple as making an assertion and receiving an affirmative response. More precisely, in conditions of common knowledge you may express to your interlocutor your readiness jointly to commit the two of you to believing some proposition as a body, and he may immediately do likewise in response, thus ensuring that the two of you are jointly committed to believe that proposition as a body. In other words, he may promptly seal the deal.

3.3 Ways of Closing the Deal: Accepting a Collective Belief Proposal

There are many ways in which one’s acceptance of a proposed collective belief might be shown. It need not involve verbal behavior. A nod of the head, for instance, might suffice, as might a smile, in the right context. If words are used a simple “Yes” may suffice, depending on its tone. Similarly, the expression “Uh-huh”, said in the right way, is a form of acceptance. Something like this is presumably the most economical way to make sure one’s acceptance is verbally marked. A hesitant “Ye-es” or a quizzical “Uh-huh” may lead to further discussion, and be intended so to lead. Or it may serve to indicate that the proposition in question is *not* yet collectively believed, though it has not been pushed off the table. In such a case a participant may describe the situation to one who asks what they collectively think by saying: “We’ve not yet decided *what* we think about that”.

In a given context silence may rightly be taken as acceptance. For instance, suppose two friends, Dan and Debbie, are out on a walk, and are engaged in conversation. Debbie is doing most of the talking and it is clear that Dan is focusing closely on what she says. She may preface some of what she says with locutions like “We both know that...”, “It is obvious that ...” and “Everyone agrees that...” Such statements, which implicitly ascribe a particular view to the hearer, seem especially likely to move the hearer to voice an objection if he takes them to be false, so it may be especially reasonable to take his silence as acceptance in this context.

It is not clear that silence should always be taken as acceptance, even given that one knows one has been heard. One who maintains a “stony silence” in the face of someone else’s assertion is likely to be thought to be refusing to negotiate any collective belief rather than to be accepting the other’s proposal for such a belief, or offering a counter-proposal.

Nonetheless, in order to avoid misunderstandings, it is probably best to make an explicitly rejecting move such as saying “I can’t speak to that”, or the more sweeping “I don’t want to talk to you”, or physically turning away. If something like this is done, the proposal of a particular collective belief has gone nowhere, at least for the time being. A rather special kind of rejecting move will be mentioned shortly.

In some contexts it may be mutually agreed that a given person's silence counts as acceptance. Thus suppose someone has been rendered temporarily or permanently incapable of speech and almost all movement, but has the ability to understand his friend's speech and tap the table with his own finger, though with great difficulty. An understanding could be established between them that, roughly, as long as he does not tap the table with his finger, he is in agreement with what is said by his friend. These two are now in a position to establish a series of collective beliefs through a series of assertions followed by silence.

3.4 Ways of Rejecting a Collective Belief Proposal

Collective belief proposals can be *rejected* in various ways such as a refusal to negotiate (as with a stony silence), or a vocal rejection of the proposal. We say something about moves of the later kind in due course. Another thing one's interlocutor may do is, in effect, substitute for one's original proposal an alternative, perhaps one more, or less, sweeping in its scope. The following short conversation involves a special type of alternative proposal:

Tom: "Sure is a beautiful day!"

Jennie: "So *you* think!"

If Tom's response to Jennie is of the accepting variety, the two of them have come collectively to believe not that *it is a beautiful day*, but that *Tom thinks* it is a beautiful day.

Here Jennie's contribution to the negotiation is simultaneously to reject the proposition proposed for collective belief and, in effect, to offer to replace it with a closely related proposition about what Tom thinks, namely: he thinks the proposition he proposed for collective belief is true.

The foregoing example makes it clear that *we* can have views about what you, on the one hand, and I, on the other, think, and so on. In other words, some collective beliefs are about the thoughts or opinions of one or more of the participants.

In some special contexts, indeed, most likely those that are not paradigmatic conversations, it may be understood that the interlocutors are only really interested in the subjective stance of one party. This may be the case, for instance, in certain situations involving a psychotherapist and client. As the client says such things as "I think that...", "I'm uncomfortable when..." "I'm worried that...", and the therapist responds with words indicative of acceptance of the subjective proposition in question such as "I see", a series of collective beliefs about what the client thinks, feels, worries about, and so on may be established.²⁶

²⁶ Cf. Gilbert (1989: 297).

3.5 A Dual Role for Certain Sentences

Even outside special contexts such as that of a psychotherapy session, people often seem to propose for collective belief a proposition that concerns a propositional attitude of their own. Thus Kerry may say: “I like our chances in the upcoming basketball game”. Another way of saying roughly the same thing would be “I think our chances of winning the upcoming basketball game are good”. The role of such sentences in the production of collective beliefs in conversation is of some interest and, we make some tentative remarks on that score here.

For the sake of a label we shall refer to these sentences as speaker propositional attitude sentences or, for short, *SPA sentences*. They are tailored to ascribe a propositional attitude to the speaker. As we shall put it, they are tailored to express *SPA propositions*. In many if not most cases the proposition to which the speaker’s propositional attitude is directed—what we shall call its *target*—is not itself an SPA proposition.²⁷ Rather, its truth or falsity is independent of the speaker’s propositional attitudes. We shall refer to such propositions as, simply, non-SPA propositions, and the sentences tailored to express them, non-SPA sentences.

We shall focus our discussion on SPA sentences expressing SPA propositions with non-SPA target propositions, such as: the sentence, said by Tom, “I think it’s a beautiful day”. This is a SPA sentence expressing a SPA proposition; *I, Tom, think* it is a beautiful day. The target of this proposition is a non-SPA target proposition—the proposition that *it is* a beautiful day.

Consider now the following conversation:

Kerry: “I like our chances in the upcoming basketball game”

Jack: “No way, our team has no depth this year”

Phyllis: “That’s wrong! With Mark Smith as our go to guy, we have a shot.”

Kerry: “He’s the one!”

Jack: “So...maybe we do have a shot”.

Kerry begins by uttering an SPA sentence. In responding as he does, however, Jack does not address the corresponding SPA proposition, but focuses on the question of their team’s chances in the upcoming game: are they good—or not? In other terms, he focuses on the non-SPA proposition that is the target of the relevant SPA proposition. Phyllis continues along the same lines, as do the others in what follows. Kerry, for her part, may well not take Jack’s response as a change of subject, but rather as a challenge to what she said.

²⁷ It seems that the target of a SPA proposition could be another a SPA proposition. Thus consider such cases as: “I think I believe him”, or “I’m not sure I really do think that”.

Things could have gone differently. In particular, Jack might have focused on the relevant SPA proposition, responding to Kerry with, for instance: “I know you do. I don’t though.”

It seems, then, that not only can SPA sentences be used as vehicles for proposing SPA propositions for collective belief. They can be—and often are—used as vehicles for proposing non-SPA propositions. Indeed, it may be that with SPA sentences of the kind under discussion the default interpretation of what is going on when one of them is uttered is that both the pertinent SPA and its target non-SPA proposition is being proposed. Thus it may be that these SPA sentences are “Janus-sentences”.²⁸

A respondent, then, has a variety of options: focus on the pertinent SPA proposition, the target non-SPA proposition, or both. In the following dialogue Kate focuses on both and adds something of her own:

Bill: I really hate the fact that the philosophy department is moving to another building.

Kate: I know, so do I. It really is a shame.

Kate’s response to Bill indicates a willingness jointly to commit them to at least the following collective beliefs:

1. Bill hates the fact that the philosophy department is moving... [The SPA proposition expressed by the SPA sentence Bill uttered; affirmed by “I know”.]
2. The philosophy department moving to a new building is a bad thing. [The non-SPA target proposition associated with Bill’s propositional attitude, affirmed by “It really is a shame”].²⁹
3. Kate feels as Bill does. [A new SPA proposition; indicated by “so do I”].

In light of the foregoing we now consider a case based on a real-life conversation, where the first speaker felt that something had gone wrong. The pertinent part of this conversation went as follows:

Tabitha: “I prefer strawberry ice-cream to any other.”

Elmo: “No, no...I prefer chocolate.”³⁰

Why did Elmo’s response seem odd to Tabitha? This is not something one can be sure of, but suppose that, as we believe, she intended to put forward a SPA proposition—something along the lines of “I think strawberry ice-cream is the best”. Then his “No, no...” would have seemed inappropriate, assuming that he had both understood her intention and that he himself intended to express a SPA

²⁸ Cf. Nowell-Smith (1954): “a given word can not only do two or more jobs at once but also *is often in the absence of counter-evidence or express withdrawal presumed to be doing two or more jobs at once*” (emphasis added). He refers to this as *the Janus principle* (p. 100).

²⁹ Note that the objective proposition in question here is not the proposition that the department is moving but the proposition that its moving is something bad—apt to be hated.

³⁰ See Gilbert (1989: ch. 5).

proposition with his next words, that is: “I prefer chocolate”. After all, *his* preferring chocolate ice-cream does not show that *she* does not prefer strawberry.

Where might Elmo have been coming from? One possibility is this: he was taking Tabitha to be intent on proposing that they collectively believe the non-SPA target proposition of the SPA proposition associated with her utterance, namely: *strawberry ice-cream is preferable to any other*. Wanting their collective belief about what was objectively preferable to reflect his personal preference, he therefore objected, indicating his willingness to participate in one collective belief but not another. Though it might have been more natural from the point of view of English conversational style to speak more along the lines of Jack and Phyllis in their responses to Kerry, omitting any reference to his own perspective, he couched his objection to the pertinent non-SPA proposition using a SPA sentence—perhaps because Tabitha herself had used such a sentence.³¹

3.6 *Implicit Collective Beliefs*

To what extent, if at all, can the parties to a paradigmatic conversation fail to establish one or more collective beliefs in the course of that conversation? In discussing this question we focus, as before, on the explicit collective belief proposals made in the course of a given paradigmatic conversation, the way these are handled by the protagonists, and the consequences of this handling.

Suppose that every collective belief proposal put forward by one party is rejected by the other party or parties. Here is a simple case:

Tom: “Sure is a beautiful day, isn’t it?”

Jennie: “No it isn’t. It’s far too hot!”

At this point in this conversation there seems to be a stalemate. It seems wrong to say either that Tom and Jennie collectively believe that it’s a beautiful day or that they collectively believe that it isn’t.

Note that even if it were true that there is *nothing* that Tom and Jennie collectively believe as a result of this conversation, that would not refute the NCB thesis. Tom’s contribution can still be viewed as a move in the negotiation of a collective belief, as can Jennie’s. He proposes the collective belief that it is undoubtedly a beautiful day, and Jennie rejects his proposal. The negotiation stalls, but it is still a negotiation.

³¹ An alternative way of interpreting this case is offered in Gilbert (1989: 296). This too was developed in accordance with the NCB thesis and accords with it. We shall not attempt to adjudicate between these two options here.

In this example, it may seem that, indeed, the parties have failed to establish any collective beliefs between them. It can be argued, however, that this is not so.³²

We have so far considered only what we are referring to as *explicit collective beliefs*, that is, beliefs established on the basis of a verbal expression of the proposition in question. It can be argued that these are not the only kind of collective beliefs that may be established in the course of a conversation. Non-explicit or *implicit collective beliefs* may also be established *whether or not an explicit collective belief is established*. Indeed, the establishment of implicit collective beliefs is quite hard to avoid.

To see this let us go back to the last example. Consider, first, the following variant of it:

Tom: “Sure is a beautiful day, isn’t it?”

Jennie: “Are you crazy? It’s the middle of the night!”

Tom might really be crazy, or there may be some other explanation for his saying what he said. Be that as it may, Jenny does not bother to consider the details of the proposition he puts forward—we shall assume—as a candidate collective belief. She attacks another proposition, something she takes to be false: that they are speaking during the day, making it clear that she is not ready collectively to believe *this* with Tom. In doing so she also makes it clear that, *a fortiori*, she is not ready collectively to believe the proposition he verbally expressed.

Without attempting a definition, and intending to invoke a broad intuitive notion of implication, we shall say that in the previous dialogue, Tom *implies* that they are speaking during the day. We allow that what one implies may or may not be matter of what is logically entailed by the proposition expressed by the sentence one utters. It is not, however, simply a matter of the speakers intentions with respect to his utterance. In the last example, Jenny attacks a proposition implied by Tom though not verbally expressed by him. In the example of which this was a variant, she attacks the proposition he verbally expresses. We return now to this example, namely:

Tom: “Sure is a beautiful day, isn’t it?”

Jennie: “No it isn’t. It’s far too hot!”

Here, too, Tom implies that he and Jennie are conversing during the day. In this case Jennie does *not* question the implied proposition, but implicitly accepts it.

³² With respect to the discussion that follows in this section, the literature on conversation contains familiar, related ideas that do not bring collective beliefs into the picture. Relevant classic texts include Stalnaker (1973, 1974, and elsewhere) and Lewis (1979). In this section we do not try to align what we say with this literature. We briefly compare and contrast the perspective of Lewis and Stalnaker with that of the NCB thesis in the third section of this article. To briefly anticipate here: the fact that the NCB fits as well with their ideas as it does lends support to the NCB, which we take to be a plausible way of elaborating their perspective.

It seems fair to say that she and Tom now have at least the following implicit collective belief: they are conversing during the day. At the same time, they have reached a stalemate as far as the collective belief that was explicitly proposed is concerned. They may be said not to know what to think, collectively, as far as it is concerned.

Clearly, it is harder than might be thought for the parties to a paradigmatic conversation to fail to establish one or more collective beliefs in the course of their conversation. Consider the last example once again: though Tom and Jennie's conversation is a stalemate in terms of the collective belief explicitly proposed, it establishes at least one collective belief, the implicit collective belief that they are conversing during the day. We take most paradigmatic conversations to be analogous: during negotiation, interlocutors may or may not form a number of explicit collective beliefs. Even if they form no explicit collective beliefs, many implicit collective beliefs are likely to be established as their conversation continues. When this happens, the parties establish a joint commitment to believe as a body the propositions in question, propositions they may not consciously have entertained.

Let us return now to the example in which Jennie makes it clear that she is neither prepared to form the explicitly proposed collective belief nor the implied proposition that the two of them are speaking during the day. Indeed, she rejects the former by rejecting the latter.³³ Here we make two points. First, we echo something said earlier: if there is nothing even in the way of implied propositions that the parties come collectively to believe in such scenarios, that does not refute the NCB thesis. We would have a stalled negotiation, but it would be a negotiation nonetheless. Second, such brief interchanges will often be only a segment in a longer conversation.

In the case at issue here, suppose Tom responds "You're right". In that case Jennie's counter-proposal wins and a new collective belief is formed. This helps to confirm the point that it is difficult, though not necessarily impossible, for a conversation to occur without the establishment of one or more collective beliefs following an explicit collective belief proposal.

3.7 Some Principles of Conversational Collective Belief Formation

Something like the following principle of conversational collective belief formation suggests itself:

³³ Something similar seems to occur when people respond to a collective belief proposal with such a retort as "Rubbish!" Here it is clear that the explicitly proposed collective belief is rejected. Though it is not clear why it is rejected, it is hard to say that any of the implied propositions are accepted.

When in the course of a conversation someone explicitly proposes a particular collective belief for the parties, *he simultaneously if implicitly proposes that all of the propositions he implies in saying what he says be collectively believed as well.*

Let us call this the implication principle.³⁴

The following related principles also suggest themselves. For the sake of simplicity we assume that only two persons are involved in the conversation. These principles can be generalized to cover larger conversational groups.

First, there is *the whole shebang principle*:

If the hearer responds by accepting a proposed explicit collective belief without questioning any of the implied propositions, the conversationalists now collectively believe both the explicitly proposed belief and all of the implied propositions.³⁵

Second, there is *the selection principle*:

when a speaker proposes a particular explicit collective belief for the parties, his hearer may explicitly reject that belief but still explicitly or implicitly accept one or more of the implied propositions for collective belief.

Thus Jennie might have responded to Tom's "Sure is a beautiful day, isn't it?" thus:

"Well, it's a day...but not a beautiful one!"

Doubtless more such principles can be articulated, and those stated here can be made more fine-grained. If they or principles much like them are correct, however, it is indeed quite hard to avoid establishing a collective belief or two as one's conversation progresses. That is in large part because of the possibility of implicit collective beliefs.

3.8 The Normative Force of Implicit Collective Beliefs

As indicated above, implicit collective beliefs as we understand them are structured similarly to explicit collective beliefs. Through the constitutive joint commitment, once the implicit belief is in place, the group members obligated to one another to act as if they were "of one mind" with respect to its truth.

Thus suppose that Xavier and Yorick have established the implicit collective belief that Bill Clinton was once the President of the United States. In what follows in their conversation they are committed and obligated to one another not explicitly to gainsay this belief or to say anything that implies its falsity. If one of them happens somehow to come up with something that says or implies that the

³⁴ In referring to this as a "principle" we mean only that it (or something like it) is a true generalization about conversational collective belief formation. It is not a normative principle, i.e. it is not a principle requiring or recommending certain conduct. Note added in response to a comment by Antonella Carassa and Marco Colombetti, personal communication 2011.

³⁵ This implies that the individual conversationalists may not be consciously aware of the content of some of the collective beliefs of the group, in particular the implicit collective beliefs.

implicit belief is not true, the other has the standing to take him to task for this, as in: “What do you mean?” said in a tone of rebuke.

In contrast with this, suppose that the conversation is just beginning, and Yorick starts with “Bill Clinton was never president”. This could well *surprise* Xavier, given that the contrary is so well known. He might then raise his eyebrows and ask “What do you mean?” Failing his having some special authority relation to Yorick, however, he would not be in a position to respond in the way described above—with a rebuke.

4 The NCB Thesis Compared with Some Other Approaches to Conversation

The NCB thesis can be fruitfully aligned, contrasted, and compared with a number of existing discussions of conversation in the literature. In this section we make a beginning in this direction, without attempting anything like a complete discussion either with respect to those contributions we do discuss, or with respect to the pertinent contributions that exist.

4.1 *Conversation and Obligation*

The NCB thesis implies that as conversations proceed the participants are highly likely to accrue a set of joint commitments along with associated obligations towards one another to act as is appropriate to these commitments. Along with these obligations come authority relations consisting of the standing to demand fulfillment of the obligations, the standing to rebuke for non-fulfillment, and so on. Intuitively, obligations of joint commitment “trump” personal inclinations, at least, as far as the issue of what one ought to do is concerned. Thus, the creation of a conversational collective belief, implicit or explicit, has important normative consequences for the parties.

Though we cannot attempt this here, an extended comparison of the perspective of the literature on what have been called “discourse obligations” with that of the NCB thesis would be of considerable interest. Clearly these perspectives share the contention that obligations of some kind are part and parcel of the situation of those engaged in conversation. That said, so-called obligations are of very different kinds.

In at least some of the discourse obligations literature the type of obligation at issue is not, or not clearly, the one at issue here. In particular, it is not a matter of obligations of the kind correlated with claim-rights and the associated standings to make related demands and rebukes—such as are the obligations of joint commitment.

Related to this point is the following. Some discourse obligation theorists have emphasized the role of “penalties” in keeping people on the track of their obligations, as in the following quotation.

The concept of *penalty* is employed here because of the analogy with everyday situations in which a person is threatened with concrete sanctions if he or she fails to fulfill a particular obligation (e.g. the obligation to return a library book by a certain date). In the context of dialogue, the negative consequences of failing to fulfill a discourse obligation are of course much less tangible and measurable; they include consequences such as irritation and negative judgments on the part of the dialog partner... (Jamieson and Weis 1995).

It is worth observing that what the NCB thesis predicts with respect to violation of the obligations associated with the collective beliefs established in conversation is that the “penalty” in question may well take the form of something genuinely punitive—an authoritative rebuke. Such rebukes lie in the province of claim-rights and their correlative obligations. In contrast, “irritation” on the part of one’s co-conversationalists or their “negative judgments”, as such, may not rise to the level of a punitive move.³⁶

4.2 *Lewis on Conversational Score and Presuppositions*

The NCB thesis can be brought to bear on the central theme of David Lewis’s classic article “Score-Keeping in a Language Game”³⁷ Lewis himself thought that his discussion in that paper and in earlier papers by Robert Stalnaker might plausibly be further elaborated in terms of Margaret Gilbert’s account of collective belief.³⁸ This and the next sub-section briefly show how such an elaboration might go.

Discussing Lewis here, we are interested in the broad outlines of his ideas about conversation rather than any particular details. In particular we focus on his idea of a “conversational score”.³⁹

Consider first the following quotation, whose points are intended to apply to any “well-run conversation” (339). This is in the opening section of Lewis’s paper, which focuses on presuppositions:

Presuppositions can be created or destroyed in the course of a conversation. This change is rule-governed, up to a point....If at time *t* something is said that requires presupposition *P*

³⁶ Thanks to Frederick Schmitt for emphasizing this point, personal communication October 24 2011.

³⁷ Lewis (1979).

³⁸ In the late 1980s: personal communication with Margaret Gilbert, after reading Gilbert (1987) around the time of its publication. Lewis acknowledges his debt to Stalnaker’s work on presupposition in footnote 1 of his paper.

³⁹ Lewis (1979 esp. p. 347).

to be acceptable, and if P is not presupposed just before t, then—*ceteris paribus* and within certain limits—presupposition P comes into existence at t. (339-340).

Later in the same paper he recapitulates with an important addition:

Say something that requires a missing presupposition, and straightaway that presupposition springs into existence...*Or at least, that is what happens if your conversational partners tacitly acquiesce—if no one says “But France has three kings!” or “Whadda ya mean?”* (339; emphasis added).

See also, later:

Presupposition evolves according to a rule of accommodation specifying that any presuppositions that are required by what is said straightway come into existence, provided that nobody objects. (347).

In saying that presupposition P “springs into existence” at t, Lewis is talking in terms of an idea of a “conversational score” such that “the components of a conversational score at a given stage are *abstract entities*” (345, emphasis added). Among these components are the presuppositions required by the said rule of accommodation.

What, one may ask, is going on between the material interlocutors as this score, abstractly conceived, develops? Here is a role for the NCB thesis. Focusing here on those presuppositions that come into existence as part of the conversational score, when this is considered in the abstract, the NCB suggests that this score *tracks the accrual of a particular type of implicit collective belief*.

Perhaps, though, we should stick to Lewis’s language and speak now of *collective presuppositions*. Prior to fixing on a definition of this particular propositional attitude, in general, and without insisting that it is a form of belief,⁴⁰ the following account of collective presupposition recommends itself.

A and B *collectively presuppose* that p if and only if A and B are jointly committed to presuppose that p as a body.

Collective presupposition is understood here along the lines of collective belief. In that case, it has no implications for the personal propositional attitudes of the participants, nor are any facts about their personal propositional attitudes sufficient for collective presupposition—even given common knowledge of those personal propositional attitudes. It has the same commitments, obligations, and standings associated with any phenomenon constituted by a joint commitment.

Lewis’s rule of accommodation for presuppositions—and a range of possible versions of it—can plausibly be rewritten in terms of collective presupposition as just defined, and incorporated within the NCB thesis. Thus consider this interchange:

⁴⁰ Stalnaker (1973: 448) refers to presupposition as a propositional attitude, which he does not equate with belief.

Joe: “The economy is so bad that Mitt Romney actually has a chance at becoming president.”

Peter: “You’re right—it’s that bad!”

According to the version of the NCB thesis now envisaged, Joe is in effect making the following proposal to Peter: “Collectively believe with me that the economy is so bad that Mitt Romney actually has a chance at becoming president, and, in so doing, collectively presuppose with me whatever propositions are required in order that that belief be true.” Then, at least by the time Peter has finished his sentence, making clear that he accepts Joe’s proposal, Joe and Peter jointly presuppose that Mitt Romney would not ordinarily be a viable candidate. In other words they are now jointly committed to presupposing as a body that Mitt Romney, etc.

More generally, the suggestion is that in making a particular collective belief proposal one concomitantly implicitly proposes all of its presuppositions for collective presupposition. We shall henceforth assume that the NCB thesis includes this point.

Evidently, things do not always go as smoothly as in the brief interchange between Joe and Peter. Implicitly proposed presuppositions may be challenged. Thus Peter might have said to Joe, referring to Mitt Romney: “What do you mean he *actually has a chance*?” He is rejecting or at least stalling with respect to Joe’s implicit proposal that they collectively presuppose that Mitt Romney would not ordinarily be a viable candidate.

4.3 *Stalnaker on Presupposing and Common Ground*

In Robert Stalnaker’s classic discussions *presupposing* is something done by an individual participant in a conversation. Thus Stalnaker’s analysandum in an important article is “A speaker presupposes that P at a given moment in a conversation....” (1973: 448). As far as we can tell Stalnaker does not consider the possibility of conversationalists collectively presupposing that such-and-such, where this is distinct from each of them personally presupposing it.

According to the NCB this leaves out a crucial dimension of conversation. In this section we briefly relate the notion of collective presupposition just outlined in our discussion of Lewis to the passage from which the above partial quotation from Stalnaker comes.

Stalnaker writes, more fully, that “a rough definition might go something like this:”

A speaker presupposes that P at a given moment in conversation just in case he is disposed to act, in his linguistic behavior, as if he takes the truth of P for granted, and as if he assumes that his audience recognizes that he is doing so. (Emphasis in the original).⁴¹

⁴¹ Stalnaker (1973).

What we would like to note in relation to this is the following.

Suppose that speaker S is a member of conversational group G, and the members of G, as such, collectively presuppose that P. Then, roughly, while this collective presupposition remains in effect, and in this conversational setting, it will be incumbent upon S (and the other members of G) *to act, in his linguistic behavior, as if he takes the truth of P for granted*. Note that the italicized words are also to be found in the quotation from Stalnaker.

We say this because *acting in his linguistic behavior as if he takes the truth of P for granted* is, roughly, what a joint commitment to presuppose that P as a body dictates.⁴² More precisely, with the provisos noted, he is to do his part along with the others in emulating a single body that believes that P. Further, this will be common knowledge among the parties who jointly committed themselves to presuppose that P in the first place. It seems just possible, then, that the conversational phenomenon to which Stalnaker is responding in his discussion is in fact the collective phenomenon of collective presupposition.

That said, his discussion of conversation overall does not give that impression, though some of the pertinent phrases—such as “shared beliefs” “common beliefs” and “common ground”—can in principle be given a collective interpretation. A proponent of the NCB thesis need not doubt, of course, that in the background of many conversations there are beliefs of the individual parties with the same content, beliefs about these beliefs, and so on, which constitute a species of “common ground”, and that these have an important role to play in what transpires. The NCB thesis contends that beyond this there develops an increasingly rich cognitive profile of *collective* beliefs and presuppositions with associated expressive obligations of the parties that act as constraints on their future linguistic behavior. This is if you like the *collective* ground beyond the *common* ground.

4.4 A Suggestion Derived from Lewis

Drawing on some remarks in Lewis’s “Scorekeeping” article, we note that in negotiating their collective beliefs conversationalists in a broad sense may be constrained by certain collectively accepted parameters, something that a full development of the NCB thesis will take into account.⁴³ Here is the pertinent quotation from Lewis with material inserted in square parentheses so that it makes our point:

The conversationalists may conform to [collective] directives, or may simply [collectively] desire, that they [individually] strive to steer certain components of the

⁴² “Roughly”: Stalnaker’s discussion of the “as if” in his account is pertinent here. See also Gilbert (1989: ch. 5), also 1987, on the behavioral requirements of a collective belief.

⁴³ We would interpret “collectively accepted” here in terms of a joint commitment to accept as a body.

conversational score [i.e. of the set of conversational collective beliefs] in certain directions. Their efforts may be cooperative, as when all participants in a discussion try to increase the amount that [they collectively] presuppose.⁴⁴ Or there may be conflict, as when each of two debaters tries to get his opponent to grant him—to *join with him in presupposing*—parts of his case, and to give away parts of the contrary case (345; emphasis added).

Note the words we have emphasized in the quotation. They suggest that Lewis may have been somewhat predisposed to the NCB thesis. To envisage one person's *joining with another in presupposing that such-and-such* may only be to envisage that each of them will presuppose. This language fits well, however, with a reading in terms of collective presupposition as that is understood here.

4.5 Why Converse? Charles Taylor on Rising Above Common Knowledge

According to the NCB thesis, everyday conversation is structured around collective belief formation. This holds off from saying anything about the purpose or function of conversation, and it has not been our intent to pursue that question. Indeed, given that different people and groups of people seek out conversation from different motives, it may be hard to fix on “the” function of conversation. We shall not pursue this issue in any detail here, but before concluding we note some related points.

A thought sometimes advanced is that the point of paradigmatic conversation is the exchange of information. This suggests that conversations are primarily about the beliefs of *one individual* being “transferred” to *another individual*. If the NCB thesis is correct, however, the personal beliefs of the pertinent individuals do not play a central role in conversation as a general practice. What centrally happens in conversation is that *what may or may not have been a belief of an individual conversationalist* becomes *the belief of a group*—that group constituted by the conversationalists in question.

Of course, people may come away from a given conversation with personal knowledge or beliefs that they did not have before. If in the context of our conversation Jack, who is known to be a sensible, well-informed person, remarks to me that “Seana is going to France”, and I say “Uh-huh”, thus sealing our collective belief, I am likely also to come away with the personal belief that Seana is going to France. Possibly, too, Jack's primary personal purpose in saying what he said was to produce this belief in me.

That said, my coming away from a situation with new personal beliefs—and, indeed, with new knowledge—is something that happens in many other instances.

⁴⁴ Here we replaced “all of them willingly” with “they collectively”. The rest of the material in square brackets in the quotation is inserted into the text without replacing anything.

People come away with new personal knowledge when reading a book or newspaper, when observing the world around them, or when reflecting on their own ideas. Conversation, then, is just one source of new personal knowledge—when it does amount to that. Whereas what happens in conversations *between* individuals—in particular the formation of a new collective belief—is not something which can happen to an individual *alone*.

Indeed, the formation of a new collective belief brings both belief and related propositional attitudes *to the collective level*, a possibility which has rarely been considered in contemporary philosophy at least until recently. As the plausibility of the NCB thesis suggests, however, if this possibility is neglected an important dimension of human life has been ignored.

Much work in philosophical discussions of human interactions generally has proceeded in terms of individual human beings' beliefs, desires, intentions, and so on. David Lewis famously and importantly added "common knowledge" to the mix. Not only can we both believe something, there can be common knowledge between us that we do. There have been many attempts to define "common knowledge" along the lines of Lewis. As far as we know, none see common knowledge as a plural subject phenomenon. It has to do, roughly, with individuals' knowledge of individuals' knowledge, or individuals' beliefs about individuals' beliefs. It certainly adds an important dimension to a situation, but there can be common knowledge without matters rising to the collective level.

Charles Taylor has offered examples that make very vivid the transition from a situation where there is common knowledge between two persons that *p* to a situation in which, in the terms of this paper, there is a collective belief that *p*.⁴⁵ The following represents the gist of these examples.

Two people are sitting next to each other on the subway. It is very hot there. Both are aware that it is hot, and it is common knowledge between them that it is hot. Thus each knows that the other does not need to be told about this for the sake of his being better informed. Let us suppose also that it is common knowledge between them that each is feeling the heat: each is visibly perspiring and so on. Yet one may turn to the other and say: "Whew, it is it hot here, isn't it?" the other then offering an accepting response. In spite of the fact that no one has new information about the current temperature, or how the other feels, something important happened when they exchanged these words.

The fact that it is hot there and that each is feeling the heat is now "out in the open", "in public space", or *entre nous*—to use Taylor's terms. He does not, to our knowledge, define these closely related terms. That said, it is clear that in the terms of this paper: it is a matter of collective belief.

Previously each individual had held a number of *personal beliefs* including the belief that it is hot where he is. With the exchange of words something comes into existence which is not reducible to facts about what each personally believes or the common knowledge between them. They are now in a position to talk about what

⁴⁵ Taylor (1985).

we (collectively) believe, as they were not before. As we have explained, this has consequences. Both are now jointly committed to act, at least for the time being, as though it is hot there. Perhaps they would have done this regardless. Nonetheless, each now is obligated to the other to do this, and each has the standing to rebuke the other if he does not so act. This standing did not exist prior to their conversation.

The conversation of Taylor's protagonists may seem trivial and far from a central case. Yet it has the virtue of making it clear that conversation can be for something other than the exchange of information. Furthermore, along the lines of the NCB thesis, it suggests that *whatever other purposes a given conversation serves, it is apt to bring the parties together precisely in the way of collective belief*. This is a very substantial effect, as we have argued. It is at least somewhat plausible then to suggest that it is this effect that paradigmatic conversations are all about. That is, *conversations are places where two individuals may express their willingness to be jointly committed to believe propositions as a body and thus effect such a joint commitment*. Among other things, any information exchanged is accidental, not essential, to the conversation *qua* conversation.⁴⁶

4.6 Summary and Concluding Remarks: Collectivizing Conversation

This paper presented and to some extent developed a thesis connecting paradigmatic conversations with collective beliefs. This was labeled the negotiation of collective belief thesis or NCB thesis. The paper began by arguing for a particular account of the phenomenon to which everyday collective belief ascriptions refer. Such ascriptions are naturally made in the context of conversations and related exchanges. That is, after or in the course of a conversation people say such things as "We think that would be great." Many attempts to explain statements of this sort have been *summative*. They have reduced the belief of a group to the beliefs of all or most individual members of the group, perhaps with additional conditions such as common knowledge of these beliefs. Such reductions do not accurately describe the phenomenon at issue. We proposed that, in the basic case, a collective belief that *p* is formed, roughly, when in conditions of common knowledge group members express their readiness to be jointly committed to believe that *p* as a body. This can be done even if no individual group member personally believes that *p*. The group members are then *jointly committed* to believe *p* as a body, and thereby constitute a *plural subject* that believes *p*.

According to the NCB thesis, paradigmatic conversation is to be understood as at least in part a negotiation between the parties over propositions put forward for

⁴⁶ On this point see also Gilbert (1989: 295).

collective belief. So negotiating is hard to avoid in the context of a paradigmatic conversation. Absent special background understandings, when one interlocutor explicitly expresses a given proposition, he simultaneously makes a proposal that this proposition be collectively believed. The other interlocutor may refuse to negotiate; otherwise he must accept or reject the proposal. If he rejects it, the explicitly proposed belief does not become a collective one. However, even with this rejection, it is likely that other collective beliefs and collective propositional attitudes of the conversationalists come into being. These include collective presuppositions understood as joint commitment phenomena.

The move to understanding conversation in terms of joint commitment in the sense at issue here is a move away from the individualism of much of the existing literature on conversation, including both the classic work of Stalnaker and Lewis discussed in section III, and important later work, such as that of Clark, who emphasizes that conversation is a “joint activity”.⁴⁷ In referring to “individualism” we have in mind that, for one thing, insofar as *commitments* are referred to by these authors these are the personal commitments of the individuals involved, albeit in combination with or in some other way related to other personal commitments.⁴⁸ Otherwise their discussions proceed in terms of the personal beliefs, goals, and so on, of the participants, including their personal beliefs about the personal beliefs of others. In our view, without recourse to the notion of joint commitment invoked in this paper, we cannot properly articulate what is going on in conversation.

Let us speak of a collective *activity* as one that involves two or more persons who act in light of a joint commitment to pursue as a body a certain goal.⁴⁹ Without specifying the details, it is natural enough to see conversation as a form of collective activity. As was once tartly observed “conversation is not a monologue”.⁵⁰ It takes at least two, and like the tango, it is something those two—or more—do *together*.⁵¹ Allowing for this important point, the NCB thesis proposes

⁴⁷ See Clark (1996). In emphasizing that conversation is a “joint activity” is Clark not showing himself to take a non-individualistic approach to conversation? Given his carefully expounded account of what he takes joint action and joint activity to be, and our construal of “individualism” (see the text immediately below), it seems not. The work of Carassa and Colombetti, cited earlier, is an exception here.

⁴⁸ Clark (1996: ch. 10) invokes a notion of “joint commitment”. This appears to be an individualistic notion in the sense noted in the text, above, and not therefore to be the sense delineated here. Cf. Carassa and Colombetti (2009a: 1841) on Clark.

⁴⁹ See Gilbert (1989: ch. 4); (1990).

⁵⁰ Gilbert [Martin] (1971: 384, also 477).

⁵¹ On acting together generally, see e.g. Gilbert (1989: ch. 4), (1990), (2006: ch. 6 and 7). This is there argued to have a joint commitment at its core—in this case a joint commitment to espouse as a body a certain goal.

that conversation consists in large part of the development of a *collective cognitive profile*. It is collective as opposed to “shared” in the sense of being attributable to each of the participants as individuals, and so on.⁵²

References

- Brennan, S.E., and H.H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1482–1493.
- Carassa, Antonella., and Colombetti, Marco. 2009a. Joint meaning. *Journal of Pragmatics* 41: 1837–1854.
- Carassa, Antonella., and Columbetti, Marco. 2009b. Situated Communicative Acts: A Deontic Approach. *Proc. CogSi* 1382–1387.
- Clark, Herbert. 1996. *Using language*. Cambridge: Cambridge University Press.
- Durkheim, Emile. 1895. In *The rules of sociological method*. ed. Catlin, G. New York: Free Press.
- Gilbert, Margaret. 1987. Modeling collective belief. *Synthese* 73: 85–204.
- Gilbert, Margaret. 1989. *On social facts*. London: Routledge and Kegan Paul; 1992. Princeton: Princeton University Press.
- Gilbert, Margaret. 1990. Walking together: A paradigmatic social phenomenon, reprinted in *Living Together*.
- Gilbert, Margaret. 1994. ‘Remarks on collective belief. In *Socializing epistemology: The social dimensions of knowledge*, ed. Schmitt, F. 235–253. Lanham, MD: Rowman and Littlefield, reprinted in *Living Together*.
- Gilbert, Margaret. 1996. *Living together: Rationality, sociality and obligation*. Lanham, MD: Rowman and Littlefield.
- Gilbert, Margaret. 2002. Belief and acceptance as features of groups. *Protosociology* 16: 35–69 (online journal: www.protosociology.de).
- Gilbert, Margaret. 2003. The structure of the social atom: Joint commitment as the foundation of human social behavior. In *Social Metaphysics*, ed. Schmitt, F. 39–64. Lanham, MD: Roman and Littlefield, 39–64.
- Gilbert, Margaret. 2004. Collective epistemology. *Episteme* 1: 95–107.
- Gilbert, Margaret. 2006. *A theory of political obligation: membership, commitment, and the bonds of society*. Oxford: Oxford University Press.
- Gilbert, Margaret. 2010a. Culture as Collective Construction *Sociologische Theorie kontrovers*, eds. Albert, Gert, and Sigmund, Steffen, *Kolner Zeitschrift für Soziologie und Sozialpsychologie*, 50: 383–393.
- Gilbert, Margaret. 2010b. Joint Commitment and Group Belief. In *Sociologische Theorie kontrovers*, eds. Albert, Gert, and Sigmund, Steffen, *Kolner Zeitschrift für Soziologie und Sozialpsychologie*, 50: 405–410.
- Gilbert, Margaret. 2013. Commitment. In *The Encyclopedia of Ethics*, ed. LaFollette, Hugh.
- Gilbert, Martin. 1971. *Winston S Churchill*, 3rd ed. Boston: Houghton Mifflin.
- Hart, H.L.A. 1961. *The Concept of Law*. Oxford: Clarendon Press.

⁵² In this paper we often refer to what we (Gilbert and Priest) believe, and so on. These references should be understood in accordance with the theory of collective belief adumbrated here, and therefore not logically to entail anything about what either Gilbert or Priest personally believes. Similarly, when we maintain, in the first footnote, that responsibility for the points made here is ours alone, neither one of us means to ascribe responsibility for any or all of the points made to her personally. Collective responsibility, however, is a topic for another occasion.

- Jameson, A., and Weis, T. 1995. How to juggle discourse obligations. *Proceedings of the Symposium on conceptual and semantic knowledge in language generation* 171–185.
- Lewis, D.K. 1969. *Convention: A philosophical study*. Harvard: Harvard University Press.
- Lewis, David K. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8: 339–359.
- Nowell-Smith, Patrick. 1954. *Ethics*. London: Penguin Books.
- Quinton, Anthony. 1975. Social objects. *Proceedings of the Aristotelian Society* 76: 1–27.
- Schiffer, Stephen. 1972. *Meaning*. Oxford: Oxford University Press.
- Simmel, Georg. 1971. *Georg Simmel: On Individuality and Social Forms*, ed. Levine, D. Chicago: University of Chicago Press. (First published 1908).
- Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophic Logic* 2: 447–457.
- Stalnaker, R. 1974. Pragmatic Presuppositions In *Semantics and Philosophy*, eds. Munitz, M.K. and Unger, P.K. New York: University Press.
- Taylor, Charles. 1985. *Human agency and language*. Cambridge: Cambridge University Press.
- Vanderschraaf, Peter and Sillari, Giacomo, Common Knowledge, *The Stanford encyclopedia of philosophy* (Spring 2009 Edition), ed. Zalta, Edward, N. September 20th, 2011, URL = <http://plato.stanford.edu/archives/spr2009/entries/common-knowledge/>.

The Single Norm of Assertion

Martin Montminy

Abstract Timothy Williamson holds that the knowledge rule of assertion, according to which one should assert only what one knows, is the single norm of assertion. I explain and defend Williamson's thesis. I identify three key features associated with the thesis: the single norm of assertion should be *constitutive*, *individuating* and *basic*. Roughly, a constitutive rule of a speech act governs every possible performance of the act. A rule of assertion is individuating if it differentiates assertion from every other speech act. Finally, a rule of assertion is basic if it does not derive from other rules governing assertion, or other speech acts. I thus construe Williamson's thesis as the claim that the knowledge rule is the only individuating, constitutive and basic rule of assertion. This thesis is compatible with the idea that assertion is governed by non-epistemic constitutive rules such as moral and prudential rules. Since these rules are not individuating, there is no need to understand Williamson's thesis as bearing solely on the *epistemic* rule of assertion. I explain how the existing arguments in favor of the knowledge rule can serve to support the thesis that this rule is also the single norm of assertion. I also respond to some objections against the knowledge rule, and criticize alternative rules.

Many authors¹ advocate the knowledge rule of assertion, according to which one should assert only what one knows. Timothy Williamson goes further and defends what he calls a *simple* account, according to which the knowledge rule is *the single norm of assertion*.² This simple account, he writes, is "an attractively simple suggestion" (2000, p. 241) that "would be theoretically satisfying, if it worked" (*ibid.*, p. 242). Williamson's "simple suggestion" has been greeted with

¹ DeRose (2009), Hawthorne (2004) and Unger (1975), to mention just a few.

² All references will be to the slightly revised version of Williamson's (1996) article "Knowing and asserting," which appears as a book chapter in his (2000).

M. Montminy (✉)

University of Oklahoma, Norman, Oklahoma, USA

e-mail: montminy@ou.edu

skepticism, even by the supporters of the knowledge rule. Keith DeRose, for instance, characterizes it as “little more than an item of faith” (2009, p. 93, n. 15).

In this chapter, I will explain and defend Williamson’s thesis that the knowledge rule is the single norm of assertion. This will require a careful examination of three key features associated with this thesis: the single norm of assertion should be *constitutive*, *individuating* and *basic*. Once this is done, the thesis will appear quite plausible, given the existing arguments in support of the knowledge rule.

1 Constitutivity

Williamson writes:

[N]ot all norms for assertion are on a par. Norms such as relevance, good phrasing, and politeness are just applications of more general cognitive or social norms to the specific act of assertion. Perhaps the norm of informativeness results from a more complex interaction between a general norm of cooperativeness and the nature of assertion as a source of information. But on this view, not all norms for assertion derive from more general norms; otherwise nothing would differentiate it from other speech acts. This chapter aims to identify the constitutive rule(s) of assertion (2000, p. 238).

This passage suggests that a constitutive rule of assertion should (among other things) not derive from other norms and should differentiate assertions from other speech acts. In other words, the *constitutive* rule(s) of assertion should be *basic* and *individuating*. I will explain each of these three features more carefully, and how each is related to Williamson’s project of identifying the *single* norm of assertion. But let us start with Williamson’s notion of constitutivity.

Williamson writes that “a rule will count as constitutive of an act only if it is essential to that act: necessarily, the rule governs every performance of the act” (ibid., p. 239). This, of course, does not mean that every assertion respects the constitutive rule(s) of assertion. A speaker who breaks a constitutive rule of assertion may still be performing an assertion. The idea is thus that every possible assertion is governed by such rule(s), even though some assertions may violate it (them).

It is worth contrasting constitutive rules with *regulative* ones. There may well be norms that govern assertion, but are not essential to it. Let us call these regulative norms. Assertions may, in a certain community or type of context, be governed by rule R, without it being the case that R is constitutive of assertion. Social rules (e.g., rules associated social institutions, rules of protocol, rules of politeness) govern assertions, but they do so as regulative, rather than constitutive, rules. In some societies, it is considered inappropriate for politicians to comment publicly on religious matters, and for religious leaders to discuss political issues. In some social groups, it is improper to talk with your mouth full. These social rules regulate assertion (in some contexts), but they are obviously not constitutive of assertion.

Clearly, for Williamson, if a rule is constitutive of an act, it is essential to it. But it is unclear whether the converse is true for him. Unfortunately, beyond an analogy with the constitutive rules of a game, Williamson does not tell us what, precisely, would be needed for constitutivity. According to him, does the fact that a rule is constitutive of a speech act entail that it is basic, or individuating, or that it is the unique rule of that speech act? Instead of trying to answer this question, I will simply stipulate that a rule is constitutive of assertion just in case it applies to every possible assertion.

2 The Simple Account

Williamson writes that an “attractively simple suggestion” (ibid., p. 241) is that there is just one rule of assertion, where the rule says:

(The C rule) One must: assert p only if p has C.

Williamson takes the C rule to be a constitutive norm of assertion, that is, it governs every possible assertion.

Williamson writes, “Furthermore, the envisaged account takes the C rule to be individuating: necessarily, assertion is the *unique* speech act A whose unique rule is the C rule” (ibid., p. 241). I take this claim to be equivalent to:

(1) Necessarily, assertion is governed by the C rule and no other rule, and no speech act other than assertion is governed by the C rule as its unique rule.

Given that ‘necessarily’ distributes over conjunction, we get the following first conjunct from (1):

(1a) Necessarily, assertion is governed by the C rule.

(1a), which is equivalent to the claim that every possible assertion is governed by the C rule, simply reiterates the claim that the C rule is constitutive.

The second conjunct concerns the idea that the C rule is the single norm of assertion:

(1b) Necessarily, no rule other than the C rule governs assertion.³

But (1b) is clearly too strong: assertion is governed by moral rules, prudential rules, the rule of relevance, etc. Hence, it is not the case that necessarily, the C rule is the unique rule governing assertion.

Perhaps the idea is that the C rule is the unique *constitutive* rule of assertion. But this assumes that moral rules, prudential rules, the rule of relevance, etc., are merely regulative rules of assertion. And this assumption is questionable. Some would argue that moral (prudential) norms hold as a matter of necessity. If so, then

³ Williamson also writes that “necessarily, assertion is a speech act A whose unique rule is ‘One must: perform A with the content p only if p has C’” (ibid., p. 241).

every possible action, and a fortiori every possible assertion, would be governed by moral (prudential) norms. Similarly, it may be that the rule of relevance, according to which a speech act must be relevant to the purpose of the conversation, holds as a matter of necessity. If so, every possible assertion would also be governed by this rule.⁴ Hence, assertion arguably has several constitutive rules. One's account of the rule of assertion had better steer clear of the controversial commitment that assertion has only one constitutive rule. Perhaps an account of what it is for a norm to be the single norm of assertion should await an explanation of what an *individuating* norm is.

3 The Individuating Condition

Let us thus turn to the last part of (1), which is meant to express the individuating condition:

(1c) Necessarily, no speech act other than assertion is governed by the C rule as its unique rule.

Unfortunately, (1c) also seems too strong. Let us assume, for ease of discussion, that the C rule is the knowledge rule:

(KR) One must: assert that p only if one knows that p .

Let us also assume that KR is the only constitutive rule of assertion.⁵ Our proposed individuating condition is thus:

(1c') Necessarily, no speech act A other than assertion is governed by the rule 'One must: perform A with the content p only if one knows that p ' as its unique rule.

But consider the speech act of conjecturing. The constitutive rule for conjecture is less demanding than the constitutive rule for assertion. Hence, if one satisfies the constitutive rule for asserting that p (i.e., if one knows that p), then one also satisfies the constitutive rule for conjecturing that p . But we can conceive of a

⁴ We can imagine a community of speakers in which "aimless" conversations are encouraged, because, say, they are thought to foster creativity. But this does not obviously constitute a counterexample to the rule of relevance: a contribution that does not address what was just said could nevertheless be relevant to the *broader* purpose of the conversation, which is to foster creativity.

⁵ As I pointed out in the previous section, this assumption is questionable. But if we did not make it, then assertion would have constitutive rules (e.g., moral, prudential, relevance rules) that are also constitutive of other speech acts. This would entail the falsehood of (1c). Making the assumption will allow me to identify an *additional* problem with (1c).

community in which some social rule makes it inappropriate to conjecture that p when one does not know that p . Imagine, for example, a community in which it is considered arrogant or gossipy to conjecture that p , unless one knows that p . Alternatively, imagine a society ruled by a tyrant who does not tolerate his subjects to conjecture in his presence, unless they know what they are conjecturing. In these communities, the rule ‘One must: conjecture that p only if one knows that p ’ would be the only rule governing conjecture.⁶ This shows that contra (1c’), a speech act A other than assertion could be governed by ‘One must: perform A with the content p only if one knows that p ’ as its unique rule.

For Williamson, a rule of assertion is individuating if it can serve to differentiate assertion from every other speech act. Consider Williamson’s discussion of the truth rule:

(TR) One must: assert that p only if p .

One problem with TR, Williamson points out, is that some speech acts other than assertion, such as conjecturing and swearing, also aim at truth. In other words, for some speech act A other than assertion, the rule ‘One must: perform speech act A with the content p only if p is true’ is constitutive.⁷ TR thus fails to be individuating.

Perhaps the individuating condition should state that no speech act other than assertion is governed by the C rule as a constitutive rule. This statement avoids the problem mentioned above: conjecture could, in some context or community, be governed by the rule ‘One must: conjecture that p only if one knows that p ’ as a *regulative* rule; however, it is not governed by it as a *constitutive* rule. This statement of the individuating condition also addresses the earlier problem concerning moral rules, prudential rules and the rule of relevance. It may be that these rules apply to every possible assertion, and are thus constitutive rules of assertion. But they are clearly not individuating rules of assertion, for they are constitutive rules of speech acts other than assertion.⁸

Unfortunately, if we adopt this statement of the individuating condition, then KR is vulnerable to the same objection as TR. The evidential norms for swearing are more demanding than those for assertion. Hence, if the rule ‘One must: perform speech act A with the content p only if one knows that p ’ is constitutive of

⁶ Recall that I am assuming that moral rules, prudential rules, etc., are merely regulative rules of assertion. This allows me to stipulate that these rules are not in place in the imagined communities.

⁷ Or perhaps I should say: if the rule ‘One must: perform speech act A with the content p only if p is true’ is constitutive of assertion, it is also constitutive of conjecture and swearing. However, as we will see in Sect. 5, there are good reasons to hold that conjecture is *not* governed by the truth rule.

⁸ Or, more carefully: if these rules are constitutive of assertion, then they are constitutive of other speech acts as well.

assertion, it is also constitutive of swearing. KR is thus not individuating, according to the proposed statement.

But the proponent of KR can point to a difference: while that rule is *basic* with respect to assertion, it is not so with respect to swearing. In other words, the basic rule of swearing would be, as Williamson (ibid., p. 244) suggests, something like ‘One must: swear that p only if one has grounds for unusual certainty that p .’ And from this rule, one can derive the less demanding rule ‘One must: swear that p only if one knows that p .’

I thus propose to state the individuating condition as follows:

(IND) No speech act other than assertion is governed by the C rule as a constitutive and basic rule.⁹

4 The Single Norm of Assertion

We are now in a position to explicate the thesis that the C rule is the single norm of assertion. This thesis holds that:

(SING) The C rule is the only individuating, constitutive and basic rule of assertion.

The knowledge account of assertion, according to which KR is the single rule of assertion, can thus be taken to hold that assertion is governed by KR as its unique individuating, constitutive and basic rule. This construal of the knowledge account allows its proponents to avoid a common objection: they can admit the possibility that assertion is governed by many constitutive rules (moral, prudential, etc.), while maintaining that only one constitutive rule of assertion, namely KR, is both basic and individuating. To repeat, saying that KR is the single norm of assertion is not saying that assertion is governed by only one norm.

It is also worth noting the existence of a *secondary norm*, which derives from KR. As Keith DeRose writes,

As it happens with other rules, a kind of secondary propriety/impropriety will arise with respect to [KR]. While those who assert appropriately (with respect to this rule) in a primary sense will be those who actually obey it, a speaker who broke this rule in a blameless fashion, by asserting something she didn’t know but reasonably thought she did know, would in some secondary sense be asserting properly, and a speaker who asserted something she reasonably thought she did not know, but in fact did know (if this is possible), would be asserting improperly in that secondary sense (DeRose 2009, pp. 93–94).¹⁰

⁹ Williamson’s objection against TR still holds on this construal of the individuating condition. The proponent of TR cannot plausibly hold that the rule ‘One must: assert that p only if p ’ is basic, while ‘One must: swear that p only if p ’ is non-basic.

¹⁰ See also Williamson (2000, p. 256).

The proponent of KR can appeal to secondary propriety to handle Gettier cases, for example. Suppose that the clock stopped on 3 PM 24 h ago. Sally looks at the clock, and on this basis, forms the belief that it is 3 PM. Although Sally does not know that it is 3 PM, it is reasonable for her to assert that it is 3 PM. In this case, the proponent of KR can say, Sally's assertion is improper, because she does not know that it is 3 PM; however, she cannot be blamed for this impropriety, for she reasonably believes that she knows that it is 3 PM.

According to KR, knowing that p is necessary to properly assert that p . But some authors hold that knowledge is also *sufficient* for proper assertion.¹¹ How should we understand this proposal? Consider the following biconditional form of the knowledge rule:

(2) One must: assert p if and only if one knows p .

This rule is clearly inadequate: one is under no obligation to assert everything one knows. The following biconditional avoids this problem:

(3) One is positioned to make a proper assertion that p if and only if one knows that p .

But (3) is also problematic, for a speaker who knows that p may improperly assert that p , because such an assertion would be immoral or imprudent, for example. We thus need to restrict the rule to *epistemic* propriety:

(BKR) One is positioned to make an epistemically proper assertion that p if and only if one knows that p .

Let us call the *sufficiency rule* the claim that one is positioned to make an epistemically proper assertion that p if one knows that p .

Is the sufficiency rule constitutive of assertion? Let me first note that none of the three main arguments in favor of KR support the sufficiency rule. As we will see in the next section, these arguments merely highlight the impropriety of asserting that p when one does not know that p . However, the sufficiency rule is *prima facie* plausible. As Jessica Brown writes, "one can criticise someone's failure to assert that p by pointing out that she knows that p . For instance, I may be fairly criticised if I fail to mention to you a fact which is highly relevant to your projects and which I know. You may criticise my failure saying, 'Why didn't you say that p ? You knew it all along'" (2010, p. 552).¹²

However, an adequate defense of the sufficiency rule is a complex matter that I cannot undertake here. First, a defender of the sufficiency rule would need to

¹¹ See DeRose (2009, Chap. 3). Hawthorne (2004) and Williamson (2000) also seem to support the sufficiency rule. Williamson writes that the knowledge account "can be roughly summarized in the slogan 'Only knowledge warrants assertion.' 'Warrant' is used here as a term of art, for that evidential property (if any) which plays the role of the property C in the correct simple account of assertion" (2000, p. 243).

¹² I should note, however, that Brown does not endorse the sufficiency rule.

address so-called “high-standards” contexts, that is, contexts in which the stakes are high and error possibilities have been raised. In such contexts, some argue, asserting that p is epistemically improper, even though the conversational participants know that p , according to a classical, invariantist account of knowledge.¹³ The defender of the sufficiency rule would also have to address other counterexamples such as those recently presented by Brown (2010) and Lackey (2011). I will thus focus solely on the necessity rule, or KR, in the rest of the Chapter.

5 Three Arguments for KR

The first argument for KR is inspired by Moore’s paradox, but with ‘know’ substituted for ‘believe.’ Assertions of the form ‘ p , but I don’t know that p ’ (hereafter, Moorean assertions) seem incoherent. Appealing to KR allows us to offer a simple explanation. A Moorean assertion is proper only if (1) the speaker knows that p and (2) the speaker knows that she does not know that p . Since knowledge is factive, (2) is the case only if the speaker does not know that p . But this contradicts (1). KR thus entails that all Moorean assertions are improper.

The second argument for KR is that ‘How do you know?’ and ‘Do you know that?’ are appropriate ways to challenge an assertion. The first challenge presupposes that a speaker who asserts that p does know that p . Such a challenge would be inappropriate if someone who asserts that p were not expected to know that p . The second challenge is also easy to understand if KR is true: this challenge questions the speaker’s warrant for the assertion that p , which, according to KR, is tantamount to knowing that p .

The third argument concerns lottery cases. Suppose you hold a ticket in a fair lottery. After the winning ticket has been drawn but before it is announced, I tell you that your ticket did not win. If I do not have any inside information, my assertion seems improper, even though it is very likely to be true. This is easy to explain if KR is true: my assertion is improper because I do not know that your ticket did not win.¹⁴

A large part of the literature on the norms of assertion has focused on these arguments and on alternative accounts of the data they are based on.¹⁵ One

¹³ See Williamson (2005) for a defense of the sufficiency rule that maintains a classical invariantist position. For those who endorse contextualism, a relativized version of the knowledge norm is required: “A speaker, S , is well-enough positioned with respect to p to be able to properly assert that p if and only if S knows that p according to the standards for knowledge that are in place as S makes her assertion” (DeRose 2009, p. 99).

¹⁴ KR should be slightly amended: there should be an appropriate causal connection between the asserting and the knowing. As Williamson points out, “if I know p and assert p , but the asserting is causally independent of the knowing, then something is wrong with the asserting” (2000, p. 268, n. 16). See Turri (2011). In what follows, I will construe KR as requiring an appropriate causal connection between assertion and knowledge.

¹⁵ See, for instance, Weiner (2005) for an account based on the truth rule, and Douven (2006), Kvanvig (2011) and Lackey (2007) for accounts based on variants of the warrant rule, according

advantage of KR is that it offers simple and plausible explanations of the data. By contrast, alternative accounts are at best convoluted, and at worst inadequate. I do not have the space to review all such accounts here, but I will discuss one interesting proposal recently made by John MacFarlane (2011). MacFarlane contends that the truth rule TR can explain everything that KR does, provided that it is combined with the following principle, which I will call MacFarlane's belief principle, or MBP:

(MBP) One ought not believe that p when one knows that one does not know that p .

MacFarlane examines the three objections raised by Williamson against TR, and claims that they can be addressed if one also adheres to MBP.

First, as we saw in Sect. 3, Williamson claims that there are speech acts other than assertion, for example conjecturing, that are governed by the truth rule. But this is far from clear, writes MacFarlane: "If conjecturing were governed by the truth rule, it would be irresponsible to make conjectures one didn't have strong reason to think were true, and it isn't" (2011, p. 85). If conjecture were governed by the truth rule, then it would be associated with the following secondary norm: one should not conjecture that p unless one has a reasonable belief that one's conjecture is proper, that is, unless one reasonably believes that p . But as MacFarlane points out, this secondary norm is clearly too demanding. MacFarlane's point is well taken. However, Williamson can still insist that the truth rule does not differentiate assertion from swearing. Furthermore, even if conjecturing is not governed by the truth rule, it surely is governed by some rule. The most plausible candidate is something like 'One must: conjecture that p only if one has some evidence that p .' But then it is hard to see how proper conjecture could require the speaker to satisfy certain evidentiary standards while proper assertion does not.

Williamson also argues that TR cannot explain why we should not assert that one's lottery ticket will not win. MacFarlane replies that it can, if we invoke MBP. According to this principle, it is unreasonable for one to believe that p when p is a lottery proposition. But if one cannot reasonably believe that p , then one should not assert that p . Doing so would go against the secondary rule 'One should not assert that p if one does not reasonably believe that p .' TR, together with MBP, thus entails that assertions of lottery propositions are (secondarily) improper.

Williamson's third objection concerns Moorean assertions, that is, assertions of the form ' P , and I don't know that p .' Can TR explain their oddity? Yes, answers MacFarlane. Given MBP, one cannot reasonably believe the proposition *that p and that one does not know that p* , for one knows that one cannot know that

(Footnote 15 continued)

to which, roughly, one should assert that p only if one's belief that p is justified (or reasonable). In my view, many of the objections Williamson (2000) raises against these kinds of account still hold. See DeRose (2009, Chap. 3) for additional criticisms.

proposition. And according to the secondary rule associated with TR, one should not assert a proposition one cannot reasonably believe.

TR, in conjunction with MBP, can thus explain a lot of the data. Unfortunately, it is hard to believe that MBP is a basic principle, and MacFarlane does not explain what he bases his acceptance of MBP on. He does however mention that this principle can be derived from the knowledge norm of belief:

(KNB) One must: believe that p only if one knows that p .

If belief is governed by KNB, one should not believe that p if one knows that one does not know that p , for one's belief that p could not possibly satisfy KNB in such a case.

There are, however, difficulties with a position that endorses both TR and KNB, which I will now highlight.¹⁶ First, pace KNB, utterances of the form 'I believe that p , but I don't know that p ' are appropriate. One could respond that in these utterances, 'I believe that p ' conveys not that one outright believes that p , but merely that one gives some credence to p . But my intuition is that one can appropriately believe that Sally is in her office, while acknowledging that one does not know it.

A further problem concerns Gettier cases. Recall the example of the clock that stopped on 3 PM 24 h ago. Sally looks at the clock, and on this basis, asserts that it is 3 PM. As we saw, KR can explain the ambivalence we experience about Sally's assertion: although the assertion violates the primary rule, namely KR, the secondary rule is respected, for Sally's belief that it is 3 PM is reasonable. But TR does not yield the same "split" verdict. According to TR, Sally's assertion that it is 3 PM is proper, since it is true. And given that her belief that it is 3 PM is reasonable, she does not violate the secondary rule either. The mixed verdict KR allows seems preferable to the unmitigated ratification entailed by TR. Furthermore, the idea that there is nothing wrong with Sally's assertion is also hard to reconcile with what KNB entails about Sally's belief. According to this norm, Sally's belief that it is 3 PM is improper, even though she reasonably takes herself to know that it is 3 PM. The combination of TR with KNB thus offers a poor treatment of Gettier cases.¹⁷

¹⁶ MBP can also be derived from the weaker norm:

(RBKB) One must: believe that p only if one rationally believes that one knows that p .

But substituting RBKB for KNB avoids none of the problems I am about to signal.

¹⁷ RBKB does better here, for Sally rationally believes that she knows that it is 3 PM. Hence, according to this norm, her belief, just like her assertion, is proper. But this should not lead us to prefer RBKB over KNB. Consider the slightly different case in which Sally rationally believes that she knows that it is 3 PM, while it is in fact 2 PM. According to RBKB, Sally's belief that it is 2 PM is proper; however, TR entails that her assertion that it is 2 PM is improper. This is odd.

Third, consider what epistemic contextualists call “high-standards” contexts. In such contexts, because the stakes are high and error possibilities have been mentioned, speakers tend to claim that they ‘don’t know’ that p . According to contextualism, such knowledge denials are true, because speakers fail to satisfy the high epistemic standards in place in their contexts. But as Keith DeRose (2009, pp. 190–193) acknowledges, the contextualist’s argument rests on the assumption that in a high-standards context, the speaker’s belief that p is appropriate, despite the fact that she does not count as knowing that p in that context.¹⁸ KNB is thus at odds with contextualism.¹⁹ Now, MacFarlane endorses not contextualism, but relativism.²⁰ But relativists hold that relative to a high-standards context, a speaker’s assertion that she does not know that p is true, not because she inappropriately believes that p , but because she does not satisfy the high epistemic standards in place in her context. Hence, relativists like MacFarlane should be reluctant to endorse KNB.

Let us finally consider the kind of account of belief KNB requires. According to the *Lockean view of belief*, believing that p is having high credence for p , higher than some threshold lower than 1 but greater than 0.5. On this view, one’s belief that p is (epistemically) appropriate just in case it’s (epistemically) appropriate for one to have a degree of confidence in p that’s sufficient for belief.²¹ On the Lockean view, one may properly believe that one’s lottery ticket will not win. Hence, the proponent of KNB cannot endorse the Lockean view. A popular alternative is the *pragmatic view of belief*, according to which “one believes p outright when one is willing to use p as a premise in practical reasoning” (Williamson 2000, p. 99). On this view, belief is a form of “inner assertion,” or “assertion to oneself.”²² It is thus natural for proponents of this view to endorse the same norm for both belief and assertion. But then, it is unclear how MacFarlane could motivate KNB, or MBP for that matter, while avoiding a commitment to KR.

¹⁸ Note, though, that for DeRose (2009, p. 186–187), personal certainty (or high confidence) rather than belief is the attitude required for knowledge.

¹⁹ Brown (forthcoming) imagines a subject in a low-standards context who properly believes that p . Sally is at the train station and, after consulting the timetable, forms the true belief that the next train is an express. She is then approached by a stranger who tells her that it is extremely important to him whether the next train is an express or not. It seems that she is not in a position to properly assert that the next train is an express, even though she can still properly believe that.

²⁰ Like contextualism, relativism holds that knowledge claims have context-sensitive truth values; however, unlike contextualism, it denies that the *content* of a knowledge claim is context sensitive. According to relativism, the epistemic standards that are relevant to evaluating a knowledge claim are those in place in the context of assessment. Hence, the same knowledge claim may be assigned different truth-values in different contexts of assessment.

²¹ See, for instance, Foley (1993, Chap. 4) and Sturgeon (2008).

²² Williamson writes: “It is plausible [...] that occurrently believing p stands to asserting p as the inner stands to the outer. If so, the knowledge rule for assertion corresponds to the norm that one should believe p only if one knows p ” (2000, pp. 255–256). See also Adler’s (2002) assertion/belief parallel.

6 Is KR Constitutive?

In this section and the next two, I will argue that KR can plausibly be regarded as constitutive, individuating and basic. The first argument for KR provides good *prima facie* evidence for the claim that KR is constitutive. It seems that Moorean assertions are *always* incoherent: it always seems improper to add ‘I don’t know that p ’ to one’s assertion that p (unless, of course, one has changed one’s mind midway through one’s utterance).

However, the literature on the norms of assertion contains several alleged counterexamples to KR. Before examining some of these counterexamples, the following case, imagined by Williamson, is worth discussing:

I shout, ‘That is your train,’ knowing that I do not know that it is, because it probably is and you have only moments to catch it. Such cases do not show that the knowledge rule is not the rule of assertion. They merely show that it can be overridden by other norms not specific to assertion. The other norms do not give me warrant to assert p , for to have such warrant is to satisfy the rule of assertion (*ibid.*, p. 256).

KR, we have seen, is just one of the many rules that govern assertion. And as I have shown, this plurality is perfectly compatible with the thesis that KR is the single norm of assertion, when that thesis is properly understood. But with a plurality of rules, conflicts may arise. Williamson’s train case is an instance of such conflict. In this case, prudential considerations move the speaker to shout ‘That is your train,’ despite the fact that doing so violates KR. The speaker’s assertion is “all-things-considered” appropriate, because the potential negative consequences of the hearer’s missing the train outweigh the impropriety associated with a violation of KR.²³

Now, despite the fact that assertion is governed by many rules, it is governed by only one *epistemic* rule. So one could hold that although KR is defeasible—an assertion may be proper even though it violates KR—the following epistemic rule allows for no exception:

(KER) One’s assertion that p is epistemically proper only if one knows that p .

KER is how a number of commentators have proposed to read the knowledge rule.²⁴ I see no contradiction in endorsing both KR and KER, provided that we make their respective statuses clear. To repeat, although KR is the single norm of assertion (in the sense given in Sect. 4), it is not the only rule governing assertion and can thus be defeated by other rules. KER, on the other hand, is the indefeasible epistemic rule of assertion.

²³ Alternatively, one could hold that because of the speaker’s lack of knowledge, the assertion really is improper simpliciter; however, given the circumstances of the assertion, the speaker can be *excused* from this impropriety.

²⁴ See, for instance, Brown (2010), DeRose (2009) and Kvanvig (2011).

Let us now look at some alleged counterexamples, starting with Jennifer Lackey's cases involving what she calls 'selfless assertions.' One such case concerns a science teacher who believes in creationism. The teacher rejects evolutionary theory, but nevertheless recognizes that this theory is well supported by evidence. She decides to teach evolutionary theory to her class, because "she regards her duty as a teacher to include presenting material that is best supported by the available evidence" (Lackey 2007, p. 599). Lackey holds that the creationist science is "not subject to criticism in any relevant sense. Indeed, [...], *qua* asserter, [she] is *appropriately subject to praise*" (ibid., p. 599). Lackey takes this case to refute KR, because the science teacher properly asserts that *p* without believing and, a fortiori, without knowing that *p*.

But like Williamson's train example, Lackey's case shows not that KR is false, but that it can be outweighed by other norms. In this case, the teacher has (at least) two very good reasons to set aside her dogmatic beliefs: there is overwhelming evidence for evolutionary theory and she was hired to teach this theory to her students. We judge the creationist teacher's assertions to her students appropriate, because we think she makes the right decision: given her social role, teaching evolutionary theory to her students is what she should do. But just like in the train example, this does not entail that her assertions satisfy all the rules of assertion: the impropriety entailed by violating KR in this case is minor compared to the impropriety that would be involved with shirking her role as a teacher. Furthermore, we can maintain that her assertions are epistemically improper and thus violate KER, for she lacks knowledge of evolutionary theory.²⁵

Lackey raises another objection against KR based on the following examples:

[A] friend asserts that she is 28, when in fact she is 27 years, 364 days old; a contractor asserts that the room is 10 × 11 feet, when in fact it is 9 feet, 11 inches by 10 feet, 11 inches; a scientist asserts that the temperature of the water is 33.65 °C when in fact it is 33.65432 °C; a friend asserts that it is 6:00 PM, when in fact it is 6:01 PM, and so on. In all of these cases, the assertions in question are false because they are not absolutely precise, but they nonetheless seem perfectly proper (ibid., pp. 607–608).

Lackey ignores the fact that many of our utterances are loose, or non-literal. When one speaks loosely or non-literally, what one means differs from what the sentence one uses literally means. Typically, when one says, 'It's 6:00 PM,' one asserts not that it is 6:00 PM, but that it is *approximately* 6:00 PM. In other words, one speaks loosely: the content of one's assertion is not exactly the same as the literal content of the sentence one uses. Furthermore, all measurements are subject to uncertainty, and a scientist who reports on a measured value usually does not mean to specify this value exactly. In fact, it is inappropriate for a scientist to claim that the temperature of the water is '33.65 °C,' without admitting a certain uncertainty or margin of error. In a scientifically rigorous context, a scientist would

²⁵ Note that the problem concerns not the teacher's epistemic position (she is assumed to know both what evolutionary theory states and that this theory is well supported by evidence), but her disbelief.

be expected to specify this margin of error, and assert, for instance, that the temperature of the water is $33.65 \pm 0.01^\circ\text{C}$. Lackey's examples thus do not show that false assertions can be proper.

Matthew Weiner (2005) argues that some instances of prediction and retrodiction are counterexamples to KR. Sherlock Holmes looks at a crime scene and exclaims,

(4) This is the work of Professor Moriarty! It has the mark of his fiendish genius.

Weiner (*ibid.*, p. 231) holds that Holmes' assertion is proper, even though he does not know the asserted proposition. But Weiner's case is unconvincing. Proponents of KR can plausibly reply that Holmes is not asserting, but conjecturing that this is the work of Professor Moriarty.²⁶

Weiner writes that this response undermines the treatment of lottery cases advanced by proponents of KR. The latter would have to hold that in uttering 'Your ticket is a loser,' one is not making an assertion either. But the intuition in the lottery case is that I am in no position to flat-out assert that your ticket is a loser. 'It's probably a loser' and 'It's very likely a loser' seem appropriate things to say, while 'It's a loser' is not, unless I do something, such as adjusting my tone of voice, to weaken the assertoric force of my utterance. Things are similar with respect to Holmes' utterance. We charitably assume that Holmes could not be presenting himself as warranted in asserting the content of (4). We thus assume that something in his tone of voice or the context indicates a weakened assertoric force.

7 Is KR Individuating?

Let us now consider the claim that KR is individuating, that is, that no speech act other than assertion is governed by 'One must: perform speech act A with the content p only if one knows that p ' as a constitutive and basic rule. Although the three arguments for KR do not directly support it, this claim is very plausible. Speech acts that do not have assertoric force (e.g., commanding, promising, wishing) are clearly not governed by a knowledge rule. Let us thus consider speech acts that do have assertoric force (e.g., guessing, asserting, swearing). Swearing and asserting both aim at truth, but, as we saw in Sect. 3, they are subject to epistemic standards of differing strengths. The *basic* evidential norm of swearing is more stringent than that of assertion. There are thus good reasons to hold that KR is an individuating rule of assertion.

²⁶ Weiner's case is under-described. Perhaps Holmes has cleverly detected a clue that clearly incriminates Moriarty, in which case he would be entitled to assert the content of (4). But if we rule out scenarios such as this one, the claim that Holmes properly asserts (4) strikes me as implausible.

To be clear, none of this means that assertion is not governed by non-epistemic constitutive rules, such as moral and prudential rules, as well as the rule of relevance. But as I pointed out earlier, these rules are not individuating. The thesis that KR is the single rule of assertion is thus well supported, if we can show that KR is basic.

8 Is KR Basic?

I now want to examine the difficult question whether KR is a basic norm of assertion. Let us first consider the possibility that KR derives from the belief rule:

(BR) One must: assert p only if one believes that p .

Kent Bach writes:

[I]t is plausible to hold that knowledge is the norm on belief. If that is correct (I won't try to defend it here), then we can derive the knowledge rule on assertion by combining the knowledge norm on belief with the belief rule of assertion. This suggests that there is really nothing special about the knowledge requirement on assertion. It has no independent significance but is, rather, the combination of the knowledge norm on belief and the belief rule on assertion (2008, p. 77).

Bach favors what we may call a 'speech-act account' of norms: types of speech acts are individuated in terms of types of expressed attitudes.²⁷ Since the role of an assertion is to express a belief,²⁸ one should not assert that p if one does not believe that p . Hence, on Bach's account, BR is the basic rule and KR derives from it, because belief itself is governed by the knowledge norm:

(KNB) One must: believe that p only if one knows that p .

I have already raised doubts about KNB in Sect. 5. But even if we grant KNB, this derivation of KR does not work: BR, together with KNB, does not entail KR. The reason is simple: we can accept both BR and KNB without accepting KR. Suppose one asserts that p , based on one's belief that p , but that belief is false or unjustified (or Gettierized). In other words, one does not know that p . If both BR and KNB hold, one's assertion is proper, even though one's belief is improper. However, since one does not know that p , one's assertion that p would be improper according to KR. The impropriety that KR assigns to the assertion is thus not predicted by BR and KNB. Hence, contrary to Bach's suggestion, KR cannot be

²⁷ See Bach and Harnish (1979, Chap. 3).

²⁸ This is, of course, just a partial account of assertion: presumably, assertion is associated with communicative intentions as well. See, for instance, Bach and Harnish (1979). Note that they construe 'express' slightly differently than I do. On their account, in asserting that p one (among other things) expresses the belief that p , where to express a belief is (roughly) to communicatively intend the hearer to take one's utterance as reason to think one has that belief. This difference does not affect the present point, since they agree that an insincere assertion is improper.

derived from BR, even if we assume KNB. Of course, a speaker who violates KR necessarily violates BR or KNB. But this does not show that KR derives from BR and KNB. The problem is that the impropriety involved in a violation of BR or KNB does not necessarily concern assertion; it may merely concern belief.

To avoid this problem we need to tie the propriety of assertion to the propriety of belief. In other words, we ought to adopt what we may call the *proper belief rule*:

(PBR) One must: assert p only if one properly believes that p .

But if we accept KNB, that is, if proper belief is knowledge, then PBR is just a different statement of KR. In such a case, KR does not derive from PBR.

KR does not derive from BR and KNB; however, it may be derivable from other principles. Consider the claim that knowledge is a norm for practical reasoning:

(KPR) One must: rely on p in practical reasoning only if one knows that p .²⁹

Hawthorne (2004, pp. 29–30) imagines a subject who is offered one penny for his lottery ticket, and reasons as follows:

1. My ticket will not win.
2. If I keep my ticket, I will win nothing.
3. If I sell my ticket, I will get a penny.
4. Hence, I should sell my ticket.

This reasoning is unacceptable, plausibly because the subject does not know the first premise. This supports the claim that practical reasoning is governed by a knowledge norm. Hawthorne and Stanley (2008) also point out that it is appropriate to criticize a person for acting on the basis of proposition p , if she did not know that p .

If KR and KPR both hold, then there is a common epistemic standard governing both assertion and practical reasoning. Following Brown (forthcoming), call this *commonality*. It would be strange if commonality were just a coincidence: if commonality is true, it is plausible to expect an explanation for it. This explanation may well entail that KR is not basic, in the sense that it is ultimately explained by some other principles governing cognitive states and activities.³⁰ But this should

²⁹ See, for instance, Hawthorne (2004) and Williamson (2005, p. 231). I take Hawthorne and Stanley's (2008) reason-knowledge principle ('Where one's choice is p -dependent, it is appropriate to treat the proposition that p as a reason for acting iff you know that p ') to be in the same spirit.

³⁰ In my (2013), I show that KR derives from KPR and what I call the *manifestation rule*, that is, 'One must: assert that p only if that assertion manifests the belief that p .' I also explain how this derivation avoids the objections Brown (forthcoming) raises against other explanations of commonality.

not necessarily trouble proponents of the knowledge account of assertion, for they can take the term ‘basic’ to have a different sense. Remember that in Sect. 3, I pointed out that TR is not an individuating norm of assertion, for it derives from KR. TR is thus not basic in the sense that it is not the strictest rule governing assertion: it is entailed by another norm governing assertion, namely KR. This is, it seems to me, the sense of ‘basic’ that is relevant to our purposes: a norm for speech act A is basic (in that sense) just in case it does not follow solely from other norms governing A. The basic norm for A is thus simply the strictest norm governing A: non-basic norms are entailed by it. This means that a speech act norm such as KR can be basic, even though it is ultimately explained by some other norm governing cognitive states and activities.

9 Conclusion

I have argued that the knowledge rule ‘One must: assert that p only if one knows that p ’ is the single norm of assertion. This, I have explained, means that the knowledge rule is the only individuating, constitutive and basic rule of assertion. It is constitutive in the sense that it applies to every possible assertion. It is basic in the sense that it does not derive from other norms governing assertion. And it is individuating in the sense that no speech act other than assertion is governed by ‘One must: perform speech act A with the content p only if one knows that p ’ as a constitutive and basic rule.

I have explained how the arguments for the knowledge rule provide good support for the claim that this rule is the single norm of assertion. I have remained neutral on three controversial issues: whether other rules governing assertion (such as moral and prudential rules) are constitutive; whether the sufficiency rule ‘One is positioned to make an epistemically proper assertion that p if one knows that p ’ is constitutive; and whether the knowledge rule is ultimately explained by a broader rule governing other cognitive activities. None of these issues need to be resolved to support the thesis that the knowledge rule is the single norm of assertion.³¹

References

- Adler, Jonathan. 2002. *Belief's own ethics*. Cambridge (MA): MIT Press.
- Bach, Kent. 2008. Applying pragmatics to epistemology. *Philosophical Issues* 18: 68–88.
- Bach, Kent, and Robert Harnish. 1979. *Linguistic communication and speech acts*. Cambridge (MA): MIT Press.
- Brown, Jessica. 2010. Knowledge and assertion. *Philosophy and Phenomenological Research* 81: 549–566.

³¹ I am grateful to Sherri Irvin for her useful comments on an earlier draft.

- Brown, Jessica. Forthcoming. Assertion and practical reasoning: Common or divergent epistemic standards? *Philosophy and Phenomenological Research*.
- DeRose, Keith. 2009. *The case for contextualism: Knowledge, skepticism, and context*, vol. 1. Oxford: Clarendon Press.
- Douven, Igor. 2006. Assertion, knowledge, and rational credibility. *Philosophical Review* 115: 449–485.
- Foley, Richard. 1993. *Working without a net*. Oxford: Oxford University Press.
- Hawthorne, John. 2004. *Knowledge and lotteries*. Oxford: Oxford University Press.
- Hawthorne, John, and Jason Stanley. 2008. Knowledge and action. *Journal of Philosophy* 105: 571–590.
- Kvanvig, Jonathan. 2011. Norms of assertion. In *Assertion: New philosophical essays*, eds. J. Brown, and H. Cappelen, 233–250. Oxford: Oxford University Press.
- Lackey, Jennifer. 2007. Norms of assertion. *Noûs* 41: 594–626.
- Lackey, Jennifer. 2011. Assertion and isolated secondhand knowledge. In *Assertion: New philosophical essays*, eds. J. Brown, and H. Cappelen, 251–276. Oxford: Oxford University Press.
- MacFarlane, John. 2011. What is assertion? In *Assertion: New philosophical essays*, eds. J. Brown, and H. Cappelen, 79–96. Oxford: Oxford University Press.
- Montminy, Martin. 2013. Why assertion and practical reasoning must be governed by the same epistemic norm. *Pacific Philosophical Quarterly* 94: 57–68.
- Sturgeon, Scott. 2008. Reason and the grain of belief. *Noûs* 42: 139–165.
- Turri, John. 2011. The express knowledge account of assertion. *Australasian Journal of Philosophy* 89: 37–45.
- Unger, Peter. 1975. *Ignorance*. Oxford: Clarendon Press.
- Weiner, Matthew. 2005. Must we know what we say? *Philosophical Review* 114: 227–251.
- Williamson, Timothy. 1996. Knowing and asserting. *Philosophical Review* 105: 489–523.
- Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford: Oxford University Press.
- Williamson, Timothy. 2005. Contextualism, subject-sensitive invariantism, and knowledge of knowledge. *Philosophical Quarterly* 55: 213–235.

From Thought Experiments to Real Experiments in Pragmatics

András Kertész and Ferenc Kiefer

Abstract The puzzle of thought experiments is a hot topic in the philosophy of science. The chapter raises the puzzle with respect to pragmatics as follows: How is it possible that thought experiments in pragmatics yield new experiential information about communication, although they are carried out entirely in one's head? The chapter shows, first, that the structure of thought experiments in pragmatics consists of a series of plausible inferences. Second, the function of thought experiments is to serve as the initial step in the process of plausible argumentation as well as to test the plausibility of rival hypotheses. Third, while on the one hand, thought experiments and real experiments may be continuous, on the other hand, the former may be also indispensable components of the latter. Fourth, these properties provide a solution to the puzzle of thought experiments in pragmatics. The key idea of the solution is that thought experiments in pragmatics cannot generate new experiential information; rather, during the process of plausible argumentation they contribute to the retrospective re-evaluation of experiential information already given.

1 Introduction

The investigation of the relationship between thought experiments and real experiments in pragmatics is motivated by the state of the art in two different fields of inquiry.

A. Kertész

Research Group for Theoretical Linguistics, Hungarian Academy of Sciences,
Debrecen, Hungary

e-mail: andraskertes@ yahoo.de

F. Kiefer (✉)

Research Institute for Linguistics, Hungarian Academy of Sciences,
Benczúr utca 33 H-1068 Budapest, Hungary

e-mail: kiefer@nytud.hu

The first is, of course, *pragmatics*. Currently, there is a growing interest in the application of experimental methods.¹ The use of real experiments in pragmatics is especially interesting, if one relates it to the *philosophical* origins of the discipline. In particular, Grice's ideas put forward in his seminal papers, Searle's widely discussed arguments for his account of illocutionary acts, and many other classical contributions to pragmatics have been based on invented 'stories' supporting the hypotheses which fertilized the development of this field. The way some authors characterise these stories highlights features which are typical of thought experiments. For example, Meibauer (2012: 769) maintains that "pragmatic evidence is more often than not construed as a *story*, the story reflecting *pragmatic intuitions* of a speaker or hearer" (emphasis added). Noveck and Sperber (2004a, b: 8) maintain that "the only source of *evidence*" which those philosophers and linguists who developed pragmatics "have ever used has been their own intuition about how an *invented* utterance *would be* interpreted in a *hypothetical situation*" (emphasis added). The features mentioned in these quotations—narrative structure, intuition manifested in stories, hypothetical situation, invented case, the stories' being used as 'evidence'—strongly resemble those features which the philosophical literature attributes to thought experiments. Therefore, we assume that the hypothetical, imaginary situations which classical contributions to pragmatics make use of are interpretable as thought experiments. Then, the historical process that led from Grice's, Searle's, Austin's, Anscombe's and others' philosophical accounts of language use to the current state of the art in pragmatics, is associated with the transition from the exclusive use of thought experiments to a richer methodology which, besides retaining thought experimentation, includes the increasing application of real experiments as well (cf., for example, Noveck and Sperber 2004a, b, Clark and Bangartner 2004, Noveck and Reboul 2008, etc.).

The second field is *the philosophy of science* which during the past two decades has seen a lively discussion on the structure and function of scientific and/or philosophical thought experiments. It has been realised that thought experiments had already been applied by the Pre-Socratics and later triggered many important insights such as Galileo's, Einstein's, Maxwell's, Schrödinger's discoveries, etc., thus significantly contributing to the development of modern science.² The problem of the experiential content of thought experiments and their relationship to real experiments is one of the central topics of current discussions in the philosophy of science. The reason why thought experiments are subject to heated debates and why they are currently at the centre of interest is that they raise a seemingly mysterious puzzle. As Horowitz and Massey (1991a, b: 1) put it, "[...] thought experiments often have *novel empirical import* even though they are

¹ For state-of-the-art overviews of current approaches to experimental pragmatics see e.g. Breheny (2011), Katsos and Cummins (2010), Noveck and Reboul (2008), Noveck and Sperber (2004a, b), Noveck and Sperber (2007), Németh T. (2006), Sauerland and Yatsushiro (2009), Meibauer and Steinbach (2011), Németh T. and Bibok (2010).

² For discussion see Brown and Fehige (2010), Rescher (2005), Sorensen (1992), Kühne (2005), Buzzoni (2008), Cohnitz (2006), Brown (1991), Gendler (2000), Cohnitz (2006).

conducted entirely inside one's *head* [...]” (emphasis added). They speak of “Kuhn's paradox”, because Kuhn (1977: 241) raised it in a pointed way.³ Norton (2004a: 44) labels the same issue “the epistemological problem of thought experiments”; Kühne (2005: 25) calls it “the paradox of thought experiments”. For the sake of later references, we formulate the puzzle as follows⁴:

The puzzle of thought experiments: How is it possible that thought experiments yield new experiential information about the world even though they are carried out in one's head?

Against this background, the aim of the present chapter is to contribute to the clarification of the nature of thought experiments in pragmatics. However, due to the extreme methodological pluralism of pragmatics, our line of reasoning will have clear limits. First, we are compelled to discuss the problems we will raise on a high level of abstraction. Second, for lack of space, we cannot capture the whole diversity of thought experiments in pragmatics; therefore, we will illustrate our main tenets by a classical example taken from the philosophical origin of pragmatics. We will not be able to discuss the similarities and differences between the example chosen other kinds of thought experiments. Third, we will thereby focus on the issue of ‘what is said’ and ‘what is implicated’, because current trends in experimental pragmatics predominantly, though not exclusively, are directly or indirectly rooted in Gricean ideas. Fourth, the issue of ‘what is said’ and ‘what is implicated’ will motivate the choice of the examples by means of which we will illustrate the relationship between thought experiments and real experiments.

In order to achieve the aim we have outlined, we need a metatheoretical framework. Therefore, our first problem to be solved will be:

(P1) What metascientific framework is suitable for the analysis of thought experiments in pragmatics?

We will start our line of reasoning in Sect. 2 by giving a brief overview of current trends in the philosophy of science which facilitate the investigation of the nature of thought experiments. From this overview, we will derive our framework, which we will apply to thought experiments in pragmatics in order to obtain the solutions to the problems (P2)–(P5) as raised below.

In Sect. 3 we will begin the application of the framework with the following problem:

³ “How, [...] relying upon familiar data, can thought experiment lead us to new knowledge or to new understanding of nature?” (Kuhn 1977: 241).

⁴ The terminology in the literature on thought experiments is vague. For example, the terms ‘empirical’ and ‘experiential’ are not defined, and very often they are used interchangeably. Since the definition of ‘empirical’ is highly controversial and we cannot discuss it here, we will use the notion ‘experiential’ in a pre-explicative sense. We will apply this term to pieces of information which are rooted in a person's ‘experiences’ whatever the latter may be. For example, the results of real experiments or communicative situations in which one has already taken part, count as experiences and statements about such experiences are called ‘experiential’.

(P2) What is the structure of thought experiments in pragmatics?

As an example, we will analyse one of Grice's classical thought experiments. Applying our metascientific framework, we will arrive at a solution to (P2).

Presupposing our solution to (P2), in [Sect. 4](#) we will deal with the function of thought experiments:

(P3) What is the function of thought experiments in pragmatics?

Based on the solution to (P2) and (P3), in [Sect. 5](#) we will relate thought experiments to real ones:

(P4) What relationship is there between thought experiments and real experiments in pragmatics?

Having solved (P4) as well, we will be in a position to consider the puzzle of thought experiments with respect to pragmatics:

(P5) How do the solutions to (P2)–(P4) make it possible that thought experiments yield new experiential information about communication even though they are carried out in one's head?

Our solution to (P5) will partly integrate some of the insights discussed in the general philosophical literature and partly go beyond this literature.

[Section 7](#) will pave the way for future research.

2 On (P1): Thought Experiments and the Philosophy of Science

2.1 Introductory Remarks

In this section, we will set out to find a suitable metascientific framework capable of providing solutions to the problems (P2)–(P5). Before doing so, it should be made clear that there seems to be agreement that striving for a generally acceptable definition of the notion of thought experiment would be a hopeless endeavour (Kühne 2005). Accordingly, we will not attempt to give a general definition of thought experiments, either. Rather, we will propose a metascientific framework, the application of which will be expected to reveal at least some constitutive properties of particular thought experiments in pragmatics.

In [Sect. 2.2](#) we will briefly summarise those trends in the philosophy of science which led to the current interest in the puzzle of thought experiments. In [Sect. 2.3](#) we will have a look at an account of thought experiments in linguistics. In [Sect. 2.4](#) we will introduce the notions of plausible inference and plausible argumentation. Based on these notions, we will summarise our solution to (P1) in [Sect. 2.5](#).

2.2 *The State of the Art in the Philosophy of Science*

As regards the history of the philosophical reflection on the nature of thought experiments, there were early attempts made at the clarification of their structure and function in the late 18th and early 19th century. Kühne starts his overview of the state of the art with Immanuel Kant's *Kritik der reinen Vernunft* (Kühne 2005: 95 ff.) and Hans Christian Ørsted's *Experimenta circa effectum conflictus electrici in acum* (Kühne 2005: 105 ff.). Ernst Mach is considered the first 20th century thinker to call attention to the significance of thought experiments from the point of view of the philosophy of science (Mach 1960, 1976). However, the then prevailing perspective of logical positivism—which, following Reichenbach's distinction, focused on the 'context of justification' and was not interested in the 'context of discovery'—did not permit the realisation of the relevance of thought experiments for theory formation (cf. Moue et al. 2006: 64 ff, Kühne 2005: 19 f., Buzzoni 2008: 12 ff.). Thereby, it was the decline of logical positivism—decisively influenced by Kuhn's (1970) work—that, among others, facilitated the emergence of a series of tendencies within the philosophy of science and epistemology which paved the way for the systematic reflection on thought experiments:

- i. Due to Kuhn's (1970) impact, the relevance of the history of science for the philosophy of science was realised and, accordingly, issues were focused on which were very different from the rational reconstruction of scientific theories within the context of justification as advocated by logical positivism.
- ii. Related to the Kuhnian turn, the investigation of the context of discovery which logical positivism had excluded from the domain of the philosophy of science was, from the 1970s on, also identified as a substantial subject matter of the philosophy of science (Nickles 1980).
- iii. Finally, besides the shift of focus in the philosophy of science, new developments in psychology, cognitive science, logic and artificial intelligence research also contributed to the emergence of the process of scientific problem solving as a new topic of metascientific reflection.

Consequently, one may "trace the 'career' of thought experiments from almost total disregard to the point where thought experiments have been seen as a prominent methodological tool for 'actual thinking'" (Moue et al. 2006: 62).

In the past decades, numerous solutions to the puzzle of thought experiments have been suggested in the philosophical literature,⁵ but in the course of current debates none of them has turned out to be significantly more convincing than the others. The main argument often invoked against most of them (e.g. against Brown 1991; Norton 2004a, b; Gooding 1990; Kuhn 1977; Sorensen 1992) in the debates is that they focus on very specific kinds of thought experiment taken from the

⁵ See Brown (1991), Kuhn (1977), Rescher (1991), (2005), Gendler (2000), Norton (2004a, b), Irvine (1991), Häggquist (1996), Cohnitz (2006), Nersessian (1993), Mišćević (1992), Cooper (2005), Gooding (1990), McAllister (1996).

natural sciences, but at the same time they commit the fallacy of hasty generalisation. In contrast, the workability of a few other approaches (Nersessian 1993; Mišćević 1992; Cooper 2005) is questionable, because they are too general and ‘empty’ insofar as they do not define the means with the help of which particular thought experiments can be analysed. Due to these problems, none of the philosophical solutions to the puzzle of thought experiments can be adapted to pragmatics as the *sole* framework for the solution of our problems (P2)–(P5).

2.3 Thought Experiments in Linguistics

Thomason (1991) discusses thought experiments in different fields of linguistics, but she does not mention examples from pragmatics. She distinguishes between two kinds: stage setting thought experiments and introspection. She characterises stage setting thought experiments as follows⁶:

1. (a) “In fact, the result of this sort of linguistic thought experiment is likely to be theory-dependent. The thought experiment will therefore not be a test of a hypothesis, but rather a stage-setting device that suggests tests that the linguist can carry out. In other words, the major role of the thought experiment is to clarify the theoretical issue, or to make it vivid, *as a first step in an argument.*”
- (b) Often the thought experiment serves to get the audience’s agreement in advance about what would count as *supporting evidence* for the theory, even if that exact kind of evidence is not going to be forthcoming.
- (c) The second step—and it is a necessary one if the argument is to be successful—is a demonstration that some real-world situation is sufficiently *similar* to the result of the thought experiment that other linguists will accept that situation as supporting evidence.” (Thomason 1991: 247–248; emphasis added).

(1)(a) defines the *function* of stage setting thought experiments. It is important to emphasise that Thomason conceives of stage setting thought experiments in linguistics as components of an *argumentation process*.

(1)(b) suggests that stage setting thought experiments *do* have evidential significance, although—in connection with their being the first, preparatory step in an argumentation process—only to a restricted extent.

Thomason does not discuss the puzzle of thought experiments explicitly, but (1)(c) seems to open the possibility of a solution. In our interpretation (1)(c) says

⁶ We will divide the quotations into smaller units which makes it easier to refer to parts of the quotation to be analysed. The units do not necessarily correspond to paragraphs in the original text.

that an *analogical inference* can be constructed, one premise of which assumes the similarity between a real-world situation and the imaginary one described in the thought experiment. Via this premise the real-world situation supports the conclusion which corresponds to a particular hypothesis.

Introspective thought experiments are different:

2. (a) “Let us turn now to the other kind of linguistic thought experiment—*the kind that involves introspection*, by the linguist or by an informant (a native speaker of some language the linguist is investigating), about the appropriateness of a particular linguistic form or construction.
- (b) Thought experiments of this type are *actual tests of hypotheses* about language structure.” (Thomason 1991: 252–253; emphasis added)

(2)(a) characterises the *structure* of this kind of thought experiment, while (2)(b) contrasts the *function* of introspective thought experiments with that of stage setting ones.

2.4 Plausible Inferences and Plausible Argumentation

We think that a possible framework applicable to the solution of (P2)–(P5) should comprise the following components:

- i. *Stage-setting versus introspective thought experiments.* We basically accept Thomason’s (1991) distinction between stage setting and introspective thought experiments in linguistics. In the course of our reasoning we will have to check whether these two kinds are really distinct, or whether there are cases in which they are combined.
- ii. *The ‘What if?’ question.* Most authors agree that thought experiments are basically *conditional*. In accordance with Cooper (2005), Gendler (1998), (2000), Rescher (2005) and others, we will assume that a thought experiment supports a hypothesis about what the world would be like *if* the imaginary situation described were real. However, we have to specify *by what particular means* the answer to the ‘What if?’ question is obtained.
- iii. *Plausible inferences.* We assume that the link between the ‘What if?’ question and the answer to this question is provided by *plausible inferences*. This assumption can be motivated by the fact that in the philosophical literature there is a group of approaches which consider thought experiments to be reconstructed as inferences (Rescher 1991, 2005; Norton 2004a, b; Irvine 1991; Häggquist 1996; Cohnitz 2006). However, most of the inferences that thought experiments consist of, lead to conclusions which, instead of being true with certainty, are only plausible to a certain extent. As Rescher (2005: 34) points out, “[...] thought experimentation [...] is not an instrumentality of *demonstration* but rather merely one of *plausibilification*: it is a tool not of *proof* but of *plausible reasoning*” (emphasis as in the original). The pioneering

works on plausible inferences were put forward by George Polya, Douglas Walton and Nicholas Rescher (cf. e.g. Polya 1948, 1954; Rescher 1976, 1987; Walton 1992, 2001; Walton et al. 2008). In the present chapter we will adopt Kertész and Rákosi's (2012) approach to plausible inferences and plausible argumentation. Below we sketch the notions we need in a highly simplified, informal manner.⁷

The essence of plausible inferences can be best understood by contrasting them with deductive and demonstrative inferences. The conclusion of a deductive inference is true *provided that* the premises are true. Demonstrative inferences have an additional feature: besides being deductively valid, they rest on premises which one *knows to be true with certainty*. Therefore, their conclusion is *certain*, too (Polya 1948, 1954). As opposed to this, the main characteristic of plausible inferences is that they infer *plausible* conclusions from the premises. They are uncertain and fallible although they have heuristic power.

With respect to the structural relation between the premises and the conclusion, there are three basic types of plausible inference. In the first type at least one of the premises is not true but only plausible. Therefore, the conclusion cannot be certain, but is only plausible, too. However, it may be the case that the premises are true, but there is no logical consequence relation between them and the conclusion. Accordingly, in the second type *latent background assumptions* are needed which work as *hidden premises* and make it possible to transform the given inference-candidate into an inference in which there is a logical consequence relation between the premises and the conclusion. Such inferences are called *enthymematic*. This transformation consists in the extension of the set of the premises *explicitly given* by such latent background assumptions.⁸ Finally, the characteristics of these two types may be present *simultaneously*: there is no logical consequence relation between the explicit premises and the conclusion, and in addition, at least one of the explicit premises or latent background assumptions is not true with certainty but only plausible.

- iv. *Plausible argumentation*. The application of the notion of plausible argumentation can be motivated, among others, by Thomason's (1991) claim that stage-setting thought experiments in linguistics are parts of *argumentative processes*. However, we introduce the notion of plausible argumentation in a specific sense related to that of plausible inference, in accordance with Kertész and Rákosi (2012). Our point of departure is the fact that thought experiments are undoubtedly tools of *problem solving*. An informational state may be

⁷ For their precise definitions and numerous applications to different linguistic theories see the work mentioned. A first attempt to apply this approach to thought experiments is Kertész (2010).

⁸ A well-known example is inductive inferences, in which the set of the explicit premises has to be supplemented by the background assumption that the cases not examined also possess the characteristics that could be found in the investigated ones. The conclusion states the presence of these characteristics as a generalisation.

problematic in three ways. First, it may be informationally overdetermined. Informational *overdetermination* means that a certain statement is made plausible by a given source or an inference while its negation is also made plausible by some other source or inference; accordingly, inconsistency emerges. Second, the informational state at issue may be *underdetermined*, if there are statements which are neither plausible nor implausible. Third, it may be both over- and underdetermined with respect to different statements.

We call the heuristic device whose application aims at the elimination of the under- and/or overdetermination, *plausible argumentation*. It consists of a sequence of plausible inferences. During the process of plausible argumentation, by means of the sequence of plausible inferences, a problematic informational state is continuously re-evaluated by the elaboration of possible solutions to the given problems, the evaluation of the alternative solutions and the comparison of the latter. Since heuristics do not necessarily lead to the solution of a problem and may fail, plausible argumentation is fallible, too.

The process of plausible argumentation is not linear. It is *cyclic* in that during the argumentation process one returns to previous informational states and *retrospectively re-evaluates* former findings. In the course of such cycles the retrospective re-evaluation of the findings is—to use Rescher’s (1987) term—*prismatic* in the sense that it is carried out from continuously changing perspectives during which new information is considered, and/or earlier findings are modified, deleted or supplemented by additional assumptions, etc. It is important to emphasise that cyclic argumentation in this sense must be distinguished from circular argumentation. Circular argumentation (*circulus vitiosus*) returns to the starting informational state, therefore it is uninformative. As opposed to this, cyclic argumentation is effective, because it feeds back to a re-evaluated information content.⁹

2.5 The Solution to (P1)

We summarise the solution to (P1) which we propose as (SP1):

SP1. The metatheoretical framework for thought experiments which is meant to provide a possible solution to (P2)–(P5) includes

- (a) ‘What if?’ question;

⁹ “The sort of ‘self-criticism’ at issue does not reflect any vicious or vitiating circularity, but in effect amounts simply to a *feedback* process that uses later, more refined stages of the analysis to effect revisionary sophistications in the materials from which earlier stages proceeded. One indeed returns to ‘the same point’ but does so *at a different cognitive level*” (Rescher 1976: 119; emphasis as in the original).

- (b) the distinction between stage setting and introspective thought experiments;
- (c) plausible inferences as the tools of providing a possible answer to the ‘What if?’ question; and
- (d) the cyclic, prismatic and retrospective re-evaluation of information by means of plausible inferences which constitute the process of plausible argumentation.

3 On (P2): The Structure of Thought Experiments in Pragmatics

3.1 An Example: Grice on ‘What is Said’ and ‘What is Implicated’

In order to show how our framework contributes to a possible solution of (P2), we choose as an example the first thought experiment in Grice (1989a) that raised the problem of the relationship between ‘what is said’ and ‘what is implicated’. Since this is even today one of the most widely discussed central quandaries of pragmatics and its consequences reach as far as present-day experimental pragmatics, it is well suited to illustrate both the impact of thought experiments and the process of plausible argumentation that leads from thought experiments to real experiments in pragmatics.

Grice’s famous thought experiment is as follows:

3. (a) “Suppose that A and B are talking about a mutual friend, C, who is now working in a bank. A asks B how C is getting on his job, and B replies, *Oh quite well, I think; he likes his colleagues, and he hasn’t been to prison yet.* At this point, A might well inquire what B was implying, what he was suggesting, or even what he meant by saying that C had not yet been to prison. The answer might be any one of such things as that C is the sort of person likely to yield to the temptation provided by his occupation; that C’s colleagues are really very unpleasant and treacherous people, and so forth. It might, of course, be quite unnecessary for A to make such an inquiry of B, the answer to it being, in the context, clear in advance. It is clear that whatever B implied, suggested, meant in this example, is distinct from what B said, which was simply that C had not been to prison yet.
- (b) It is clear that whatever B implied, suggested, meant in this example, is distinct from what B said, which was simply that C had not been to prison yet.
- (c) I wish to introduce, as terms of art, the verb *implicate* and the related nouns *implicature* (cf. *implying*) and *implicatum* (cf. *what is implied*). [...]” (Grice 1989a: 24.)

At first sight (3) seems to be a stage setting thought experiment in the sense of Sect. 2.3.¹⁰ Namely, (3) is the first step in the argumentation process starting with the section entitled *Implicature* in Grice's paper in that it raises the theoretical problem of the distinction between 'what is said' and 'what is implicated'. In this respect, the informational state at which (3) appears is underdetermined and thus problematic in the sense of (iv) in Sect. 2.4. In addition, the thought experiment in (3) aims at convincing the reader that this distinction and the introduction of the notion of implicature are well-motivated. Finally, the thought experiment is expected to be supplemented by further steps in later stages of the argumentation process which partly continue the argumentation initiated by the thought experiment, and partly re-evaluate its outcome retrospectively.

Our framework also suggests that Grice's thought experiment seeks an answer to the following 'What if?' question:

4. What would be the case if the imaginary state of affairs described in (3)(a) were actual?

(3)(a) describes a hypothetical, imaginary situation which is analogous to the well-known and much analysed paradigmatic examples of classical thought experiments discussed in the philosophy of science. It gives the following answer to (4):

5. If the imaginary state of affairs described in (3)(a) were actual, then the hypothesis would be plausible that the distinction in (3)(b) and the notion of implicature in (3)(c) are adequate.

We assume that in accordance with (SP1)(c) thought experiments in pragmatics can be decomposed into a series of plausible inferences. For the sake of illustration, we present a very simplified and merely partial reconstruction of the plausible inferences underlying the quotation in (3), emphasising that Grice's argumentation in the first and second paragraph of the section entitled *Implicature* (Grice 1989a: 24–25) is considerably more complicated. The relationship between (3)(a) and (3)(b) rests on the following inference¹¹:

6. Premises:

- (a) It is certain that by uttering *Oh quite well, I think; he likes his colleagues, and he hasn't been to prison yet* B said that C was not in prison yet.

¹⁰ However, we will see later that this impression needs to be refined.

¹¹ Latent background assumptions are set within '<' and '>'. Moreover, we accept 'the weakest link principle' which says that the conclusion of a plausible inference takes its plausibility value from that of the 'weakest' explicit premise or latent background assumption (see e.g. Rescher 1976: XI f., 15, Kertész and Rákosi 2012). We also remark that the expressions 'it is plausible that' and 'it is certain that' merely indicate a very simplified representation of plausible statements. Kertész and Rákosi (2012) uses numerical scales of plausibility values whose introduction in the present paper would be beside the point.

- (b) It is plausible that by uttering *Oh quite well, I think; he likes his colleagues, and he hasn't been to prison yet* B implied, suggested, meant that “C is the sort of person likely to yield to the temptation provided by his occupation, that C's colleagues are really very unpleasant and treacherous people, and so forth” (Grice 1989a: 24).
- (c) <It is plausible that “what someone has said” is “closely related to the conventional meaning of the words (the sentence) he has uttered”.> (Grice 1989a: 25)
- (d) <It is plausible that what someone implies or suggests or means by uttering something, is distinct from the conventional meaning of the words he utters.>
- (e) <It is plausible that if the conjunction of (a), (b), (c) and (d) is given, then “whatever B implied, suggested, meant in this example, is distinct from what B said, which was simply that C had not been to prison yet” (Grice 1989a: 24)>.

Conclusion:

It is plausible that “whatever B implied, suggested, meant in this example, is distinct from what B said, which was simply that C had not been to prison yet” (ibid.).

It is the plausible inference in (7) that underlies the relationship between (3)(b) and (3)(c):

7. Premises

- (a) <It is plausible that if two things have to be distinguished, and at least one of them has no name, then a new name for the latter has to be introduced.>
- (b) It is plausible that what B implied, suggested, meant in this example, is distinct from what B said [...].
- (c) <It is certain that what B implied, suggested, meant in this example, has no name.>

Conclusion:

It is plausible that a new name for what B implied, suggested, meant in this example, has to be introduced.¹²

To show how these inferences shape the structure of the thought experiment itself as well as that of the process of plausible argumentation, which they are part of, we add a few comments. Let us first reflect on the relevant properties of the plausible inferences in (6) and (7):

- i. It is important to realise that except for (6)(a) the explicit premises and the latent background assumption of the inference in (6) are merely plausible. For example, the fact that (6)(c) and (d) are not true with certainty, but only plausible, is witnessed by the discussions that have taken place on the relationship between ‘what is said’ and ‘what is implicated’ during the past two decades (see also Sect. 5).

¹² The pattern of both inferences corresponds to that of plausible *modus ponens*: {it is plausible that if A then B; A is plausible} \Rightarrow B is plausible.

- ii. Both inferences are enthymematic. Beside the explicit premises, they also include latent background assumptions which have not been introduced explicitly, but without which the conclusion could not have been obtained. Consequently, (6) and (7) belong to the third type of plausible inference in that they are enthymematic and at least one of their explicit premises or latent background assumptions is not true with certainty. It is worth remarking that “[o]ne problem with enthymemes is that reasonable people can have differences of opinion on what the implicit assumptions are supposed to be. Filling in the missing parts of an enthymeme may depend on interpreting the natural language text in which the argument was put forward [...]” (Walton et al. 2008: 189).¹³ This problem with enthymemes may be one of the reasons why Grice’s distinction between ‘what is said’ and ‘what is implicated’ raised so much controversy and triggered so many interpretations. (In Sect. 5 such interpretations will be rendered as cyclic processes of the retrospective re-evaluation of (3)).
- iii. Moreover, (7) exemplifies how latent methodological rules shape the process of plausible argumentation and contribute to drawing plausible conclusions.

In accordance with (iv) in Sect. 2.4, we also have to assume that the plausible inferences in (6) and (7) are components of a process of plausible argumentation. In this respect, (6) and (7) have the following characteristics:

- i. At that stage of the argumentation process at which the thought experiment is carried out in Grice’s article, it is not clear in which sense Grice uses the word ‘say’. Therefore, the plausibility of (6)(c) is low. ‘Say’ is elucidated only in the next paragraph: “In the sense in which I am using the word *say*, I intend what someone has said to be closely related to the conventional meaning of the words (the sentence) he has uttered” (Grice 1989a: 25). That is, in a new ‘miniature’ cycle within the argumentation process the latent background assumptions (6)(c) and (6)(d) are retrospectively re-evaluated through the prism of the new information and, as a consequence, their plausibility increases. Nevertheless, the plausibility of (6)(d) is still lower than that of (6)(c), because there is no similar, relatively explicit feedback clarifying what is meant by ‘imply or suggest or mean’. These notions are retrospectively re-evaluated indirectly through the prism of further thought experiments in Grice (1989a), the introduction of the Cooperative Principle and the maxims as well as the argumentation in Grice (1989a) and (1989c).
- ii. In Grice (1989a) the above quotation is immediately followed by a series of further miniature thought experiments with the structure we have indicated. A new sub-cycle of the plausible argumentation process starts with a ‘What if?’ question; the plausibility of a possible answer to this question is assumed; and plausible inferences are used to obtain this answer, while these inferences are

¹³ Of course, this applies to our reconstruction of Grice’s argumentation in (6), (7), (9) and (10), too.

related to each other in the course of the cyclic, prismatic and retrospective re-evaluation of previously introduced information. Thus, the main issues which Grice's paper discusses—i.e. the distinction between 'what is said' and 'what is implicated', the Cooperative Principle, the conversational maxims, different examples of conversational implicature, etc.—are not considered by a single thought experiment, but rather, by a complicated system of thought experiments.¹⁴

- iii. In discussing features of thought experiments in general, Sorensen (1992: 89) remarks that “[m]any thought experiments function as *reminders*” (emphasis added; see also Cohnitz 2006: 81). In Sect. 2.3 we quoted Thomason, who, in (1)(c), requires the *similarity* between the imaginary situation in the thought experiment and a real one. Independently of Sorensen's and Thomason's stance, Meibauer attributes the same feature to the stories pragmatists make use of:

8. “Part of the appeal of pragmatics to a wider audience certainly has to do with our willingness to reflect about stories. In the process of reflecting about stories, we *compare* the story told with *our own experiences*, i.e. a set of contexts we remember, *sharing* some properties with the story told.” (Meibauer 2012: 771–772; emphasis added)

Accordingly, if we identify such stories with thought experiments, then we may assume that the reader, for example of the thought experiment in (3), *compares* the imaginary communicative situation described there with her previous experiences made in real communicative situations. The structure of the comparison may be reconstructed as an analogical inference (which is a sub-type of plausible inference):

9. Premises:

- (a) <It is plausible that the imaginary communicative situation as described in (3)(a) is similar to real communicative situations one has already experienced.>
- (b) <It is plausible that in the imaginary communicative situation as described in (3)(a) whatever the speaker implied, suggested, meant is distinct from what she said.>

Conclusion:

It is plausible that in the real communicative situations one has experienced, whatever the speaker implied, suggested, meant is distinct from what she said.

Clearly, the conclusion of (9) retrospectively re-evaluates previous experience through the prism of the imaginary situation presented in the thought experiment.

¹⁴ A detailed analysis of this structure, using the terminology of the framework we proposed above, would transgress the limits of the present paper.

iv. After having introduced the Cooperative Principle and the conversational maxims, Grice (1989a: 29) explains them, maintaining that “it is just a well-recognized *empirical fact* that people do behave in these ways” (emphasis added). This is crucial for two reasons. First, it is not mentioned again in the later argumentation. Second, in the reconstruction of every further thought experiment in later sub-cycles, with the help of which Grice illustrates different aspects of conversational implicature, the empirical hypothesis that people behave in accordance with the Cooperative Principle and the maxims must be considered *as a latent background assumption* of an enthymematic plausible inference thus integrating *experiential content* into the thought experiment. As a simple example we refer to Grice’s (1989a: 31) analysis of the story in (3)(a), the aim of which is to illustrate how conversational implicature is worked out with the help of the Cooperative Principle and the maxims, among other factors. A simplified reconstruction of one of the plausible inferences involved in the argumentation on p. 31 may be as follows:

10. Premises:

- (a) <It is plausible that if (3)(a) is given and it is a well-recognised empirical fact that people behave according to the Cooperative Principle and the maxims, then B implicates that C is potentially dishonest.>
- (b) It is certain that (3)(a) is given.
- (c) <It is plausible that it is a well-recognised empirical fact that people behave according to the Cooperative Principle and the maxims.>

Conclusion:

It is plausible that B implicates that C is potentially dishonest.

3.2 *The Solution to (P2)*

The above (simplified and partial) reconstruction of the structure of (3) illustrates a possible solution which our metascientific framework as summarised in (SP1) gives to (P2):

SP2. A thought experiment in pragmatics of the kind illustrated in (3) has the following relevant structural properties:

- (a) It is a process of plausible argumentation beginning with a ‘What if?’ question and terminating with the answer to this question.
- (b) The answer is elaborated by a series of plausible inferences during the process of prismatic cycles of plausible argumentation retrospectively re-evaluating previously accepted explicit premises and latent background assumptions.
- (c) The plausible inferences applied may be enthymematic.

- (d) The process of plausible argumentation may also involve the comparison of the imaginary communicative situation with one's own experiences of similar real communicative situations.
- (e) Hypotheses based on experience and considered empirical may be introduced in one of the sub-cycles of the argumentation process, and may be used as latent background assumptions of an enthymematic plausible inference in a later subcycle.

Although we have obtained (SP2) solely on the basis of (3), for the time being we assume that it is sufficiently abstract to serve as a starting point for capturing other kinds of thought experiments as well (see also Sects. 5.4.1 and 5.4.2). We assume that (SP2) is the 'genus proximum' which may be supplemented by 'differentia specifica' characteristic of particular subtypes of pragmatic thought experiment. Nevertheless, here we leave open the question, whether, and if so, how (SP2) may be modified in the light of further analyses of thought experiments used in current pragmatic research.

4 On (P3): The Function of Thought Experiments in Pragmatics

The structural properties as summarised in (SP2) pave the way for the fulfilment of the following functions of thought experiments:

- i. The philosophical literature distinguishes, among others, between *constructive* and *destructive* thought experiments (see e.g. Peijnenburg and Atkinson 2003; Brown and Fehige 2010, etc.), and this distinction appears to be compatible with our metascientific framework. A thought experiment in pragmatics is constructive if in a process of plausible argumentation it increases the plausibility of a particular hypothesis about the subject matter of the argumentation at issue; it is destructive if it is intended to reduce the plausibility of the rival hypothesis. Grice's (1989a) *stage setting* thought experiments in his seminal paper are both constructive and destructive. On the one hand, among other things, Grice argued *against* "a common mistake" which "arises from the inadequate attention to the nature and importance of the conditions governing conversation" (Grice 1989a: 24). On the other hand, he argued *for* the plausibility of the hypothesis that there are "general conditions that, in one way or another, apply to conversation as such, irrespective of its subject matter" (ibid; on Grice's aims see also e.g. Neale 1992; Saul 2002 etc.).
- ii. Nevertheless, (3) is not a clear case of a stage setting thought experiment. It includes the description of English sentences while it is presupposed that the reader of the paper understands their conventional meaning in the same way as Grice does. Therefore, *semantic intuition* is indispensable for the workability of the thought experiment.

- iii. In addition, *pragmatic intuition* with respect to the context in which the utterance is uttered is needed for the workability of the thought experiment, too (see also Sect. 5 on semantic and pragmatic intuition). Meibauer (2012) emphasises that the stories—which we have interpreted here as thought experiments—are rooted in the author’s pragmatic intuition in this sense.
- iv. This means that (3), besides being a stage setting thought experiment, can also be seen as *testing* alternative hypotheses by the use of *semantic and pragmatic intuition* involved in the thought experiment at issue. Thus, a thought experiment may be both stage setting and introspective, and the function of stage setting and testing may be *simultaneously* present. Therefore, Thomason’s distinction is not mechanically applicable to thought experiments in pragmatics.¹⁵
- v. Accordingly, the thought experiment we analysed functions as *evidence* for the hypothesis in (5) proposed as an answer to the ‘What if’ question in (4) both in the sense of (1)(b) and (2)(b).¹⁶

Thus, we have obtained the following solution to (P3):

- SP3. The function of a thought experiment in pragmatics of the kind exemplified in (3) may be
- (a) to work as a stage setting device in order to provide the initial step of a plausible argumentation process by
 - increasing the plausibility of a hypothesis about the subject matter of the argumentation at issue, and/or
 - reducing the plausibility of rival hypotheses;
 - (b) to test the plausibility of alternative hypotheses about communication on the basis of semantic and pragmatic intuition;
 - (c) to serve as evidence for or against the alternatives both in the sense of (1)(b) and (2)(b).

Just as in the case of (SP2), the question remains open, to what extent (SP3) as the ‘genus proximum’ should be supplemented by ‘differentia specifica’ in order to capture different kinds of pragmatic thought experiments.

¹⁵ For examples supporting this insight in the practice of current pragmatic research see Németh T. (2008), (2010).

¹⁶ Here we use the notion of ‘evidence’ pre-explicatively. For its explication within the present metascientific framework, see Kertész and Rákosi (2012).

5 On (P4): The Relationship Between Thought Experiments and Real Experiments in Pragmatics

5.1 *Introductory Remarks*

The philosophical literature on thought experiments devotes considerable space to the question of in what cases thought experiments fail (see for example Peijnenburg and Atkinson 2003; Cooper 2005; Thomason 1991; Cohnitz 2006). If we relate the suggestions discussed there to the framework we introduced in Sect. 2.4, we will obtain a straightforward answer to this question with respect to pragmatics. In particular, a thought experiment fails if the process of plausible argumentation, which it is part of, does not fulfil its function as specified in (SP3). Nevertheless, since the process of plausible argumentation is dynamic in that it involves the continuous re-evaluation of the information during the cycles, the failure of a thought experiment may be only temporary and may be remedied in later cycles.

In Sect. 5.2 we will discuss an example of the potential failure of Grice's thought experiment quoted in (3). Section 5.3 will show how in the pragmatic literature attempts have been made to correct the potential failure by the use of further thought experiments. Then, in Sect. 5.4 we will illustrate in what way real experiments are connected to thought experiments in the process of plausible argumentation. Finally, in Sect. 5.5 we will derive our solution to (P4) from our comments on these examples.

For lack of space, we will not present a precise reconstruction of the plausible inferences which constitute the argumentation process, but we will restrict the discussion to illustrating the ways in which Grice's thought experiment may fail and how its potential failure, mediated by many further thought experiments, eventually motivates the use of real experiments.

5.2 *From Grice's Thought Experiment to Grice's Circle*

Early criticisms had already pointed out many alleged shortcomings of the answer that Grice gave in (5) to the 'What if' question in (4). These criticisms also affect indirect consequences of (5) inferred during the plausible argumentation process in Grice (1989a), such as his approach to speaker's meaning, the Cooperative Principle, the maxims, the way implicatures work, etc.

Basically, the philosophical literature on thought experiments discusses three ways in which a thought experiment may fail. These can be exemplified with respect to (3) as follows:

- i. The first reason why a thought experiment may fail is that the answer given to the ‘What if?’ question is *unworkable* (see also Cooper 2005). As regards pragmatics, for example, the early criticisms of Grice (1989a) in Kempson (1975), Sadock (1978) and Kiefer (1979) can be interpreted in such a way:
11. (a) “[...] the theory is [...] *vacuous* and therefore of *no* explanatory power” (Kiefer 1979: 57)
 - (b) [the Gricean maxims] do not even account for *the type of communication* they have been devised for.” (Kiefer 1979: 60)
 - (c) “[...] “almost *any meaning* can be worked out on the basis of almost *any meaning*” (Kiefer 1979: 70).¹⁷
- ii. A second kind of failure is when a thought experiment leads to *contradictory* conclusions (Cooper 2005; Thomason 1991; Peijnenburg and Atkinson 2003). Grice’s work has often been charged with making this mistake. For example, Green (1996: 91) claims that the maxim of manner is contradictory in a self-refuting way:
12. “Insofar as *perspicuous* and *prolixity* are unnecessarily obscure expressions (compared to *clear* and *verbosity* or *too many words*), the statement of the maxim and its third submaxim violate the first submaxim; submaxim (3) violates itself as well with the obscure and repetitious paraphrase; and submaxim (4) is not ordered well with the others, bringing to mind infuriating recipes like that in (i).”

This is a good example of how in the course of the argumentation process thought experiments can—indirectly, through many plausible inferences, as remote consequences—*lead to* contradictions. Through a sequence of further thought experiments and plausible inferences, Grice’s seminal thought experiment quoted in (3) led, among others, to the hypothesis that the Cooperative Principle and the conversational maxims are adequate. In contrast, one of the latter, in turn, led to the self-refuting contradictions unveiled in (12).

- iii. The philosophical literature mentions *circular argumentation* as a third possibility (Peijnenburg and Atkinson 2003; Thomason 1991; Cohnitz 2006). For lack of space, in what follows, we will illustrate only this type in somewhat greater detail.

‘Grice’s circle’ emerges as an immediate consequence of the thought experiment with the help of which Grice made the first step toward introducing his account of implicature. Levinson defines the circle as follows:

13. (a) “Grice’s account makes implicature dependent on a prior determination of ‘the said’. The said in turn depends on disambiguation, indexical resolution, reference fixing, not to mention ellipsis unpacking and generally narrowing. But each of these processes, which are prerequisites to

¹⁷ See also Searle (1969: 43–45).

- determining the propositions expressed, may themselves depend crucially on processes that look indistinguishable from implicatures.
- (b) Thus, what is said seems both to determine and to be determined by implicature. Let us call this *Grice's circle*.
- (c) It should be clear that this is not a minor point in Gricean exegesis. It is a circle that equally afflicts any theory that seeks to make a semantics/pragmatics distinction play a crucial role in the general theory of meaning. The 'said' can be taken to be truth-conditional content—the proposition expressed, the output of the process of semantic interpretation; the proper domain of a theory of linguistic meaning. The 'implicated' can be taken more generally than I am taking it here, to include all the processes of pragmatic inference; it is the proper domain of a theory of communication." (Levinson 2000: 186; emphasis as in the original)

(13)(b) states that 'what is said'—i.e. truth-conditional meaning—is the input to 'what is implicated', and 'what is implicated' is the input to 'what is said'. Nevertheless, Levinson goes further in (13)(c) where he raises the general problem of the relationship between truth-conditional semantics and pragmatics. Since the propositional content of utterances is the subject matter of truth-conditional semantics, while implicature belongs to the realm of pragmatics, the circular relation between semantics and pragmatics is at stake, too.

The fundamental relevance of the circle cannot be overlooked. It undermines the answer which Grice gave in (5) to the 'What if?' question in (4) via the thought experiment we quoted in (3) and which included the distinction between 'what is said' and 'what is implicated'. The reason why the circle undermines (5) is that "the pragmatic enterprise that concedes that pragmatics intrudes into semantics (read: truth-conditional content or propositional forms) is circular, hence a *definitionally impossible enterprise*" (Capone 2006: 650). Thus, if the circle cannot be resolved, then the thought experiment in Grice (1989a) fails exactly in the sense the literature deems the emergence of circularity as the failure of thought experimentation in general.

5.3 Possible Resolutions of Grice's Circle via New Thought Experiments

Without striving for a comprehensive overview, we illustrate possible resolutions of the circle proposed by a few examples (for a survey of the different standpoints see Ariel 2008: 261–308 and Buchanan 2010).

Bach (1994) introduced the notion of *impliciture* and elaborated on it in a series of later papers (see e.g. Bach 1999, 2001, 2007). This notion is intended to capture the middle ground between 'what is said' and 'what is implicated' and in this way the circle does not emerge. 'Impliciture' applies to cases in which the speaker has

not been fully explicit. There are two kinds of implicature. The first is when an utterance does not express a complete proposition, and therefore needs completion so that a complete proposition can be produced. In the second case, the utterance expresses a proposition, but the speaker communicates another proposition. In both cases, the speaker intends the hearer to understand something which has not been expressed explicitly in the utterance.

One of the most influential resolutions of the circle is Recanati's, which the author put forward in widely cited publications (e.g. Recanati 1989, 2001, 2004, etc.). In his view "there is no such thing as 'what the sentence says' (in the standard sense in which that phrase is generally used)" (Recanati 2001: 87). Namely, 'what is said' incorporates optional contextual elements. It includes saturation (i.e. the fixing of reference, the specification of indexicals and the disambiguation of expressions) as well as optional processes such as *free enrichment* (Recanati 2001: 79). Thus, the notion of 'what is said' is a pragmatic notion at the outset, for almost everything one says includes pragmatic enrichment.

Levinson's own approach distinguishes between three levels of meaning: sentence meaning, utterance-type meaning and utterance-token meaning. Sentence meaning is basically assumed to be conventional and semantically underdetermined. Utterance-token meanings are calculated on the basis of particularised conversational implicatures which are highly context-dependent and may include encyclopaedic information as well. Utterance-type meanings result from inferences based on generalised conversational implicatures. However, in Levinson's approach, generalised conversational implicatures are assumed to be default inferences that are relatively independent of the context, although in specific contexts they may be cancelled. They may contribute to the truth-conditional content of utterances. These distinctions boil down to the insight that "[t]here is every reason to try and reconstrue the interaction between semantics and pragmatics as the intimate interlocking of distinct processes, rather than, as traditionally, in terms of the output of one being the input to the other" (Levinson 2000: 242; see also Huang 2010: 627). Consequently, the circle can be avoided.¹⁸

Carston's (2002a, b, 2004) relevance-theoretic resolution denies the assumption that there are generalised conversational implicatures as default inferences and claims that all inferences are warranted by contextual relevance. She assumes only a single principle, namely, relevance—in the sense of relevance theory—which is responsible for all kinds of conversational implicature as well as the contribution of pragmatics to truth-conditional meaning. In her view "[t]he mechanism that mediates the inferences from logical form to communicated propositions is one of '*mutual parallel adjustment*' of explicatures and implicatures, constrained by the

¹⁸ Carston (2002a) criticises Levinson's approach because, according to her, he does not offer an overall solution to Grice's circle. Rather, he merely assumes that his approach to generalised conversational implicature, which he developed independently of the circularity issue, may soften its harmfulness. Carston also remarks that there are particularised conversational implicatures which influence propositional meaning, while others do not.

comprehension strategy” (Carston 2002a: 139; emphasis added). Consequently, there is no sequentiality, no input/output, and, consequently, no circle.

These approaches illustrate the status of thought experiments within the overall process of plausible argumentation focusing on the relationship between ‘what is said’ and ‘what is implicated’. Namely, all four approaches make use of *numerous thought experiments* in elaborating their own resolution of the circle. Nevertheless, even if one admits that they involve possible solutions to the problem of the relationship between ‘what is said’ and ‘what is implicated’, the problems they actually tackle are *different* both from those raised by Grice’s thought experiment and from what Levinson explicated as Grice’s circle. *None* of them are resolutions of *the* Gricean circle, but they are resolutions of *prismatically and retrospectively re-evaluated* versions of the latter. New thought experiments motivate, for instance, the application of ‘prisms’ such as the notions ‘generalised conversational implicature as default’, ‘implicature’, ‘free enrichment’, ‘explicature’, ‘mutual parallel adjustment’ by means of which the problems are retrospectively re-evaluated. Liedtke (2011: 46) claims that

14. “[...] the protagonists (and antagonists) of the debate seem to operate on different levels. Looking for intuitions concerning the limits of what-is-said and elaborating conceptual tools for the description of the different types of reasoning processes do not seem to conflict necessarily, rather they might ‘cooperate’ in the search for an adequate account of layer-specific aspects of meaning.”

(14) can be easily interpreted as referring to the same process which we described above.

From the mechanism of plausible argumentation it also follows that Grice’s thought experiment did not fail in some ‘absolute’ sense. What was conceived of as a failure is merely a *temporarily existing* problematic informational state in particular cycles within the process of plausible argumentation. It disappears during the later cycles, while the latter, in turn, lead to further problematic states triggering new thought experiments that are subject to further re-evaluations and so on.

5.4 *From Thought Experiments to Real Experiments in Pragmatics*

5.4.1 **Example 1: The Cyclic, Retrospective and Prismatic Connection Between Thought Experiments and Real Experiments in Pragmatics**

Works dealing with experimental pragmatics often mention that armchair linguistics or stories or intuitions, which we interpreted here as thought experiments,

are predecessors of real experiments (e.g. Clark and Bangarter 2004; Gibbs 2004; Noveck and Sperber 2004a, b, 2007). In order to illustrate how our framework handles this issue, we choose Noveck and Sperber (2007) as an example.

The authors' aim is to show how *the limits* of intuitions as manifested in the stories we interpreted as thought experiments can be overcome by real experiments. In our terminology, they point out that thought experiments, which are simultaneously stage setting and introspective, fail, because they cannot provide a workable solution to a specific sub-problem—i.e. that of scalar implicatures—of the problem of the relationship between 'what is said' and 'what is implicated'. In addition, they also argue that real experiments provide *crucial evidence* with the help of which decisions can be made between alternative approaches that interpret the utterances at issue in the same way, but at the same time draw different conclusions with respect to the cognitive processes underlying these interpretations. These are decisions which the thought experiments they carried out were, due to their alleged failure, unable to trigger.

The authors raise the question of the relationship between 'what is said' and 'what is implicated' with respect to scalar implicatures. They argue against Levinson's generalised conversational implicature (abbreviated as GCI) theory and for an account in terms of relevance theory. Relevant features of their line of argumentation can be summarised as follows.

- i. They start with a series of *thought experiments* the function of which is to set the stage for subsequent steps of their argumentation process. The authors put forward their own hypothesis which they contrast with Levinson's. In Levinson's view scalar implicatures are paradigmatic cases of generalised conversational implicatures. They are default inferences carried out in one step and require neither the context nor the entire mechanism of implicatures. As opposed to this view—and in accordance with what we have already mentioned in Sect. 5.3—Noveck and Sperber's experiments serve to illustrate that in accordance with relevance theory the interpretation of scalar implicatures is the outcome of mutual parallel adjustment between explicature and implicature, and that linguistic expressions do not encode but rather indicate the speaker's meaning.
- ii. It is important to remark that these thought experiments fall within the scope of the very abstract characterisations in (SP2) and (SP3) which are merely meant to provide the 'genus proximum' of pragmatic thought experiments. However, the thought experiments applied by Noveck and Sperber also show clear differences from Grice's thought experiment in (3).¹⁹ For example, one substantial difference is that in numerous plausible inferences related to the thought experiments the authors (unlike Grice) continuously change one of the explicit

¹⁹ The issue of the relationship between 'philosophical' and 'scientific' thought experiments is an important one, but discussing it would transgress the limits of the present paper.

premises or latent background assumptions in order to discover the consequences of the changes. In this way, during a series of plausible argumentation cycles, they consider several alternative hypotheses from which, by the use of further plausible inferences, they try to select the most plausible ones.

iii. In this way, the thought experiments eventually lead to two rival hypotheses with respect to the speed of interpretation in cases in which an enriched interpretation is not contextually primed. The rival hypotheses are as follows:

15. The GCI theory (in the form Noveck and Sperber retrospectively re-evaluate it) predicts that:

- (a) the literal interpretation of the scalar term is slower when it is enriched by default and then context-sensitively cancelled;
- (b) the enriched interpretation is computed by default and is hence faster.

16. Relevance Theory predicts that:

- (a) the literal interpretation without enrichment is faster;
- (b) the context-sensitively enriched interpretation is slower.

iv. At this point *real experiments* enter the scene, whose aim is to support the plausibility of (16) and to reduce that of (15). Within the first sub-cycle of the real experimental argumentation cycle the authors report on three real experiments in which children's reasoning was investigated with respect to scalar inferences. The results of the experiments were published in Noveck (2001). With respect to the decision between the two hypotheses in (15) and (16), Noveck and Sperber's evaluation of the outcome of these developmental experiments was double-sided:

17. (a) "If children had been found to perform scalar inferences by default, this would have been strong evidence in favour of the GCI theory approach. However, taken together, developmental data suggest that, for children, enriched interpretations of scalar terms are not default interpretations

- (b) This data is not knock down evidence against GCI theory, because it is compatible with two hypotheses: 1) scalar inferences are not default interpretations for adults either (even if adults are more likely to derive them because they can do so with relatively less effort and because they are more inclined to invest effort in the interpretation of an utterance given their greater ability to derive from it cognitive effects). Or, 2) in the course of development, children become capable and disposed to perform scalar inferences by default. The first hypothesis is consistent with the relevance theory approach while the second is consistent with the GCI approach.

(c) To find out which approach has more support, further work had to be done with adults." (Noveck and Sperber 2007: 203)

17(a) claims that the data obtained by these experiments increase the plausibility of the hypothesis of relevance theory and reduce that of GCI theory.

17(b) re-evaluates this conclusion by taking further information into account which may speak both for assumption 1) and assumption 2) in the quotation. Thus, a local inconsistency between (17)(a) and (b) arises which corresponds to informational overdetermination and is problematic.²⁰

17(c) indicates that, in order to make the decision between the alternatives, another sub-cycle within the plausible argumentation process is to be carried out which once again re-evaluates this informational state through the prism of the information provided by future experiments.

This example illustrates that in current research, thought experiments and real experiments may be *directly connected*. But this connection involves much more than the fact that “pragmatic intuitions about exemplary utterances might be backed by an experimental account that might strengthen the evidence of certain claims about the limits of what-is-said [...]” (Liedtke 2011: 46). Namely, the connection between the thought experiments and the real ones is, as opposed to what is assumed by Liedtke, *not linear*, but cyclic. In the terminology of our metascientific framework, the relation between thought experiments and real ones can be characterised as the transgression between two argumentation cycles in the following way:

- i. The argumentation process based on the plausible inferences, which thought experiments consist of, is continued by a new argumentation cycle based on real experiments.²¹ In this way, during the real experimental cycle the findings provided by thought experiments are made use of as *stage setting* devices without which real experiments could not have been designed.
- ii. The results of the real experiments retrospectively *re-evaluate* the conclusions obtained by the plausible inferences in the thought experiments in that they increase the plausibility of (16)(a) and (b) and decrease that of (15)(a) and (b). The results of the thought experiments are viewed through the prism of the new pieces of information, which include, among others, the experimental design and the outcome of the real experiment.

²⁰ See (iv) in Sect. 2.4 for these notions.

²¹ In fact, it is not the experiment itself, but the experimental report that is to be considered as plausible argumentation. The experimental report is for example part of Noveck and Sperber (2007) and thus the argumentation cycles it consists of are comparable to the argumentation cycles including the thought experiments. In this context we remark that, first, it would be necessary to ask questions with respect to pragmatic experiments in analogy to (P2) and (P3)—that is, what kind of structure they have and what their functions are. However, for lack of space we refer to Kertész and Rákosi (2012) where these questions are raised and partially answered in an extensive case study within the framework of plausible argumentation. Second, the philosophical literature also discusses the fact that every real experiment has to be carried out in thought before it is carried out in reality (see e.g. Buzzoni 2008).

5.4.2 Example 2: Thought Experiments as Parts of Real Experiments in Pragmatics

There is also another important kind of relationship between thought experiments and real experiments: thought experiments may be substantial *components* of what pragmatists treat as real experiments. Again, the illustration will be taken from the controversy about the problem of ‘what is said’ and ‘what is implicated’.

The reason why despite its well-known shortcomings we choose Gibbs and Moise (1997) as an illustrative example is that, on the one hand, it was the first attempt to examine experimentally whether people distinguish systematically between ‘what is said’ and ‘what is implicated’.²² On the other hand, the discussion which the paper raised is still used as a point of departure in current research (see e.g. Larson et al. 2009; Liedtke 2011).

In one of the experiments the authors report on the experimentees were asked to read stories like the following:

18. “Bill wanted to date his co-worker Jane. Being rather shy and not knowing Jane very well, Bill asked his friend, Steve, about Jane. Bill didn’t even know if Jane was married or not. When Bill asked Steve about this, Steve replied *Jane has three children.*” (Gibbs and Moise 1997: 61; Gibbs 2004: 64)

Next, they had to choose the best paraphrase of the final expression from these three options:

19. (a) Jane has at least three children.
 (b) Jane has exactly three children.
 (c) Jane is already married.

(19)(a) is the minimal interpretation of ‘what is said’, (19)(b) is the pragmatically enriched interpretation and (19)(c) is the particularised conversational implicature. The result of the experiment was that the majority of the participants chose the enriched interpretation. Consequently:

20. “These findings show that pragmatics strongly influences people’s understanding of both what speakers say and communicate. It appears that Grice’s examples of generalised conversational implicatures are not implicatures at all but understood as part of what speakers say. More generally, the Gibbs and Moise (1997) findings suggest that the distinction between saying and implicating is orthogonal to the division between semantics and pragmatics.” (Gibbs 2004: 64)

From the point of view of the present paper, this example illustrates the following:

²² Gibbs and Moise (1997) was fiercely criticised in Nicolle and Clark (1999); see also Gibbs (1999) answer.

- i. Sorensen (1992: 76-86) and Cohnitz (2006: 80-83) characterise stories in real experiments similar to (18) as thought experiments ‘embedded’ into a real experiment.
- ii. From the point of view of the experimenter, there is an underlying ‘What if?’ question the answer to which is expected to support the plausibility that the hypothesis in (20) is adequate. This process can be reconstructed as a series of plausible inferences in analogy to those in Sect. 3.
- iii. (18) also shows characteristics of introspective thought experiments. It is built on two kinds of intuitions. It involves the semantic intuition of both the experimenters and the experimentees who have to judge the conventional meaning of the italicised sentence. In addition, it involves pragmatic intuition, which concerns the context presented (Meibauer 2012).
- iv. A thought experiment like (18) may be a necessary component of a real experiment without which the latter could not have been carried out.
- v. As in the preceding subsection, it should be pointed out here, too, that due to their high level of abstraction, (SP2) and (SP3) are assumed to apply to such thought experiments as well. Nevertheless, they may have additional features (‘differentia specifica’) making (18) differ both from (3) and from the thought experiments applied in Noveck and Sperber (2007). The question of the exact nature of these differences has to be left open here.

The latest literature abounds in reports on real experiments tackling the problem of the relationship between ‘what is said’ and ‘what is implicated’. Many of them include thought experiments similar to (18) (see e.g. Breheny et al. 2006; Guasti et al. 2005; Katsos and Cummins 2010; Larson et al. 2009; Liedtke 2011, etc.).

5.5 *The Solution to (P4)*

Now we can summarise our solution to the problem (P4):

- SP4. (a) Although thought experiments may fail, they are indispensable for the success of pragmatic inquiry.
- (b) There may be a cyclic, prismatic and retrospective connection between thought experiments and real experiments within the overall process of plausible argumentation in that
- the latter might remedy the potential failure of the former;
 - thought experiments may work as stage setting devices for real experiments; and
 - real experiments may increase or decrease the plausibility of a hypothesis obtained by means of a series of plausible inferences initiated by a thought experiment.

- (c) Thought experiments may correspond to subcycles of the plausible argumentation process involving the real experiment as manifested in the experimental report.

6 On (P5): The Puzzle of Thought Experiments in Pragmatics

Having illustrated the structure and function of thought experiments as they appear in a classical contribution to pragmatics as well as some aspects of the relationship between thought experiments and real experiments, we now turn to our solution of the problem (P5). The key idea of the solution follows from (SP2)–(SP4): Thought experiments in pragmatics do not provide new experiential information directly; they do so *indirectly*. On the one hand, new experiential information is obtained from sources which are outside the thought experiment: for example, from real experiments or personal experiences with real communicative situations. On the other hand, thought experiments may *feed* experiential information rooted in such sources *into* particular cycles of the plausible argumentation process. Thereby, they cannot generate new experiential information, but they contribute to the re-evaluation of experiential information already given.²³ This key idea boils down to the following solution to (P5):

- SP5. Thought experiments in pragmatics *may* contribute to feeding new experiential information into the process of plausible argumentation indirectly by means of
- (a) the latent background assumptions of plausible inferences as a result of which the information content of the conclusion may transcend that of the explicit premises;
 - (b) the comparison of an imaginary communicative situation with experiences of similar communicative situations in the real world;
 - (c) the cyclic, prismatic and retrospective connection between thought experiments and real experiments;
 - (d) the integration of thought experiments into real experiments.

Let us add the following comments to (SP5):

Ad (SP5)(a). In (SP2)(c) we pointed out that in pragmatic thought experiments enthymematic plausible inferences as introduced in (iii) in Sect. 2.4 may be present. In these cases the information content of the conclusion, due to the latent background assumptions, goes beyond the information content of the explicit premises. Thus, those thought experiments which are built on enthymematic inferences, *may* yield *new* knowledge—nevertheless, this knowledge is not new in

²³ This claim is in accordance, for example, with Gähde (2000) and Rescher (2005).

some ‘absolute’ sense, but is only new *relative to the explicit premises*. Even in this case, the new information need not be experiential, though it *may* be. If it is, then the conclusion may contain experiential information not inherent in the explicit premises—as we have seen in (SP2)(e) and in (10). But the conclusion cannot include more experiential information than the latent background assumptions. Thereby, the source of the experiential information in the latent background assumptions is something external to the thought experiment and, accordingly, the corresponding cycle of the argumentation process at issue—for example, personal experience as witnessed by the example in (10). Thus, in certain cases a thought experiment containing at least one enthymematic inference can feed experiential information into the argumentation process indirectly, through the latent background assumptions of plausible inferences.

Ad (SP5)(b). From (SP2)(d) it follows immediately that through the comparison between an imaginary communicative situation which was referred to in the thought experiment and real communicative situations already experienced, insights can be obtained with respect to the way real communication works. This claim is in accordance with Sorensen’s (1992: 89) and Thomason’s (1991: 248) stance. Accordingly, the conclusion of the analogical inference in (9) contains experiential information which, nevertheless, has already been included in the explicit premises and/or the latent background assumptions. However, the sources on which the explicit premises and/or the latent background assumptions lie, are outside the thought experiment. Here again, the effect of the thought experiment is indirect in that through such an analogical inference it adds experiential information—which has already been given quite independently of the thought experiment—to the argumentation process.

Ad (SP5)(c). On the one hand, in agreement with (SP3)(a), the second claim of (SP4)(b) says that thought experiments in pragmatics may set the stage for real experiments. Thus, they *indirectly* contribute to the results of the latter. This finding is in agreement with the claim that “[t]hought experiments anticipate the results of real experiments and in this way they inductively extend our knowledge” (Buzzoni 2008: 96). On the other hand, the new experiential information obtained by carrying out the real experiment may retrospectively re-evaluate the outcome of the thought experiment through the prism of the experimental design, the materials and the results themselves. In this case the real experiment works as an *experiential* source on the basis of which the plausibility of the hypotheses the thought experiment leads to can be judged. Consequently, the outcome of this judgement may be experiential—but only indirectly, obtained in the course of the cycles of retrospective re-evaluations through the prism of the real experiment.

Ad (SP5)(d). As (SP4)(c) maintains, thought experiments may be components of real experiments. In accordance with Sorensen (1992: 76–86) and Cohnitz (2006: 80–83), they work as new data sources and may support the experiential findings obtained by the real experiment.

The way and the extent to which some or all of these factors are present and interact depends on the particular properties of the plausible argumentation process at issue and cannot be generalised.

7 Concluding Remarks

By proposing possible solutions to the problems (P1)–(P5), we seem have achieved the aim of the present paper. Yet, our findings as summarised in (SP1)–(SP5) raise further problems among which the following are especially important:

- i. Throughout our line of reasoning we reflected on the philosophical literature on thought experiments in order to demonstrate that our findings partly integrate some of its insights, and partly go beyond the latter. In our comments on (SP5) we cited the philosophical literature which our findings are in agreement with. However, in spite of this, our approach differs substantially from any other account of thought experiments we are familiar with. In particular, our claim that thought experiments in pragmatics are pieces of the *dynamic* processes of plausible argumentation in the way we characterised it, goes *far beyond* the approaches discussed in the philosophical literature. Accordingly, the question arises, in what way and to what extent more detailed applications of Kertész and Rákosi (2012) would shed fresh light on this process.
- ii. (SP2)–(SP5) also show that in pragmatics thought experiments are still *indispensable* tools of research and cannot be simply substituted by real experiments. Nevertheless, their role with respect to experimental pragmatics is certainly much more sophisticated than we could indicate here and needs, therefore, further scrutiny.
- iii. Our argumentation theoretic solution to the puzzle of thought experiments in pragmatics raises the question of whether it can be applied to thought experiments outside pragmatics. If so, then pragmatics may *contribute to* the solution of a methodological puzzle which is a hot topic in the contemporary philosophy of science.
- iv. Both thought experiments and real experiments are sources of the *data* on which hypotheses rest (see Kertész and Rákosi 2012). In this respect, it is important to touch briefly on the current discussion on data in linguistics (see, e.g. Borsley (2005), Penke and Rosenbach 2007, Kepsers and Reis (2005), Sternefeld (2007), Featherston and Winkler (2009), Winkler and Featherston (2009), etc.). The main findings of the present paper may contribute to the discussion by calling attention to the necessity to *integrate* different data sources and data types in pragmatics (see e.g. Clark and Bangartner 2004; Meibauer 2012, etc.). As we have seen, the integration of stage setting thought experiments, introspective thought experiments and real experiments as different *data sources* is well-motivated. These sources yield different *data types* each of which seems to be relevant for theory formation in pragmatics; consequently, semantic intuition, pragmatic intuition and experimental data should be used in combination. How these data sources can be integrated and how these data types interact is a central and still open methodological question of current research in pragmatics.

- v. Our paradigmatic example which served as a point of departure for solving (SP2)–(SP5) was a *single* thought experiment, taken from classical philosophical pragmatics. Throughout our line of reasoning we referred to the diversity of thought experiments and indicated some of the properties which might distinguish the thought experiments mentioned in Sect. 5.4 from our paradigmatic example in (3). However, we did not systematically go into the question of what the ‘*differentia specifica*’ are that may define particular types of pragmatic thought experiment.

Due to these problems, whose discussion must be left for future research, any hasty generalisation of our solutions to the problems (P2)–(P5) would be misguided.

Acknowledgments András Kertész’ contribution to the present paper was supported by the Research Group for Theoretical Linguistics of the Hungarian Academy of Sciences as well as the projects OTKA K 77823 and TAMOP 4.2.1./B-09/1/KONV-2010-0007. We are grateful to Enikő Németh T. and an anonymous reviewer for helpful comments on earlier drafts.

References

- Ariel, Mira. 2008. *Pragmatics and grammar*. Cambridge: Cambridge University Press.
- Bach, Kent. 1994. Conversational implicature. *Mind and Language* 9: 124–162.
- Bach, Kent. 1999. The myth of conventional implicature. *Linguistics and Philosophy* 22: 327–366.
- Bach, Kent. 2001. You don’t say? *Synthese* 128: 15–44.
- Bach, Kent. 2007. Regressions in pragmatics (and semantics). In *pragmatics*, ed. Noël Burton-Roberts, 24–44. Basingstoke: Palgrave Macmillan.
- Borsley, Robert D ed. 2005. Data in theoretical linguistics. *Lingua* 115: 1475–1665.
- Breheny, Richard. 2011. Experimentation-based pragmatics. In *Foundations of Pragmatics*, eds. Wolfram Blublitz and Neal R. Norrick, 561–586. Berlin and New York: Mouton de Gruyter.
- Breheny, Richard, Napoleon Katsos, and John Williams. 2006. Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100: 434–463.
- Brown, James Robert. 1991. *The laboratory of the mind: Thought experiments in the natural sciences*. London and New York: Routledge.
- Brown, James Robert, and Yiftach Fehige. 2010. Thought experiments. *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/thought-experimen/>. Accessed 13 June 2011.
- Buchanan, Ray. 2010. A puzzle about meaning and communication. *Noûs* 44: 340–371.
- Buzzoni, Marco. 2008. *Thought experiment in the natural sciences: An operational and reflexive-transcendental conception*. Würzburg: Königshausen & Neumann.
- Capone, Alessandro. 2006. On Grice’s circle (a theory-internal problem in linguistic theories of the Gricean type). *Journal of Pragmatics* 38: 645–669.
- Carston, Robyn. 2002a. Linguistic meaning, communicated meaning and cognitive pragmatics. *Mind and Language* 17: 127–148.
- Carston, Robyn. 2002b. *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.
- Carston, Robyn. 2004. Truth-conditional content and conversational implicature. In *The semantics/pragmatics distinction*, ed. Claudia Bianchi, 1–36. Stanford: CSLI Publications.

- Clark, Herbert H., and Adrian Bangartner. 2004. Changing ideas about reference. In *Experimental pragmatics*, eds. Ira A. Noveck, and Dan Sperber, 25–49. Basingstoke: Palgrave Macmillan.
- Cohnitz, Daniel. 2006. *Gedankenexperimente in der Philosophie*. Paderborn: Mentis.
- Cooper, Rachel. 2005. Thought experiments. *Metaphilosophy* 36: 328–347.
- Featherston, Sam, and Susanne Winkler eds. 2009. *The Fruits of Empirical Linguistics, Vol. 1: Process*. Berlin and New York: de Gruyter.
- Gähde, Ulrich. 2000. Gedankenexperimente in Erkenntnistheorie und Physik: strukturelle Parallelen. In *Rationalität, Realismus, Revision: Vorträge des 3. internationalen Kongresses der Gesellschaft für Analytische Philosophie vom 15. bis zum 18. Sep 1997 in München*, ed. Julian Nida-Rümelin, 457–46. Berlin and New York: de Gruyter.
- Gendler, Tamar Szabó. 1998. Galileo and the indispensability of scientific thought experiment. *British Journal for the Philosophy of Science* 49: 397–424.
- Gendler, Tamar Szabó. 2000. *Thought experiments: On the power and limits of imaginary cases*. New York: Garland Publishing.
- Gibbs Jr, Raymond W. 1999. Speaker's intuitions and pragmatic theory. *Cognition* 69: 355–359.
- Gibbs Jr, Raymond W. 2004. Psycholinguistic experiments and linguistic pragmatics. In *Experimental pragmatics*, eds. Ira A. Noveck, and Dan Sperber, 50–71. Basingstoke: Palgrave Macmillan.
- Gibbs Jr, Raymond W., and Jessica F. Moise. 1997. Pragmatics in understanding what is said. *Cognition* 62: 51–74.
- Gooding, D. 1990. *Experiment and the making of meaning*. Dordrecht: Kluwer.
- Green, Georgia M. 1996. *Pragmatics and natural language understanding*. Mahwah: Erlbaum.
- Grice, Paul. 1989a. Logic and conversation. In Grice 1989d, 22–40.
- Grice, Paul. 1989b. Further notes on logic and conversation. In Grice 1989d, 41–57.
- Grice, Paul 1989c. Meaning. In Grice 1989d, 213–223.
- Grice, Paul. 1989d. *Studies in the way of words*. Cambridge: Harvard University Press.
- Guasti, Maria Teresa, Gennaro Chierchia, Stephen Crain, Francesca Foppolo, Andrea Gualmini, and Luisa Meroni. 2005. Why children and adults sometimes (but not always) compute implicatures. *Language and cognitive processes* 20: 667–696.
- Hägglund, Sören. 1996. *Thought experiments in philosophy*. Stockholm: Almqvist & Wiksell International.
- Horowitz, Tamara, and Gerald J. Massey (eds.). 1991a. *Thought experiments in science and philosophy*. Savage: Rowman & Littlefield.
- Horowitz, Tamara, and Gerald J. Massey. 1991. Introduction. In: eds. Horowitz and Massey, 1–28.
- Huang, Yan. 2010. Neo-Gricean pragmatic theory of conversational implicature. In *The Oxford handbook of linguistic analysis*, eds. Bernd Heine, and Heiko Narrog, 607–631. Oxford: Oxford University Press.
- Irvine, Andrew D. 1991. Thought experiments in scientific reasoning. In eds. Horowitz and Massey, 149–165.
- Katsos, Napoleon, and Chris Cummins. 2010. Pragmatics: From theory to experiment and back again. *Language and Linguistics Compass* 4(5): 282–295.
- Kempson, Ruth M. 1975. *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.
- Kepser, Stephan and Marga Reis eds. 2005. *Linguistic evidence. Empirical, theoretical and computational perspectives*. Berlin and New York: de Gruyter.
- Kertész, András. 2010. On stage setting thought experiments in cognitive linguistics: A case study. In *Von Katastrophen, Zeichen und vom Ursprung der menschlichen Sprache*, ed. Cornelia Stroh, 7–30. Bochum: Brockmeyer.
- Kertész, András, and Csilla Rákosi. 2012. *Data and evidence in linguistics: A plausible argumentation model*. Cambridge: Cambridge University Press.
- Kiefer, F. 1979. What do conversational maxims explain? *Linguisticae Investigationes* 3: 57–74.
- Kuhn, Thomas S. 1970. *The structure of scientific revolutions*. Chicago: The University of Chicago Press.

- Kuhn, Thomas S. 1977. A function for thought experiments. In *The essential tension: Selected studies in scientific tradition and change* by Thomas S. Kuhn, 240–265. Chicago: University of Chicago Press.
- Kühne, Ulrich. 2005. *Die Methode des Gedankenexperiments*. Frankfurt am Main: Suhrkamp.
- Larson, Meredith, Ryna Doran, Yaron McNabb, Rachel Baker, Matthew Berends, Alex Djalali, and Gregory Ward. 2009. Distinguishing the *Said* from the *Implicated* using a novel experimental paradigm. In eds. Uli Sauerland and Kazuko Yatsushiro, 74–93.
- Levinson, Stephen C. 2000. *Presumptive meanings. The theory of generalized conversational implicature*. Cambridge: MIT Press.
- Liedtke, Frank. 2011. The impact of literal meaning on what-is-said. In eds. Jörg Meibauer and Markus Steinbach, 43–62.
- Mach, Ernst. 1960. *The science of mechanics*. La Salle: Open Court.
- Mach, Ernst. 1976. On thought experiments. In *Knowledge and error*, ed. Ernst Mach, 137–47. Dordrecht: Reidel.
- McAllister, James. 1996. The evidential significance of thought experiment in science. *Studies in History and Philosophy of Science* 27: 233–250.
- Meibauer, Jörg. 2012. Pragmatic evidence, context, and story design. An essay on recent developments in experimental pragmatics. In *Converging data sources in cognitive linguistics*, eds. András Kertész, Monika Schwarz-Friesel and Manfred Consten, 768–776. Amsterdam etc.: Elsevier. [= Special Issue of *Language Sciences* 34/6].
- Meibauer, Jörg, and Markus Steinbach. 2011. *Experimental pragmatics/semantics*. Benjamins: Amsterdam & Philadelphia.
- Miščević, Nenad. 1992. Mental models and thought experiments. *International Studies in the Philosophy of Science* 6: 215–226.
- Moue, Aspasia S., Kyriakos A. Masavetas, and Haido Karayianni. 2006. Tracing the developments of thought experiments in the philosophy of natural sciences. *Journal for the General Philosophy of Science* 37: 61–75.
- Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509–559.
- Németh T., Enikő. 2010. How lexical-semantic factors influence the verbs' occurrence with implicit direct object arguments in Hungarian. In *The role of data at the semantics-pragmatics*, eds. Enikő Németh T. and Károly Bibok, 305–348.
- Németh, T., Enikő. 2006. Review of Noveck and Sperber eds. 2004. *Intercultural Pragmatics* 3: 111–122.
- Németh, T., Enikő. 2008. Verbal information transmission without communicative intention. *Intercultural Pragmatics* 5: 153–176.
- Németh, T., Enikő, and Bibok Károly (eds.). 2010. *The role of data at the semantics–pragmatics interface*. Berlin & New York: Mouton de Gruyter.
- Nersessian, Nancy. 1993. In the theoretician's laboratory: Thought experiments as mental modelling. In: *Proceedings of the 1992 biennial meeting of the philosophy of science association*, Vol. 2, 291–301. Michigan: Philosophy of Science Association.
- Nickles, Thomas. 1980. Scientific discovery, logic and rationality. In *Scientific discovery, logic, and rationality*, ed. Thomas Nickles, 1–59. Dordrecht: Reidel.
- Nicolle, Steve, and Billy Clark. 1999. Experimental pragmatics and what is said: A response to Gibbs and Moise. *Cognition* 69: 337–354.
- Norton, John D. 2004a. Why thought experiments do not transcend empiricism. In *Contemporary debates in philosophy of science*, ed. Christopher Hitchcock, 44–66. Oxford: Blackwell.
- Norton, John D. 2004b. On thought experiments: Is there more to the argument? In: *Proceedings of the 2002 biennial meeting of the philosophy of science association* [= *Philosophy of Science* 71: 1139–1151].
- Noveck, Ira A. 2001. When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition* 78: 165–188.
- Noveck, Ira, and Anne Reboul. 2008. Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences* 12: 425–431.

- Noveck, Ira, and Dan Sperber (eds.). 2004a. *Experimental pragmatics*. Basingstoke: Palgrave Macmillan.
- Noveck, Ira A., and Dan Sperber. 2007. The why and how of experimental pragmatics. In *Pragmatics*, ed. Noël Burton-Roberts, 184–212. Basingstoke: Palgrave Macmillan.
- Noveck, Ira and Dan Sperber. 2004. Introduction. In eds. Noveck, Ira and Dan Sperber, 1–23.
- Peijnenburg, Jeanne, and David Atkinson. 2003. When are thought experiments poor ones? *Journal for General Philosophy of Science* 34: 305–322.
- Penke, Martina and Anette Rosenbach eds. 2007. *What Counts as Evidence in Linguistics?* Amsterdam and Philadelphia: John Benjamins.
- Polya, George. 1948. *How to solve it*. Princeton: Princeton University Press.
- Polya, George. 1954. *Patterns of plausible inference*. London: Oxford University Press.
- Recanati, François. 1989. The pragmatics of what is said. *Mind and Language* 4: 295–329.
- Recanati, François. 2001. What is said. *Synthese* 128: 75–91.
- Recanati, François. 2004. *Literal meaning*. Cambridge: Cambridge University Press.
- Rescher, Nicholas. 1976. *Plausible reasoning*. Assen and Amsterdam: Van Gorcum.
- Rescher, Nicholas. 1987. How serious a fallacy is inconsistency? *Argumentation* 1: 303–316.
- Rescher, Nicholas. 2005. *'What if?': Thought experimentation in philosophy*. New Brunswick: Transaction Publishers.
- Rescher, Nicholas. 1991. Thought experimentation in presocratic philosophy. In eds. Horovitz and Massey, 31–41.
- Sadock, Jerrold M. 1978. On testing for conversational implicature. In *Syntax and semantics 9: Pragmatics*, ed. Peter Cole, 281–297. New York: Academic Press.
- Sauerland, Uli, and Kazuko Yatsushiro. 2009. *Semantics and pragmatics*. Basingstoke: Palgrave Macmillan.
- Saul, Jennifer. 2002. What is said and psychological reality: Grice's project and relevance theorist's criticisms. *Linguistics and Philosophy* 24: 347–372.
- Searle, John R. 1969. *Speech acts. An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Sorensen, Roy A. 1992. *Thought experiments*. Oxford: Oxford University Press.
- Sternfeld, Wolfgang ed. 2007. *Data in generative grammar [= Theoretical Linguistics 33/3: 269–413]*.
- Thomason, Sarah G. 1991. Thought experiments in linguistics. In eds. Horovitz and Massey, 247–257.
- Walton, Douglas N. 1992. *Plausible argument in everyday conversation*. Albany: State University of New York Press.
- Walton, Douglas N. 2001. Abductive, presumptive and plausible arguments. *Informal Logic* 21: 141–169.
- Walton, Douglas, Chris Reed, and Fabrizio Macagno. 2008. *Argumentation schemes*. New York: Cambridge University Press.
- Winkler, Susanne and Sam Featherston eds. 2009. *The fruits of empirical linguistics. Vol. 2: Product*. Berlin and New York: de Gruyter.

What Makes a Property “Semantic”?

Michael Devitt

Abstract This paper is concerned with the semantics-pragmatics dispute. It is common to distinguish the “semantic” properties of an utterance from its “pragmatic” properties, and what is “said” from what is “meant”. What is the basis for putting something on one side rather than the other of these distinctions? Such questions are usually settled largely by appeals to intuitions. The paper rejects this approach arguing that we need a theoretical basis for these distinctions. This is to be found by noting that languages are representational systems that scientists attribute to species to explain their communicative behaviors. We then have a powerful theoretical interest in distinguishing, (a), the representational properties of an utterance that arise simply from the speaker’s exploitation of a linguistic system from, (b), any other properties that may constitute the speaker’s “message”. I call the former properties “semantic”, the latter, “pragmatic”. The semantic ones are constituted by properties arising from linguistic conventions, disambiguations, and reference fixings. Devitt foreshadows an argument in a forthcoming book, *Overlooking conventions: the trouble with linguistic pragmatism*, that many of the striking examples produced by linguistic pragmatists exemplify semantic rather than pragmatic properties. This argument counts against the popular pragmatist theses of “semantic underdetermination” and “truth-conditional pragmatics”. It is very much in the spirit of the tradition that pragmatists reject.

Perspectives on Pragmatics and Philosophy. A. Capone, F. Lo Piparo, and M. Carapezza, eds. Springer.

M. Devitt (✉)
Graduate Center, City University of New York, 365 Fifth Avenue,
New York, NY 10016, USA
e-mail: mdevitt@gc.cuny.edu

1 Introduction

Perhaps the most exciting development in recent philosophy of language has been the debate surrounding a movement that emphasizes the “pragmatic” features of language over the traditional “semantic” ones. The movement is often known as “linguistic contextualism” but I think Neale’s (2004) term for it, “linguistic pragmatism”, is more apt.¹ The movement’s seminal work is probably Sperber and Wilson’s *Relevance* (1986/1995). In this paper I propose a basis for settling this “pragmatics versus semantics” dispute and foreshadow my argument to settle it on that basis.

The folk seem to distinguish what a person *says*, or *literally says*, in an utterance from what the person *means*, from the intended message of the utterance. Spoonerisms provide entertaining examples: thus, when Spooner *said* that a student was hissing his mystery lectures what he *meant* was that the student was missing his history lectures.² Paul Grice emphasizes a distinction along these lines between “what is said” and what is “implied, suggested, meant” (1989: 24), giving many interesting examples; thus, a philosopher writes a reference in which he *says* that a student’s English is excellent and his attendance regular but what the philosopher *means*, his “conversational implicature”, is that the student is no good at philosophy (p. 33). Many would say that what is said is a “semantic” matter but what is merely meant is a “pragmatic” matter.

According to a traditional view, stemming from Grice, a large part of “what is said” by an utterance is constituted by the conventional (linguistic) meanings of the expressions in the language employed by the speaker. These meanings are “known”, in some sense, by a competent speaker of the language simply in virtue of her being competent. They are said to be “encoded”. However, those meanings do not usually exhaust what is said, for two reasons. (1) An expression will frequently be ambiguous: more than one meaning is conventionally associated with it. If an expression is ambiguous, what is said when it is used will be partly determined by which of its meanings the speaker has in mind, by which convention she is participating in. (2) An utterance may contain indexicals (and tenses), deictic demonstratives, or pronouns, the references of which are not fully determined simply by conventions. The reference of a “pure” indexical is partly determined by facts about the speaker: ‘I’ refers to whoever is the speaker, ‘here’ to his spatial location, ‘now’ to his temporal location. It is natural to say that the reference of a demonstrative like ‘that’ or a pronoun like ‘it’ is determined by what the speaker has in mind in using the term. What is it for the speaker to have *x* in

¹ The term, and the related term “linguistic pragmatist”, are loose but convenient. In using these terms I do not mean to suggest, of course, that those in the movement agree about everything.

² I have earlier placed quite a bit of trust in ordinary ascriptions of saying that: “So it is likely that, at least, we ought to ascribe to tokens that are thought and uttered the properties that we do ascribe and hence that those properties are meanings” (1996: 71). I acknowledged the apparent folk distinction between saying that and meaning that but did not make much of it (p. 59 n.).

mind in using the term? It is for that use to be the expression of a part of a thought that refers to x .³ And according to the best theory, in my view (2004: 290–5), that part of the thought refers to x because it is related in a certain sort of causal-perceptual way to x . (It follows from this, note, that the reference of the term is determined by *a mental state of the speaker*. The context external to the speaker’s mind plays a reference-determining role only to the extent that relations to that context partly constitute the mental state.)

It is taken for granted by almost all that “what is said” involves disambiguation and reference determination as well as the conventions of the language employed, as well as what is strictly encoded. The controversy is over whether anything else determined in context, and if so what else, is involved. And over whether the constitution of what is said is “semantic” or “pragmatic”. Pragmatists want to include much more into what is said and think that their enlarged what is said is partly “pragmatic”. This is their theses of “semantic underdetermination” and “truth-conditional pragmatics”.

Pragmatists are led to their theses by a range of interesting phenomena. Here are some common examples⁴:

- (1) I’ve had breakfast.
- (2) You are not going to die.
- (3) It’s raining.
- (4) The table is covered with books.
- (5) Everybody went to Paris.

Taken literally, (1) seems to say that the speaker has had breakfast sometime in the past and yet, in context, it likely means that she has had breakfast this morning. Similarly, (2) seems to attribute immortality to the addressee but, in context, will mean something like that he will not die from that minor cut. Although (3) does not say so explicitly it surely means that it is raining in a certain location. According to the standard Russellian account, (4) makes the absurd claim that there is one and only one table and it is covered with books. Yet it is surely being used to say that a certain table is so covered. And (5) seems to say that every existing person went to Paris and yet the message it surely conveys is that everyone in a certain group went to Paris. Examples like these are taken to show

³ Some prefer to say that the reference is determined by what the speaker “intends to refer to”. This can be just a harmless difference but it may not be. Having x in mind in using the term simply requires that the part of the thought that causes that use refers to x . In contrast, for a speaker *literally* to intend to refer to x , given that intentions are propositional attitudes, seems to require that she entertain a proposition containing the concept of *reference*. So she can’t refer without thinking about reference! This would be far too intellectualized a picture of referring. Uttering and referring are intentional actions, of course, but it seems better to avoid talking of intentions when describing them.

⁴ Taken from François Recanati, who has a helpful discussion of them (2004: 8–10). My (2013b) is a critical discussion of Recanati’s position on the semantics-pragmatics issue.

that a deal of “pragmatic” enrichment is needed to get from what is “semantically” determined to “what is said”.

The debate has yielded many theories and a bewildering array of distinctions, and terminology. The many uses of “semantic”, “pragmatic”, and “what is said” are particularly troubling. One wonders immediately which terms and distinctions are appropriate. And this leads to a deeper question: How should we get to the truth of the matter on the semantics-pragmatics dispute? This methodological question receives a clear answer from the literature: we should rely on our intuitions. I shall start this paper by rejecting that answer. I shall go on to propose an alternative methodology. I shall use this to motivate some terminology and distinctions, in particular a notion of *what is said*. This notion is (or is close to) the traditional one that the pragmatists oppose. I shall conclude by foreshadowing an argument for a controversial view on the substantive issue. In *Overlooking Conventions: The Trouble With Linguistic Pragmatism* (forthcoming), I argue that the linguistic pragmatists have made the wrong response to many of the striking phenomena that they have noticed. In particular, I argue that there is no interesting “semantic underdetermination” or place for “truth-conditional pragmatics”. My view is that the striking phenomena are best accommodated within a view that is in the spirit of the tradition.

2 The Role of Intuitions⁵

As noted, the received view among linguistic pragmatists is that we should find the truth about language by consulting intuitions. Thus, consider what some leading figures have to say. Stephen Neale gives the following sweeping endorsement of the role of intuitions:

Our intuitive judgments about what *A* meant, said, and implied, and judgments about whether what *A* said was true or false in specified situations constitute the primary data for a theory of interpretation, the data it is the theory’s business to explain. (2004: 79).

Jason Stanley and Zoltan Szabó claim that “accounting for our ordinary judgements about the truth-conditions of various sentences is the central aim of semantics” (2000: 240). Robyn Carston thinks that the various criteria in the pragmatics literature for placing “pragmatic meanings” into “what is said”, “in the end, ... all rest ... on speaker/hearer intuitions” (2004: 74). François Recanati claims that “‘what is said’ must be analysed in conformity to the intuitions shared by those who fully understand the utterance” (2004: 14).

This enthusiasm for consulting intuitions is not confined to the linguistic pragmatists, of course: intuition mongering is the *modus operandi* of the

⁵ I draw on a fuller discussion in Devitt (2012).

philosophy of language in general. A famous passage from Saul Kripke’s *Naming and Necessity* is often cited in support:

Of course, some philosophers think that something’s having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don’t know, in a way, what more conclusive evidence one can have about anything, ultimately speaking. (1980: 42).

A similar practice is also to be found in linguistics. Noam Chomsky claims that ‘linguistics...is characterized by attention to certain kinds of evidence...largely, the judgments of native speakers’ (1986: 36). Liliane Haegeman, in a popular textbook, says that ‘all the linguist has to go by...is the native speaker’s intuitions’ (1994: 8).

It is clear from these claims, and other similar ones, that intuitions are commonly thought to provide *the evidence* or, at least, *the main evidence* for theories of language. Yet, often, claims of this sort seem to suggest a stronger view: *the very task* of theories of language is to explain or systematize competent speaker’s intuitions about language.

Such a view of the task is very puzzling. For, the obvious way to describe the task of the theory of language is to *explain the nature of language*, to explain properties like *meaning*, *truth*, *reference*, and *grammaticality*, real properties of linguistic expressions playing some sort of explanatory role. If we start from this view, surely as good a starting place as one could have, why take the task to be to capture *ordinary intuitions about* such properties, intuitions that must simply reflect folk theory? Nobody would suppose that the task of physics, biology, or economics was to explain folk intuitions, why should the situation be any different in linguistics and semantics?

So let us set aside the idea that the study of linguistic intuitions is the task of theories of language and turn to the more plausible view that these intuitions are evidence for those theories.

Why do philosophers have this view of the evidential value of intuitions? One answer seems to be that the intuitions are a priori. Michael McKinsey thinks that it is “fairly clear” that “the principle that the meanings of words are knowable a priori...is taken for granted by most philosophers of language and by many linguists” (1987: 1). I think McKinsey may be right about this. If so, the situation is very regrettable. Even if one supposes that we have a priori knowledge of logic and mathematics, which I don’t think we should (1998, 2011a), it is a giant leap to suppose that we have it of meanings. I have argued that we should not take the leap (1996: 52–3; 2012).

If we don’t take the a priori leap, what else could justify the ubiquitous reliance on intuitions in the philosophy of language? Philosophers provide little in the way of an answer. As Jaakko Hintikka remarks, talking of philosophy generally: “One searches the literature in vain for a serious attempt to provide” a justification for the appeal to intuitions (1999: 130). In similar vein, Timothy Williamson remarks: “there is no agreed or even popular account of how intuition works, no accepted explanation of the hoped-for correlation between our having an intuition that P and

its being true that P”. He describes this as “a methodological scandal” (2007: 215). Still, we can look to linguistics for a possible answer for *semantic* intuitions. (Perhaps we should see this answer as an explication of the obscure a priori answer.)

The common linguistic view of intuitive judgments is expressed in passages like the following:

it seems reasonably clear, both in principle and in many specific cases, how unconscious knowledge issues in conscious knowledge...it follows by computations similar to straight deduction. (Chomsky 1986: 270).

I have described the common view as follows: linguistic competence, all on its own,

provides information about the linguistic facts....So these judgments are not arrived at by the sort of empirical investigation that judgments about the world usually require. Rather, a speaker has a privileged access to facts about the language, facts captured by the intuitions, simply in virtue of being competent... (2006a: 483–4; 2006b: 96).

On this view, intuitive syntactic judgments are, “noise” aside, “the voice of competence”. Let’s call this thesis “VoC”.⁶

Stephen Stich has suggested that philosophers of language might be guided by linguistics in seeking a justification for the authoritative role given to semantic intuitions (1996: 40; see also Hintikka 1999; Williamson 2007). Stich’s particular concern is with referential intuitions. He suggests that philosophers may think that speakers derive their referential intuitions from a representation of referential principles. So, just as, according to VoC, the true grammar that linguists seek to discover is already represented in the mind of every speaker, so too, according to this suggestion, are true semantic theories of reference, meaning, and the like. Semantic intuitions, like syntactic ones, are the result of something like a deduction from a represented theory. Thus, speakers have “Cartesian access” to linguistic facts simply in virtue of being competent.

I have argued that VoC is highly implausible and urged instead that we should see intuitions about language, like intuitions about anything, as ordinary empirical judgments (1994, 1996, 2006a, b, c; see also Kornblith 1998).⁷ The view I have urged is, briefly, as follows:

Linguistic intuitions do not reflect information supplied by represented, or even unrepresented, rules in the language faculty. Rather, they are empirical central-processor responses to linguistic phenomena differing from other such responses only in being fairly immediate and unreflective. (2006b: 10).

Like all empirical responses, these intuitive judgments are theory-laden.⁸

⁶ My attribution of VoC to linguists has been surprisingly challenged (Collins 2008a: 17–19; Fitzgerald 2010; Ludlow 2011: 69–71). I have responded (2010a: 845–7; 2013d: sec. 4).

⁷ For an exchange on VoC, see Rey 2013 (sec. 3.1) and Devitt 2013d (sec. 6).

⁸ This is not to say that they are the result of *theorizing* or should count as *theoretical* (cf. Mišcević 2006: 539; Devitt 2006c: 595). They are like “observation” judgments. As such, they are theory-laden in the way we commonly think observation judgments generally are.

On this modest view—modest because it makes do with cognitive states and processes we were already committed to—the importance of competence to intuitions is not as *a source of information* but as *a source of data*.⁹ The competent speaker has ready access to a great deal of linguistic data because she and her competent fellows produce them day in and day out. So she is surrounded by tokens that may, *as a matter of fact*, refer to so and so, be true in such and such circumstances, say that so and so, be ambiguous, and so on. So she is in a position to have well-based opinions about language by reflecting on these tokens. This is not to say that she will reflect. Indeed, a totally uneducated person may reflect very little and hence have few if any intuitive judgments about her language. Still it is clear that the normal competent speaker with even a little education *does* reflect on linguistic reality just as she reflects on many other striking aspects of the world she lives in. And this education will usually provide her with the terms and concepts of folk semantics and linguistics, at least. As a result she is likely to be able to judge in a fairly immediate and unreflective way that a token *does* refer to so and so, *is* true in such and such circumstances, *is* grammatical, *is* ambiguous, and so on. Such intuitive opinions are empirical central-processor responses to linguistic phenomena. They have no special authority: although the speaker’s competence gives her ready access to data it does not give her Cartesian access to the truth about the data.

Still, are these intuitions likely to be right? I think we need to be cautious in accepting them: thinking about language is notoriously hard and the folk are a long way from being experts. Nonetheless, it does seem to me that their “simplest” intuitions, involving syntactic and semantic vocabulary that we suppose the folk have mastered well enough, are quite likely to be right. So we can often be confident about judgments that a name “refers” to *x*, that this pronoun must “refer to the same thing” as that name, and that this expression is “ambiguous”. In sum, we have good reason to suppose that the core judgments of folk linguistics, reflecting the “linguistic wisdom of the ages”, are good evidence for linguistic theories.

This having been said, the intuitions that philosophers and linguists should prefer are the ones that they do, as a matter of fact, mostly prefer: those of philosophers and linguists themselves. For, they are much more expert. This is particularly so when we get beyond the simple cases to fanciful ones like Kripke’s Gödel (1980) and Putnam’s Twin Earth (1975).¹⁰

To say that intuitions, whether those of philosophers or of the folk, are evidence is not to say that they are the only, or even the primary, evidence. Science in general does not rest on the evidence of people’s intuitions, no more should the theory of language. Biology proceeds not by consulting people’s intuitions about

⁹ The sense of “data” here is the same as in “primary linguistic data”. So the data provided by competence are *linguistic expressions* (and the experiences of using them) *not any observational reports* about those expressions.

¹⁰ I argue this in a critical discussion (2011b) of Machery et al. (2004).

living things but rather by seeking direct evidence from living things themselves. The theory of language should proceed similarly. So, instead of relying on the indirect evidence of linguistic intuitions, the theory should seek direct, less theory-laden, evidence by studying what the intuitions are *about*, the linguistic reality itself. This reality is to be found in linguistic behavior (usage). I shall say more about this in [Sect. 3](#).

In sum, the received view that intuitive metalinguistic judgments should be the evidence, or at least the main evidence, in the philosophy of language should be rejected. Reliance on intuitions is a methodological flaw of linguistic pragmatism.¹¹

A general moral to be drawn from this is that the claims of any theorist about language should not be assessed simply by consulting intuitions. These intuitions reflect empirical theories, largely folk theories, that are themselves open to question. The study of language is empirical and theoretical through and through. Applying this moral to the semantics-pragmatics dispute, we should not take the traditional Gricean notion of *what is said* for granted: we need to argue that the notion is theoretically interesting and, even if it is, that our intuitions about its application are probably right. And we should make the same theoretical demands on all the notions that pragmatists have introduced; for example, Sperber and Wilson's notion of an *explicature* (1995), and many people's notions of *the proposition expressed*. We should look critically at the *theoretical motivation* for these notions. I turn now to that task.

3 The Theoretical Motivation

What theoretical point is there to the traditional distinction between what is said and what is meant? And, what theoretical point is there to the similar distinctions made by many theorists? I aim to be very careful in answering these questions because the truth of the matter about the semantics-pragmatics dispute turns on the answers.

We need to start our answers by considering why we are interested in language in the first place. Our theoretical interest in language comes from our theoretical interest in thoughts and their communication.

It is a piece of folk wisdom that people have thoughts, which is to say that they have beliefs, desires, and other such "propositional attitudes", mental states with intentional contents or meanings. So the folk are "intentional realists". I think the case for this realism is very strong (2006b: 125–7). We need to posit thoughts to people for at least two reasons: to explain people's behaviors and to explain the

¹¹ Elsewhere (2013a) I argue that linguistic pragmatism has two other methodological flaws: confusing the metaphysics of meaning with the epistemology of interpretation; and accepting Modified Occam's Razor. I think (forthcoming) that these flaws have been significant in causing the mistaken thesis of "semantic underdetermination".

way they use others as a guide to a largely external reality. Thus ascribing to Mark the belief that it is raining explains both Mark’s picking up an umbrella and how those present gain information about the weather (by assuming that he is reliable about such matters). We clearly have a great theoretical interest in the details of this process of explaining behavior and learning from each other.

It is a familiar piece of folk psychology that, without any involvement of language, we can sometimes use our insight into other minds and knowledge of the world to figure out what a person thinks. Thus, we might come to our view that Mark believes it is raining from observing him putting on a raincoat. And he might deliberately communicate his belief to us, without using language, by pointing upwards meaningfully as he puts on the raincoat.

We have a similar theoretical interest in the inner states of other organisms and their communication. We posit these states to explain behavior and to explain how one organism can communicate “information” to another. There is much debate in cognitive ethology and comparative psychology about these matters. There is no presumption that an organism’s learning from another must involve a language. At one extreme, chemical detectors may sometimes do the job. At the other extreme, the idea is seriously entertained that this learning should sometimes be explained by attributing to an organism something like human insight into other minds.

So, we don’t always have to posit languages to explain this learning. Still, scientists often do. What are they thus positing? What is a language? It is a system of representations or symbols that is constituted by a set of governing rules, and that a group of organisms use to communicate with each other. Most such languages are not very interesting because they simply communicate information about the animal’s own current state; for example, that the animal is hungry, or wants a mate. The interesting ones are the ones known as “referential”, ones that convey information about the environment. The honey bee provides a famous, and very surprising, example. The bee uses a “waggle dance” to communicate the direction and distance of a food source. Gunnison’s prairie dogs provide another example: they have a system of barks that convey information about which sort of predator is threatening and about the characteristics of a particular predator of that sort. Clearly, the whereabouts of food is a pressing concern for the bee, the presence and nature of a predator, for the prairie dog. A bee that has returned from a food source has reliable information about the former, a prairie dog that has observed a predator, the latter. Their languages enable them to communicate this valuable information.¹²

The rules of the bee’s language are very likely entirely innate. The rules of the prairie dog’s language seem to be partly learned and, perhaps we should say, “conventional”: its alarm calls vary a bit from colony to colony; and when an experimenter used a plywood model to simulate a new sort of predator, the prairie

¹² And it is worth noting that sometimes we are confident that an animal has a language because we have *taught* it one; think of some dolphins and primates that have been taught surprisingly complex languages.

dogs introduced a new call (Slobodchikoff 2002). In any case, whether a language used to communicate information is innate or conventional, we have a powerful theoretical interest in that language and its rules. Serious scientists work to discover the natures of the symbols in these representational systems, to discover their *meanings*.

Karl von Frisch is a notable example. He won a Nobel Prize for his discoveries about the bee's dance. I shall simplify by ignoring what he discovered about how the dance conveys the distance of the food source, attending only to what it conveys about direction. Von Frisch found the following remarkable rule:

To convey the direction of a food source, the bee varies the angle the wagging run makes with an imaginary line running straight up and down...If you draw a line connecting the beehive and the food source, and another line connecting the hive and the spot on the horizon just beneath the sun, the angle formed by the two lines is the same as the angle of the wagging run to the imaginary vertical line. (Frank 1997: 82).

In hypothesizing that a certain behavior in members of a species involves a symbol that represents something in their language, we are supposing that the behavior was produced *because*, in some sense, it involves that symbol representing something in their language; and it is *because* of what the symbol represents in their language that other members of the species respond to the behavior as they do. And it is because of what it represents that the symbol plays its striking role in the life of an organism.

Evidence for such hypotheses is to be found, of course, in regularities in behavior. Thus, von Frisch's hypothesis was offered as an explanation of his many painstaking observations of the bee's behavior. But is it the *best* explanation? For some time it was not obvious that it was. A rival hypothesis was that a bee heading off in the direction of the food source was not responding to information communicated by a bee's dance but rather was following an odor trail left by other bees. But this rival did not stand up to ingenious experiments. The consensus is now that the hypothesis that the bee is using the language described by von Frisch is indeed the best explanation of the bee's behavior (Dyer 2002; Riley et al. 2005; Vladusich et al. 2006).¹³

Return to humans. It is a truism that they have languages which they use to communicate "messages": as the folk say, "language expresses thought". This idea seems irresistible once one has accepted intentional realism, accepted that humans have thoughts (2006b: 127–8). As Fodor, Bever, and Garrett say, "there is much to be said for the old-fashioned view that speech expresses thought, and very little to be said against it" (1974: 375). So, just as the bees and the prairie dogs have

¹³ For more on this issue see Devitt (2006c) (pp. 585–586) responding to Smith (2006) (pp. 440–441).

representational systems used to communicate the contents of inner states to each other, so do we.¹⁴ The evidence for this in our behavior seems overwhelming.¹⁵

Consider again our example of Mark and the ascription to him of the belief that it is raining. Suppose that the people present ascribe this belief on the basis of his production of the sound /It is raining/. According to the rules of English, this sound means that it is raining (ignoring here, for convenience, pragmatists’ worries about the “underdetermination” of such weather reports). If the people assume that Mark is being literal and straightforward, they will take that meaning to be the message the speaker intentionally communicates, his “speaker meaning”. As a result, they have evidence of his thoughts. Taking him to be sincere in his expression, they conclude that he has a belief with that meaning (content), the belief that it is raining. In this way, language is an extraordinarily effective way of making the thoughts of others accessible to us, thoughts that otherwise would be largely inaccessible; and of making our thoughts accessible to others, often in the hope of changing their thoughts and hence their behavior. Even though, as we noted, the thoughts of others are sometimes accessible to us without language, they mostly are not.

Just as we have a powerful theoretical interest in the languages of bees and prairie dogs, we have one in human languages and their rules: we need to know about the natures of the representations used to communicate in these systems.¹⁶

The rules of human languages are largely conventional.¹⁷ Indeed, it is a truism that expressions in a language (like English) have their meanings by convention. As David Lewis points out at the beginning of his classic, *Convention*, it is a “platitude that language is ruled by convention” (1969: 1). Where do these conventions come from? On rare occasions they are established by some influential people stipulating that a certain form has a certain meaning. However, following Grice (1989) and Schiffer (1972), I think that the conventions associated with a linguistic form—a sound, an inscription, etc.—in a community typically come from the regular use in the community of that form to convey certain parts of messages. That regular use of it in utterances with a certain speaker meaning, leads, somehow or other, to that form having that meaning conventionally in the language of that community.

This being so, we find evidence of linguistic conventions, hence evidence of the rules that constitute a human language, in such regularities in behavior. We should

¹⁴ Language is used for purposes other than communication: to muse, to make notes, to try out a line for a poem, and so on. Communication is the most striking use of language but, contrary to “intention-based semantics”, it is not essential to language. What is essential to (human) language is the expression of thought.

¹⁵ Nonetheless, this view of language is rejected by Chomskians. They see a human language as an internal state not a system of external symbols that represent the world. I argue that this is deeply misguided (Devitt 2006a: chs 2 and 10; 2006c; 2008a, b).

¹⁶ Some philosophers and linguists, impressed by the great difference between a human language and the representational systems of other animals, resist calling those systems “languages”. I can see no theoretical point to this resistance. In any case, the point is merely verbal.

¹⁷ I say “largely” because I do not reject the Chomskian view that some syntax is innate. The qualification should be taken as read in future.

posit a conventional rule for a human language, just as we posit an innate rule for the bee's, if it provides the best explanation of the observed behavior.

I said earlier (Sect. 2) that the primary evidence for our theories of language should not be the indirect evidence of intuitions but rather the direct evidence provided by linguistic reality itself. The regularities in linguistic behavior, or usage, that provide evidence of linguistic conventions are to be found in this reality. How are we to gather this evidence? I have urged elsewhere (2011b, c) that it be gathered by the informal observation, and the scientific study, of the corpus of the language. Furthermore, it can be gathered using “the technique of elicited production”: we can conduct experiments in which situations are constructed or described and people are prompted to see what they say and understand in those situations.

Conventions should loom very large in our view of human language in general¹⁸ and in our response to linguistic pragmatism in particular. What can be said about them? I have just talked of regularities in usage causing conventions and providing evidence for them. This is not to say that the regularities constitute conventions. Since linguistic conventions can be created by stipulation and agreement, there can be conventions without regularities. And there can be regularities without conventions, regularities that are best explained otherwise. So, what then is a linguistic convention and what is it to participate in one? Lewis (1969) and Schiffer (1972) have a lot to say about this, of course, but their accounts, particularly requirements of “common knowledge” (Lewis) or “mutual knowledge*” (Schiffer), seem too intellectualized.¹⁹ And this appearance is reinforced by the thought that we probably have to see the behavior of some nonhuman animals as conventional; see above on the prairie dogs. We seek a less intellectualized view.

Where there is a linguistic convention for an expression in a community, members of that community share a communicative disposition to associate the expression with a certain speaker meaning, a certain part of messages. But this sharing is not enough. The sharing has to be partly explained by appropriate causal relations between the speakers' communicative dispositions as Lewis and Schiffer made clear. If we back away from intellectualized accounts it is, of course, hard to say precisely what causal relations are needed. Still, the center of what has to be said is that any speaker has her disposition *because other speakers have theirs*; there is a certain sort of dependency of the disposition of each member of the

¹⁸ In stark contrast, Chomsky thinks that the “regularities in usage” needed for linguistic conventions “are few and scattered” (1996: 47; see also 1980: 81–83). Furthermore, such conventions as there are do not have “any interesting bearing on the theory of meaning or knowledge of language” (1996: 48). I think these views are very mistaken (2006b: 178–89; see also 2006c: 581–2, 598–605; 2008a: 217–229).

¹⁹ Stephen Laurence (1996, 1998) rejects “convention-based semantics” largely on the basis of criticisms of Lewis' account of conventions in general and of linguistic conventions in particular. These criticisms are partly of the highly intellectualized nature of Lewis' account. I sympathize with those criticisms. I reject the rest (2006b: 180).

community on the dispositions of others. And for a person to participate in this convention is for her to act on such a dependent disposition.

This account does not go nearly as far as one would like. How concerned should we be about that? Not very, in my view.²⁰ We may lack a satisfactory complete account of conventions but what we do have does not leave conventions in any way mysterious. And we *need* a notion of convention to explain much human behavior including, as just noted, human linguistic communication. So even though not fully explained, the notion is much needed and not mysterious. Science is full of such notions. Indeed, it *must* be since, as they say, at any time “explanations must stop somewhere”.

Return to the bee’s language. Our theoretical interest is not simply in the rules of this language, discovered by von Frisch, but in the representational properties of any *particular* dance that is governed by those rules. For, it is in virtue of the dance having those properties that it plays its striking role in the lives of the bees that are present: it is because the dance has those properties that the one bee produces it and the other bees respond to it with certain flight behaviors. *And the dance does not get all its explanatory representational properties, all its meaning, simply from the rules of the bee’s language.* Thus, a dance at a certain angle to the vertical represents a particular direction of the food source *only partly* in virtue of being governed by the rule described above, *only partly* in virtue of what is encoded. The dance represents that direction also partly in virtue of the actual direction of the spot on the horizon beneath the sun from the beehive at the time of the dance. Those two factors constitute the message about direction that the dance communicates; they constitute “what the bee says” by that particular dance “utterance”. And it is important to note that the second factor, part of what constitutes the representational property or meaning that we are theoretically interested in, is provided “in the context of utterance”. The linguistic rule “demands” that this contextual “saturation”²¹ be taken into account in constituting the property that interests us; it demands that the “slot” for the direction of the spot on the horizon be filled.

Similarly with human languages, the linguistic rules for any word with an indexical or demonstrative element demand saturation in the context. These rules differ from those of the bees in being conventional not innate, but that is beside the point. Thus, suppose that I am right in my earlier claim that, according to a linguistic rule, a token of the demonstrative ‘that’ refers to whatever object is linked to it in the appropriate causal-perceptual way. So that rule captures the convention for expressing the demonstrative part of a thought, the convention for expressing that particular sort of meaning in an utterance. It is simply in virtue of being governed by that rule that a demonstrative has its encoded meaning. A

²⁰ John Collins thinks otherwise. He criticizes my previous discussions along these lines by claiming that “the relevant notion of convention remains wholly opaque” (2008b: 245; see also 2008a: 35) and hence, by implication, unacceptable. I have responded (2008b).

²¹ ‘Saturation’ is Recanati’s neat term for what is similarly demanded by a human language (2004: 7).

consequence is that any token that is governed by this rule has its representational property partly in virtue of the object that is as a matter of fact linked to it in that causal-perceptual way. So it has that property partly in virtue of something that is not encoded. Just as with the bee, a linguistic rule “demands” a certain contextual saturation in constituting the property of theoretical interest, the representational property that plays a striking role in the lives of the humans that are present. It is because of what the demonstrative refers to as a result of saturation that a speaker produces it and that hearers respond to it with behaviors that are likely to involve, in some way, the object of reference.

We have earlier noted (Sect. 1) that what is said by a human utterance, according to the tradition, is often partly constituted by whatever determines the reference of indexical or demonstrative elements. We have just seen the theoretical need for a notion with this property. The tradition has got it right here.

We have earlier noted also that, where an expression is ambiguous in the language employed by the speaker then, according to the tradition, the constitution of what is said takes account of which of the expression’s meanings the speaker has in mind: it takes account of the speaker’s “disambiguation” of the expression. This contribution to what is said is also motivated by our interest in the representational properties or meanings that play the striking role. Ambiguity arises when more than one linguistic rule governs a symbol. With human languages, this multiplicity arises from multiple conventions for the symbol; the symbol encodes more than one meaning in the language.²² The multiplicity demands disambiguation to arrive at the representational properties that are of theoretical interest. Thus, we are interested in *which* of the conventionally possible representational properties or meanings of ‘bank’ is actually instantiated in a particular utterance of ‘I went to a bank’; we are interested in *which* of the two conventions for ‘bank’ the speaker is participating in. A language is constituted by a system of linguistic rules. It is because speakers exploit those rules that the language is such an effective device for communicating with those who share the language. The explanatory role of a particular symbol depends on which rule has been exploited. The tradition is right again.

So, our theoretical interest in the representational properties that an utterance has simply in virtue of the speaker’s exploitation of a language motivates an interest in three sorts of properties. (1) Our interest starts with encoded properties, ones that the utterance has solely as a result of the conventional linguistic rules of the language, rules for using certain physical forms to express certain parts of thought contents. But we are interested also in two other sorts of properties. (2) Where a symbol in an utterance is ambiguous and so covered by more than one linguistic rule, we have an interest in which meaning the speaker has in mind, hence which rule is in fact governing the symbol in this utterance. (3) Where a symbol is governed by a rule that demands saturation—for example, an indexical,

²² Multiplicity could, in principle, arise from multiple innate rules for a symbol but, so far as I know, there is no example of this in nature.

tense, or demonstrative—we are interested in the results of that saturation. I call any property, like (1) to (3), that is simply a language-exploiting representational property of an utterance part of “what is said”, thus staying close to the tradition. Waiving my usual resistance to talk of propositions (1996: 210–15), I also sometimes call what is said, “the proposition said”.²³

Does any sort of property of an utterance other than (1) to (3) go into what is said? Does our interest in language-exploiting representational properties motivate an interest in any other sort? I know of no reason for supposing so and will assume not.

My notion of *what is said* is (or is close to) the traditional one that the linguistic pragmatists oppose. The usual case for this notion rests on appeals to intuitions. I certainly think that the notion is intuitively appealing, but that is not my case for it. I claim to have provided a theoretical motivation for the notion.

In *Overlooking Conventions* (forthcoming), I go further in the spirit of the tradition by arguing, controversially, that much of the pragmatist phenomena fall under this notion of what is said. I shall say a bit more about this argument in Sect. 6.

So, we have theoretically motivated a notion of *what is said* by an utterance. We have earlier motivated the familiar talk of “thoughts” and have identified the “message” of an utterance as the thought that the speaker intentionally conveys, what the speaker means by the utterance.²⁴ Waiving my resistance again, we could also call this, “the proposition meant”. Earlier examples (Sect. 1) from Spooner and Grice show vividly that what is meant can differ dramatically from what is said. Spooner said that a student was hissing his mystery lectures and the philosopher said that a student’s English is excellent and his attendance regular: those, respectively, are the properties that these utterances have simply in virtue of the speaker’s exploitation, in the context, of the conventional linguistic rules of the language. Yet what Spooner meant, the message he intended to convey, was that the student was missing his history lectures. Similarly, the philosopher, and the message that the student was no good at philosophy.²⁵

If the linguistic form used by a speaker to convey a thought expresses that thought according to the conventions of the language, then the message will be what is said. If not, the message might be quite different from what is said, as Grice’s example illustrates. Something important follows from that possible difference: the fact that *p* is *what is said* by an utterance does not entail that *p* is *meant* by the utterance (does not entail that *p* is the utterance’s *message*). This is a consequence of the notions of saying and meaning that I claim to have

²³ This needs qualification because what is said using a sub-sentential is sometimes only a fragment of a proposition (forthcoming).

²⁴ A speaker might intend to convey more than one message. I am simplifying by setting this possibility aside.

²⁵ I am ignoring, for convenience, the distinction between sayings that are statings or assertions and sayings that are mere rehearsings. It is convenient to ignore this because whether the saying is a stating or a rehearsing, its language-exploiting content is the same.

theoretically motivated. In contrast, many theorists have notions of saying and meaning that have the consequence that saying that p entails meaning that p . I can see no theoretical motivation for this.²⁶

Grice calls the message conveyed in his example a “conversational implicature”. There are other ways in which what is meant can depart from what is said, what I call “pragmatic enrichments” and “pragmatic impoverishments”.

The notions of what is said and what is meant that I have explained are at the absolute center of my response to linguistic pragmatism.

4 Terminology

The naming of what is said raises the vexed matter of terminology. We need to address this matter thoroughly to avoid merely verbal disputes. In addressing it, we should introduce terms only for notions that have been theoretically motivated. And, given the confusing array of terminology already in the debate, we should try to stay as close as possible to familiar uses of familiar terms. I take it that my uses of ‘what is said’ and ‘the proposition said’, and ‘the message’ and ‘the proposition meant’, meet these criteria. But how should we use ‘semantic(s)’ and ‘pragmatic(s)’?

Many philosophers favor a narrow use of ‘semantics’ according to which it refers to the study of only the conventional linguistic rules of a language, rules that determine what is encoded in the language. Related to this, these philosophers take ‘semantic property’ to refer only to those properties of expression *types* fully determined by those linguistic rules, properties “known” by every competent speaker of the language. There can be no objection to this usage. Our theoretical interest in what is said by an utterance (in the sense just explained) motivates an interest in any property that constitutes what is said. And our discussion has shown that properties determined by linguistic rules constitute a large part of what is said. So we have a strong motivation for identifying and naming those properties and their study. Indeed, von Frisch won a Nobel as a result of his interest in the rules of the bee’s dance. However, we also need a term for the study of what is said itself. And we need a term for *all* the properties that constitute it, properties of types (1) to (3), not just for the ones determined by the rules of the language, properties of type (1). ‘Semantics’ and ‘semantic’ seem, respectively, ideal for these jobs too. Semantics in this wide sense includes semantics in the narrow sense because the

²⁶ Of course, it would not be the case that a certain proposition was said by an utterance, any more than it would be the case that a certain proposition was meant by it, were it not for the fact that the speaker intentionally produced that utterance. So we might ordinarily say that the speaker “meant” both propositions. But this ordinary way of speaking does not provide a theoretical motivation for treating the distinct items, what is said and the message, as parts of the one “meaning”. (There is a subtle issue about the relation of the intentional act of uttering to what is said; see Devitt 2013c: note 15 and accompanying text).

conventional properties of an expression type are among the ones that constitute what is said by any particular use of that type in an utterance.

Kent Bach is one who favors the narrow uses of ‘semantics’ and ‘semantic’ and rather deplors their wide uses “for features pertaining to truth and reference” (1994: 133 n. 9).²⁷ Nathan Salmon is another, and launches a polemic against “the speech-act centered conception” which takes “semantics as concerned with what a speaker *says* or *asserts* in uttering a declarative sentence” (2005: 321). In response, I emphasize first that this *is* only a verbal matter and hence not worthy of much passion.²⁸ Second, there is a custom, as both Bach and Salmon acknowledge, of using these terms widely. Third, this wide use relates the semantics of language nicely to what Jerry Fodor has neatly called “psychosemantics” (1987), the study of thought meanings (contents). For those meanings are truth-referential. In any case, I use the terms widely unless they are accompanied by the qualification ‘narrow’.

I call the *properties* that constitute what is said, “semantic”. So it is appropriate to also call “semantic” the *processes* in a speaker that are responsible for utterances having those semantic properties, processes of convention participation and reference fixing. However, there seems to be little occasion to do so.

Finally, we must turn to the *really* vexed matter of ‘pragmatics’ and ‘pragmatic’. These terms have distressingly many uses in the literature, as Bach nicely demonstrates (1999). One common way to use ‘pragmatics’ is for the study of communication or, as it is usually called, “interpretation”, the study of the way that a hearer comes to understand the message conveyed by a speaker’s utterance: “the goal of pragmatics is to explain how the hearer’s task...can be carried out” (Sperber and Wilson 1995: 179; see also: 10). And the processes that are thus studied are called ‘pragmatic’. I go along with this usage. And I take it that the terms so used also cover processes in the speaker of choosing among ways to communicate a message.

A word of warning about “interpretation”. Like many English nouns this one is ambiguous between a process and a product. So it can refer to the process of interpreting the meaning of an utterance or it can refer to the meaning assigned by that process. In calling the theory of interpretation “pragmatics”, as I have just done, I am taking it to be concerned with the process of interpreting. If we took it to be concerned with the meaning assigned by the process then it would be akin to

²⁷ My (2013c) is a detailed discussion of Bach’s position on the semantics-pragmatics issue. I argue that his notion of “what is said” is not theoretically motivated and that his methodology for deciding what counts as “semantic” (in his sense) is faulty, which has the conservative effect of keeping out new meanings.

²⁸ Salmon is used to hearing this response and is not impressed: “‘It’s all just terminology’ is the last refuge of the speech-act centered conception” (2005: 327). Kepa Korta and John Perry’s view (2007: 98–99) that “something has gone awry” in Cappelen and Lepore’s account of “semantic minimalism” (2005: 143–145) seems to rest simply on an insistence that ‘semantics’ be used narrowly.

semantics, a very different matter.²⁹ Because of this ambiguity, I do not talk of the theory or study “of interpretation”.

So we have defined a use of ‘pragmatic’ to describe certain processes, the ones involved in interpreting communications. And we have defined a use of ‘pragmatics’ to describe the study of those processes. Understood in this way, pragmatics and what is pragmatic are *obviously* very different from semantics and what is semantic: pragmatics is about certain *processes* (mostly) in a hearer, semantics is about certain *properties* of an utterance. Of course, those semantic properties play a big role in the pragmatic processes: the hearer’s grasp of what is said is a key part, at least, of her process of figuring out the message communicated (Sect. 5). Still, understood in this way, there is no way in which semantics and pragmatics are *rivals*. What then are we to make of the pragmatics versus semantics dispute?

The answer is that it is common to use ‘pragmatic’ also to describe *properties of an utterance that are not semantic*, in some preferred sense, and yet contribute to a message conveyed; for example, Gricean implicatures and certain enrichments and impoverishments. And it is common to use ‘pragmatics’ also to describe the study of these non semantic properties. So, understood in this way, “pragmatics” and “semantics” offer different accounts of the properties of utterances and *are* rivals of some sort (see, for example, Bach 1999; Carston 2007). It is only when ‘pragmatics’ is understood in this way that the dispute over where to draw the line between pragmatics and semantics is intelligible. For want of better terms, I shall reluctantly go along with this common usage, with “*nonsemantic*” understood in *my sense*, of course.

So ‘pragmatics’, as commonly used and as used by me, is ambiguous, referring sometimes to the study of communication and sometimes to the study of these pragmatic properties of utterances. This ambiguity does not get the attention it deserves.³⁰

Many use ‘pragmatic’ to describe more properties of an utterance than the ones that are nonsemantic in my sense: it is used to describe all the properties of what is said other than conventional properties of expression types, other than encoded *narrow-semantic* properties (e.g., Sperber and Wilson 1995: 36, 217; Recanati 2004: 8; Carston 2007). Using the term in this way to describe what we might call “wide-pragmatic” properties is unfortunate because it conceals the theoretically crucial distinction between the nonconventional properties of an utterance that constitute what is said and those that constitute what is meant *but not said*. For the former properties, but not the latter, are the ones the utterance has simply in virtue of the speaker’s exploitation of her language. They are ones that make the utterance an instance of her linguistic system just as, in the case of a bee’s dance, its

²⁹ Stephen Neale’s “theory of interpretation” seems to cover both processes and products (Neale 2004: 71–90).

³⁰ This is presumably related somehow to the confusion of the metaphysics of meaning with the epistemology of interpretation that I discuss elsewhere (2013a).

property of representing the direction of the food source, a property it has partly in virtue of the actual direction from the beehive of the spot on the horizon beneath the sun, makes the dance an instance of the bee’s linguistic system. So, in my usage, pragmatic properties are not part of what is said.

It is important to note that although some properties of an utterance may be pragmatic in one sense, *all* the hearer’s processes of understanding the utterance’s message are pragmatic in the other sense, whether they involve grasping pragmatic properties or semantic properties (what is said). And all of these processes are studied by pragmatics as the theory of communication (and none of them by pragmatics as the theory of pragmatic properties).

5 Communication

It has always been clear that we can use our insight into other minds to successfully convey messages without using language; Mark’s pointing upwards meaningfully as he puts on the raincoat is an example. And Grice gave us vivid examples of our using our insight into other minds, *together with what is said*, to successfully convey messages that differ from what is said; thus, the philosopher conveyed the message that the student was not good at philosophy without using any conventional form for expressing that thought and by saying something entirely different. In these cases the message conveyed is not the representational property of the symbol produced but something that the speaker supposes that the hearer can infer in part from her awareness of that property.

Where communication involves a language, the representational properties of symbols in the language should certainly be center stage in a theory of communication among the organisms. For, it is largely, even if not entirely, because symbols of the language have those properties, and because the organisms have the capacity to exploit those properties in sending and receiving messages, that communication occurs.³¹ So pragmatics, as the theory of communication, must start with what Stanley and Szabó call the “descriptive” problem of semantics, the assigning of “semantic values” to expressions of the language in question (2000: 223). (They contrast the descriptive problem with the “foundational” one concerned with in virtue of what expressions have those values.). But the theory of communications requires much more than this description of the representational properties of the symbols in that language.

Consider communication among the bees, for example. We start with the theory of their language provided by von Frisch, together with the obvious fact that the bees are competent in that language. But then we need to know how bees can use that competence to send and receive particular messages. How does a sender bee use its competence to turn its “knowledge” of the direction of the food and of

³¹ See Pagin and Pelletier (2007: 35) for a nice statement of this sort of view.

the direction of the spot on the horizon beneath the sun into a dance at an angle that, according to the rules of its language, conveys the direction of the food? How does a receiver bee use its competence to turn its “knowledge” of that angle and of the direction of the spot on the horizon into the message about where the food is? We know sadly little of the answers.

The point is even more striking with communication among humans, of course. For this communication involves not only features analogous to those in bee communication but also disambiguations, pragmatic enrichments and impoverishments, and implicatures, so admirably demonstrated by the pragmatists. But, to emphasize, pragmatics, as the theory of this communication, starts with the semantics of the language, with what is said, but goes much further. And whereas the focus of the semantics is on the speaker because the speaker determines both what is said and the message, the focus of pragmatics is on the hearer because the task of receiving a message is more difficult than the task of sending one.

6 The Semantics-Pragmatics Dispute

My methodology for tackling the semantics-pragmatics dispute starts from the view that a language is a representational system posited to explain communication.³² From this start I provide a theoretical motivation for a sharp distinction between two sorts of properties of an utterance. On the one hand there are its (wide-) semantic properties. These are properties that an utterance has simply in virtue of the speaker’s exploitation of the conventional rules of her language. They include, but are not limited to, an utterance’s encoded narrow-semantic properties, ones that are fully determined by the rules governing the symbol type of which the utterance’s symbol is a token. Semantic properties constitute what is said, the proposition said. On the other hand, there are the utterance’s other properties which, perhaps along with semantic ones, contribute to the message conveyed, the proposition meant. These are the utterance’s pragmatic properties.³³

From my perspective, semantics is concerned with the representational properties that symbols have in virtue of being uses of a language, the properties that constitute what is said. These properties contribute to conveying the message of an utterance. Other factors may also contribute to conveying the message but these are not the concern of semantics. So the key issue for the semantics-pragmatics dispute is the nature of those linguistic representational properties. The symbols have those properties largely in virtue of being part of a representational system of conventional rules. So the key issue comes down to: What are the conventions that constitute the system?

³² And to explain some non-communicative uses of language; see note 14.

³³ As Manuel García-Carpintero says: “the semantics/pragmatics divide is in my view ultimately about which meaning-properties are constitutive of natural languages and which are not” (2005: 43).

How do we answer this key question? As noted in [Sect. 3](#), we look for evidence from regularities in behavior (usage). Is this expression regularly used to express a certain speaker meaning? If so, is this regularity best explained by supposing that there is a convention of so using the expression? We take properties to be semantic if that is the best explanation of regularities in behavior.

In [Sect. 3](#), I argued that what is said is constituted by three sorts of properties: briefly, those arising from (1) convention, (2) disambiguation, and (3) saturation. And I assumed that what is said is constituted by no other sorts of properties. This view of what is said by an utterance is part of the traditional view that is rejected by the linguistic pragmatists. Another part of that view, implicitly at least, is that what is thus said is, with perhaps the occasional exception, the literal truth condition (meaning, content) of the utterance: what is said in that sense is standardly the message conveyed. The pragmatists have mounted a formidable challenge to this view. They use a range of striking examples, illustrated by utterances (1) to (5) in [Sect. 1](#), to argue that the literal truth conditions of utterances are mostly, if not always, underdetermined by properties (1) to (3); the truth conditions are “semantically underdetermined”; a new theoretical framework is called for. In *Overlooking Conventions* (forthcoming) I confront this challenge by arguing, controversially, that many of the pragmatists’ examples exemplify properties of sorts (1) to (3). There are more of such properties than we have previously noted: much more of the content of messages should be put into the convention-governed what is said—into semantics - than has been customary. The pragmatists have made the wrong response to their examples. So I am urging a view that is very much in the spirit of the tradition. No new framework is called for.

The argument for this view requires, of course, a discussion of the striking examples. This cannot be attempted here. However I shall summarize one such discussion that is already published, my discussion of definite descriptions ([1981](#); [1997a, b](#); [2004](#); [2008c, d](#)).

Utterances containing referential uses of descriptions, like the example in [Sect. 1](#),

(4) The table is covered with books,

are among the pragmatists’ favorite examples of a pragmatically enriched what is said. Pragmatists from Grice onwards have denied that these uses of descriptions should be explained semantically by taking descriptions to have a referential meaning as well as the generally accepted Russellian quantificational meaning. Instead, pragmatists tend to think that referential uses are best explained pragmatically as conversational implicatures (or something similar). I have presented a number of reasons for thinking that such pragmatic explanations are wrong. The most important is “The Argument from Convention”.³⁴ The basis for the thesis that descriptions have referential meanings is not simply that we *can* use them referentially for, as the pragmatists point out, we can use any quantifier referentially. The basis is rather that we *regularly* use descriptions referentially. Indeed,

³⁴ This is Neale’s apt name for the argument ([2004](#): 71). The Argument from Convention is also to be found in Reimer ([1998](#)).

the vast majority of uses of descriptions are referential. Furthermore, there is no good pragmatic explanation of this regularity. The best explanation is a semantic one: there is a convention of using descriptions referentially.³⁵

Definite descriptions illustrate previously unnoticed semantic properties of all three sorts. Concerning (1), a referential use of a description exemplifies an overlooked referential convention. Concerning (2), since there is also a quantificational convention for descriptions, what is said when a description is used referentially is partly constituted by disambiguation. Concerning (3), the convention for referential descriptions demands saturation in the context.

The view of the semantics-pragmatics issue that I am urging, in the spirit of the tradition, goes against pragmatist theses of “semantic underdetermination” and “truth-conditional pragmatics”. We all accept, of course, that semantic conventions do not alone determine what is said, do not alone determine truth-conditional content: disambiguation and saturation are also required. But pragmatists think that there is much more semantic underdetermination than this. Relatedly, many urge truth-conditional pragmatics. According to that thesis, the meaning of the sentence in an utterance does not alone yield a truth-conditional content (even after disambiguation and reference fixing); it needs to be pragmatically supplemented and can be so in indefinitely many ways yielding indefinitely many truth conditions.³⁶ A consequence of my view is that there are no interesting theses of semantic underdetermination and truth-conditional pragmatics.³⁷ Semantic properties alone constitute the theoretically interesting notion of what is said.

7 Conclusion

How should we discover the truth about language? The paper rejects the received answer that we should do so by consulting our intuitions about language. Those intuitions reflect empirical theories, largely folk theories, that are themselves open to question. In answering the question, we should start by considering what a language is and why we sometimes suppose that organisms have one. A language is a representational system posited to explain communication. This gives a powerful theoretical interest in the natures of the representations yielded by this system. For human languages, I argue, these natures are constituted by three sorts of properties, briefly, those arising from convention, disambiguation, and saturation: these three sorts of properties constitute “what is said” by an utterance, “the

³⁵ In my view, Felipe Amaral (2008) has driven the final nail into the coffin of the received Russellian view of descriptions with his discussion of “Kripke’s Test”.

³⁶ Bezuidenhout (2002); Recanati (2003, 2004); Searle (1980); Travis (1996, 1997, 2006).

³⁷ Truth-conditional pragmatics has a further problem, I argue (2013b): it cannot explain how the conventionally constituted property of a sentence in an utterance constrains a truth conditional content without determining one; nor how the property could allow indefinitely many truth conditions.

proposition said”. This amounts to a theoretical motivation for a fairly traditional Gricean notion of *what is said*.

Evidence of the properties that constitute what is said is to be found in behavior. In particular, the regular use of an expression with a certain speaker meaning provides evidence of a convention of so using it and hence of what makes up the language.

The theoretical motivation provides a sharp distinction between two sorts of properties of a linguistic utterance: its “semantic” properties which constitute what is said and the proposition said; its other, “pragmatic”, properties which, perhaps along with semantic ones, contribute to “the message” conveyed, “the proposition meant”.

The “semantic” theory of the representational properties of symbols in a language is one thing, the “pragmatic” theory of communication among users of that language, another. Yet the semantic theory should be center stage for the pragmatic one. For, it is largely, even if not entirely, because symbols of the language have those “semantic” properties, and because the organisms have the capacity to exploit those properties in sending and receiving messages, that communication occurs. And whereas the focus of the semantic theory is on the speaker because the speaker determines both what is said and the message, the focus of pragmatic theory is on the hearer because the task of receiving a message is more difficult than the task of sending one.

Implicitly at least, the traditional view is that what is said by an utterance, in something like the sense I have motivated, is standardly the literal truth-conditional content communicated. I have foreshadowed an argument that confronts the challenge that the linguistic pragmatists have posed to this view. The argument is that many of the pragmatists’ striking examples exemplify properties arising from convention, disambiguation, and saturation and so go into constituting my semantic what is said. This counts against the popular pragmatist theses of “semantic underdetermination” and “truth-conditional pragmatics”. And it is very much in the spirit of the tradition that pragmatists reject.³⁸

The Graduate Center, The City University of New York

References

- Amaral, Felipe. 2008. Definite descriptions are ambiguous. *Analysis* 68: 288–297.
- Bach, Kent. 1994. Conversational implicature. *Mind and Language* 9: 124–162.
- Bach, Kent. 1999. The semantics-pragmatics distinction: what it is and why it matters. In *The semantics-pragmatics interface from different points of view*, ed. Ken Turner, 65–84. Oxford: Elsevier.

³⁸ The first version of this paper, under the title, “Pragmatics versus Semantics”, was delivered at an International Conference, “Meaning”, University of Erfurt (Germany), in September 2009. Later versions, sometimes under the title “Linguistic Pragmatism”, were delivered at many universities. I am grateful for comments at these talks.

- Bezuidenhout, Anne. 2002. Truth-conditional pragmatics. *Philosophical Perspectives* 16: 105–134.
- Cappelen, H., and E. Lepore. 2005. *Insensitive semantics: a defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell Publishing.
- Carston, Robyn. 2004. Truth-conditional content and conversational implicature. In *The semantic/pragmatics distinction*, ed. Claudia Bianchi, 65–100. Stanford: CSLI Publications.
- Carston, Robyn. 2007. Linguistic communication and the semantics/pragmatics distinction. *Synthese* doi. doi:[10.1007/S:1122900791918](https://doi.org/10.1007/S:1122900791918).
- Chomsky, Noam. 1980. *Rules and representations*. New York: Columbia University Press.
- Chomsky, Noam. 1986. *Knowledge of language: its nature, origin, and use*. New York: Praeger Publishers.
- Chomsky, Noam. 1996. *Powers and prospects: reflections on human nature and the social order*. Boston: South End Press.
- Collins, John. 2008a. Knowledge of language redux. *Croatian Journal of Philosophy* VIII, 3–43.
- Collins, John. 2008b. A note on conventions and unvoiced syntax. *Croatian Journal of Philosophy* VIII, 241–247.
- Devitt, Michael. 1981. Donnellan's distinction. *Midwest Studies in Philosophy, Volume VI: The Foundations of Analytic Philosophy*, eds. Peter A. French, Theodore E. Uehling Jr., and Howard K. Wettstein, 511–524. Minneapolis: University of Minnesota Press.
- Devitt, Michael. 1994. The methodology of naturalistic semantics. *Journal of Philosophy* 91: 545–572.
- Devitt, Michael. 1996. *Coming to our senses: a naturalistic program for semantic localism*. Cambridge: Cambridge University Press.
- Devitt, Michael. 1997a. Meanings and psychology: a response to Mark Richard. *Nous* 31: 115–131.
- Devitt, Michael. 1997b. Responses to the Maribor papers. In *The Maribor papers in naturalized semantics*, ed. D. Jutronic, 353–411. Maribor: Pedagogška fakulteta Maribor.
- Devitt, Michael. 1998. Naturalism and the a priori. *Philosophical Studies* 92, 45–65. Reprinted in Devitt 2010b.
- Devitt, Michael. 2004. The case for referential descriptions. In *Reimer and Bezuidenhout 2004*: 280–305.
- Devitt, Michael. 2006a. Intuitions in linguistics. *British Journal for the Philosophy of Science* 57: 481–513.
- Devitt, Michael. 2006b. *Ignorance of language*. Oxford: Clarendon Press.
- Devitt, Michael. 2006c. Defending *Ignorance of language*: responses to the Dubrovnik papers. *Croatian Journal of Philosophy* VI, 571–606.
- Devitt, Michael. 2008a. Explanation and reality in linguistics. *Croatian Journal of Philosophy* VIII, 203–231.
- Devitt, Michael. 2008b. A response to Collins' note on conventions and unvoiced syntax. *Croatian Journal of Philosophy* VIII, 249–255.
- Devitt, Michael. 2008c. Referential descriptions and conversational implicatures. *European Journal of Analytic Philosophy* 3: 7–32.
- Devitt, Michael. 2008d. Referential descriptions: a note on Bach. *European Journal of Analytic Philosophy* 3: 49–53.
- Devitt, Michael. 2010a. Linguistic intuitions revisited. *British Journal for the Philosophy of Science* 61: 833–865.
- Devitt, Michael. 2010b. *Putting metaphysics first: essays on metaphysics and epistemology*. Oxford: Oxford University Press.
- Devitt, Michael. 2011a. No place for the a priori. In *What place for the a priori?*, eds. Michael J. Shaffer and Michael L. Veber, 9–32. Chicago and La Salle: Open Court Publishing Company. Reprinted in Devitt 2010b.
- Devitt, Michael. 2011b. Experimental semantics. *Philosophy and Phenomenological Research* LXXXII, 418–435.
- Devitt, Michael. 2011c. Whither experimental semantics? *Theoria* 72: 5–36.

- Devitt, Michael. 2012. The role of intuitions. In *Routledge companion to the philosophy of language*, eds. Gillian Russell, and Delia Graff Fara, 554–565. New York: Routledge.
- Devitt, Michael. 2013a. Three methodological flaws of linguistic pragmatism. In *What is said and what is not: The semantics/pragmatics interface*, eds. Carlo Penco and Filippo Domaneschi, 285–300. Stanford: CSLI Publications.
- Devitt, Michael. 2013b. Is there a place for truth-conditional pragmatics? In a book symposium on François Recanati’s *Truth-Conditional Pragmatics*. *Teorema* XXXII: 85–102.
- Devitt, Michael. 2013c. Good and bad Bach. *Croatian Journal of Philosophy*.
- Devitt, Michael. 2013d. Linguistic Intuitions are not ‘the Voice of Competence’. In *Philosophical Methodology: The Armchair or the Laboratory?* ed. Matthew Haug, 268–293. London: Routledge.
- Devitt, Michael. Forthcoming. *Overlooking conventions: the trouble with linguistic pragmatism*.
- Dyer, F.C. 2002. The biology of the dance language. *Annual Review of Entomology* 47: 917–949.
- Fitzgerald, Gareth. 2010. Linguistic Intuitions. *British Journal for the Philosophy of Science* 61: 123–160.
- Fodor, Jerry A. 1987. *Psychosemantics: the problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
- Fodor, Jerry A., T.G. Bever, and M.F. Garrett. 1974. *The psychology of language: an introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill Book Company.
- Frank, Adam. 1997. Quantum honey bees. *Discover*, 80–87.
- Grice, Paul. 1989. *Studies in the way of words*. Cambridge MA: Harvard University Press.
- Haegeman, Liliane. 1994. *Introduction to government and binding theory*, 2nd edn. Oxford: Blackwell Publishers. 1st edn. 1991.
- Hintikka, J. 1999. The emperor’s new intuitions. *Journal of Philosophy* 96: 127–147.
- Kornblith, H. 1998. The role of intuition in philosophical inquiry: an account with no unnatural ingredients. In *Rethinking intuition: the psychology of intuition and its role in philosophical inquiry*, eds. M.R. DePaul, and W. Ramsey, 129–141. London: Rowan and Littlefield Publishers.
- Korta, Kepa, and John Perry. 2007. Radical minimalism, moderate contextualism. In *Preyer and Peter* 2007: 94–111.
- Kripke, Saul A. 1980. *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Laurence, Stephen. 1996. A Chomskian alternative to convention-based semantics. *Mind* 105: 269–301.
- Laurence, Stephen. 1998. Convention-based semantics and the development of language. In *Language and thought: interdisciplinary themes*, eds. Peter Carruthers, and Jill Boucher, 201–217. Cambridge: Cambridge University Press.
- Lewis, David K. 1969. *Convention: a philosophical study*. Cambridge, MA: Harvard University Press.
- Ludlow, Peter. 2011. *The philosophy of generative linguistics*. Oxford: Oxford University Press.
- McKinsey, M. 1987. Apriorism in the philosophy of language. *Philosophical Studies* 52: 1–32.
- Machery, Edouard, Ron Mallon, Shaun Nichols, and Stephen P. Stich. 2004. Semantics, cross-cultural style. *Cognition* 92: B1–B12.
- Miščević, Nenad. 2006. Intuitions: The Discrete Voice of Competence. *Croatian Journal of Philosophy* VI, 523–548.
- Neale, Stephen. 2004. This, that, and the other. In *Reimer and Bezuidenhout* 2004: 68–182.
- Pagin, Peter, and Jeffrey Pelletier. 2007. Context, content, and composition. In *Preyer and Peter* 2007: 25–62.
- Preyer, Gerhard, and Georg Peter (eds.). 2007. *Context-sensitivity and semantic minimalism*. Oxford: Oxford University Press.
- Putnam, H. 1975. The meaning of ‘meaning’. In *Mind, language and reality: philosophical papers Vol. 2*, 215–271. Cambridge: Cambridge University Press.
- Recanati, François. 2003. The limits of expressibility. In *John Searle*, ed. Barry Smith. 189–213. Cambridge: Cambridge University Press.
- Recanati, François. 2004. *Literal meaning*. Cambridge: Cambridge University Press.

- Reimer, Marga. 1998. Donnellan's distinction/Kripke's test. *Analysis* 58: 89–100.
- Reimer, Marga, and Anne Bezuidenhout 2004. eds. *Descriptions and beyond*. Oxford: Clarendon Press.
- Rey, Georges. 2013. "The possibility of a naturalistic cartesianism regarding intuitions and introspection". In *Philosophical Methodology: The Armchair or the Laboratory?* ed. Haug, 243–267. Matthew London: Routledge.
- Riley, J.R., U. Greggers, A.D. Smith, D.R. Reynolds, and R. Menzel. 2005. The flight of honey bees recruited by the waggle dance. *Nature* 435: 205–207.
- Salmon, Nathan. 2005. Two conceptions of semantics. In *Semantics versus pragmatics*, ed. Zoltan G. Szabó. Oxford: Clarendon Press, 317–328.
- Schiffer, Stephen. 1972. *Meaning*. Oxford: Clarendon Press.
- Searle, John. 1980. The background of meaning. In *Speech act theory and pragmatics*, eds. John Searle, Ferenc Kiefer, and Manfred Bierwisch, 221–232. Dordrecht: Reidel.
- Slobodchikoff, C.N. 2002. Cognition and communication in prairie dogs. In *The cognitive animal: empirical and theoretical perspectives on animal cognition*, eds. Colin Allen, and Gordon M. Burchardt, 257–264. Cambridge MA: MIT Press.
- Smith, Barry C. 2006. Why we still need knowledge of language. *Croatian Journal of Philosophy* VI, 431–456.
- Sperber, Dan, and Deirdre Wilson. 1995. *Relevance: communication and cognition*, 2nd edn. 1st edn. 1986. Oxford: Blackwell Publishers.
- Stanley, Jason, and Zoltán G. Szabó. 2000. On quantifier domain restriction. *Mind and Language* 15, 219–261.
- Stich, S.P. 1996. *Deconstructing the mind*, New York: Oxford University Press.
- Travis, Charles. 1996. Meaning's role in truth. *Mind* 100: 451–466.
- Travis, Charles. 1997. Pragmatics. In *Companion to the philosophy of language*, eds. B. Blake, and C. Wright, 87–107. Oxford: Blackwell.
- Travis, Charles. 2006. Insensitive semantics. *Mind and Language* 21: 39–49.
- Vladusich, Tony, Jan M. Hemmi, and Jochen Zeil. 2006. Honeybee odometry and scent guidance. *The Journal of Experimental Biology* 209: 1367–1375.
- Williamson, T. 2007. *The philosophy of philosophy*. Oxford: Blackwell Publishing.

What is a Context?

Steven Gross

Abstract In this chapter, I examine a variety of criteria for context individuation, with an eye towards identifying and distinguishing various theoretical projects in semantics and pragmatics for which they would be appropriate. The discussion is organized around six possible desiderata for contexts: (1) that they metaphysically determine proposition expressed, (2) that they be non-trivial (in a sense explained in the text), (3) that they provide epistemic illumination of how hearers understand what proposition was expressed, (4) that they be (in various senses) finite, (5) that they be non-intentionally characterized, and (6) that they be context-insensitively characterized. Particular attention is directed to the conceptions of context developed by Lewis, Stalnaker, and Davidson(ians).

Keywords Context • Semantics • Pragmatics • Lewis • Stalnaker • Davidson

1 Introduction

Language users occasionally utter sentences. In attempting to explain this behavior, theorists about language often advert to various features of these occasions under the rubric: *the context of utterance*. What is a context?

Let's narrow our scope somewhat. Sentences are uttered as parts of various speech-acts, and features of the occasion may be invoked to explain what determines *which* speech-act is performed. Suppose, though, we just assume it given

This paper originally appeared as Chapter II of my 1998 Harvard Ph.D. Dissertation, which was subsequently published as Gross (1998/2001). Many thanks to Alessandro Capone for the invitation to reprint it here.

S. Gross (✉)
Johns Hopkins University, Baltimore, US
e-mail: sgross11@jhu.edu

that the speech-act in question is assertion. Now, a typical instance of assertion involves (1) an utterance of a sentence with a certain standing meaning in a language, in order thereby (2) to express a proposition, and perhaps in addition (3) to communicate pragmatically some further proposition(s). Features of the occasion of utterance may be invoked in explaining facts at any of these three levels—the pre-semantic, semantic, and post-semantic, as they might be called.¹ But suppose we focus just on level (2): what determines what proposition is expressed, given the meaning of the sentence uttered. (It's clear that natural language sentences are pervasively context-sensitive in this sense: what propositions speakers express in uttering them is not determined by their meaning alone. 'The kettle is black' can be used to express a variety of propositions, depending for instance on the kettle in question, the time of utterance, whether the layer of soot matters (the speakers might be photographers composing a black-and-white still-life, or they might be kettle collectors wondering whether, beneath its sooty exterior, this is a rare *white* 1924 model XYZ), and more.) Our focus thus narrowed, how ought we to characterize and individuate contexts?

The primary point of this chapter is to argue that this is a bad question—at least until we specify what we want a notion of context *for*. Contexts are just (representations of) bundles of facts. There's no sense to inquiring into their proper individuation conditions without an idea of what questions we hope to answer in part through adverting to context. Only given an idea of the task to which they will be put can we debate what information they should encode. We thus can say what a context is only after clarifying our theoretical aims.

Our preliminary narrowing has already illustrated how context may be invoked to help explain different aspects of linguistic behavior. In what follows, we shall see that even *after* this narrowing there remains room for different notions of context, corresponding to different questions one might want answered. Since the proper characterization of context depends on one's particular explanatory ends, certain debates about context lose their point: conceptions of context are not in *competition*, if they are meant to serve different ends. Showing this is fruitful not only because it helps clarify—indeed dissolve—certain disputes, but also because, in forcing a clarification of these varying explanatory ends, it elicits a sharper delineation of the varying projects aimed at those ends.²

I begin by listing six *desiderata* one might want an account of context to satisfy.³ I then turn to three accounts of context (identified with Lewis, Stalnaker,

¹ Cf. Kaplan (1989: 575–576). The first level is “pre-semantic” in that what must be explained is *what* sentence (with what meaning) has been uttered; that sentence's semantic properties are only then put to use at the next level, where propositional content is fixed.

² I shall assume, or stipulate, that contexts don't shift across utterances of individual sentences. (Stanley and Williamson (1995) argue that quantifier domains, and thus contexts, shift in the utterance of such sentences as 'Everyone converted someone' (suppose we've been talking about what the evangelists did upon entering the park). But why not say instead that the context includes a sequence of domains?)

³ I don't claim that this list is exhaustive.

and Davidson) to see how they fare. (Assessing the theoretical potential of these three conceptions of context in particular is thus a secondary aim of this chapter.) What will be crucial to note is that whether these *desiderata* ought to be accepted—whether they place legitimate constraints on an account of context—depends on what work one’s notion of context is supposed to do. Without a clear conception of this, one cannot assess whether a failure to satisfy some *desideratum* should count as an *objection*. We shall see that the three conceptions of context are addressed to three different theoretical tasks, which require satisfying three different sets of *desiderata*.

2 Six *Desiderata*

The first *desideratum* is the minimal demand that the context, given the meaning of the sentence, *determine* or *fix* the proposition expressed. What was said ought to *supervene* on context plus sentence meaning. If this weren’t the case, then one and the same sentence, with the same meaning, could be uttered in one and the same context, and yet express a different proposition. But then, presumably, there was some further factor that in part determined what proposition was in fact expressed. We should then include this factor—or something that determines *it*—in the context. That is, we want contexts to be (if you will) *complete* with respect to the task of determining the proposition expressed, given the meaning of the sentence uttered. I say that this is the minimal demand on contexts, because it simply amounts to a restatement of the task on which I suggested above we focus our sights.

Second, we might require that contexts not be *trivial*. Contexts are trivial if they are characterized in a way that begs the questions we want to answer. For example, a trivial way of satisfying the demand for determination is to characterize contexts in terms of *what proposition is expressed*. What proposition is expressed will of course *supervene* on itself, but what is wanted presumably is some characterization of propositional determination in independent terms. Perhaps it could be argued that no such account is forthcoming, at least none meeting other suitable constraints. But until such an argument is compellingly made, the explanatory aspiration embodied in the first *desideratum* will seem reasonable and worthy of pursuit. Similarly for whatever other explanatory aspirations we might have: if there is some question we want answered, we will want it answered, if possible, in a non-trivial manner.

Third, we might require that contexts be characterized in a manner that is *epistemically illuminating*. For example, we might require that the features constituting context be among those language users actually track in understanding one another—or, at least, that they be such that they *could* in principle enable such understanding, given our cognitive capacities. This would require both that the features be epistemically accessible and that their bearing on propositional content, given sentence-meaning, be apparent (at least upon reflection) to normal speakers

of the language. Suppose what proposition is expressed on any given occasion happens to supervene on some complex microphysiological state of the speaker. Contexts comprising just this state would satisfy the determination requirement in an epistemically *unilluminating* way, since this state would neither be (in general) epistemically accessible to hearers capable of understanding what is said nor such that those hearers would or could (in general) grasp its bearing on content (even if the state *were* epistemically accessible to them).

Fourth, we might require that contexts be *finite*, that they comprise a finite number of features. (If contexts are represented as n-tuples, the requirement would thus be that n be finite.) This requirement could stem from various sources. For instance, it might arise as a constraint imposed by the form semantic theorizing might take. If our task is to *articulate* all the ways contextual features may affect propositional content, its completion might require finiteness.⁴ Or a finiteness constraint might be motivated by learnability considerations.

Now, it may be noted that finiteness seems all too easy to satisfy: one can always replace an infinite sequence with a finite one that contains one object—viz., the infinite sequence itself. So, we might want rather to require that contexts be *finitely grounded*, that they comprise a finite number of features that themselves comprise a finite number of features, and so on for a finite number of levels. There are two other relevant refinements perhaps worth mentioning here. One might want to require that the number of features be not only finite, but also *fixed*, so that no change in the set of relevant utterances (e.g., in anticipation of possible changes in the language) would require the addition of contextual features. Further, one might require not only that contexts be finite, or finitely grounded, but in addition that they be *manageable*—that is, that the list of features be reasonably short. A finite, but utterly unwieldy list might be practically useless, at least for some purposes; and its lack of perspicuousness might render the account of which it is a part unilluminating.⁵

⁴ I hedge because I am simplifying. Perhaps we should allow an infinite number of contextual features so long as they are recursively specifiable—such as: the first demonstrated object, the second demonstrated object, the third demonstrated object, etc. (Suppose we are considering sentences of the form ‘I’ll have one of those, and one of those, and one of those,....’)

⁵ One might also require that each feature admit only a finite number of values. It is arguable, in fact, that such a limit follows from the requirement of epistemic illumination, given our finite cognitive capabilities. But, as will emerge below, it is also arguable that we are capable of being in an infinite number of intentional states at once. Now, if one allows features capable of countably infinite variation, it’s possible to code an arbitrarily large finite number of features into a single feature. And continuous variation enables the coding of countably many—indeed, continuum-many—features into one. But note that such codings will typically not preserve epistemic illumination, since so characterized the bearing of context on content will be opaque to normal speakers. (The reduction of features by coding can still be accomplished even if the variation within a feature is finite: the single feature into which the others are coded will just have to admit as much variation as the product of the variation the multiple features allow. But again the point about the failure to preserve epistemic illumination will often apply.)

Fifth, we might require that contexts be characterized in *non-intentional* terms. Like the finiteness requirement, this constraint might have a variety of motivations. For instance, one's semantic theorizing might be intended to subserve a broader reductionist project. Intentional characterizations will be precluded if one's naturalizing strategy, for example, is to reduce the mental to the linguistic and then the linguistic to the non-intentional.⁶ Or the requirement might stem from a desire to advert only to relatively *epistemically uncomplicated* contextual features—in particular, features not subject to the epistemological holism of the mental. Finally, at least in the case of certain intentional terms, the constraint may simply follow from the previously mentioned fear of triviality, as we shall see below.

The last possible *desideratum* I will mention is that the characterization of context be itself context-insensitive. The worry behind this constraint is that a theory that adverts to context-sensitively characterized contexts will itself be context-sensitive, so that what the theory itself says will vary across contexts—or, rather, there won't be any *one* theory put forward in these terms. It doesn't follow just from this that a context-sensitive characterization of context precludes a *general* theory, one that is about utterances and contexts generally. For perhaps from a distinguished context (the contribution to content of context-sensitive terms fixed by that context), a theory can be put forth that achieves this generality. But if so, then it is straightforward to supply context-insensitive replacements—e.g., by stipulating, as a new technical sense for the terms, that they context-insensitively contribute to content what the terms in their non-technical sense context-sensitively contribute in the distinguished context. This *desideratum*, then, requires that this—or something equivalent—can be done.⁷

3 Three Conceptions of Context

Let me now introduce three conceptions of context found in the literature.⁸ The first is due to David Lewis, the second to Robert Stalnaker. The third—perhaps not due to any one person in particular—is the notion of context required by what I consider the best version of (neo-)Davidsonian semantics. The Lewis and

⁶ Sometimes this strategy permits the reduction of *public* language to the mental, but then argues that the mental must be reduced to something with linguistic structure, a “language of thought.” (See, e.g., Field 1978.) The constraint would thus apply only to any contextual features invoked at that level.

⁷ Appeals to context are criticized in Hornstein (1984: 144–145) on account of the context-sensitivity of ‘context.’ Unfortunately, he simply claims that there is such context-sensitivity, and that semantical theories presuppose that there isn't, without exploring the conceptions of context such theories in fact deploy

⁸ Again, there's no claim to exhaustiveness here.

Stalnaker conceptions of context are easily introduced. Getting the Davidsonian conception on the table will require a few words about his semantic program.

A context, for Lewis, is an ordered triple: the “location—time, place, possible world—where a sentence is said.”⁹ According to Stalnaker, a context is a speaker’s presuppositions, “the propositions whose truth he takes for granted as part of the background of the conversation.”¹⁰ Stalnaker favors representing such a context as the set of possible worlds compatible with what is presupposed; but this is a further step, independent of the basic conception of context.

On the Davidsonian conception, a context comprises the features adverted to by a truth-theory that assigns truth-conditions for the sentences of a given language.¹¹ These truth-conditions are captured by “T-sentences” derivable within the theory. For context-sensitive sentences, which possess *variable* truth-conditions, such T-sentences take the following “normal form”¹²:

(u)(x)(y)... (z)[If uR(S, x, y,..., z), then {u is true iff F(x, y,..., z)}].

(In something closer to English: For any utterance of a sentence S and any contextual features x, y,..., and z, if those features obtain on the occasion of utterance (and meet whatever other relevant constraints there might be), the utterance is true if and only if x, y,..., and z satisfy F.) The features constituting a context would be the values of variables x, y,..., and z. For example, consider the context-sensitivity arising from the variable reference of ‘it’ in ‘It’s green’ (ignore other sources of context-sensitivity). Its T-sentence would be:

(u)(x)If u is an utterance of ‘It’s green’ in which ‘it’ refers to x, then (u is true iff x is green).

The T-sentence requires that contexts be at least individuated according to *what object is referred to by ‘it’ on that occasion*. The T-theory as a whole commits one to individuating contexts according to whatever features are thus adverted to within it.¹³

How well do these three conceptions of context fare in satisfying the various *desiderata* listed above?¹⁴

⁹ Lewis (1980: 79). As we will see, Lewis advocated a different conception of context in earlier work. Lewis’ later notion of context should not be confused with what he calls an index, which basically provides a circumstance of evaluation.

¹⁰ Stalnaker (1978: 84).

¹¹ See, e.g., Davidson (1967).

¹² Such conditionalized T-sentences are found in Burge (1974). I follow Higginbotham (1988) in speaking of normal forms here. Context-sensitivity and truth-theoretic accounts of semantic competence are further discussed in Gross (1998/2001: Chap. III)

¹³ We needn’t say that, since different features may appear in the T-sentences for different sentences, these sentences require different conceptions of context; for we may conglomerate these features into one conception of context and consider some of them suppressed for simplicity’s sake when they are irrelevant.

¹⁴ Begging the reader’s indulgence, I won’t discuss the six *desiderata* in the order presented above, but will rather let expository convenience determine the order in each case.

(a) *Lewis*

Lewis' conception of context as location clearly satisfies *desiderata* 1, 2, 4, and 5. What proposition is expressed certainly supervenes on the utterance's having occurred at that place, time, and world¹⁵; these three features are clearly non-trivial—at least so far as content-determination goes—in that they don't invoke what requires explanation; they are finite in number, finitely grounded, presumably fixed, and, being only three, manageable; and they are obviously non-intentional.

Whether *desideratum* 6 is satisfied—whether such contexts can be context-insensitively characterized—is, however, not so obvious. There are two issues here. First, specifications of place and time are context-sensitive as to their relevant *extent*: how extensive are the place and time that in part constitute the speaker's location? I won't dwell on this, however; perhaps a reasonable stipulation will settle the matter. Second, it is unclear whether we can context-insensitively refer to an arbitrary world. Location in space–time does not pose this problem, since we may avail ourselves of a coordinate system (dubbed with a name—once an origin, orientation, and unit magnitude are stipulated¹⁶). But context-insensitively picking out a location in the space of possible worlds is not so easy. We may provide *our* world the context-insensitive name '@,' and likewise take care of some others we have first context-sensitively picked out. But we seem to lack the means to refer context-insensitively to an arbitrary world.¹⁷

But is this a problem? Lewis has, after all, context-insensitively characterized *what contexts are* (supposing the first concern addressed): they are places, times, and worlds.¹⁸ This suffices to individuate contexts, even if we lack the means context-insensitively to refer to arbitrary contexts. This second problem, then, is really only a problem if there is some reason Lewis need be committed to our having such an ability. We thus see that there is a distinction to be drawn between context-insensitively characterizing what contexts are and context-insensitively characterizing (arbitrary) particular contexts in a manner sufficient to single them out. Identifying what suffices to determine propositional content (given sentence-meaning) requires only the first. Even if your aspirations include articulating a theory that enables the context-insensitive specification, for an arbitrary utterance in a context, of the proposition it would express, this in fact would suffice. Articulating the *theory* doesn't require the ability to refer context-insensitively to arbitrary contexts: it doesn't even require the ability to refer to arbitrary contexts *simpliciter* (whether context-sensitively, or context-insensitively).

¹⁵ Indeed, what sentence is uttered with what meaning supervenes on the utterance's Lewisian location as well!

¹⁶ Of course, there would be no bar to employing context-sensitive terms in thus fixing the name's referent.

¹⁷ There are more worlds than we could name, in any case.

¹⁸ There are indeed questions about how *worlds* are individuated. (Lewis himself takes up the matter in his (Lewis 1986: 69–81).) But, arguably, providing an answer in this case—unlike in the case of a place and a time's *extent*—requires not a stipulation as to the use of a context-sensitive term, but rather a clarification of the individuation criteria with which a context-insensitive expression—viz., 'possible world'—is associated.

However, if you lack this ability, you'll only be in a position to *apply* the theory to utterances entered in contexts to which you could refer.

The failure of Lewis' conception of context to satisfy *desideratum* 3—the requirement that contexts be epistemically illuminating—is more obvious. Lewisian location is certainly knowable (minimally, we can demonstratively identify our current spatial, temporal, and modal whereabouts). But from the fact that P supervenes on Q, it doesn't follow that finite agents, who as it happens can recognize that Q, can thereby learn that P. And indeed knowledge of these features, combined with knowledge of what sentence was uttered with what meaning, does not in general enable normal competent speakers to recover propositional content. Knowing that 'I love you,' for instance, was uttered in our world at precisely noon on January 12, 1879 in room 17 of the Dakota off Central Park obviously does not enable me to know what proposition was expressed.¹⁹ Lewis' conception of context, though it may help answer questions requiring only the *determination* of propositional content, seems ill-suited to addressing how speakers succeed in *understanding* what is said.²⁰

But is this a *criticism* of Lewis? Not if addressing how speakers succeed in understanding what is said just isn't his aim—or at least not his aim in advancing this conception of context. So what is his aim? Lewis tells us that his advancing this conception of context subserves his attempt to provide “part of a systematic restatement of our common knowledge about our practices of linguistic communication [...] that] assign[s] semantic values that *determine* which sentences are true in which contexts.”²¹ Providing that part that covers our *understanding* of what others say is simply not the task he sets himself here. Indeed, in other work, he seems to hold that at least part of this task requires something like Stalnaker's conception of context.²²

(b) *Stalnaker*

Let's turn then to Stalnaker's conception of context as the set of propositions the speaker presupposes to be mutual among participants in the conversation. The first point to note is that for Stalnakerian contexts generally to satisfy *desideratum* 1—for them to determine propositional content, given sentence-meaning—it is arguable that an assumption of normality must be made. A context is normal—or, as Stalnaker puts it, “non-defective”—when the speaker's presuppositions are in fact the same as her audience's.²³ When the context is defective, it *can* be that it is rather the hearer's presupposition set that determines the proposition expressed—at least this is a defensible position in some cases.²⁴ The most likely such cases are

¹⁹ That it might conceivably suffice for some super-intelligent being with extraordinary powers is neither here nor there.

²⁰ Thus Stalnaker (1998: 15), referring to Montague (1968) and Lewis (1970, 1980), comments that “nothing was explicitly said in the theory about the epistemic status of such a context.”

²¹ Lewis (1980: 79) (emphasis added).

²² See Lewis (1979).

²³ Stalnaker (1978: 85).

²⁴ But only in some cases. The point is not that the hearer's presuppositions are what matter in *all* defective contexts. This is the point of the qualifier 'generally' above.

ones in which the hearer's presuppositions are very reasonable, but some of the speaker's are not. For instance, if Tommy and Suzy are playing with Suzy's marbles, and Tommy declares "All the red ones are mine," it is unclear that the fact that he has in mind rather the marbles he left at home should suffice to render them the domain of discourse relevant to the assessment of the proposition expressed (perhaps as opposed to the proposition he *meant*).²⁵ It is unreasonable of Tommy to expect Suzy to share the relevant background beliefs in this case; Suzy's most likely assumption—that he is falsely claiming ownership of the perceptually salient red marbles, the ones before them that they've been playing with for the past hour—is much more reasonable. Perhaps one should conclude in such cases rather that it is in some sense indeterminate what proposition was expressed.²⁶ Then there would be no objection to maintaining that Stalnakerian context determines propositional content (given sentence-meaning), since cases in which there's no (determinate) proposition expressed at all would just be irrelevant to the claim. But if one is willing to allow that propositional content can in some cases be determined by what the *hearer* presupposes, then the claim holds only if normality is assumed.²⁷

To satisfy the second constraint—that the context be non-trivial—the speaker's presuppositions must not include the proposition that she is asserting that P, since among what wants explaining is that (the hearer can grasp that) this is the case. Of course, the speaker *believes* that this is what she is doing, but normally she can't expect the *hearer* to believe this prior to his grasping what the speaker says.²⁸ Once the assertion is made, it can indeed be reasonably expected that the speaker's having asserted that P will be mutually presupposed. This will typically be among the *effects* on context the speaker intends her speech act to have. But this just means that the speaker's having asserted that P will be conversationally available as background for *further* assertions. There seems no reason, then, to doubt the non-triviality of Stalnakerian contexts.

²⁵ I borrow the example from Gauker (1997).

²⁶ Gauker (1997: 28–29) argues against this possibility, claiming that if the utterance expresses no determinate proposition, then the process by which they can clear up their confusion, which seems to take the would-be claim as its basis, is rendered unintelligible. This argument, however, is unconvincing. There's plenty of content around for the process to utilize, even if the utterance itself fails to express a proposition—e.g., the content of the intentions with which he spoke and of the beliefs he had about what was presupposed.

²⁷ Alternatively, one might hold that what determines propositional content (given sentence-meaning) is the set of *reasonable* presuppositions. In normal contexts, this just will be the *speaker's* presuppositions.

²⁸ There are indeed cases in which the hearer may also (correctly) believe that the speaker will assert, and then is asserting, that P—she might know her conversational partner all too well. But then her excellent prediction will call for some other explanation. What's more, from the fact that the speaker and hearer both *believe* something, it doesn't follow that they presuppose it to be mutual. Agreement in belief does not yet amount to common knowledge. There indeed can be cases in which there *is* common knowledge that the speaker is expressing the proposition that P (e.g., in some cases of expressing gratitude or anger, or in prayer), but I would consider these, not cases of assertion proper, but of acknowledgment. Cf. Matthews (1980).

Now, in advertent to a speaker's presuppositions, one obviously violates the non-intentionality constraint—or, rather, what follows is that this conception of context is not suited for projects that motivate this *desideratum*. This is not to say that Stalnakerian contexts can't be deployed by someone with naturalistic ambitions, but only that one's strategy must be of the appropriate sort: Stalnaker himself advocates an approach that would explain linguistic intentionality in terms of mental intentionality, and then the latter in terms of a broadly speaking causal story.²⁹ This involvement with the mental, however, may well *limit* the epistemic illumination forthcoming from an approach that utilizes such contexts. Before examining this, it will be useful to discuss *desideratum* 4—the finiteness constraint.

Stalnakerian contexts are trivially finite; for they comprise a single feature—viz., *what the speaker presupposes*. Here is a case where it's more interesting to ask instead whether the context is *finitely grounded*. The proper answer to this question is clouded by questions concerning the proper analyses of tacit belief and common knowledge.³⁰ I certainly act as if I believe that I am under 10 ft tall. Do I tacitly *believe* it? What about my being under 11, 12 ft tall, etc.? If I do tacitly believe all these things, don't I also tacitly believe that you believe them, and that you believe that I believe them? So, aren't they among my presuppositions? If so, then presupposition sets need not be finitely grounded. And even if I don't have all those tacit beliefs (nor all those presuppositions) about my height, consider some arbitrary P that I *do* presuppose on some occasion of utterance—for instance, that I am speaking. Well, if it's a presupposition, then you presuppose it also. And *that* we both presuppose it will also be included among what we take for granted. So, we'll each presuppose that we each presuppose it, etc. The mutuality built into presupposition can seem to ensure an indefinite iteration of presuppositions. So, if this proves correct, we must again conclude that presupposition sets are not finitely grounded.

These brief remarks are intended merely to mark the issue, not to convince.³¹ At the very least, however, it would seem that the number of propositions in a

²⁹ See Stalnaker (1984: chap. I).

³⁰ The literature is large, but see, e.g., Lycan (1986) on tacit belief, and Smith (1982) on common knowledge.

³¹ Perhaps I believe that, for all $n > 10$, I am under n feet tall. But do I believe each instance? Note that there are finite numbers so large that there exist no numerals referring to them in any physically realizable numeration scheme. (Suppose x is such a number. You might think you could refer to x exploiting a base x numeration scheme, according to which '10' names x . But the world is *ex hypothesi* is "too small" to contain tokens of all the distinct types needed to name the numbers between 0 and x . So such a scheme couldn't be introduced.) It would likewise seem that I am only capable of grasping so many iterations of presupposition. But, on the other hand, perhaps tacit belief (and presupposition) doesn't require that I be capable of *occasionally* entertaining those thoughts (or that they satisfy the constraints imposed by being represented in a Language of Thought). Two further questions. Does mutuality really ensure—or require—an indefinite iteration of presuppositions? And, even if one concedes for the reasons advanced in the text that Stalnakerian contexts are not finitely grounded, should this infinity be thought truly *bothersome*, given that both examples involve recursively characterizable totalities? (Cf. my comment above on multiple demonstrations.)

presupposition set could be enormously large.³² Does this threaten *unmanageability*? That depends on what work the theory employing this notion of context is supposed to do. Let's return to the question of epistemic illumination to see one way this worry could be pressed.

Stalnakerian contexts certainly satisfy the conditions necessary for being epistemically illuminating. It is obvious that the attitudes adverted to are epistemically accessible, and that their bearing on propositional content is apparent to actual speakers. If explaining our ability to understand one another is among our goals, it is clear that presupposition sets to this extent fit the bill in a way that Lewisian locations can't. But why might this conception of context limit the epistemic illumination approaches employing it may provide? Well, suppose we wanted some sense of *how* hearers, on the basis of their knowledge of sentence-meaning and context, (could reasonably) form a belief as to the proposition expressed. A standard thought is to try to articulate rules that could take one from sentence-meaning and context to propositions.³³ But now the worry is that too much has been included in our contexts to make this feasible. Sure, not *all* the speaker's attitudes are there—but still the presupposition set can seem to comprise such a large, unwieldy, holistically intertwined body of information as to render the task intractable. "He's crazy," my conversational partner says. Among the presuppositions of the context are that we have been talking about her husband, that her child is visually salient, that her husband is visually salient, that she is pointing at her child, that a dog just barked, that she is gesturing vaguely with her head over her shoulder, etc. What rule will take me from this mess of facts (and many more) to the conclusion—obvious to anyone minimally competent and present—that, in uttering 'He's crazy,' she referred to her dog (who, by barking, had interrupted her story about her equally crazy husband) while she simultaneously indicated to her husband, by pointing to her child, that he should keep an eye on the baby? Now, merely pointing out the complexity of our cognitive abilities in this fashion of course does not constitute an argument that such rules could not be articulated. But it at least serves to indicate that constructing such a

³² One might try to keep the numbers down by considering only the presuppositions *relevant* to understanding the particular assertion in question. But this threatens to run one afoul of the triviality constraint, if there exists no way of identifying which presuppositions are relevant without advertent to the proposition expressed. And if there *does* exist some way, then it would seem explanatorily preferable to construct a notion of context on *that* basis.

³³ I speak of rules, instead of functions, because we might want more than functions in the set-theoretic sense—i.e., sets of n-tuples. This need not be simply because, as theorists engaged in explanatory projects, we seek illumination, not just correlation, and thus prefer correlations that somehow pattern the data. The motivation might be rather, or further, that language use—the *explanandum* in question—is conscious rational activity, one that requires those who engage in it to address reasonable demands for justification. We might thus require that our representations of the relation between input and output "organize" the chart in a manner that articulates how actual speakers grasp this relation. Part of what theorists want here to comprehend, after all, is how speakers comprehend each other.—As perhaps goes without saying, however, I'm broaching issues that can't be adequately addressed in a brief note.

theory would require something approaching a complete theory of rational psychology. We needn't place bets here on *that* project. Suffice it to say that *if* attempts to codify "general intelligence" are subject to in principle limits, then it seems that semantic projects that utilize Stalnakerian contexts could be as well. Semantics pursued in this fashion would then be *as* hard as pragmatics.

Of course, the proper response might be that this constitutes no *objection* to Stalnakerian contexts. It merely provides a reason not to employ such contexts in *that* explanatory project.³⁴—Assuming, that is, that such a project can be carried out at all: there's even less reason to object if it can't. If it can, however, or if we at least possess no grounds at present for thinking it can't, then what conception of context *ought* one employ? A possible moral to draw from our examination of *Lewisian* contexts is that there's no avoiding intentionally characterized contextual features if one aims at the sort of epistemic illumination described above. But now we see that a profligate aversion to mental states threatens intractability. The Davidsonian conception can be viewed as an attempt to find a middle ground suitable to the project of articulating semantic rules. It adverts to intentionally characterized contextual features, but attempts to let in only those directly relevant to the calculation of propositional content. What features of context must be tracked in order to calculate the semantically relevant contextual features *themselves* is a question, according to this view, that semanticists needn't address.³⁵ The larger questions of rational psychology can thus be left to one side, if this approach is correct. Before turning to this conception of context, however, it remains briefly to remark upon the context-sensitivity of Stalnakerian contexts.

Again, we must distinguish the question of context-insensitively characterizing what contexts are from that of context-insensitively characterizing particular contexts. Regarding the former, there is, first, the vagueness, and consequent context-sensitivity, of 'presupposes;' and, second, the arguably context-sensitive individuation conditions for propositions, the objects of the speakers' presuppositions—i.e., the context-sensitivity of 'same proposition as.' Regarding the latter, there is further the question of whether the contents of others' attitudes can even in principle be context-insensitively characterized. Many accounts of attitude ascription posit context-sensitivity of one sort or another that could not be easily removed.³⁶ In addition, it is arguable that there are some essentially indexical

³⁴ I should note explicitly that this is *not* the project Stalnaker himself intends his conception of context to subserve. Stalnaker's discussions of the role of context in language use typically attempt:

to describe the structure of discourse in a way that abstracts away from the details about the mechanisms and devices that particular languages may provide for doing what is done in discourse... to get clear about what language is for—what it is supposed to do. (Stalnaker 1998: 4).

³⁵ It is possible that Stalnaker (1984: 40) loses track of this point.

³⁶ See Richard (1997) for a survey.

contents expressible only using context-sensitive means.³⁷ In any case, as we have seen, the violation of finite groundability (or at least manageability) would seem to leave us unable in at least some cases to characterize *exhaustively* a Stalnakerian context, whether context-insensitively or not.

(c) *Davidsonians*

The Davidsonian conception of context, recall, is that required by a truth-theory for a language: it comprises the parameters to which the theory must advert in specifying the variable truth-conditions of context-sensitive sentences. The determination and epistemic illumination constraints—*desiderata* 1 and 3, respectively—will clearly be satisfied *if* such a theory can be constructed. First, given such a theory, we may take the relevant T-sentence to provide—or at least to be determined by—the uttered sentence’s meaning. Then, if one plugs in the relevant contextual features, the biconditional’s right-hand side will yield what was said.³⁸ Second, a successful theory will advert to epistemically accessible contextual features whose bearing on propositional content is explicitly exhibited in the T-sentence, which would thus seem better to fit the above ideal of a semantical rule. *Whether* such a theory can indeed be constructed for an actual human language is, however, a large question. We’ll touch on some possible problems below, and elsewhere I discuss some more.³⁹

Now, as is apparent from my earlier example, Davidsonian contexts may include intentionally characterized features: recall the feature *what the speaker referred to in uttering ‘it’* invoked in providing a T-sentence for ‘It’s green.’⁴⁰ (So, again, this conception of context is not available if one is pursuing certain reductionist strategies.) Two worries might arise here. First, it may be worried that the threat of triviality looms. While there’s no reason to think that the contextual features adverted to will include *what proposition is expressed*, we see that other arguably semantic phenomena *will* be invoked, as *reference* just was in our example. The worry, note, doesn’t so much concern the invocation of intentionality *per se*, as the invocation of a certain *sort* of intentionality, specifically a kind

³⁷ See Perry (1979) for discussion.

³⁸ I pass over such complications as the Foster Problem. See Foster (1976).

³⁹ See Gross (1998/2001: Chaps. III and IV).

⁴⁰ The claim that Davidsonian semantics *can’t* invoke intentionally characterized contextual features crops up in one form or another surprisingly often. (See, e.g., Stalnaker 1984: 40). Sometimes the motivation seems to be one of those that will be discussed in the text: naturalistic restrictions, a fear of triviality, or a fear of epistemic entanglement in the holism of the mental. But other times, the motivation is much more inchoate. It’s worth noting, however, that no one questions the appropriateness of *who the speaker is* as a contextual feature (how else can one specify truth-conditions for sentences containing the first-person pronoun?)—and yet it’s obviously intentionally characterized (parrots, message machines, and perhaps foreign-language interpreters *qua* interpreters are not *speaking* in the relevant sense—they lack the proper intentions). Well, perhaps it’s not always obvious what’s obvious (at least, until it’s pointed out that it is) if even Stalnaker (1984: 40) can list *who the speaker is* as a paradigm example of a *non-intentionally* characterized contextual feature.

of *linguistic* phenomenon. Once again, the proper response is to clarify what questions Davidsonian context is supposed to help answer. It is certainly interesting to ask, for instance, what determines a speaker's reference on a particular occasion of utterance of a context-sensitive referring expression (and what enables hearers to grasp the reference). But that the Davidsonian approach doesn't answer *that* question won't constitute an objection if it doesn't *aspire* to answering it. Of course, there *would* be room for objecting if it begs off answering too many questions (or doesn't take on enough), or if it separates questions that seem to go together explanatorily. But this doesn't seem to be the case. In fact, the Davidsonian approach effects a natural division between our standing knowledge of language and the occasional knowledge we must in addition bring to bear in order to understand particular utterances. The truth-theory aims only to capture the former (which does, however, involve specifying what *role* one's occasional knowledge—whatever it is—plays on particular occasions). Tackling the latter can then be reasonably left for pragmatics to handle, to the extent that it can.

This division of labor also provides an answer to the second worry. Above I said that Davidsonian context can be understood as aiming for a middle ground between the non-intentional Lewisian locations and the profligately intentional Stalnakerian presupposition sets. The second worry, then, is whether such a middle ground can in fact be found—whether, in particular, Davidsonian contexts can avoid entanglement in the holism of the mental. Now we see how this indeed *is* possible. For Davidsonian contexts need advert only to those intentionally characterized features that need be directly plugged into the theory's variables. Their values, in other words, can be simply taken as *given*, as far as the truth-theory is concerned: *what* fixes those values in the first place becomes irrelevant.

Although not all of the speaker's presuppositions need be invoked on this conception of context, still there might remain the worry that the number of intentionally—and indeed non-intentionally—characterized contextual features required might prove quite large. Let's turn, then, to *desideratum* 4—the finiteness constraint.

I begin by recalling some well-known pre-history. Lewis, in his early paper "General Semantics,"⁴¹ specified eight contextual features ("a list," he later noted, "that was long for its day"⁴²) and, wondering whether more were needed, discussed others one might add.⁴³ Cresswell replied to the "agonies of conscience" he saw Lewis enduring "in trying to decide whether [he had] taken account of enough," by suggesting, with the support of one example, that "it is impossible to lay down in advance what sort of thing is going to count as a context

⁴¹ Lewis (1970: 195).

⁴² Lewis (1980: 87).

⁴³ Lewis (1970: Appendix).

of use... there is no way of specifying a finite list of contextual coordinates.”⁴⁴
Lewis accepted the point:

Contexts have countless features... the dependence of truth on context [is] surprisingly multifarious... Cresswell rightly complained... [there are] indefinitely many features of context.⁴⁵

And it is not uncommon for those who agree simply to refer back to this exchange.⁴⁶

But not everyone did agree. Lycan, for example, replied:

If English speakers and hearers in fact have all the context-interpretive skills requisite to negotiating all the discoverable factors, then (however many and various these factors may be) there is some definite [sc. finite] number of them, even if as theorists we cannot say in advance what this number is. Therefore, we can be confident that the addition of members to our “indices” would come to an end at some point...⁴⁷

And *perhaps* Lewis conceded *this* point as well in writing that “[w]e could wait for the end of linguistic inquiry, and define our indices then”—though he added that “the less patient of us may prefer another solution [viz., adverting to his locations].”⁴⁸—Here ends my historical narrative.

⁴⁴ Cresswell (1972: 8). The example is:

I have no pain, dear mother now,
But Oh! I am so dry;
Just fetch your Jim another quart
To wet the other eye. [Anon.]

which, he says, requires at least coordinates for the speaker, the time, and the type of previous drink.

⁴⁵ Lewis (1980: 87).

⁴⁶ This is then sometimes taken to show why we need to give up so-called indexical approaches to context and replace them with, e.g., presupposition sets. But note that Stalnaker’s conception of context can itself be construed as an example of the indexical approach: for presupposition sets may be viewed as just contexts that have one index (viz., *what the speaker presupposes*). The question, that is, is not one of whether to have indices, but of what those indices should be. Note that Lewis’ original approach includes presuppositions as an index. So, when one adds to a Stalnakerian context further elements—to track, say, reference across a discourse, etc. (yielding what is called a discourse structure)—you get something that again looks very much like those bad ol’ index kind of contexts!

⁴⁷ Lycan (1984: 51).

⁴⁸ Lewis (1983: 230). (Though I suggest that this remark may be read as providing a possible response to Lycan, Lycan is not actually mentioned in this passage. —The Lewis remark was also published a year before Lycan’s, though this of course does not entail that the one was *written* before the other.)

Now, while I won't go so far as to defend Lycan's argument,⁴⁹ I do want to defend his conclusion—or at least to suggest that no good argument has been given against it. The first point to note is just an application of the point I stressed above. Many have been overly impressed by the multifariousness of contextual features on which—to use Lewis' language—“truth sometimes depends,”⁵⁰ features that are “relevant” to the determination of extension.⁵¹ But this language is a bit unclear. Yes, just about anything could be relevant—if what is meant, for instance, is that it could play some causal or evidentiary role either in the determination of content or in a hearer's coming to grasp what was said. But it doesn't follow that an interesting semantic theory must account for all aspects of these processes. All the Davidsonian need do is isolate those features of context that provide the truth-theory the contextual information it needs—again, however these features are themselves in turn determined. The semanticist, that is, needs to identify those features of context that, as it were, sum up the rest of the contextual situation in a manner sufficient to yield truth-conditions, given the semantic rules embodied in the truth-theory. In sum, from the infinity of possibly “relevant” contextual features, it doesn't follow that Davidsonian contexts must be infinite.

Still, one may attempt to argue, by appeal to cases, that even Davidsonian contexts must advert to an infinite number of features. But note that if a language's lexicon and syntax is finite, then the infinity would have to arise already from a single term or rule.⁵² Is there reason to think the axiom for any single lexical item, for instance, would require an infinity of contextual features? Predications of 'green' might be contextually sensitive to the relevant part of the subject of predication, the color's etiology, facts about the viewing conditions, etc.⁵³ But could this list go on forever? Some argument would be needed to give us reason for thinking so. There is much complexity, to be sure (surprisingly, perhaps—though in retrospect maybe this is what we should have expected), but from complexity doesn't follow infinity. It is sometimes suggested, however, that there is an open-endedness to our language use, in that as new situations arise, so will new ways of deploying our existing words: mountains are high, but so are tones—and there's no end of fields to which the high/low contrast could come to be

⁴⁹ The movement from our being able to negotiate these features to there being a finite number of them may seem a *non sequitur*, but perhaps it's instead enthymematic, tacitly presupposing elements of the “metatheory” Lycan earlier (1984: 13–18) lays out. This threatens, however, to render the argument question-begging.

⁵⁰ Lewis (1980: 79).

⁵¹ Lewis (1983: 230).

⁵² I ignore the results of indefinitely iterating the same source of context-sensitivity. Sure, in theory you could, in a single sentence, keep on referring to items via the demonstrative ‘that’ as long as you liked. But this could be reasonably represented as a sequence, and thus as a *single* contextual feature. This sort of infinite grounding seems harmless, since it could be recursively characterized. (And, in any case, in *practice* you *couldn't* keep on demonstrating as long as you'd like.)

⁵³ Cf. Gross (1998/2001: Chap. I) and the work cited therein by Travis (1981) and Szabó (1995).

applied. Similarly, there may be no end to the kinds of contextual feature that could become relevant to assessing predications of ‘green.’⁵⁴ The proper response to this, however, is that semantics (or at least this chapter of it) need only be in the business of describing how our language is *now*. It needn’t as well account for all the ways it might expand over time.⁵⁵ The worry might remain that, although the number of contextual features admitted on the Davidsonian conception has not been shown to be infinite, it has been shown to be quite large—and therefore perhaps *unmanageable*. But here the Davidsonian may reasonably shift the burden back: if this is to become a full-fledged objection, the demand for manageability must both be clarified and motivated. After all, as mentioned, at this late date we perhaps should not be surprised by the *complexity* of our linguistic competence.

I turn, finally, to *desideratum* 6—that contexts be context-insensitively characterized. If among the posited contextual features are propositional attitudes, then context-insensitively characterizing particular contexts will face the same obstacles encountered by Stalnakerian presupposition sets. In addition, the values of many non-intentional features would seem to resist context-insensitive characterization. Consider the sentence ‘That style of play is risky.’ A relevant contextual feature would seem to be the contextually relevant style of play. But unlike with times and places, we seem to lack a general context-insensitive way to identify arbitrary styles of play. Though we may be able to draw attention to aspects of the particular style, we may lack the resources to single it out in any way other than demonstratively.⁵⁶ And so we wouldn’t be able to context-insensitively refer to an arbitrary context. But perhaps this needn’t bother us, so long as we can context-insensitively refer to those contexts to which we have some reason to refer. And, in any case, for all that’s been said, perhaps we might have no reason *not* to allow ourselves to refer to contexts using context-sensitive terms.

More interesting is the question whether we can characterize what contexts *are* context-insensitively. For if we can’t, then we are blocked from context-insensitively stating the truth-conditions of the sentences of our target language. The T-sentences for context-sensitive sentences, recall, have the following form:

$$(u)(x)(y)\dots(z)[\text{If } uR(S, x, y, \dots, z), \text{ then } \{u \text{ is true iff } F(x, y, \dots, z)\}].$$

⁵⁴ See Ross (1981), whose discussion of many interesting examples is marred *inter alia* by a failure to mark various of the distinctions described above.

⁵⁵ This might require a more fine-grained individuation of languages than is normal in everyday speech. But that needn’t affect the theory’s explanatory interest. Note, by the way, that in the text I speak of ways a language (in the coarse-grained sense) *might* change over time. This is because *actual* language change is of course finite.

⁵⁶ For any given case, we may *introduce* a context-insensitive name—perhaps, ‘Kasparov’s style.’ But if there are *many* styles, we may be unable to introduce enough names to cover them all. We also could refer to the style via the context-insensitive description ‘the style displayed at such-and-such time at such-and-such location.’ (Let’s not worry about specifying the world.) But this description doesn’t present the style in a way that enables an arbitrary speaker to grasp what style it is. (Cf. the remarks at the close of Gross (1998/2001: Chap. I) on ineliminability and content individuation.)

For example, the T-sentence capturing the variable reference of ‘it’ in ‘It’s green’ would be:

(u)(x)If u is an utterance of ‘It’s green’ in which ‘it’ refers to x, then (u is true iff x is green).

The T-sentence thus adverts to the contextual feature *what ‘it’ refers to*. If this contextual feature is itself context-sensitively characterized, then the T-sentence as a whole will be as well. But if the T-sentence is context-sensitive, it does not—considered independently of any particular context of utterance—express a proposition at all; and so, in particular, it doesn’t specify the sentence’s truth-conditions. The Davidsonian semanticist must therefore either allow that the content of the truth-theory itself can vary across contexts of utterance, or she must ensure that she uses only context-insensitive terms in characterizing what the relevant contextual features are. If ‘refers’—or any other expression used to characterize a contextual feature—is subject itself to context-sensitivity (even if only through being subject to vagueness), then taking the latter tack will require stipulating this context-sensitivity away. Note that this requires that the individuation conditions of the objects quantified over be fixed as well.⁵⁷

4 Conclusion

This completes my examination of these three conceptions of context. We have seen that each satisfies some of the six *desiderata*, while violating others. And in each case, we have seen as well how these violations needn’t constitute an objection, so long as one is clear about what projects each conception may and may not be suited to subserve. Not clearly delineating these projects leads only to confusion. It is my hope that this examination not only helps sort out what is—or ought to be—at issue in discussions of context, but also sheds some light on the nature of these varying projects themselves.

I close by addressing one reason why some of the issues I have tried to separate are sometimes lumped together—viz., a not unreasonable desire to treat language and linguistic phenomena *uniformly*. It would seem a laudable goal to seek an explanatory framework that could accommodate *all* of the projects we have discussed. The desire for such motivates searching for a single conception of context suitable for all our linguistic needs. I have not argued that there could not be such a thing—beyond, that is, gesturing towards the possible intractability of certain areas.⁵⁸ But note that

⁵⁷ This amounts to removing whatever context-sensitivity there may be in the ‘what’ of, for instance, ‘what ‘it’ refers to.’

⁵⁸ Note that this is just as much a worry at the pre-semantic as at the post-semantic (or, pragmatic) level. This is in part because judgments at one level can affect the other. We may reject a disambiguation, for instance, on account of the untoward pragmatic implications this would force. Indeed, all three levels—the pre-semantic, semantic, and postsemantic—are holistically intertwined epistemologically speaking.

even if there *were* a single conception of context suitable in principle for all projects aimed at explaining linguistic behavior, it would not follow that adverting to such contexts would yield the most perspicuous explanation no matter the issue. For example, employing Stalnakerian contexts, and the highly complex rules they would require, would certainly shed less light on specifically what we know in knowing a particular language than would the use of Davidsonian contexts. (I'm supposing here that both approaches could pan out in principle, and that Stalnaker's would thus be suitable to a greater range of questions than the Davidsonian's.) Various specifically *semantic* facts, explicitly highlighted on the Davidsonian approach, would get lost in the complex semantic/pragmatic brew the Stalnakerian approach would yield. We would thus still have use for more than one conception of context.⁵⁹

References

- Burge, Tyler. 1974. Demonstrative constructions, reference, and truth. *Journal of Philosophy* 70: 205–223.
- Cresswell, Max. 1972. The world is everything that is the case. *Australasian Journal of Philosophy* 50: 1–13.
- Davidson, Donald. 1967. Truth and meaning. *Synthese* 17: 304–323.
- Field, Harry. 1978. Mental representation. *Erkenntnis* 13: 9–61.
- Foster, John. 1976. Meaning and truth theory. In *Truth and meaning*, eds. G. Evans, and J. McDowell, 1–32. Oxford: Oxford University Press.
- Gauker, Chris. 1997. Domains of discourse. *Mind* 106: 1–32.
- Gross, Steven. 1998/2001. *Essays on Linguistic Context-Sensitivity and its Philosophical Significance*. New York: Routledge.
- Higginbotham, James. 1988. Contexts, models, and meanings: A note on the data of semantics. In *Mental representations: The interface between language and reality*, ed. R. Kempson, 29–48. Cambridge: Cambridge University Press.

⁵⁹ Throughout this chapter I focus on conceptions of context suitable for various *explanatory* purposes. Let me mention here another reason why one might advert to context, one discussed in Gross (1998/2001: Chap. V). Disagreement can arise, on some occasion of utterance, both about *what* was said, and about *whether* anything was said at all. Many context-sensitive sentences can give rise to such disputes (if the sensitivity to context be subtle enough), but my primary interest is in such disputes as may arguably arise from the use of philosophically loaded context-sensitive terms—such as (perhaps) ‘exists’ and ‘object.’ Assuming the parties to the dispute are both competent speakers of the language in question, and thus know the meaning of the words used, it is natural that they turn, in order to adjudicate the dispute, to the relevant contextual features that do (or do not, as the case may be) enable the utterance of those words on that occasion to express a proposition. Now, is there a conception of context that might suit this situation, that might provide them with a neutral way to settle their disagreement? Well, certainly Lewis’ won’t help. But adverting to Davidsonian contexts might just force the dispute back to the features it comprises: if two people disagree over whether an utterance of ‘He’s hungry’ expressed a proposition, they probably disagree over whether ‘he’ succeeded in referring on that occasion. Finally, using Stalnakerian contexts threatens again to leave us with only general intelligence to go on; and stubborn disputants will disagree over whose intelligence should be trusted. It thus looks unlikely that there’s a way of individuating contexts especially well-suited for the epistemic task of adjudicating disputes over content.

- Hornstein, Norbert. 1984. *Logic as grammar*. Cambridge: MIT Press.
- Kaplan, David. 1989. Afterthoughts. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein, 565–614. Oxford: Oxford University Press.
- Lewis, David. 1970. General Semantics. *Synthese*, 22, 18–67. Cited as reprinted in *Philosophical Papers: Volume I*, D. Lewis, 189–229. Oxford: Oxford University Press (1983).
- Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8: 153–166.
- Lewis, David. 1980. Index, context, and content. In *Philosophy and grammar*, ed. S. Kanger, and S. Öhman, 79–100. Dordrecht: D. Reidel.
- Lewis, David. 1983. Postscript to “General Semantics.” In *Philosophical Papers: Volume I*, D. Lewis, 230–232. Oxford: Oxford University Press.
- Lewis, David. 1986. *On the plurality of worlds*. Oxford: Blackwell.
- Lycan, William. 1984. *Logical form in natural language*. Cambridge: MIT Press.
- Lycan, William. 1986. Tacit Belief. In *Belief: Form, content, and function*, ed. R. Bogdan, 61–82. Oxford: Oxford University Press.
- Matthews, Gareth. 1980. Ritual and the religious feelings. In *Explaining emotions*, ed. A. Rorty, 339–353. Berkeley: University of California Press.
- Montague, Richard. 1968. Pragmatics. In *Contemporary philosophy: A survey*, ed. R. Klibansky, 102–122. Florence: La Nuova Italia Editrice.
- Perry, John. 1979. The problem of the essential indexical. *Nous* 13: 3–21.
- Richard, Mark. 1997. Propositional attitudes. In *A companion to the philosophy of language*, eds. B. Hale, and C. Wright, 197–226. Oxford: Blackwell.
- Ross, John. 1981. *Portraying analogy*. Cambridge: Cambridge University Press.
- Smith, Neil (ed.). 1982. *Mutual knowledge*. London: Academic Press.
- Stalnaker, Robert. 1978. Assertion. *Syntax and Semantics*, 9, 315–332. Cited as reprinted in R. Stalnaker, *Context and Content*, 78–95. Cambridge: MIT Press (1999).
- Stalnaker, Robert. 1984. *Inquiry*. Cambridge: MIT Press.
- Stalnaker, Robert. 1998. On the representation of context. *Journal of Logic, Language and Information* 7: 3–19.
- Stanley, Jason, and Timothy Williamson. 1995. Quantifiers and context-dependence. *Analysis* 55: 291–295.
- Szabó, Zoltan. 1995. *Problems of compositionality*. Dissertation: MIT Ph.D.
- Travis, Charles. 1981. *The true and the false*. Amsterdam: John Benjamins.

Implicature, Inference and Cancellability

Michael Haugh

Abstract The standard position in pragmatics to date has been that cancellability is useful way of differentiating implicatures from logical implications, semantic entailments and the like. In recent years, however, there has been considerable debate as to whether implicatures are in fact always cancellable, or indeed whether they are cancellable at all, amongst linguistic pragmaticians and language philosophers. In this chapter, it is suggested that cancellability encompasses a range of actions that play out in different ways depending on whether we are analysing inferences that can lead to implicatures or the implicatures themselves. In this way, we can see how analysts have often underplayed the contingency of inferences as well as the inherent indeterminacy of implicatures in such debates. It is concluded that cancellability should be the subject of further empirically-driven analyses in order to provide a solid foundation for the theorization of implicature.

1 Introduction

In introducing the theory of conversational implicature, Grice made three key claims, each of which has subsequently become the subject of considerable debate in pragmatics. First, Grice claimed that the inferences underlying conversational implicatures are made in a principled way. In other words, conversational implicatures are made available to recipients by speakers with reference to normative assumptions about the nature of cooperative interaction (cf. Bianchi this volume). Grice formalised this claim in postulating the Cooperative Principle and four

M. Haugh (✉)

School of Languages and Linguistics, Griffith University, Brisbane, Australia
e-mail: m.haugh@griffith.edu.au

related conversational maxims. Second, conversational implicatures are generated through ordinary reasoning (as opposed to strict logical reasoning), either once-off reasoning in a particular context (nonce inferencing), or regularised or conventionalised reasoning across contexts (default inferencing). Third, since conversational implicatures arise through ordinary as opposed to strictly logical reasoning, these inferences are defeasible, from which it follows that a *putative* implicature can always be cancelled. Grice argued that a conversational implicature can be cancelled in two different ways. Explicit cancellation involves adding a subsequent cancellation clause that either retracts the implicature (e.g., “not that I mean [to imply *x*]”) or blocks it from arising in the first place (e.g., “in fact [not *x*]”). Contextual cancellation involves finding situations in which the putative implicature would not arise (Grice 1989: 44).

The standard position in pragmatics to date has thus been that implicatures are cancellable through explicit and/or contextual cancellation, and that the “cancellability test” is useful way of differentiating implicatures from logical implications, semantic entailments and the like (see, e.g., Blome-Tillmann 2008: 156). Cancellability, in other words, has often been assumed to be one of the necessary, albeit not sufficient, tests for identifying implicatures. In recent years, however, there has been considerable debate as to whether implicatures are always in fact cancellable, or indeed whether they are cancellable at all amongst linguistic pragmaticians and language philosophers.¹ Much of this debate has arisen as a consequence of the conflation of the different orders of language description that these two types of cancellation represent, as Jaszczolt (2009) has pointed out. Explicit cancellation is arguably of primary importance to those analysing implicatures in actual discourse while contextual cancellation is more pertinent to those engaging in analyses of individual utterances. In this chapter, I attempt to further reframe the terms of this debate by proposing that “cancellability” plays out in different ways depending on whether we are analysing inferences that can lead to implicatures or the implicatures themselves. An implicature is a type of speaker meaning that goes beyond what is literally said. More specifically, implicatures pertain to separate individual, additional thoughts with their own pragmatic force (Haugh 2002: 128–130; Haugh and Jaszczolt 2012: 96). Inference, in contrast, refers to the cognitive processes by which participants figure out (speaker) meaning beyond what is said or encoded. While many scholars treat implicatures as essentially synonymous with non-logical/pragmatic inference (Carston 2002; Levinson 1983, 2000; Sperber and Wilson 1995), other scholars insist that assimilating implicature to inference constitutes a conceptual and analytical error (Bach 2006; Horn 2004, 2012). In this chapter I favour the latter position. A key difference between the two is arguably that while speakers can be

¹ It is worth noting that this debate has also been extended to the relevance theoretic concept of explicature (see Burton-Roberts 2010; Capone 2006, 2009, 2010; and Carston 2002, 2010 for further discussion). However, those debates are only touched upon here to the extent they make reference to the (neo-)Gricean notions of particularised and generalised conversational implicatures.

held accountable for (or committed to) implicatures to varying degrees, neither speakers nor recipients are held accountable for all the inferences they make when participating in interaction (unless they make available such inferences to others through saying or implying). It is thus argued here that a clear distinction between inference and implicature needs to be maintained if we are to succeed in unravelling some of the recent controversy over the cancellability (or not) of implicature.

I start this chapter by outlining three of the key challenges to the validity of cancellability test for implicatures that have been raised thus far. The first of these is claim that implicatures may in some instances be entailed and thus cannot strictly speaking be cancelled since such cases involve logical inferences that are not defeasible. The second is the argument that something which is speaker intended, such as a particularised conversational implicature cannot be “unintended”, and thus such implicatures cannot, by definition, be cancelled. The third involves cases of so-called cancellation-resistant implicatures, where explicit cancellation phrases are either non sequitur or end up reinforcing the implicature. In the course of discussing these various challenges I suggest that analysts have tended to under-estimate the contingency of implicatures and over-estimate their degree of determinacy. Building on this discussion, I next focus on cancellability vis-à-vis inference, suggesting that blocking and suspension are two of the key ways in which inferences might be “cancelled”. I also argue that the blocking or suspension of pragmatic inferences is more productively conceptualised as a consequence of their inherent contingency rather than defeasibility, in contrast to what is traditionally assumed in theorising implicature. I then discuss cancellability vis-à-vis implicature, and propose that we are dealing here with two quite different social actions that can arise as a consequence of the use of cancellation phrases in interaction, namely, correction/repair and reinforcement. The former type of action encompasses denying, retracting or clarifying what has been implied, and is arguably a function of the degree of indeterminacy of the implicature in question, as well as the extent to which a speaker can be held to be implying something. Whether an implicature can be plausibly or legitimately denied, retracted or clarified by the speaker is thus proposed to be an empirical question, which requires recourse to an account of the ways that speakers make implicatures available to recipients through appeals to normative frames of interpretation, as well as issues of salience and sequential placement. The latter type of action, namely, reinforcement arises as a function of the way in which implicatures can be strengthened or bolstered without redundancy through cancellation phrases. I conclude by suggesting that rather than being abandoned, as suggested by Carston (2002, 2010), cancellability ought to be the subject of further analyses that acknowledge the critical distinction between forms of “cancellation” that apply to inference as opposed to implicature.

2 The Cancellability Debate in Pragmatics

A number of apparent challenges to the cancellability of implicatures have been raised in recent years. In this [Sect. 1](#) briefly discuss each of these challenges in turn.

2.1 *Implicature and Entailment*

One challenge to the efficacy of the cancellability test has been the observation that in some cases, implicatures can also be treated as entailments (Bach 2006: 24; Carston 2002: 138–140; Higashimori and Wilson 1996: 122; Vicenti 1998; Wilson and Sperber 1986: 61). Entailment is a type of semantic relation between propositions (p and q) where p entails q ($p \Vdash q$) if and only if the truth of p guarantees the truth of q (Huang 2007: 16; Levinson 1983: 174). An implicature, in contrast, is generally assumed to rely on a pragmatic relation, p implies q ($p \text{+>} q$), which is non-truth conditional (i.e., the truth of q is independent of p). However, if an implicature is also entailed then it would appear the implicature cannot strictly speaking be cancelled, because the truth of the implicature is no longer independent of the truth of what is said in such cases. A number of examples have been proposed that illustrate instances where what is implied (by the speaker) is also entailed (by the utterance).

- (1) Peter: Would you like to listen to my Rolling Stones record?

Mary: No. I'd rather hear some music.

+> Peter's Rolling Stones record is not music.

\Vdash Peter's Rolling Stones record is not music.

(Higashimori and Wilson 1996: 122).

- (2) Adam: Does John drink slivovitz?

Bob: He doesn't drink any alcohol.

+> John does not drink slivovitz.

\Vdash John does not drink slivovitz.

(Carston 2002: 139; Wilson and Sperber 1986: 61).

- (3) Andy: Nobody has ever long-jumped over 28 ft.

Bill: What'ya mean? Bob Beamon long-jumped over 29 ft way back in 1968.

+> Somebody has long-jumped over 28 ft.

\Vdash Somebody has long-jumped over 28 ft.

(Bach 2006: 24).

In the above examples, what is implied by the speaker also logically follows from what is said by analytic rule. In other words, we have instances of implicated entailments. Carston (2002) argues that such examples demonstrate that “non-contradictory cancellability is not a necessary property of implicatures” (p. 140). However, as she latter admits, “the concept of entailment and the concept of implicature belong to different explanatory levels, in fact different sorts of theory—the one a static semantic theory which captures knowledge of linguistic meaning, the other an account of the cognitive processes and representations involved in understanding utterances” (Carston 2004: 644). Building on this point, I would suggest that there are two important differences between analyses of entailment and implicature that bear on the issue of cancellability.

First, implicatures are not required to be formally determinate in the same way as entailments. Higashimori and Wilson’s (1996) gloss of example (1), for instance, brushes over all sorts of possible indeterminacy in the interpretation of Mary’s response. For a start, while they claim that Mary implies the Rolling Stones is “not music”, clearly we’re dealing with something more than that. Mary’s response could be interpreted as her implying it is “not real music”, “not good music”, “not pleasant to listen to music” and so on and so forth. It also remains indeterminate as to who this evaluation is attributed to. Is it just Mary who thinks this, or is she implying that “others” (e.g., people of good taste) might also share the same belief? Such issues of indeterminacy in interpretation lie outside the scope of formal analyses of entailments, but they are critical for the interpretation of implicatures. For this reason they bear on issues of cancellability, as such nuances in interpretation allow for the speaker (in this case Mary) to subsequently deny particular interpretations on the part of the recipient (Peter) on the basis that she didn’t *say* that which has been specifically attributed to her. As Bach (2006) points out, “it’s not what the speaker says but that he [or she] says it (or even that he [or she] puts it a certain way) which carries the implicature” (p. 24). In that sense, this is what gives Mary an edge over Peter in determining what has been implied. Peter may, of course, have his doubts, and so this denial might be disputed, but the very fact that what has been implied can be disputed means there is more to the cancellability issue than first meets the eye.

Second, the interpretation of implicatures is inevitably contingent on context, including what comes before and after the utterance in question, as well as the presuppositions of speakers (and recipients). Entailments, on the other hand, involve a strictly defined and fixed set of assumptions. In example (2), for instance, Bob could subsequently add something like “except for the odd tot of slivovitz” without contradiction of what is said, thereby “removing” the entailment relation between his utterance and what would have been implied without this proviso (i.e. “John does not drink slivovitz”). In that case, what is implied is something along the lines that John’s habit of drinking slivovitz (albeit not often) is quite striking in light of his general abstinence from drinking alcohol.

Bach’s (2006) example of an implicated entailment seems more difficult to remove, in part because it arises not only through Bill’s assertion, but Bill’s assertion *in response to* Andy’s assertion. However, once again the interpretation

of this putative implicature can be shown to be contingent on subsequent discourse. Bill might add something like “although if you mean since Beamon’s jump then you’re right” if he believes (falsely) that no one has jumped over 28 ft since Beamon’s 29 ft jump in 1968, and moreover attributes this (false) belief to Andy. All in all, this suggests that the application of formal types of analyses to natural language usage is somewhat fraught, as ordinary discourse involves all sorts of contingencies that are not accounted for within strictly formalistic approaches, including the analysis of entailment relations.²

2.2 *Implicature and Speaker Intentions*

A second challenge to the claim that implicatures are cancellable is more rhetorical in nature. Based on the general assumption that implicatures are speaker-intended, a number of scholars have argued that once something is intended it cannot be “unintended”, and thus particularised conversational implicatures, at least, are not cancellable by definition (Burton-Roberts 2006, 2010; Capone 2009, 2010). Burton-Roberts (2010), for instance, argues that “There is no question, with PCI [particularised conversational implicature], of the implicature in any sense ‘arising’ independently of the speaker’s intention. In short, a PCI is only ever an *actual* implicature. My claim, then, that PCIs (as *actual*) must be interpreted as *intended* and so cannot be cancelled” (p. 151, original emphasis). Capone (2006, 2009, 2010) is more circumspect in allowing for cases of weak commitment or vagueness in some examples of implicatures. He nevertheless argues that in instances where there is “strong intentionality”, the resulting implicatures are not cancellable:

When a strong intentionality is projected, it can no longer be retracted. Implicature can only arise if intended and recognized as intended, but then it should be impossible to cancel an implicature: how would it be possible to withdraw/cancel what was intended to be implicated and was recognized as intended? An implicature could only be withdrawn/cancelled if it were not intended. But then it shouldn’t be an implicature (since implicatures by definition are intended); in other words there would be no implicature to cancel. (Capone 2009: 59; 2010: 502)

One problem with this line of argument, however, is that it is somewhat circular. It is assumed that implicatures are intended “by definition”, with this conceptualisation of “implicature as intended” being subsequently used to undermine the idea that implicatures can be cancelled. However, while some argue that Grice regarded conversational implicatures as speaker-intended (cf. Bianchi this volume; Davis 1998), there are good reasons to believe the relationship between speaker (intended) meaning and implicature is not quite as straightforward as might first

² The same limitations also face appeals to strict logic-based analyses that are implicit to the notion of defeasibility as will be discussed further in Sect. 3.

appear. Grice himself originally made a distinction between ‘implicature’ and ‘implicatum’. The former referred to the *process* of meaning (in the sense of intending) something in addition to what is said, and was thus restricted to the speaker. The latter referred to the *product* of that process, which could arise from inferences by either the speaker or the recipient (Haugh and Jaszczolt 2012). The term ‘implicature’ gradually came to refer to both notions, however, thereby obscuring the important distinction between speaker-implicatum and recipient-implicatum (cf. ‘utterer-implicature’ and ‘audience-implicature’) (Saul 2002; Horn 2012). Retaining such a distinction means recognising not only that speakers may be taken to be implying something that is contrary to their claimed intention (Cummings 2005: 20–21; Haugh 2008a, b; Bach 2001, 2006; Capone 2010: 502; fn. 3), but also that speakers (and writers) can intentionally leave the interpretation of what has been implied open to the recipient (Jaszczolt 1999: 85; Haugh 2011). If we can admit that “the subsumption of conversational implicature within the category of speaker meaning is not entirely straightforward” (Horn 2012), then it becomes an empirical question as to whether an implicature is “cancellable” without contradiction. After all, are we trying to analyse only what a particular theory recognizes as a legitimate object of study or what is demonstrably recognizable (to participants) in the world? If we assume the latter position, the above argument is open to empirical question.

It is also worth noting that implicatures are not always as determinate as is often assumed by analysts. A number of classic examples of implicature originally noted by Grice turn out to be more indeterminate than those taking the stance that implicatures are speaker intended (and thus by definition not cancellable) allow for. Capone’s (2009) discussion of Grice’s classic example of an implicature arising through a testimonial for a candidate applying for a philosophy job is somewhat illuminating in that respect.³ The example as given by Grice is reproduced below.

(4) Dear Sir, Mr. X’s command of English is excellent, and his attendance at tutorials has been regular. Yours, etc. (Grice 1989: 33)

While Grice argued that what was implicated here was that A thinks “Mr. X is no good at philosophy” (Grice 1989: 33), Capone’s discussion indicates that there are at least two possible implicatures here: first, the writer does not support the candidate’s application for the job, and second, the write thinks the candidate is not a good (or a poor) philosopher. These two implicatures are both indeterminate. As Capone (2009) points out, “though A could deny not having supported his student, he certainly could not assert his support” (p. 57). In other words, the first implicature is indeterminate as to whether the writer is strongly against the

³ Another point worth noting in passing is that in discussing this example in regards to its cancellability, Capone (2009) makes reference to the notions of “retract”, “revise”, “deny”, “disintegrate” and “unimplicate”, although these are clearly not equivalents. This suggests that “cancellability” encompasses different kinds of processes and actions, as discussed further in Sects. 3 and 4.

candidate's application or just indifferent, although *normatively* we might assume the former interpretation. The second implicature is also more indeterminate than often recognised. To be labelled as "not a good philosopher" is not the same thing as being labelled "a poor philosopher". There is a difference of degree here. In implying a broadly negative (or at least not positive) evaluation, the writer is exploiting this indeterminacy. The writer could argue, for instance, that he is not implying that Mr. X is a poor philosopher, but simply not saying he is a good philosopher. There are limits to the plausible deniability of implicatures, of course. However, this is not so much a function of what is written down by the writer, as a function of the activity type of writing testimonials, where readers have legitimate expectations constraining what can be meant by such a letter. It could be argued, then, that while in normal circumstances (i.e., when writing testimonials) the writer may be taken as making available, by conversational implicature, something to the effect that the candidate is not a good philosopher, and furthermore that he does not support the candidate's application, these conversational implicatures are not necessarily synonymous with what the writer thought he had implied (speaker-implicatum) or how the reader interpreted this letter (recipient-implicatum), due to their inherent indeterminacy.

Burton-Roberts (2010) offers another example of an implicature that is allegedly not cancellable.

(5) Max: Do you ever speak to Charles?

Ann: I never speak to plagiarists.

+> Charles is a plagiarist.

(Burton-Roberts 2010: 151).

The first thing to note here is that once again there are at least two implicatures that potentially arise here. The first is an implicated presumption (i.e. Charles is a plagiarist), while the latter is an implicated conclusion (i.e. I don't speak to Charles). The latter putative implicature can be straightforwardly cancelled with an "except for" clause. The former potential implicature appears more resistant to cancellation, but this is only if it is assumed that the implicated presumption is determinate. On closer examination, however, it appears that there is in fact real indeterminacy in regards to the speaker's degree of certainty about this suggestion. The speaker is not strictly speaking committed to the claim that Charles is definitely (cf. "is") a plagiarist, although Max might have (normatively) taken her to be so committed. Thus, the implicature can indeed be removed by Ann simply adding something like "And I think Charles might be one", thereby indicating her uncertainty about such a claim. Or she might, in a more convoluted manner, subsequently claim that "I am not saying that Charles is a plagiarist for sure, mind you, just that I think he probably is one". Max could, of course, dispute this claim. But therein lies the point: implicatures are not determinate, which is why they can be disputed in the first place.

Following Grice's (1975: 58) characterisation of implicatures as indeterminate, Sperber and Wilson (1995) argue that:

there has been a tendency in modern pragmatics to treat all implicatures along these lines: as fully determinate assumptions for which the speaker is just as much responsible as if she had asserted them directly. On this approach, utterance comprehension consists in the recovery of an enumerable set of assumptions, some explicitly expressed, others implicitly conveyed, but all individually intended by the speaker. (p. 195)

Although the degree of indeterminacy of implicatures is likely to be gradable along some kind of scale (from strongly through to weakly indeterminate),⁴ they are nevertheless always indeterminate in some respects simply because they are not said. Claiming that implicatures are inevitably indeterminate is a distinct issue from the question of whether something has been implied or not. In some instances, participants would argue that clearly something has been implied, and in such cases it is not plausible for the speaker to deny having meant something in addition to what is said. But due to the inherent indeterminacy of implicatures this something is never entirely fixed. Ultimately, then, it makes little difference whether or not one subscribes to a definition of implicatures as only encompassing those that are speaker-meant (or intended), as it is through appeals to "normal circumstances" that we take something, which is more or less indeterminate, to be implicated, as Grice originally argued (Wilson and Sperber 1986: 71). On this view, their degree of "cancellability" thus remains an empirical question.

2.3 *Cancellation-Resistant Implicatures*

While it has been suggested thus far that arguments against the efficacy of the cancellability test are not as straightforward as what might first appear, there are nevertheless instances where the implicatures appear to simply defy explicit cancellation.

Sarcasm or irony has been analysed as arising through implicature, following Grice's (1989: 34, 53–54) discussion of irony. Grice claimed with reference to the following example, for instance, that a conversational implicature can be generated through flouting the first maxim of Quality.

- (6) (X, with whom A has been on close terms until now, has betrayed a secret of A's to a business rival. A and his audience both know this)
A: X is a fine friend. (Grice 1989: 34)

Grice suggests that what is implicated here is a proposition "contradictory of the one he purports to be putting forward" (Grice 1989: 34) (i.e., X is not a fine

⁴ Cf. Sperber and Wilson (1995) who claim that implicatures vary in their degree of determinacy, that is, from strongly determinate (strong implicatures) through to weakly determinate (weak implicatures).

friend). Grice (1989) also notes that irony involves “the expression of a feeling, attitude or evaluation” (p. 53), the exact content of which is generally fairly indeterminate, although it can be understood as broadly “negative” in the case of sarcasm.

However, it has been pointed out that irony generated through implicature resists explicit cancellation (Huitink and Spender 2004; Weiner 2006). While Jaszczolt (2009) points out that this does not mean that such instances “fail” the cancellability test per se, since contextual cancellation is always another possibility, the existence of implicatures that are resistant to explicit cancellation does nevertheless throw up interesting questions for analysts. In the examples given by both Huitink and Spender (2004) and Weiner (2006), an attempt to cancel the irony ends up simply reinforcing it, because the cancellation phrase itself is interpreted ironically. In the following example, for instance, we can see how the cancellation phrase does not necessarily retract, deny or clarify what is implicated here, but rather can reinforce the ironic stance of the speaker.

(7) Mr. X’s command of English is excellent and his attendance at tutorials has been regular. He is a brilliant philosopher (Huitink and Spender 2004: 8; cf. Grice 1989: 33).

They suggest here this is because “the cancellation phrase itself is interpreted non-literally...as such, the second sentence confirms the implicature of the first sentence” (Huitink and Spender 2004: 8). It appears that the likelihood of a cancellation phrase being interpreted ironically is a function of the way in which the prior implicature arose (i.e. whether in an ironic or non-ironic interpretative frame), and the degree of salience of an ironic interpretation relative to the topic and activity type in question (cf. Huitink and Spender 2004: 10).

Another type of cancellation-resistant implicature arises in instances where utterance-final connectives reinforce the inference leading to implicature such that it becomes difficult, if not impossible, to plausibly retract or deny it. In the following excerpt from a telephone conversation there is an implied refusal that is reinforced through the deployment of an utterance-final *kara* (‘so’).

(8) (Jun and Yuki are both university students. Yuri has called up to ask Jun about notes from a class she missed)

1. Yuri: *Senshū no nōto misete moraitainde, kyō kaettemo ii desu ka?*

‘Can you show me last week’s notes and let me borrow them today?’

2. Jun: *Senshū no nōto, a, hito ni kashiteru kara...*

‘Last week’s notes, oh, I have lent them to someone so...’

3. Yuri: *A, sō nan desu ka.*

‘Oh, is that right?’

(adapted from Haugh 2008c: 436)

Yuri first asks Jun whether she can borrow some notes from the previous week's class (turn 1). Jun responds by saying he has already lent the notes to someone else. The utterance-final *kara* ('so') in turn 2 is exploited by Jun to indicate this is a reason for what is implied here, namely, that he is not able to give the notes to Yuri today. Yuri claims understanding of this implied refusal in her response in turn 3. We can thus see here that utterance-final conjunctives can be exploited to reinforce implicatures in light of the fact that conjunctive *kara* ('so', 'because') is canonically attached in Japanese to subordinate clauses to indicate a cause or reason for an action or attitude that is described in the main clause. Through the deployment of such conjunctives the speaker also retains greater control over the subsequent inferences made by the recipient, which means such implicatures are not explicitly deniable. The speaker could have subsequently claimed that he was thinking of notes for another class, and that in fact he does have the notes A is asking for, thereby correcting or repairing the prior implicature that he cannot lend the notes to B. However, this is not because the implicature is "cancellable", as it still holds from what was previously said by Jun. It is simply that the speaker has indicated he is no longer committed to what has been implied. In this case, then, the inference leading to the implicature is not so much defeasible as contingent on prior and prospective utterances (cf. Haugh 2008c: 445–446).

In the course of this discussion, I have alluded to the claim that much of the current debate about the cancellability test has unnecessarily conflated implicatures with both speaker (intended) meaning and inference. It has been suggested that implicatures are not as determinate, while inferences are more contingent, than is often assumed by analysts. There are, nevertheless, instances where implicatures do indeed appear to defy straightforward "cancellation". In the following two sections, I go on to suggest that such issues may be more productively examined by carefully distinguishing between cancellation vis-à-vis inference as opposed to implicature.

3 "Cancellability" vis-à-vis Inference

Inference is a form of reasoning whereby we can either draw consequences that a proposition has by virtue of truth-conditions (leading to entailments), or those which seem necessary, permissible or reasonable for one to draw (leading to implicatures, among other things) (Woods 2010: 218). The former type encompasses logical inferences that are not strictly open to error. The latter type encompasses inferences that are generally characterised as defeasible, which means they always allow for the possibility of error. In other words, they may not seem necessary, permissible or reasonable in certain situations. It is commonly assumed the different kinds of pragmatic inference leading to implicatures are cancellable because of their inherent defeasibility (Levinson 1983, 2000), but strictly speaking only non-monotonic inferences can be "cancelled" in the sense of

either being blocked or suspended.⁵ In this section, it is suggested that up on closer analysis, “cancellability” appears to involve two different processes in relation to inferences, namely, blocking and suspension.

Blocking inferences encompasses instances where a potential inference, which could conceivably follow from something that has been said, is not allowed through by the speaker. Blocking is thus prospective or anticipatory in nature. It generally occurs in cases where there is an unwanted inference, that is, something that *could* be inferred from what is said, but which the speaker does not wish to commit to, or be held accountable for, meaning. Blocking inferences thus contrasts with “contextual cancellation” or “non-arising” where “meanings are neither intended nor inferred: they are just ‘conceivable’ in a thought experiment” (Jaszczolt 2009: 261). A speaker can block unwanted inferences that the recipient might be expected to otherwise draw through various connectives that implicitly orient to what could have been inferred, including “but *x*” (examples 9 and 10), “if not *x*” (example 11), “in fact/actually *x*” (example 12) and so on.

(9) A: Will Sally attend the meeting?

B: Her car broke down, but she’ll still be there. (Huitink and Spenader 2004: 9).

(10) A: Smith doesn’t seem to have a girlfriend these days.

B: He has been paying a lot of visits to New York lately, but I don’t think he has a girlfriend there. (Huitink and Spenader 2004: 10; Grice 1989: 32).

(11) Some, if not all, the people left the building.

(12) Some, in fact all, the people said they liked the food. (Jaszczolt 2009: 261).

In example (9), we can see how the second phrase blocks the (unwanted) inference that Sally will (probably) not attend the meeting. In example (10), in contrast, we can how the second phrase blocks the (unwanted) inference that Smith has a girlfriend in New York, although it does not block the inference that Smith *might* have a girlfriend there but rather reinforces it. In example (11), the inference that not all of the people left the building is blocked by the intervening “if not” phrase, whereby the speaker implies epistemic uncertainty instead; that is, he or she is not sure whether or not everyone has left the building.⁶ This contrasts with example (12), where the inference that not all the people said they liked the food is blocked by the intervening “in fact” phrase. In each instance, the connective itself

⁵ We can further distinguish between defeasible inferences that are monotonic and those that are non-monotonic. Monotonicity refers to the property of inferences whereby the addition of new information does not reduce the set of what is known, and so no matter what further premises are added to a given premise set (provided none of the existing ones are erased), the set of conclusions can only increase. Non-monotonicity, on the other hand, refers to inferences where their present reasonability may be lost upon the addition of new information.

⁶ Cf. Levinson (1983: 115; fn. 15) who analyses this as an instance where the speaker is not committed to the truth or falsity of the implicature (so-called “suspension” of implicature).

orients to a possible inference that is subsequently blocked by indexing a contrast with what could have been inferred.

Suspension, on the other hand, involves removing the speaker's commitment to an inference that has already likely been drawn, and is thus retroactive in orientation. In the following excerpt from a recording of a telephone conversation between friends, for instance, we can observe how a potential default inference is suspended by a subsequent nonce inference.

- (13) (Cameron has called up Steve on the phone).
18. C: Whadaya doing tonight?
19. S: I dunno, what are you gonna do?
20. C: Oh I've been invited to a party that I don't wanna really want to go to.
21. Hope to find an excuse.
22. S: Oh right. I see.
23. Um, haveya seen Lethal Weapon? (adapted from Haugh 2009: 105).

In this short interchange, Cameron's initial question in line 18 is in the form of a standard pre-invitation (Schegloff 2007: 129). In other words, in a minimal context, namely, two friends talking on the phone, the speaker standardly implies through this utterance-type that he is going to subsequently invite the recipient to do something together, if the preparatory condition that has been invoked is met (i. e. that the other person is available). Steve orients to this possible invitation (and thus a *putative* short-circuited implicature) by inquiring about what the invitation might involve (line 19). This default inference, however, is subsequently suspended in Cameron's next turn where he outlines the preparatory conditions for another kind of invitation, namely, an invitation from Steve (lines 20–21), thereby implying that he would like Steve to invite him somewhere (and so provide the means by which he can be excused from going to a party he doesn't want to attend). This latter implicature, through which Cameron solicits an invitation from Steve, arises from a nonce inference drawing from specific contextual information. The change in Steve's understanding of the upshot of Cameron's prior utterance in line 18 in light of this qualification (in lines 20–21) is evident from his use of 'oh' and 'right' (line 22), through which he receipts 'new' information and understanding respectively (Heritage 1984; Schegloff 2007: 118), and his subsequent issuing of an invitation sequence (data not shown) starting with a standard pre-invitation (line 23) (cf. Haugh 2009: 104–107). In this example, then, we can see how the response of the recipient (Steve) helps constrain what is inferred amongst the potential interpretations initially afforded by the speaker's (Cameron) first utterance, an understanding which is subsequently qualified by the first speaker (Cameron). It is the contingency of this inferential work that allows for the suspension of the default inference that was initially drawn by the recipient (Steve).

The current analysis can also be taken to suggest that the blocking or suspension of pragmatic inferences could be more productively conceptualised as a consequence of their inherent contingency rather than defeasibility. One reason for

this is that defeasibility presumes a “logical” basis for analysing inference that arguably carries with it unnecessary baggage for the analysis of natural language use. While some pragmatic inferences may seem akin to logical inferences, the interpretation of natural language use is not bound by the formal restrictions of logical analysis. Thus, invoking (pseudo-)logical constraints on the analysis of the inferential work underlying the interpretation of implicatures is misleading as it involves a different order of analysis and language description. A second reason for favouring contingency over defeasibility is the observation that entailments can be implicated. In such cases, the inference leading to the implicated entailment cannot be regarded as defeasible (i.e. open to error), because inferences underpinning entailments are logical (i.e. follow analytic rules), and thus are not open to error strictly speaking. They are nevertheless contingent on what precedes and follows the utterance in question, as the implicated entailment can be “removed” through an “except for” clause, for instance, as was discussed in [Sect. 2.1](#). The non-defeasibility of some instances of pragmatic inference also became apparent in the course of the discussion of various instances of cancellation-resistant implicatures ([Sect. 2.3](#)). Thus, while we may admit that implicatures are always contingent, this is not because the underlying inferences are necessarily always defeasible, but rather because pragmatic inferences are always anticipatory and retroactive in nature (Haugh 2012: 263).

4 “Cancellability” vis-à-vis Implicature

An implicature is commonly conceptualized as a species of speaker meaning that pertains to separate individual, additional thoughts with their own pragmatic force, as previously noted. It was also pointed out that different types of implicatures can be distinguished on the basis of the distinction between conventionalized reasoning across contexts (default inference), and once-off reasoning in particular contexts (nonce inference). The former generate default implicatures, while the latter generate nonce (or particularised) implicatures, with a further distinction possible between default implicatures that arise across all contexts unless otherwise cancelled (generalised implicatures), and those that only require a minimal context in which to arise (short-circuited implicatures). Finally, it has also been noted that we can draw a distinction between speaker-implicatum and recipient-implicatum (or ‘utterer-implicature’ and ‘audience-implicature’).⁷ Since implicatures are “products” of inferring, then, they can vary both in terms of who is doing the inferring (i.e. speaker or recipient), and in relation to what kind of

⁷ We might add a third category, namely, “emergent implicature” (cf. Haugh 2008a, 2009), where implicatures are interactionally achieved across speaker-recipient dyads, but further consideration of this must await another time, as whether one accepts the validity of such a notion does not substantively impact on the points being made here.

inference is involved (default or nonce). Moreover, since they are also “products” of implying they are recognisably social or normative in nature. An upshot of taking this position is that the cancellation of implicatures can be conceptualised as a kind of speech act or social action in itself (Borge 2009: 152). And furthermore, in constituting a social action, cancellation may be therefore be taken literally or non-literally (Jaszczolt 2009: 264). In examining instances of the “cancellation” of implicatures in discourse it appears there are two broad categories of social action involving explicit cancellation.

The first is correction or repair. An explicit cancellation activates a kind of “not-frame” which presupposes there is something there to cancel in the first place. In this sense, then, we cannot talk of an implicature being removed or no longer existing. And for this reason a distinction between “cancellation” and “correction” of implicatures is difficult to maintain (cf. Bultinck 2005: 35; Jaszczolt 2009: 266–267). Instead, we can talk of implicatures being open to denial (e.g., claims that the implicature was not speaker intended), retraction (e.g., claims that the implicature is no longer relevant or applicable to the current discourse), or clarification (e.g., claims that what was meant is different to how it was understood or might have been understood by the recipient), among other things. All of these actions involve modulation, in some form or another, of the degree of the speaker’s commitment to, or accountability for, the implicature in question. The plausibility or legitimacy of the actions of denial, retraction or clarification is a function of the degree of indeterminacy of the implicature in question as well as its status vis-à-vis primary and secondary meanings. This means that “cancellability” is a gradable property, dependent, in part, on the speaker’s perceived intentions (Haugh and Jaszczolt 2012: 96). Perceived intentions are, in turn, a function of the degree of salience of the implicature in question, as well as the sequential placement of the utterance that gives rise to it (i.e. whether it is closely linked to a prior utterance and thus dependent on how the prior speaker’s utterance is understood). In other words, the degree of plausibility of denying, retracting or clarifying an implicature is a function of what is normatively made available by the speaker in saying *x*. In cases where something has clearly been implied (i.e. implicature as primary meaning), the speaker can usually only clarify what was implied. In cases where it is open to interpretation whether something has been implied (i.e. implicature as secondary meaning), then implicatures may be more readily denied or retracted as well. In all cases, however, correction/repair actions are always open to dispute by participants as to what exactly was implied, or the degree of speaker commitment to what was taken to be implied, given the existence of both speaker and recipient implicatum.

In the example below taken from an episode in *Seinfeld*, for instance, the two characters argue about whether something has been implied by Elaine’s question. The conversation begins when Elaine reports that her colleague, who is a former alcoholic, mistakenly took Jerry’s drink at a party they had all attended, and has since started drinking again, leading to his dismissal.

(14) (“The red dot”, *Seinfeld*, Season 3)

Elaine: Dick was fired.

Jerry: You mean to tell me if I had put that drink six inches over to the right, and none of this would have happened.

Elaine: You knew he was an alcoholic. Why’d you put the drink down at all?

Jerry: What are you saying?

Elaine: I’m not saying anything.

Jerry: You’re saying something.

Elaine: What could I be saying?

Jerry: Well you’re not saying nothing you must be saying something.

Elaine: If I was saying something I would have said it.

Jerry: Well why don’t you say it?

Elaine: I said it.

Jerry: What did you say?

Elaine: Nothing. (sighs) It’s exhausting being with you.

Here Jerry and Elaine use *saying* not so much in the sense of uttering but rather in the sense of *meaning* something, more specifically, something that is not said but implied. We have here, then, an example of an attempted denial of an implicature, although what exactly is being denied is never made clear. Notably, while Jerry appears to appeal to what would be normatively interpreted as being made available via implicature by Elaine’s question (“Well you’re not saying nothing you must be saying something”), Elaine invokes her own intention to deny implying anything (“If I was saying something I would have said it”).⁸ Disputes about attempts by speakers to deny, retract or clarify implicatures can be, of course, much more consequential or serious than in this example, as discussed with reference to media disputes about what was implied by a prominent religious figure in Australia (Haugh 2008b).

The second broad category of social action involving explicit cancellation phrases has already been alluded to, namely, reinforcement, where the implicature is reiterated without redundancy. Weiner’s (2006) example of a cancellation-resistant implicature illustrates rather nicely how what can appear on the surface to be an attempted cancellation, is in fact reinforcing the sarcasm.

(15) (Alice and Sarah are in a crowded train. Alice, who is obviously able-bodied, is sprawled across two seats, and Sarah is standing).

Sarah: I’m curious as to whether it would be physically possible for you to make room for someone else to sit down. Not that you *should* make room; I’m just curious.

(Weiner 2006: 128)

Here we can see how the second utterance, which is formulated as a denial that she is implying she wants Alice to move, does not involve correction or repair but rather reinforcement of a sarcastic stance on the part of Sarah. It is also worth noting that while it is obvious that the speaker is not committed to what is said here (i.e. Sarah is clearly not curious), the implied feeling, attitude or evaluation is quite

⁸ Cf. Capone (2009) on social versus individual intentionality.

indeterminate, albeit obviously negative. However, in constituting a highly salient, primary meaning, this latter implicature cannot be readily retracted or denied, although it is nevertheless open to clarification.

5 Concluding Remarks

These debates have highlighted a number of apparent contradictions for both (neo-) Gricean and relevance theoretic approaches to implicature: not all implicatures are necessarily speaker intended, and inferences should not be treated as synonymous with implicatures. It has been suggested here that current debates about cancellability may be advanced by acknowledging such distinctions. It has also been argued that any analysis of cancellability vis-à-vis inference and implicature should recognise both the inherent contingency of inferences, and the inherent indeterminacy of implicatures. The question of whether inferences or implicatures are “cancellable” is thus an empirical question. This is not to deny the utility of debates around cancellability that test the internal coherence of particular theoretical stances, but rather to suggest that a more nuanced understanding of “cancellability” can help us to better understand the underlying nature of implicatures.

It has also become apparent in the course of this discussion that cancellability does not provide a reliable means of distinguishing between semantic and pragmatic contributions to meaning, as has already been noted (Carston 2010; Jaszczolt 2009). The theoretical relevance of cancellability itself for pragmatics has thus come into question. This has led researchers to two quite different conclusions. On the one hand, Carston (2010) has suggested that it simply be abandoned. On the other hand, Jaszczolt (2009) has suggested that it needs to be the subject of further empirically-driven analyses. In this chapter, I have taken the latter route. I have suggested that a careful examination of how cancellability plays out in the case of different examples of inferences and implicatures remains useful in advancing our understanding of meaning beyond what is said, and the inferences that give rise to it.

References

- Bach, Kent. 2001. You don't say? *Synthese* 128: 15–44.
- Bach, Kent. 2006. The top 10 misconceptions about implicature. In *Drawing the boundaries of meaning. Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, eds. Betty Birner and Gregory Ward, 21–30. Amsterdam: John Benjamins.
- Blome-Tillmann, Michael. 2008. Conversational implicature and the cancellability test. *Analysis* 68: 156–160.
- Borge, Steffen. 2009. Conversational implicatures and cancellability. *Acta Analytica* 24: 144–154.
- Bultinck, Bert. 2005. *Numerous meanings: the meaning of English cardinals and the legacy of Paul Grice*. Oxford: Elsevier Science.

- Burton-Roberts, Noel. 2006. Cancellation and intention. *Newcastle Working Papers in Linguistics* 12(13): 1–12.
- Burton-Roberts, Noel. 2010. Cancellation and intention. In *Explicit Communication. Robyn Carston's Pragmatics*, eds. Belén Soria and Esther Romero, 138–155. London: Palgrave Macmillan.
- Capone, Alessandro. 2006. On Grice's circle (a theory-internal problem in linguistic theories of the Gricean type). *Journal of Pragmatics* 38: 645–669.
- Capone, Alessandro. 2009. Are explicatures cancellable? Towards a theory of the speaker's intentionality. *Intercultural Pragmatics* 6: 55–83.
- Capone, Alessandro. 2010. What can modularity of mind tell us about the semantics/pragmatics debate? *Australian Journal of Linguistics* 30: 497–520.
- Carston, Robyn. 2002. *Thoughts and utterances. The pragmatics of explicit communication*. Oxford: Blackwell.
- Carston, Robyn. 2004. Relevance theory and the saying/implicating distinction. In *Handbook of pragmatics*, eds. Laurence Horn, and Gregory Ward, 633–656. Oxford: Blackwell.
- Carston, Robyn. 2010. Explicit communication and 'free' pragmatic enrichment. In *Explicit Communication. Robyn Carston's Pragmatics*, eds. Belén Soria and Esther Romero, 217–285. London: Palgrave Macmillan.
- Cummings, Louise. 2005. *Pragmatics a multidisciplinary perspective*. Edinburgh: Edinburgh University Press.
- Davis, Wayne. 1998. *Implicature. Intention, Convention, and Principle in the Failure of Gricean Theory*. Cambridge: Cambridge University Press.
- Grice, H.P. 1975. Logic and conversation. In *Syntax and Semantics, Volume 3. Speech Acts*, eds. Peter Cole and Jerry Morgan, 41–58. New York: Academic Press.
- Grice, Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.
- Haugh, Michael. 2002. The intuitive basis of implicature: relevance theoretic *implicitness* versus Gricean *implying*. *Pragmatics* 12: 117–134.
- Haugh, Michael. 2008a. The place of intention in the interactional achievement of implicature. In *Intention, common ground and the egocentric speaker-hearer*, eds. Istvan Kecskes, and Jacob Mey, 45–85. Berlin: Mouton de Gruyter.
- Haugh, Michael. 2008b. Intention and diverging interpretations of implicature in the 'uncovered meat' sermon. *Intercultural Pragmatics* 5: 201–228.
- Haugh, Michael. 2008c. Utterance-final conjunctive particles and implicature in Japanese conversation. *Pragmatics* 18: 425–451.
- Haugh, Michael. 2009. Intention(ality) and the conceptualisation of communication in pragmatics. *Australian Journal of Linguistics* 29: 91–113.
- Haugh, Michael. 2011. Practices and defaults in interpreting disjunction. In *Salience and defaults in utterance processing*, eds. Kasia M. Jaszczolt, and Keith Allan, 193–230. Berlin: Mouton de Gruyter.
- Haugh, Michael. 2012. Conversational interaction. In *The Cambridge handbook of pragmatics*, eds. Keith Allan, and Kasia M. Jaszczolt, 251–274. Cambridge: Cambridge University Press.
- Haugh, Michael, and Kasia M. Jaszczolt. 2012. Speaker intentions and intentionality. In *The Cambridge handbook of pragmatics*, eds. Keith Allan, and M. Kasia Jaszczolt, 87–112. Cambridge: Cambridge University Press.
- Heritage, John. 1984. A change of state token and aspects of its sequential placement. In *Structures of social action*, eds. Maxwell J. Atkinson, and John Heritage, 299–345. Cambridge: Cambridge University Press.
- Higashimori, Isao, and Deirdre Wilson. 1996. Questions on relevance. *UCL Working Papers in Linguistics* 8: 111–124.
- Horn, Laurence. 2004. Implicature. In *Handbook of pragmatics*, eds. Laurence Horn, and Gregory Ward, 3–28. Oxford: Blackwell.
- Horn, Laurence. 2012. Implying and inferring. In *The Cambridge handbook of pragmatics*, ed. Keith Allan, and M.Kasia Jaszczolt, 69–86. Cambridge: Cambridge University Press.
- Huang, Yan. 2007. *Pragmatics*. Oxford: Oxford University Press.

- Huitink, Janneke, and Spenader, Jennifer. 2004. Cancellation resistant PCIs. In *Proceedings of the ESSLLI 2004 workshop on implicature and conversational meaning*, eds. Bart Geurts and Rob van der Sandt, 8–13. Nancy, France: University of Nijmegen.
- Jaszczolt, K.M. 1999. Default semantics, pragmatics, and intentions. In *The Semantics/Pragmatics interface from different points of view*, ed. Ken Turner, 199–232. Oxford: Elsevier.
- Jaszczolt, K.M. 2009. Cancellability and the primary/secondary meaning distinction. *Intercultural Pragmatics* 6: 259–289.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, Stephen. 2000. *Presumptive meanings. The theory of generalised conversational implicature*. Cambridge: MIT Press.
- Saul, Jennifer. 2002. Speaker meaning, what is said, and what is implicated. *Nous* 36: 228–248.
- Schegloff, Emanuel. 2007. *Sequence organization in interaction*. Cambridge: Cambridge University Press.
- Sperber, Dan, and Deirdre Wilson. 1995. *Relevance. Communication and cognition*, 2nd edn. Oxford: Blackwell.
- Vicente, Begona. 1998. Against blurring the explicit/implicit distinction. *Revista Alicantina de Estudios Ingleses* 11: 241–258.
- Weiner, Matthew. 2006. Are all conversational implicatures cancellable? *Analysis* 66: 127–130.
- Wilson, Deirdre, and Dan Sperber. 1986. Inference and implicature. In *Meaning and interpretation*, ed. Charles Travis, 45–75. Oxford: Basil Blackwell.
- Woods, John. 2010. Inference. In *The pragmatics encyclopedia*, ed. Louise Cummings, 218–220. London: Routledge.

Grice, Conversational Implicature and Philosophy

Siobhan Chapman

Abstract The importance of Grice's theory of conversation and in particular his account of conversational implicature (Grice in *Studies in the way of words*. Harvard University Press, Harvard, pp. 22–40, 1975; 1978) in the development and current concerns of pragmatics is almost impossible to exaggerate. Whether or not they agree with the details or even the broader framework of Grice's theory, pragmaticists generally acknowledge the significance of his attempt to give a formalised account of the differences between what our words literally mean and what we intend to communicate in using them. But Grice himself was a philosopher, not a linguist; his work was deeply rooted in the philosophical preoccupations of the mid twentieth century, and he never used the word 'pragmatic' in his writings as it is used in present day linguistics. This chapter will address the contrast between Grice's philosophical motivations in developing his account of conversational implicature, and the linguistic framework in which it has subsequently generally been discussed. It will do so by considering the philosophical context in which Grice was working and some of the specific philosophical problems to which he applied his notion of conversational implicature. It will begin with a review of the dichotomy in twentieth century analytic philosophy that can be summarised as a distinction between 'ideal language' and 'ordinary language' philosophy, and will discuss Grice's work as an attempt to demonstrate some fundamental misconceptions in both positions. In doing so, it will compare Grice's work on conversational implicature with the near contemporary work by Austin on speech acts (Austin in *How to do things with words*. Oxford University Press, Oxford, 1962a; *Sense and sensibilia*. Oxford University Press, Oxford, 1962b). Austin's work shares a number of apparent similarities with Grice's, but reveals some significant philosophical differences, particularly with regard to the

S. Chapman (✉)
University of Liverpool, Liverpool, UK
e-mail: src@liverpool.ac.uk

nature of ‘literal meaning’ and the role of theorising. This chapter will then offer an exegesis of Grice’s conception of conversational implicature, including the distinction which seemed necessary to him between ‘generalised’ and ‘particularised’ conversational implicatures, and some of the properties that he identified as characteristic of conversational implicatures. Following this exegesis, this chapter will consider some of the applications and extensions of the concept of conversational implicature suggested by Grice and, in some cases, developed further by his later commentators. Grice found his concept to be fruitful in addressing a range of established philosophical problems, including the viability of his own earlier work on ‘non-natural meaning’ (Grice in *Studies in the way of words*. Harvard University Press, Harvard, pp. 213–223, 1957), the contrasting claims of realism and skepticism (Grice in *Studies in the way of words*. Harvard University Press, Harvard, pp. 147–153, c. 1946–1950; pp. 154–170, c. 1953–1958; pp. 224–247, 1961), apparent differences between logic and natural language (Grice in *Studies in the way of words*. Harvard University Press, Harvard, pp. 3–21, 1967a; pp. 58–85, 1967b) and the debate over Russell’s logical account of definite descriptions (Grice in *Studies in the way of words*. Harvard University Press, Harvard, pp. 269–282, 1981). Grice introduced the technical term ‘implicature’ into his philosophy of language. It has subsequently become part of the defining terminology of present day pragmatics and is a central concept both for those working in a broadly neo-Gricean framework (Horn in *A natural history of negation*. University of Chicago Press, Chicago, 1989; *Pragmatics*. Palgrave, Basingstoke, pp. 158–183, 2007; Levinson in *Presumptive meanings: The theory of generalized conversational implicature*. M.I.T. Press, Cambridge, 2000) and for those working in relevance theory (Sperber and Wilson in *Relevance*. Blackwell, Oxford, 1995; Carston in *Thoughts and utterances*. Blackwell, Oxford, 2002). Recent interest in pragmatics has focussed on the division that Grice drew between ‘what is said’ and ‘what is implicated’, and on the viability of that distinction. This has been a focus of study in both theoretical pragmatics (Borg in *Minimal semantics*. Oxford University Press, Oxford, 2004; Recanati in *Literal meaning*. Cambridge University Press, Cambridge, 2004; Cappelen and Lepore in *Insensitive semantics: a defense of semantic minimalism and speech act pluralism*. Blackwell, Oxford, 2005) and experimental pragmatics (Gibbs, *Brain and Language* 68: 466–485, 1999; Glucksberg in *Experimental pragmatics*. Palgrave, Basingstoke, pp. 72–93, 2004; Breheny et al., *Cognition* 100: 434–463, 2006). This chapter will conclude with an assessment of the significance of Grice’s delineation of ‘what is said’ as a defining opposite of ‘what is implicated’. For Grice himself, although it was central to his original philosophical motivations, this remained one of the most troublesome and least successful aspects of his philosophy of language. For pragmaticists it has proved one of the most enduring, challenging and intriguing topics of debate.

1 Introduction

The focus of this chapter is Grice's theory of conversation, and in particular the notion of 'conversational implicature' that he expounded and developed during the course of his 1967 lecture series on 'Logic and Conversation'. It is almost impossible to overstate the importance of this notion to pragmatics, in terms both of its development as a field of linguistic inquiry and of its current concerns. But Grice's work was rooted in the philosophical debates of the mid twentieth century. It developed out of his own preoccupation with certain philosophical problems, particularly those relating to the nature of meaning and the relationship between logic and natural language. And in later writings he considered its utility in application to further philosophical problems.

This chapter begins with an overview of the philosophical background that was Grice's intellectual home. In particular, the school of 'ordinary language philosophy' with which Grice was closely associated both informed his work and, increasingly, motivated his search for an alternative philosophy of language. [Section 3](#) offers an exegesis of the theory of conversation, with a particular focus on the nature and properties of conversational implicature. Subsequent sections address some of the specific philosophical problems that were Grice's own targets in developing the theory of conversation, and in relation to which he claimed varying degrees of success. The chapter will conclude with a brief overview of some of the ways in which the notion of conversational implicature has informed the development and current state of linguistic pragmatics, with particular reference to the controversial and highly influential Gricean notion of 'what is said'.

2 'Ideal' and 'Ordinary' Language Philosophy

During the middle decades of the twentieth century, Grice was generally identified with the Oxford-based school of thought known as 'Ordinary Language Philosophy' (OLP). Grice himself was reluctant to categorise groups of philosophers in terms of approach or methodology, and even more reluctant to align himself with any single school or movement. Nevertheless, the identification seems justified, especially since Grice himself acknowledged his allegiance to OLP in one of the few contemporary commentaries on it (Grice [1958](#)).

OLP can perhaps best be understood as a reaction against certain tendencies in contemporary analytic philosophy. The emergence of analytic philosophy has been described as a time at which philosophy took a 'linguistic turn' (Rorty [1967](#)). The very general trend was away from metaphysical speculation and knowledge claims based on individual intuition, towards empirical observation and knowledge claims based on rigorous analysis, which often meant analysis of the language in which that observation was expressed. But that is not by any means to say that early analytic philosophers embraced natural language as a object of serious study, in

the way that would seem natural to a present day linguist. Indeed when many of the most influential early analytic philosophers turned their attention to natural language they found it sadly lacking in rigour and philosophical utility. The outcome of their analysis was often to reveal natural language to be flawed and misleading in relation to the true nature of knowledge and thought. Analysis might be the way forward for philosophy, but it was to be analysis of 'ideal' language, the artificial languages of logic and mathematics, rather than of messy and imperfect natural language.

For instance, Frege, in an article that can be seen as foundational to analytic philosophy, distinguished between the 'reference' of a word in the world and its 'sense', or the identifying property of that reference. In relation to proper names, Frege admits that in natural language sense may vary from one individual speaker to another, but he specifies that this 'ought not to occur in a perfect language' (Frege 1892: 58). When Russell and Whitehead published their formidable *Principia Mathematica* they prefaced it with a defence of using symbolic notation rather than natural language for the purposes of philosophical analysis. The structure of ordinary language was not a good guide to the structure of thought, they argued: 'Thus, "a whale is big" and "one is a number" both look alike, so that the eye gives no help to the imagination' (Russell and Whitehead 1910: 2). The leading philosopher of syntactic analysis, Carnap, restricted his attention to the artificial languages appropriate to logic and to science because of 'the unsystematic and logically imperfect structure of the natural word-languages (such as German or Latin)' (Carnap 1937: 2). Carnap and his ideas were highly significant in the development of the school of logical positivism, the movement within analytic philosophy that was perhaps most inimical of all to natural language. Certainly logical positivism was deeply concerned with the analysis of language, but one of the primary aims of this analysis was to distinguish the 'meaningful' from the 'meaningless', or what could legitimately be said within the confines of empirical discussion from what could not. Meaningful sentences were those capable of being assigned the value 'true' or 'false', either because of inherent properties of their meaning, or because of identifiable processes of verification in the world of experience. As a consequence, many sentences that would be used without controversy in everyday usage, including those to do with moral judgement, emotional response and metaphysical belief, were dismissed as 'mere pseudo-proposition[s]' (Ayer 1946: 48).

Against this background, in the years immediately surrounding the Second World War, J. L. Austin emerged as a champion of ordinary language. Austin was certainly not the first philosopher to pay attention to natural language and the ways in which it is used in everyday life to describe our experiences and knowledge of the world. But he was the first to found an entire philosophical approach on ordinary language, and to elevate its analysis to a philosophical method in its own right. Austin argued not just that ordinary language deserved serious philosophical attention, but also that philosophers would be best served by basing their investigations of any topic on a careful consideration of the terminology available to discuss that topic in ordinary, non-philosophical discourse. The philosopher should

always begin 'by examining what we should say when, and so why and what we should mean by it' (Austin 1956). This attitude was foundational to OLP, which was based in Oxford and centred around the figure of Austin himself.

The most direct confrontation between the methods and assumptions of ideal and those of ordinary language philosophy was that between Russell and Strawson. Their topic was the correct analysis of sentences that contain definite descriptions that do not relate to entities in the actual world: using Russell's own much-cited example, sentences such as 'the king of France is bald'. For Russell, this was a classic example of a natural language expression that was misleading as to the logical structure of the thought it expressed. Despite appearances, this sentence was not in fact a simple declarative sentence in which a property denoted by the predicate is assigned to an entity denoted by the subject. That is, we are deceived by the superficial grammar of the language if we believe it is structurally similar to an example such as 'Louis is bald'. It is not possible to analyse 'the king of France' as the subject of the sentence, because the sentence in fact has no subject, but is logically composed of a complex set of propositions concerning existence and uniqueness. In effect, the logical form of Russell's example specifies that the king of France exists and exists uniquely and is bald. Since at least one of these constituent propositions is false (there is no king of France) the logical form as a whole, and therefore the sentence that expresses it 'is plainly false' (Russell 1905: 484).

Some decades after Russell published this analysis, Strawson challenged him on the grounds that he had failed to take account of how people would actually use and respond to such expressions. What Russell said about propositions and falsity might work well enough for logic, but it could not be transferred across to an account of natural language. In everyday conversation, people are very unlikely to encounter an example such as 'the king of France is bald'. But if they did they would not be likely, as Russell's analysis seemed to predict, to judge it to be a false statement. Rather they would feel that something had gone wrong: that the statement should not even legitimately be made. For Strawson, this was because the propositions about existence and uniqueness were not logical components of the meaning of Russell's example. Rather, they were 'presuppositions': necessary conditions for either the example or its negation ('the king of France is not bald') meaningfully to be stated. In the absence of a present king of France, Russell's example was not false but rather failed to be either true or false. Russell's attempt to carry rules from logic over to natural language were misguided 'for ordinary language has no exact logic' (Strawson 1950: 344).

In his own writings, Austin focussed his attention on attempting to demonstrate that previous philosophers had been led into error by their failure to pay proper attention to ordinary language. He argued against the assumption that words have fixed meanings that are intuitively available to the philosopher and that can therefore be input into philosophical analysis. An examination of ordinary usage reveals 'that there is *no* simple and handy appendage of a word called "the meaning of (the word) x"' (Austin 1940: 30, original emphasis). Rather, the philosopher must be aware of a complex set of potentially interrelated but nevertheless distinct uses of any word. Failure to attend to the potentials of ordinary

language had led philosophers into what Austin described as the ‘descriptivist fallacy’ (Austin 1950): the mistaken belief that language is primarily a tool for expressing facts about the world that can be judged to be ‘true’ or ‘false’. People in fact have many different ways of reacting to what they hear that do not draw on a simple dichotomy between these two truth values. People do a lot more things with words than simply offering descriptions of the world.

It was partly in response to the descriptive fallacy that Austin developed what has become known as his ‘theory of speech acts’ (Austin 1962a). In doing so, he proposed to consider not just the ‘forms’ of language, which had been the traditional preoccupation of philosophers, but also its ‘functions’, or the various different things it could be used to do. These included offering descriptions of the world, but they also included a host of other functions, such as placing bets, making promises, issuing requests and asking questions. Austin distinguished between the ‘locutionary’ and the ‘illocutionary’ act performed whenever someone speaks. The locutionary act, something close to the conventional meaning of the words used, may often not coincide with the illocutionary act, which concerns what the speaker was actually doing in producing those words, and which generally relies on intention as well as convention. ‘Could you pass the salt?’ might have the locutionary form of a question, but when we examine its function in terms of the illocutionary act, we see that it is actually performing a request.

There is an inherent tension within Austin’s work around the notion of ‘literal’ or ‘conventional’ meaning, and it is in part to this tension that Grice was responding when he began to voice disquiet with what Austin was doing, and by extension with OLP. Austin’s insistence that philosophical inquiry must start with close attention to the range of ordinary usage and the resultant realisation that words had no single fixed meaning seemed to call for the abandonment of ‘meaning’ as a category distinct from ‘use’. Yet his own philosophical theory of speech acts relied on a distinction between a recognisable conventional component of meaning, represented by the locutionary act, and a separate function once that meaning was put into use in the illocutionary act. Grice commented very little publicly on speech act theory, certainly not on Austin’s version of it, but he did pick up on a closely related dilemma. Austin was inclined to consult likely contexts in which certain words would be used in order to learn about their meaning. But at the same time he was reluctant to commit himself to the question of whether these apparent features of meaning actually contributed to making the sentences in which the words were used true or false. Grice commented that ‘I am very much afraid that he was trying to have his cake and eat it’ (Grice 1967a: 13).

Grice, then, was becoming aware of a fundamental problem with OLP. Certainly, Austin had drawn attention to the tendency of philosophers of ideal language to overlook how words are actually used and therefore to assume that it is safe to consider them as having a single, fixed meaning that is available to intuitive inspection. But in attempting to replace the ‘ideal’ with the ‘ordinary’ picture of language, Austin had failed fully to take account of the differences between

meaning in general and meaning on particular occasion, or between the conventions of language and norms of usage.¹

3 Grice's Theory of Conversation

By the middle of the 1960s, Grice had set himself a tough challenge. He was to counter the dismissive, sometimes almost contemptuous attitude to ordinary language apparent in some branches of recent and contemporary analytic philosophy, by demonstrating that everyday uses of language were amenable to serious philosophical scrutiny. But he was to do so without succumbing to what he saw as the fallacy that had beset some of his colleagues such as Austin: that of elevating the use of language to a place of philosophical preeminence at the expense of the linguistic system underlying it. That is, he was to propound a philosophy of language that took seriously what he described as 'what a particular speaker or writer means by a sign on a particular occasion', while at the same time accounting for 'the standard meaning of the sign' (Grice 1957: 216–217).

There was more. Grice's aspirations for his philosophy of language were not just that it should acknowledge that what we mean by the words we use and what those words themselves literally mean may often be two different matters. He wanted not just to describe this difference, but actually to explain it. That is, he wanted to establish a theoretical apparatus to model how speaker meaning is related to and can be derived from linguistic meaning when language is used for the purposes of communication. This apparatus was to be specific enough to be expressed explicitly, but general enough to be capable of explaining the relationship between the literal and the intended across the full range of types of language use, from obviously figurative examples such as metaphor and irony to less clearly defined instances of suggestion, intimation and implication. The lack of overarching explanations was one of the features he had objected to in Austin's philosophy of language, with its emphasis on listing and classifying types of usage and even types of words. In contrast, Grice had found the new approach to the study of language apparent in the early work of Chomsky refreshing and inspiring. What he found most noteworthy in Chomsky's work was the fact that the latter had demonstrated how 'a region for long found theoretically intractable by scholars (like Jespersen) of the highest intelligence could, by discovery and application of the right kind of apparatus, be brought under control' (Grice 1986: 59–60). Chomsky's subject matter was syntax whereas Grice's was the relationship between linguistic and speaker meaning, but the goal of controlled theoretical explanation was the same.

¹ For further discussion of Grice's relationship to OLP, see Chapman (2005, 2010) and Lüthi (2006).

Grice worked on his theory over a number of years during the 1950s and 1960s, presenting early versions of it and intimations of ‘work in progress’ during his teaching at Oxford and in various lectures and papers. The version that is now generally recognised as the definitive account of ‘the theory of conversation’, was that presented during the 1967 William James lecture series at Harvard. These were collectively entitled ‘Logic and Conversation’, and the second lecture, by far the most widely republished, cited and quoted, bears the same name. In this, Grice proposed to ‘inquire into the general conditions that, in one way or another, apply to conversation as such, irrespective of its subject matter’ (Grice 1975: 24). A clear picture of these general conditions was to provide a mechanism for explaining the myriad ways in which speaker meaning could differ from linguistic meaning.

In this second lecture Grice introduced the term ‘implicature’ to describe those elements of speaker meaning that go beyond or differ from literal meaning. He translated the question of the relationship between general meaning and particular meaning into a question of the relationship between ‘what is said’ and ‘what is implicated’. The former term received rather cursory treatment: ‘In the sense in which I am using the word *say*, I intend what someone has said to be closely related to the conventional meaning of the words (the sentence) uttered’, allowing for the assignment of reference to obviously referring expression such as pronouns, and the fixing of a specific meaning for any inherently ambiguous word. The rest of the lecture was dedicated to elaborating the notion of ‘what is implicated’, and for the vast majority of cases this involved the general conditions of conversation that were Grice’s focus, because in the vast majority of cases implicatures were dependent on and could be explained in terms of these conditions.

Grice dealt with those implicatures not included in this majority, which he labelled ‘conventional implicatures’, in a paragraph. Because this chapter is concerned with the nature and impact of conversational implicature we can afford to do the same, while acknowledging that conventional implicature has proved to be a concept of significance and complexity beyond what Grice seems to have imagined. There are a limited number of words which introduce implicatures when they are used simply because of their conventional meaning, without the need for any general conditions of conversation. Here is Grice’s example:

1. He is an Englishman; he is, therefore, brave.

Grice’s contention is that what is literally said here amounts to no more than the conjunction of a statement that he is an Englishman and a statement that he is brave. The meaning that is undoubtedly conveyed that being brave is a consequence of being an Englishman is to be explained as an implicature of (1) rather than part of what is said. Grice’s reason for this is that the ‘consequence’ meaning does not contribute to the truth conditions. (1) would be straightforwardly false in a situation in which the referent of ‘he’ was not an Englishman or was not brave. But someone who disagreed with the idea that being English necessarily entailed being brave would be hard pressed to say that the speaker uttering (1) had said something literally false. We might add that the speaker could be accused of saying something misleading, inaccurate or inappropriate. Conventional

implicature has proved controversial in subsequent philosophical discussion but has recently enjoyed something of a renaissance.²

In order to explain conversational implicatures, which depend not just on conventional meaning but on ‘general features of discourse’ (Grice 1975: 26), Grice considered conversation, one type of human social interaction, as an essentially collaborative or cooperative process. The nature of conversation was summed up by Grice in a supposed injunction to speakers, a ‘rough general principle’ which in fact remained Grice’s only formulation and therefore the canonical version of ‘the Cooperative Principle’: ‘Make your conversational contribution such as is required at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged’ (ibid.). Grice proceeded to elaborate a more detailed treatment of cooperation, in which various maxims express specific expectations that participants bring to a conversation. These maxims are distributed between four general categories: Quantity, Quality, Relation and Manner. The maxims of Quantity legislate that an appropriate amount of information should be included in each conversational contribution: neither more nor less. The maxims of Quality put the onus on speakers to try to make their contributions true ones. The single maxim of Relation enjoins relevance to the current conversational topic and purpose. The category of Manner comprises a disparate and possibly open-ended list of maxims. Unlike those belonging to the other three categories, these are concerned not with the nature of the information included in a contribution but with the form in which it is expressed. Here is Grice’s own formulation of the maxims:

Category of Quantity

1. Make your contribution as informative as is required (for the current purposes of the exchange).
2. Do not make your contribution more informative than is required.

Category of Quality

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Category of Relation

1. Be relevant.

Category of Manner

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief.
4. Be orderly.

² See, for instance, Karttunen and Peters (1979), Gazdar (1979), Levinson (1983) and Potts (2005).

Some examples of conversational implicature, perhaps the least interesting and certainly those to which Grice devoted the least attention, follow straightforwardly from the simple assumption that the speaker is observing the maxims. A is standing next to an immobilised car. B approaches.

2. A: I am out of petrol.

B: There is a garage round the corner.

B has not literally said that the garage is open, or that it is the sort of garage that sells fuel, but A is highly likely to access these two pieces of information as conversational implicatures of what B has said, on the reasonable assumption that B is following the maxim of Relation.

More interesting examples depend on situations in which the maxims are not being straightforwardly observed in some way. It should be a check on any naive claim that Grice was legislating for how people ought to behave properly in conversations, or presenting an idealised and unrealistic account of conversational behaviour, that he introduces the notion of conversational implicature with a discussion of the various ways in which 'A participant in a talk exchange may fail to fulfil a maxim' (Grice 1975: 30). The maxims express norms of interactional expectations and can be used to explain what happens when those expectations are breached just as much as when they are fulfilled. To begin with, speakers may not attend to the maxims at all. They may simply 'violate' them, as in the case of lying when the speaker ignores the maxims of Quality. In such cases the speaker is intent on misleading the hearer. More overtly, speakers may 'opt out' of the requirements of cooperation, explicitly indicating their unwillingness or inability to behave cooperatively, and thereby revealing an awareness of what the conversational norms are. Grice's rather prim example is of a speaker who says, 'I cannot say more; my lips are sealed' (*ibid.*). Grice does not elaborate on this example, but it is easy to identify it as an explicit opting out from the requirements of Quantity. The speaker is aware of the onus on her to give as much information as would be appropriate, and is indicating that she will not be doing so, for reasons that are outside of the cooperative principle and that overrule it.

A speaker may fail fully to comply with the requirements of a maxim for reasons that are internal to the cooperative principle itself. This occurs in cases where the demands of one maxim come into conflict with the demands of another or, in Grice's terminology, when the speaker is faced with a 'clash' of maxims. In Grice's example, it is clear to both A and B that, for the purposes of planning a holiday, A needs to know the actual town where C could be found:

3. A: Where does C live?

B: Somewhere in the South of France.

B is not complying with the demands of the first maxim of Quantity, which requires him to give as much information as is appropriate. But he is doing so not because he is being uncooperative, but because he could not give a fuller answer without contravening the more pressing requirements of Quality. A will appreciate

that B is not answering the question fully because he does not know in which town C lives.

Finally in the list of reasons for failing to fulfil a maxim, Grice outlines the type of situation in which, he explains, conversational implicatures most characteristically occur. Not infrequently, speakers apparently act in direct contravention of the requirements of a maxim, in situations where they are clearly not trying deliberately to deceive, they are not explicitly indicating that they have reasons to opt out, and there is no apparent clash of maxims constraining their behaviour. In such cases, hearers are generally reluctant to dismiss what they hear as surprisingly and disappointingly uncooperative. Rather, confidence in the shared commitment to cooperation is so strong that they will look for an interpretation of what they have heard that is cooperative. That is, while the contribution is blatantly uncooperative at the level of 'what is said', there is a shared commitment to understanding it as cooperative at the level of 'what is implicated'. In such cases the speaker has 'flouted' or 'exploited' the maxim. She has depended on the stability and ubiquitousness of the cooperative principle to signal to the hearer that, when some aspect of it is apparently being disregarded, there must be an implicated meaning which is what she really intends to convey and which is in compliance with the maxim in question.

Any of Grice's maxims are capable of being exploited in order to give rise to an implicature. An example relating to one maxim from each category, all suggested by Grice himself, with suffice for purposes of illustration.

The first maxim of Quantity is exploited in cases where a speaker or writer blatantly gives far less information than would patently be appropriate to the hearer or reader's needs. A letter of reference for a candidate applying for a post in a department of Philosophy runs as follows:

4. Dear Sir, Mr X's command of English is excellent and his attendance at tutorials has been regular. Yours, etc.

It is obvious to the recipient that there is inadequate information about Mr X here, and that this must also have been obvious to the writer. If the writer had anything good to say about Mr X's philosophical abilities it would have been cooperative to do so. The writer has therefore implicated, but has avoided explicitly saying, that Mr X is no good at philosophy.

Many 'figures of speech' result from a flouting of the first maxim of quality; a speaker says something that she cannot possibly believe to be true, thereby implicating something that she does believe to be true. This is the format for the Gricean explanation of metaphor. Metaphors are straightforwardly false at the level of 'what is said'; they may even involve a category error, as in:

5. You are the cream in my coffee.

What is said is that 'you' and 'the cream in my coffee' are identical, which is blatantly false. The implicature is that there is some resemblance or similarity between 'you' and 'the cream in my coffee', and at least in the speaker's opinion this is true.

If a speaker says something that is blatantly irrelevant, thereby flouting the maxim of Relation, she may well be taken to implicate that what has just been said is inappropriate in some way, or that it is time to change to subject. The implicature is likely to be highly relevant in the context, as in the following exchange:

6. [*At a genteel tea party*]

A: Mrs X is an old bag.

B: The weather has been quite delightful this summer, hasn't it?

There is an assortment of maxims of Manner, and therefore a range of ways in which the requirements of Manner may be flouted. In relation to the maxim 'be brief', the reviewer of a concert who might have written 'Miss X sang "Home Sweet Home"' but instead chooses the wordier and more complicated (7) may be taken to have implicated that there was something seriously amiss in Miss X's performance.

7. Miss X produced a series of sounds that correspond closely with the score of "Home Sweet Home".

After presenting a number of such flamboyant examples, Grice concedes that all the implicatures he has discussed so far belong to the subcategory of 'particularised conversational implicatures'; implicatures that rely entirely on the context in which the utterance that carries them is produced. In a different context an utterance such as 'The weather has been quite delightful this summer, hasn't it?' would carry no implicature of inappropriateness or of a need to change the subject. A second class of implicature, 'generalised conversational implicatures' (frequently abbreviated to 'GCIs') are more closely associated with the form of words used. That is, a GCI will arise whenever a particular expression is used, unless there is something in the context to prevent it doing so. This crucial final clause distinguishes GCIs from conventional implicatures. Conventional implicatures occur whenever a word conventionally associated with them is used. GCIs occur by default: that is, in the absence of any contextual reason to prevent them.

GCIs are given relatively short shrift in the second lecture in the 'Logic and Conversation' series. This is despite the fact that, as we will see, they were central to what Grice was trying to achieve in developing his theory of conversation. It is as if for Grice they were, at least temporarily, overshadowed by their more glamorous particularised counterparts. He dwells on just one group of examples, involving implicatures that arise by default from the use of an indefinite article, such as 'a'.

8. X went into a house yesterday and found a tortoise inside the front door.

The speaker of (8) will be understood as communicating the fact that the house in question was not X's own. But this aspect of meaning does not seem to be present every time the indefinite article is used.

9. I have been sitting in a car all morning.

The question of whether it was the speaker's own car or not is left open.

10. I broke a finger yesterday.

The speaker of (10) is most likely to be taken to communicating that it was her own finger she broke, not someone else's.

Grice comments that 'I am inclined to think that one would not lend a sympathetic ear to a philosopher who suggested that there are three senses of the form of expression *an X*' (Grice 1975: 38). These three sense would have to be roughly equivalent to 'just any X', 'an X remotely related to someone in the context' and 'an X closely related to someone in the context'. Rather, it is better to see the very common but not ubiquitous interpretation of 'an X' as meaning 'not some specific X' as a GCI arising by default as a result of the first maxim of Quantity. If the speaker had been in a position to be more informative (for instance to say 'X's house' in example 8) then it would be more cooperative to do so; the implicature is that the speaker is not able to be more specific. Grice admits that it is not an easy task to explain why the default implicature does not arise in examples such as (9) and (10), and suggests that it may be explicable in terms of what would be expected to be the most informative and cooperative contribution in each case. Certainly, what Grice posits as a GCI of the indefinite article in (8) above is not invariant across contexts of use.

Before we turn to a consideration of how the apparatus of Grice's theory of conversation has been applied to philosophical problems, it is worth pausing briefly over some features of conversational implicatures to which Grice draws attention. These features are shared by particularised and generalised implicatures, and distinguish them from conventional implicatures and from 'what is said'. They therefore go some way towards demonstrating what is distinctive about conversational implicature.

Firstly, conversational implicatures are all cancelable. An element of what is communicated that is related to conventional meaning cannot be cancelled without a sense of inconsistency or contradiction. But an element that is conversationally implicated can be explicitly cancelled by what the speaker says next. The speaker of (8) above could go on to add 'in fact it was X's own house' without being accused of contradicting herself. Implicatures can also be cancelled by context. If X is generally known to be the landlord of many different houses, and to spend his days going round visiting these, then the potential implicature that the house was not X's own might be cancelled straight away.

Secondly, re-expressing what is said in different words does not alter what is implicated. As Grice explains it, 'If we call this feature nondetachability, one may expect a generalized conversational implicature that is carried by a familiar, nonspecial locution to have a high degree of nondetachability' (Grice 1975: 39).

All conversational implicatures are also what has become known as 'calculable'. Conventional meaning is something you have access to if you are a competent speaker of the language. But conversational implicature results from a process of calculating or working out, drawing on information from both conventional meaning and the maxims of conversation. Grice did not claim that participants in a conversation go through a deliberate and conscious process of

calculation every time there is a conversational implicature. On the contrary, we often understand implicatures immediately and intuitively. Grice's point is that it must be possible, perhaps after the event, to reconstruct a process of reasoning by which the implicature could have been reached. As he says, 'even if it can in fact be intuitively grasped, unless the intuition is replicable by an argument, the implicature (if present at all) will not count as a conversational implicature; it will be a conventional implicature' (Grice 1975: 31). The ingredients that are input into this process of reasoning, or argument, are: what is said, the maxims of cooperation and the context of utterance, including what has been said before and any relevant aspects of background knowledge. These ingredients are all available, and mutually known to be available, to both the speaker and the hearer. The hearer is therefore able to establish what the speaker knows must be added to literal meaning in order to maintain the assumption that the conversation is a cooperative one. The hearer can be confident that this is what is being implicated.

4 Philosophical Applications

From the outset of his lectures on 'Logic and Conversation', Grice made it clear that he had set himself a two-fold task: 'to determine how any such distinction between meaning and use is to be drawn, and where lie the limits of its philosophical utility' (Grice 1967a: 4). That is, he planned both to develop the theory of conversation itself, and also to consider its application to quite an ambitious range of established philosophical problems. In this section we will consider some of those problems. Some were fully discussed by Grice himself during the course of the lectures on 'Logic and Conversation' or subsequently, while others have been addressed by subsequent commentators.

Grice's style of philosophising was characteristically tentative and open-ended. He was reluctant to claim to have found definitive answers to philosophical problems, and often hesitant even about aligning himself to one particular side in a philosophical debate. What he tended to offer instead were suggestions as to how his philosophy of language might be applicable in certain cases, or how it might be used to defend a particular side in a debate, if anyone were inclined to do so. But behind this rather hesitant presentation are to be found some significant contributions to a range of philosophical topics. These contributions all rely on Grice's serious recognition and formalised account of the distinction between what words conventionally mean and what they may either generally or particularly mean when they are used by speakers. The detailed workings of implicature, as set out in 'Logic and conversation', do not always feature prominently in Grice's responses to various philosophical problems. But they are essential to what he did in applying his theory of conversation.

4.1 *Non-Natural Meaning*

Grice's theory of conversation developed from his own thinking about meaning, and in particular the meaning of natural language expressions used for communicative purposes. One of the philosophical problems it was trying to address was, therefore, a problem of his own making, or at least one that had arisen from critical responses to his earlier attempts to deal with the nature of meaning. His article 'Meaning' had been published in 1957, but had been completed a decade before that. In this, Grice proposes that linguistic meaning can be seen as a form of 'non-natural meaning' (or, to use Grice's own abbreviation, 'meaning_{NN}'). It is non-natural because, unlike for instance the symptoms that might be said to 'mean' a disease, or the smoke that might be said to 'mean' fire, there is no necessary or causal connection between a linguistic expression and what it is used to communicate. Rather, on any communicative occasion there is some agent who might be said to 'mean something by' the linguistic expression used.

The central mechanism involved in 'meaning something by' for Grice was that of intention on the part of the speaker. Speaker intention is part of other philosophical accounts of meaning, including that put forward in speech act theory. But Grice made a bold attempt to make it the central and defining criterion of what is communicated when a particular linguistic expression is used and, by extension, of the actual meaning of the expression in question. In order for communication successfully to be achieved, it was necessary that a speaker intend an utterance to achieve a certain communicative effect, and also that a hearer recognise that intention. If a hearer were to reach an interpretation without realising that the message had been intentionally communicated to him, the situation would not count as one of successful communication.

In Grice's formalisation of this procedure, with 'A' representing a speaker and 'x' an utterance: "A meant_{NN} something by x" is (roughly) equivalent to "A intended the utterance of x to produce some effect in an audience by means of the recognition of this intention" (Grice 1957: 220). He contends that being able to explain what A meant_{NN} by x on any particular occasion may be the starting point for explaining what x means_{NN} 'timelessly'. The psychological concept of intention, which explains what people mean by what they say on particular occasions, is by extrapolation also to explain the meanings of linguistic expressions themselves. His tentative suggestion is that "x means_{NN} (timeless) that so-and-so" might at first shot be equated with some statement or disjunction of statements about what "people" (vague) intend (with qualifications about "recognition") to effect by x' (ibid.). Timeless meanings of expressions exist as entities separate from meanings on any particular occasion of use, but are derived from them and therefore are ultimately dependent on speakers and their intentions.

Grice's subsequent work on meaning attempted to address a number of issues that confronted his psychological account in 'Meaning'. We will consider two of them here: one which Grice was already aware of when he wrote 'Meaning' and which 'Logic and Conversation' could be said to have handled particularly well,

and one that was discussed by various critics among Grice's philosophical contemporaries, that he attempted to address in 'Logic and Conversation', but that remained with him throughout the development of his philosophy.

Firstly, there is the question of exactly how far 'what A meant_{NN} by x on any particular occasion' may differ from 'what x means_{NN} (timeless)'. Grice's contention that the latter must always follow from the former would seem to suggest firstly that the two must be closely related and secondly that whatever a speaker can mean_{NN} in using an expression must have potential implications for what that expression means_{NN}. But Grice strikes a note of caution in relation to these consequences. In the closing paragraphs of 'Meaning' itself he adds a caveat to his account of meaning_{NN}. He suggests that 'only what I may call the primary intention of an utterer' can be seen as relevant to the meaning_{NN} of an utterance (Grice 1957: 221). By extension, only this 'primary intention' can be seen as relevant to questions of what the expression uttered means_{NN} timelessly. As Grice explains, a speaker producing an utterance may have more than one intention: 'if I utter x, intending (with the aid of the recognition of this intention) to induce an effect E, and intend this effect E to lead to a further effect F, then insofar as the occurrence of F is thought to be dependent solely on E, I cannot regard F as in the least dependent on recognition of my intention to induce E' (ibid). Further pieces of information (F) derived in context from pieces of information more directly communicated (E) cannot be expected to be explained in terms of the meaning_{NN} of the expression uttered.

In subsequent work, Grice formulated more precisely the potential differences between particular and timeless meaning. In an article written in the middle of the 1950s, he draws attention to a distinction between 'what a given expression means (in general) or what a particular person means *in general* by a given expression, and (b) what a particular speaker means, or meant, by that expression on a particular occasion' (Grice c. 1953–1958: 167, original emphasis). His point is that these two phenomena may sometimes be quite different from each other. He suggests metaphor and irony as two types of language use where this is apparent. In each case, in order to understand what the speaker has communicated on a particular occasion you need first to know the 'standard' meaning of the words used. Here, then, as in his earlier identification of 'primary' intentions, Grice is acknowledging that not all aspects of what is communicated by using an expression in context can legitimately be seen as relevant to the meaning of the expression itself. His theory of conversation was in part an attempt to address just this issue: to give a systematic explanation of how 'meaning in general' and 'meaning on a particular occasion' may differ and how we can understand the latter based on our knowledge of the former. In terms of the theory of conversation, for instance, Grice's examples of metaphor and of irony would be explained as particularised conversational implicatures derived from a deliberate flouting of the maxims of Quality.

However, the concept of 'meaning in general' remained more elusive and more problematic than the apparent solutions offered by implicature might seem to suggest. A number of contemporary critics of 'Meaning' picked up on the fact that,

unlike for instance the ideal language philosophers, Grice was attempting to offer an account of meaning that made no reference to the linguistic system underlying communication: that is, to the conventions of the language itself. For instance, Searle singled Grice out as an example of a philosopher who ignored the fact that ‘Meaning is more than a matter of intention, it is also at least sometimes a matter of convention’ (Searle 1969: 45). Searle argued that a theory of meaning such as Grice’s should accommodate the fact that the conventional rules of the language itself had a part to play in constraining what speakers could use their words to achieve.

Grice did incorporate a notion of convention into the development of his account of meaning in ‘Logic and Conversation’. As we have seen, his scanty description of ‘what is said’ is that it is ‘closely related to the conventional meaning of the words (the sentence) [the speaker] has uttered’ (Grice 1975: 25). But he never gave up on the idea that this aspect of meaning should ultimately be explicable in terms of speaker’s communicative intentions. In the later lectures in the ‘Logic and Conversation’ series he attempts to explain how speaker intention is still prior to conventional meaning, of the type that constitutes ‘what is said’, by introducing the specific notion of ‘M-intentions’. M-intentions are a specific type of intentions that are constrained by the conventions of the rules of the linguistic system that speakers are using to communicate. With ‘U’ standing for utterer, ‘x’ for the utterance, ‘S’ for a sentence in the language in question and ‘p’ for the proposition conveyed, “U said that p” can be explained as “U did something x (1) by which U centrally meant that p [and] (2) which is an occurrence of a type S part of the meaning of which is ‘p’” (Grice 1969: 88).

In Grice’s philosophy of language, from ‘Meaning’ through to ‘Logic and Conversation’ and on throughout the rest of his life, speaker intention was ideally to take primary place in explaining meaning, both the meaning conveyed on the particular occasions when expressions are used and the meanings of those expressions themselves. In ‘Logic and Conversation’ Grice modified the nature of the relevant intentions so that they must be directed towards the conventions of a particular linguistic system. But those conventions themselves should be explicable in terms of speaker’s intentions, or perhaps in terms of what he described in ‘Meaning’ as ‘what “people” (vague) “intend”’. The competing claims of intention and convention in explaining meaning were never fully resolved for Grice; he readily admitted this, noting much later in life that the relationship between word meaning and speaker meaning was the single question ‘that has given me most trouble’ (Grice 1987: 349). This is not an idiosyncratic problem personal to Grice, and nor is it startling that he never settled it to his own satisfaction or that of his critics. As Petrus points out ‘the debate between intentionalists and conventionalists goes on to this day’ (Petrus 2010: 16). Grice’s aspiration in this matter was clear and remained fundamentally unchanged from ‘Meaning’ to ‘Logic and Conversation’: that linguistic meaning should be shown ultimately to depend on speaker intention.

4.2 *Idealism, Skepticism and Realism*

During the course of his career, Grice employed his philosophy of language in a number of separate papers that addressed questions concerning our knowledge of the world around us: how it is that we establish knowledge about the world, and how confident we can be in that knowledge. Grice generally described this topic in terms of the debate surrounding ‘skepticalism’, the epistemological stance that we cannot know for sure the existence of physical objects. However, because of the previous philosophical works that he addressed, his contributions are also relevant to the closely related debate between idealism, the philosophy that all existence is ultimately psychological or spiritual, and realism, the opposing claim that we can be confident in the existence of material objects external to ourselves and our own consciousness.

Grice’s intervention in this debate is difficult to summarise and difficult also to characterise as support for a single epistemological stance. It is clear from Grice’s writings in this area as a whole that he was more inclined to support a realist than a skeptical point of view, but he does not explicitly endorse realism.³ He does use his ideas about language to support a realist position but also, sometimes in the same article, he proposes to see what they could offer to the alternative point of view. He gives the impression of wanting to play fair philosophically—to let skepticism have a chance to see what it can do with his new tool. Much of his writing in this area is from relatively early in his career, after he had begun to take seriously the distinction between meaning in general and meaning on a particular occasion of use, but before the lectures on ‘Logic and Conversation’. It is almost as if Grice is trying his philosophy of language on for size, or putting it through its paces in relation to various theoretical positions to find out what, potentially, it could offer to philosophy.

The literature relevant to the philosophical debate concerning the certainty of our knowledge of the external world is vast, stretching at least as far back as Plato and Aristotle. But two contributions are particularly germane to an understanding of the ways in which Grice intervened in the debate. These are G. E. Moore’s response to idealism, and the commentary on this by Norman Malcolm. Moore belonged to the philosophical generation previous to Grice. He was a near contemporary of Russell, and philosophically active from the beginning through to the middle of the twentieth century. Moore was keen to refute the idealism that had dominated Continental and British philosophy at the end of the nineteenth century. In this he was endorsing what he himself called a ‘common sense view of the world’, namely that we experience material objects precisely because those material objects exist, independent of us and of our ability to experience them. Probably his most famous defence of the common sense view of the world was in a paper presented to the British Academy in 1939, during which he held up his two

³ In this he could be seen as being in agreement with Austin, who said of realism that: ‘This doctrine would be no less scholastic and erroneous than its antithesis’ (Austin 1962b: 3–4).

hands as proof of the existence of two human hands and therefore, by extension, of material objects. In fact, Moore did not rely on the production of his hands alone, but on his ability to gesture using first his right and then his left hand while saying 'here is one hand' and then 'here is another' (Moore 1939: 146). His own certain knowledge of the truth of what he was saying could be used as premise in an argument the conclusion of which was that two human hands currently exist. The certainty of that knowledge is, he admits, not something that he can prove, but something he is nevertheless prepared to state categorically: 'How absurd it would be to suggest that I did not know it, but only believed it, and that perhaps it was not the case!'. In a slightly earlier paper Moore had argued that it was not the proper business of philosophy to offer proof of common sense statements such as these, but instead to explain how they should be analysed: that is, to explain the real meaning of such statements, or how we know them to be true. The problem in providing such an analysis, which Moore admits is a severe one, can be explained in terms of the 'gap' between the material objects we describe with confidence and the actual experience on which we base those descriptions. What we experience, as philosophers had noted, are a series of 'sense data', or individual impressions that we receive through our senses concerning colour, shape, texture and so on. Moore suggests that the proper analysis of the proposition 'this is a human hand' would be something along the lines of: 'There is a thing, and only one thing, of which it is both true that it is a human hand and that *this surface* is a part of its surface', although he does not suggest that such an analysis goes very far towards solving the problems posed by the distinction, and the relationship, between sense data and actual material objects (Moore 1925: 55).

Moore made many references to the ways in which we ordinarily describe the world, and endorsed the validity of those descriptions. Norman Malcolm went further, arguing that Moore's implicit method of arguing against philosophical statements with which he was not in agreement was to demonstrate that these statements were incompatible with ordinary language. According to Malcolm, Moore appeals 'to our language-sense; to make us feel how queer and wrong it would be to say, when we sat in a room seeing and touching chairs, that we believed there were chairs but did not know it for certain, or that it was only highly probably that there were chairs.' (Malcolm 1942: 354). Malcolm attributed to Moore a faith in ordinary language on the grounds that: 'no ordinary expression is self-contradictory' (Malcolm 1942: 359).

Grice's paper 'Common sense and skepticism' was based on classes he taught in Oxford, and was not published until it appeared in *Studies in the Way of Words* in 1989. At that time, Grice dated it as 'c1946-1950', putting it very much contemporary with his original formulation of his ideas on non-natural meaning. He addresses Malcolm's argument against skepticism, namely that the statements of skepticism are inconsistent with ordinary language, and that ordinary language is never self-contradictory or logically absurd. Grice proposes to offer his services in constructing a defence for the skeptic against this argument. He is keen to point out that he is not himself a skeptic, but that 'I do think that the skeptical position is liable to be somewhat cavalierly treated as hopeless' (Grice c. 1946-1950: 149).

As a champion of the skeptic, Grice challenges the claim, which he attributes to Malcolm, that a self-contradictory statement is equivalent to something that we would never say. For instance, ‘It would no doubt be possible to fill in the gaps in “The—archbishop fell down the—stairs and bumped—like—,” with such a combination of indecencies and blasphemies that no one would ever use such an expression’ (Grice c. 1946–1950: 150). But the expression itself is not self-contradictory. Grice also challenges the related claim that self-contradictory statements are never used descriptively. The example he suggests here is: ‘Might I not, as a result of miscalculation, say “there are eight lots here, each containing eight eggs; so there are sixty-two eggs”’. A little later, as he developed his theory of conversation, Grice would be in a position to present more compelling evidence in favour of the claim that we do sometimes say things that are simply self-contradictory. Malcolm had suggested that if on some occasion we feel a light mist falling and are asked whether it is raining: ‘we might reply “Well, it is and it isn’t”. We should not say that the phrase, used in this connection, is self-contradictory’ (Malcolm 1942: 359). A later Gricean response to this would be that the phrase certainly is self-contradictory, at the level of what is said. But saying something self-contradictory is not incompatible with, may in fact be a trigger for, implicating something perfectly acceptable.

‘G. E. Moore and philosopher’s paradoxes’ also remained unpublished until 1989. Grice dated this as ‘c. 1953–1958’, putting it at the time in which he was developing his ideas of the differences between literal and intended meaning, although before the introduction into his philosophy of the term ‘implicature’. Grice again cautions against too easy a dismissal of the skeptic’s position on the ground that it goes against ordinary language. He does not give up on the idea of using ordinary language to defeat the skeptic, but he does so by looking in more detail than Malcolm had at ways in which language is used. He draws attention to the distinction between what an expression means in general and what a speaker may mean by that expression on a particular occasion. His examples are of metaphor and of irony. “I have run out of fuel” means something different in general from what a particular speaker might use it to mean. “Jones is a fine fellow” means in general that Jones has many fine qualities, but may be used ironically to mean that Jones is a scoundrel. In such cases, ‘a proper understanding of what he says involves knowing the standard use of the sentence in question’ (Grice c. 1953–1958: 168). The skeptic who explains away everyday statements about material objects by saying that people do sometimes just say things that are self-contradictory ‘is accepting just such a divorce between the *general* meaning of a sentence and its *particular* meaning on particular occasions as that which I have been maintaining to be inadmissible’ (Grice c. 1953–1958: 169). Grice has replaced what he sees as Malcolm’s unsatisfactory argument against skepticism based on the unexamined claim that ordinary language is never self-contradictory, with a more subtle argument that takes account of ways in which meaning in particular may diverge from meaning in general, and of the limits of such divergencies.

Grice discusses the viability of the philosophical term ‘sense data’ in his 1961 article ‘The causal theory of perception’. According to the causal theory of

perception, what we have direct access to are individual sense data, but we have reason to believe that these are caused by objects in the material world. As we have seen, Moore drew on the notion of sense data when he attempted to describe how philosophy was to offer an analysis of the statements about the external world that we can be said to know with confidence. As Grice points out sense data are also necessary to skeptics, if they are to claim that, contra the causal theory of perception, we have no reason to believe that anything external to our own minds is causing these sense data. Grice reiterates the view that: ‘I think a skeptic might complain that though his worries may well prove dissoluble, he ought at least to be able to state them’ (Grice 1961: 229).

In the years since Moore had relied on sense data to explain the real meaning of statements about the external world, the term had rather fallen out of philosophical favour. This was in part due to Austin, who derided ‘sense data’, along with other technical philosophical terms, on the grounds that they play no part in ordinary language.⁴ Grice notes that those who wish to defend sense data generally do so by drawing attention to ‘such supposedly standard locutions as “So-and-so looks Φ (e.g., blue) to me,” “It looks (feels) to me as if there were a Φ so-and-so,” “I seem to see something Φ ”, and so on’ (Grice 1961: 226). Defenders of sense data, Grice implies, propose that the correct analysis of such expressions is “I am having a Φ sense datum”. But once again, Grice points out that arguments that rely on ‘what we say’ are not always as straightforward as they seem. It is true that such expressions are in use, but not true that they could be used appropriately in every situation in which we describe our perceptions of the world. In fact, such turns of phrase seem appropriate only in situations where there is reason to believe that there is some doubt, or some reason to believe that someone may want to deny, that the object has the property described. That is, the expressions can legitimately be used only in situations where the D-or-D (doubt or denial) condition applies. As Grice persuasively points out: ‘there would be something at least *prima facie* odd about my saying “That looks red to me” (not as a joke) when I am confronted by a British pillar-box in normal daylight at a range of a few feet’ (Grice 1961: 227).

The necessity of the D-or-D condition is a potential problem for the sense data theorist, who needs to be able to claim that sense data are involved whenever objects are perceived. But if an expression such as “So-and-so looks Φ (e.g., blue) to me” can be applied only in very specific circumstances, and fails to make a statement that is either true or false on other occasions, then it is not possible to argue that expressions describing perception can routinely be analysed as making reference to sense data. Grice proposes an alternative explanation, one that would rescue the sense data theorist. He proposes that expressions such as “So-and-so looks Φ (e.g., blue) to me,” are always true, and therefore technically at least applicable, when we are recounting an experience of perception. But in cases where the D-or-D condition does not hold they are nevertheless very misleading, because they introduce what Grice calls an ‘implication’ that there is some doubt

⁴ Especially in Austin (1962b).

or potential denial. This is an explanation that Grice goes on to defend, partly because it is ‘in line with the kind of thing I was inclined to say about other linguistic phenomena which are in some degree comparable’ (Grice 1961: 229). In the light of Grice’s subsequent work, we can of course recognise the D-or-D condition as a conversational implicature of the use of an expression such as “So-and-so looks Φ (e.g., blue) to me”. Such an expression is much longer and more complex than the possible alternative “So-and-so is Φ ”, so following expectations both of Quantity and of Manner it carries the implicature that the alternative does not apply, or cannot straightforwardly be used. Saying “So-and-so looks Φ (e.g., blue) to me” in a context where the D-or-D condition was not fulfilled does not after all constitute a misuse of the expression itself, although it does constitute a serious misuse of language in general. The sense data theorist is able to maintain that sense data are potential components of the analysis of every report of perceptual experience.

4.3 *Logic and Natural Language*

As successive generations of students of linguistics have discovered to their consternation, ‘Logic and conversation’ begins with a string of logical symbols. This is because Grice’s theory of conversation is intimately connected with his interests in classical logic. In the introduction to the William James lectures, Grice describes logic as ‘an area of special interest to me’ (Grice 1967a: 8). Given this, it is perhaps surprising that Grice has comparatively little to say about the specifics of logic. Having introduced that string of logical symbols at the start of ‘Logic and conversation’, he does not return to them in that lecture. One later lecture in the series, ‘Indicative conditionals’ is devoted to some, but only some, of the relevant elements of logic. But the contribution that Grice’s theory of conversation makes to logic, and particularly to our understanding of the relationship between logic and natural language, is significant. Its success in this area is in fact one of the major triumphs of a theory that is much more often discussed in relation to the intricacies of everyday interaction.

At the start of ‘Logic and conversation’ Grice contrasts ‘the formal devices - \sim , \wedge , \vee , \supset , $(\forall x)$, $(\exists x)$, (ιx) (when these are given a standard two-valued interpretation)’ and ‘what are taken to be their analogues or counterparts in natural language—such expressions as *not*, *and*, *or*, *if*, *all*, *some (or at least one)*, *the*’ (Grice 1975: 22). According to the ‘standard two-valued interpretation’ of the logical devices, their meaning can be explained in terms of their effect on the truth or falsity of the propositions in which they occur. But these truth-functional properties do not seem to be straightforwardly transferable to expressions of natural language involving their supposed counterparts. Grice draws attention to two opposing philosophical positions on this, which he labels those of ‘formalist’ and ‘informalist’ philosophers. The formalists ascribe the discrepancy between natural language and logic to ‘an imperfection of natural languages’, and take it as an

indication that philosophers should concentrate on ideal, artificial language instead (Grice 1975: 23). The formalists argue that natural language should not be judged against the standards of logic, because it simply is a different system that does different things, and should be respected in its own right. Grice's daring claim is that neither the formalist nor the informalist position is defensible simply because they are both based on a shared erroneous assumption. The natural language expressions in question do not in fact differ from their logical counterparts. Grice's solution will be to explain the apparent differences as conversational implicatures associated with the expressions in use, enabling him to retain a logical account of these expressions at the level of 'what is said'.

Before reviewing Grice's treatment of one of these expressions, and considering how the others might be similarly explained, it is worth identifying briefly where the supposed differences between the logical and the natural language expressions lie. Natural language 'not' might appear to be fairly close in meaning to logical negation, symbolised by \sim . Logically, negating a true proposition must result in a false proposition, and negating a false proposition must result in a true proposition. In natural language too, if we take a true proposition, such as (11), and add 'not' we get a false proposition, (12):

11. London is the capital city of England.
12. London is not the capital city of England.

But there is a problem. Since a positive proposition and its negative counterpart will always have different truth values, they should be incompatible with each other; in effect, combining p with $\sim p$ should always lead to a statement that is self-contradictory, and therefore simply false. But we do sometimes use natural language expressions that apparently correspond to this logical pattern, and use them successfully. In the last section we considered an example suggested by Norman Malcolm where someone asks whether it is raining or not. It would be perfectly possible to reply using (13), although logic would seem to suggest that this should be just absurdly false:

13. It is and it isn't.

Logical conjunction, symbolised by \wedge , might seem to have a lot in common with natural language 'and'. The logical conjunction of two propositions has a determinate effect on the truth or falsity of the resultant complex proposition. If two true propositions are combined together, the resultant complex proposition will be true, but if one or both of the propositions are false the result will be a false complex proposition. But there are features of natural language 'and' which are just not shared by logical conjunction. As Grice himself explains in the introduction to the William James lectures, 'it has been suggested that because it would be incorrect or inappropriate to say "He got into bed and took off his trousers" of a man who first took off his trousers and then got into bed, it is part of the meaning, or part of one meaning, of "and" to convey temporal succession' (Grice 1967a: 8).

Philosophers had for a long time noticed a potential ambiguity in the meaning of the word 'or', an ambiguity which could not be tolerated in logic and was

therefore not to be attributed to the relationship of logical disjunction, represented by \vee . 'Or' is very often used exclusively: that is to express the meaning that one or other of two alternatives must be the case, but that they cannot both be the case. Many ordinary uses of 'or', such as that in (14) below, have this interpretation.

14. The dog is in its kennel or it is in the kitchen.

But sometimes 'or' is used in a way that is quite compatible with the truth of both the apparent alternatives, as in example (15).

15. When you are eighteen you can drink in a bar or you can get married.

Philosophers had, in general, noted that this sort of 'inclusive' use of 'or' is equivalent to logical disjunction, but that despite this 'or' is often used in everyday interaction with the exclusive meaning.

Natural language 'if...then' is in many ways similar to the logical relation of implication, $p \supset q$. But whereas logical implication can be wholly defined in terms of truth conditions, the same does not seem to hold for 'if...then'. Logical implication specifies that an expression $p \supset q$ is true if both 'p' and 'q' are true, or in any situation in which 'p' is false. But such an account cannot be extended easily to 'if...then'. Examples such as (16) below seem to convey not just the truth relations between the two parts, but also some sense of causation or consequence. For this same reason (17) seems distinctly odd or problematic, despite the fact that it contains two true simple propositions, so that logically as a whole it should be unproblematically true.

16. If it is raining then we will hold the party indoors.

17. If London is the capital city of England then tigers are carnivores.

Grice discusses examples of this last type in some detail in relation to conversational implicature, although he proposes his analysis of them in a very tentative and circumspect manner. In 'Indicative conditionals' he responds to some of the things that previous philosophers have said about 'if...then' expressions. One such philosopher was Grice's then Oxford colleague P. F. Strawson, the ordinary language philosopher who had challenged Russell's account of definite descriptions on the ground that 'ordinary language has no exact logic'. Grice does not mention Strawson by name in 'Indicative conditionals', but he had studied his work on logic closely, and had recently lectured on his *Introduction to Logical Theory*. In this book, Strawson pays close attention to the expressions of ordinary language, and to what he sees as their divergences from logic. He makes a series of relevant observations, for instance that 'They got married and had a child' is not equivalent to 'They had a child and got married' and that 'in certain verbal contexts, "either...or..." plainly carries the implication "and not both...and..."', whereas in other contexts, it does not' (Strawson 1952: 80 and 92). In relation to 'if...then' expressions, he notices that logical implication does seem to contribute to their interpretation, but not fully to explain it; 'a statement of the form " $p \supset q$ " does not entail the corresponding statement of the form "if p then q"', but 'a statement of the form "if p then q" does entail the corresponding statement of the

form “ $p \supset q$ ” (Strawson 1952: 83). The extra component of meaning of ‘if p then q’, not shared by ‘ $p \supset q$ ’, includes the sense that a ‘sound or reasonable step’ in reasoning would take us from p to q.

For Strawson, although the logical expressions had much in common with their supposed natural language counterparts, they could not be seen as equivalent. He is rather imprecise about the exact nature of the differences between the two, sometimes describing them as an ‘implication’ and sometimes as a ‘logical’ difference (Strawson 1952: 88 and 80). In the introduction to his William James lectures, Grice accuses Strawson of unclarity:

It is difficult to determine whether he thought that such a conditional as ‘If I am now in Oxford, it is raining in Australia’ (which, read materially, would be true, though the required connection between antecedent and consequent is presumably not present) is false, or is inappropriate (out of order, lacking an assignable truth-value). (Grice 1967a: 12)

In ‘Indicative conditionals’ Grice puts forward his own answer to the question. A conditional such as that about Oxford and Australia above is logically perfectly true. Its oddness can be explained in terms of a conversational implicature of a particular type of relationship between antecedent and consequent, which Grice labels the ‘Indirectness Condition’. He points out that, if this aspect of meaning were to be accepted as a conversational implicature, it would pass the required tests by being nondetachable and cancelable. Example (18) below introduces the Indirectness Condition, but so too does paraphrases of it such as (19) and (20).

- 18. If Smith is in London, he is attending the meeting.
- 19. Either Smith is not in London, or he is attending the meeting.
- 20. It is not the case that Smith is both in London and not attending the meeting.

The Indirectness Condition can be explicitly canceled, as in (21):

- 21. I know just where Smith is and what he is doing, but all I will tell you is that if he is in the library he is working.

It can be canceled by context in an example such as (22), spoken in the context of a game of bridge:

- 22. If I have a red king, I also have a black king.

Grice’s proposal is that, in cases in which it is not canceled, the Indirectness Condition arises as a conversational implicature as a result of a clash of the requirements of Quantity and of Quality. A straightforward denial of the proposition ‘p’ or a straightforward assertion of the proposition ‘q’ would be more informative than a statement of the form ‘if p then q’. If the speaker were in a position to make either of these alternative statements, it would be more cooperative to do so. ‘An infringement of the first maxim of Quantity, given the assumption that the principle of conversational helpfulness is being observed, is most naturally explained by the supposition of a clash with the second maxim of Quality (“Have adequate evidence for what you say”), so it is natural to assume

that the speaker regards himself as having evidence only for the less informative statement' (Grice 1967b: 61–62).⁵

Grice himself went little further with the task he had apparently set himself of revealing the error of both formalist and informalist philosophers by showing that the various natural language expressions were in fact equivalent in terms of 'what is said' to their logical counterparts, however much they may sometimes differ in terms of 'what is implicated'. Subsequent work by linguists in the Gricean tradition has done much to demonstrate how the project could be taken further.⁶ A quick return to earlier examples can show how this could be the case. Combining a positive proposition with its own negation is indeed self-contradictory; 'it is and it isn't' is simply necessarily false. But anyone saying something so blatantly false will be taken to be flouting the first maxim of Quality; the hearer will be reluctant to understand what he hears as simply false and therefore uncooperative, but will instead seek an explanation at the level of 'what is implicated' that will maintain the assumption of cooperation. What is implicated is something such as that there is precipitation but it is so slight that in the speaker's opinion it could hardly legitimately be labelled 'rain'. Grice's example concerning the man getting into bed and taking his trousers off is literally true if both events occurred. The sense of oddness or strangeness in our response to it is a result of an implicature of sequence dependent on the category of Manner, particularly the maxim 'be orderly'. We expect, other things being equal, that speakers will relate events in the chronological order in which they occurred. Natural language 'or' could be associated unambiguously with the logical meaning of inclusive disjunction. The fact that it is often used in everyday interaction with an exclusive meaning can be explained in terms of a conversational implicature based on expectations of Quantity. 'p and q' is a lot more informative, if it is compatible with truth, than 'p or q'. The hearer infers that, since the speaker has chosen the latter and not the former, she is implicating that 'p and q' does not apply.

Grice was suggesting conversational implicature as a device for sustaining the system of logical constants as the definitive account of the meanings of the problematic natural language expressions. It enabled him to sweep aside the confrontation between the formalists and the informalists, on the grounds that there really was no disparity between logic and natural language to be dealt with. Philosophers on both sides had made the mistake of taking some common features of the use of a set of natural language expressions and assuming that these must be features of their meaning. Grice's solution also allowed him to resist the need to admit that such everyday expressions of natural language were simply multiply ambiguous, with a number of different meanings apparent on different occasions of use. Without an account in terms of implicature, 'and' simply meant 'and then' on

⁵ 'The principle of conversational helpfulness' was a term used by Grice in early version of his theory of conversation. He later changed it to 'the cooperative principle', but the earlier term has survived in the published version of this lecture.

⁶ See, for instance McCawley (1981), Levinson (1983) and Horn (1989).

some occasions and ‘or’ simply meant ‘not both’ on some occasions, but not on others. Grice was able to retain a simple semantic system where such items were assigned a single meaning, with local difference explained in terms of principles of language use that were outside of the lexicon itself. That is, he was able to maintain the principle that he labelled ‘Modified Occam’s Razor’: ‘Senses are not to be multiplied beyond necessity’ (Grice 1978: 47). For Grice, logic is not irrelevant to explaining natural language; indeed in many cases it explains ‘what is said’. It is simply that logic on its own cannot be expected to explain every aspect of what is communicated between speakers in context. Other principles of a different order are required.

4.4 Definite Descriptions

A short while after his ‘Logic and Conversation’ series, in one of a set of lectures delivered in 1970, Grice offered his own contribution to the debate between Russell and Strawson on definite descriptions. As discussed in [Sect. 2](#), Russell had argued that the apparent grammatical structure of an example such as ‘the king of France is bald’ was misleading as to its logical structure. It was not actually of subject/predicate form at all. Rather, its meaning was a set of propositions relating to existence, uniqueness and baldness; in the absence of a current king of France it was simply false. Strawson argued that in ignoring how people would actually respond to such an example in use, Russell had valued classical logic, with its division of all propositions into the ‘true’ and the ‘false’, over the realities of natural language. The existence of the king of France was not a logical entailment of the example for Strawson but a presupposition, a necessary condition for it to be legitimately used at all. In the absence of a king of France Russell’s example was not false, but failed to be either true or false.

The published version of Grice’s response to the debate between Russell and Strawson, which did not appear until 1981, was entitled ‘Presupposition and conversational implicature’. It is very tentative in its presentation, but there is no doubt that Grice intends his contribution to rank on Russell’s side. In some ways this might appear to be a surprising allegiance. Grice was generally identified with ‘ordinary’ rather than the ‘ideal’ language philosophy, and his work was characterised by a regard for the ways in which people use language in actual communicative situations, rather than for formal or artificial language. But in fact Grice’s defence of Russell was very much in line with what he had done elsewhere in his philosophy, particularly in his use of conversational implicature. In defended classical bivalent logic he was doing something very similar to what he had done in relation to the apparent gap between logical constants and their natural language equivalents. He was maintaining that logic did in fact relate closely to conventional meaning or to ‘what is said’, while denying that this meant it could tell us everything about how expressions will be used and understood in context. Implicitly, he was criticising Strawson for noticing features of how certain

expressions are generally used, and assuming that these features must play a part in describing the literal meaning of those expressions.

So Grice proposed to adopt Russell's analysis of 'the king of France is bald' as consisting logically of three separate propositions, which Grice identified as 'A: There is at least one king of France', 'B: There is at most one king of France' and 'C: Whatever is king of France is bald' (Grice 1981: 273). Russell's original example, 'the king of France is bald', he labels 'D'. So on the interpretation Grice is advocating, 'D' is logically equivalent in meaning to 'ABC'. Grice acknowledges one of Strawson's chief objections to Russell's analysis: that the existence of the king of France, an apparent logical entailment of D, seems to be shared by the negation of D, that is by 'the king of France is not bald'. A negative statement would not normally be expected to share the entailments of its positive counterpart. But Grice differs from Strawson in claiming that, while the positive and the negative counterparts do both convey the idea that there is a king of France, they do so in rather different ways. For Strawson, the existence of the King of France is a presupposition of both the positive and the negative versions. For Grice, the commitment to the existence of the king of France is part of the logical meaning of 'the king of France is bald', as Russell claimed, but it is a conversational implicature of the use of 'the king of France is not bald'.

Before describing how the existence of the king of France can be explained as a conversational implicature of 'the king of France is not bald', Grice advances some evidence as to why such an explanation can be justified. He turns to some of the general properties or characteristics of conversational implicatures that he had identified. Firstly, he claims that the supposed conversational implicature can be explicitly cancelled. 'If I come on a group of people arguing about whether the king of France is bald, it is not linguistically improper for me to say that the king of France is not bald, since there is no king of France' (Grice 1981: 270).⁷ He claims that the supposed implicature can also be canceled by context. Imagine there is an ongoing dispute about whether or not a government 'loyalty examiner' exists. If Grice himself were known to doubt the existence of such a person, he 'could perfectly well say to a plainly loyal person, *Well, the loyalty examiner will not be summoning you at any rate*, without, I would think, being taken to imply that such a person exists' (Grice 1981: 271). Furthermore the supposed implicature cannot be detached by reexpressing the example in different words, provided that the definite description 'the king of France' is retained. 'It is not the case that the king of France is bald', 'It is false that the king of France is bald' and 'It is not true that the king of France is bald' all retain the alleged implicature.

In order to understand Grice's argument in favour of the analysis of 'the king of France is not bald' in terms of conversational implicature, we need to start with his

⁷ In some later discussion of such examples, it has been claimed that saying this would in fact involve the speaker in a logical contradiction if the negation is interpreted logically, but not if it is interpreted as 'metalinguistic', registering an objection to the form of words used because of the problematic presupposition they introduce (see for instance Horn 1985). But Grice interprets his example as the explicit cancelation of an implicature, giving rise to no logical contradiction.

account of the meaning and the use of the positive version, ‘the king of France is bald’, or ‘D’. In the Russellian picture that Grice is supporting, ‘D’ is equivalent in literal meaning to ‘ABC’. This therefore raises the question of why in a particular context a speaker would choose the form ‘D’, rather than explicitly spelling the commitments out in the form ‘ABC’. Grice argues that the difference between these two alternatives can be understood in terms of the manner in which the information is presented. In effect, in the form ‘D’, the existence of the king of France is part of what is literally said, but it is not presented as something that is likely to be subjected to denial. Grice proposes an addition to his category of Manner: the maxim ‘Frame whatever you say in the form most suitable for any reply that would be regarded as appropriate’ (Grice 1981: 273). The denial of some part of what is said is one significant way in which someone might reply. The difference between saying ‘D’, then, and saying ‘ABC’ is not in the logical content of what the speaker has said, but in the implicature that neither ‘A’ nor ‘B’ are likely to be something the hearer will want to deny. This might be because ‘A’ and ‘B’ can be regarded as common knowledge between speaker and hearer. Or it could simply be that neither ‘A’ nor ‘B’ are likely to be subject to challenge. As Grice points out:

It is quite natural to say to somebody, when we are discussing some concert, ‘My aunt’s cousin went to that concert’, when we know perfectly well that the person we are talking to is very likely not even to know that we have an aunt, let alone know that our aunt has a cousin. So the supposition must be not that it is common knowledge but rather that it is noncontroversial, in the sense that it is something that we would expect the hearer to take from us (if he does not already know). (Grice 1981: 274)

For Grice, Strawson’s reaction that a contemporary use of ‘the king of France is bald’ is problematic is explained by the fact that although part of its literal meaning is that the king of France exists, which is simply false, the form of words chosen strongly implicates that this aspect of meaning is not likely to be disputed. Strawson noticed that ‘the king of France is not bald’ seems to share some of the alleged entailments of its positive counterparts. In Grice’s explanation, these are not in fact entailments of the negative sentence at all but merely conversational implicatures. The hearer of this example can be expected to notice that the speaker is speaking as if she were responding negatively to ‘D’: fulfilling the expectations of a response to ‘D’ or ‘accepting the first two conjuncts and rejecting the third’ (Grice 1981: 276).

In ‘Presupposition and conversational implicature’ Grice’s motivation was to provide support for a classical bivalent logic in the face of intuitive responses to language in use that seemed to argue against it. He was therefore concerned with ‘presupposition’ chiefly as this term was used by Strawson to challenge Russell over the interpretation of his ‘the king of France’ example. However, in the 20 years that had elapsed between the publication of Strawson’s ‘On referring’ and Grice’s presentation of the lecture that was to be published as ‘Presupposition and conversational implicature’, presupposition had been discussed, particularly in the newly emerging discipline of linguistics, as a much more wide-spread phenomena. As Grice himself acknowledged, ‘In recent years, linguists have made it increasingly difficult for philosophers to continue to keep their eyes glued to a

handful of stock examples of (alleged) presupposition, such as the king of France's baldness and the inquiry whether you have left off beating your wife' (Grice 1981: 279). He admitted that not all such cases, including for instance factive verbs, might prove amenable to being explained in terms of conversational implicatures.

Since 1970 the literature on presupposition has expanded even more dramatically. Grice's insights have been used in this literature, but often by those who, unlike Grice himself, endorse a presuppositional account of the data. For instance, Burton-Roberts presents a semantic account of presupposition while allowing for the relevance of pragmatic factors, such as questions as to what pieces of information are and what are not likely to be easily accepted or, in Grice's terms, treated as non-controversial. For Burton-Roberts, presuppositions are 'implicit commitments which, because they are implicit are not (presented as) subject to debate. In this they contrast with the explicit commitments one enters into in making assertion, which are (presented as) subject to debate' (Burton-Roberts 1989: 137). Bezuidenhout suggests that all of Grice's arguments have echoes in subsequent literature on presupposition, and that 'Grice's insights can be recast in a presuppositionalist framework' (Bezuidenhout 2010: 91).⁸ In the face of the wide array of other types of data that appear to behave in a similar way, it seems that a broadly presuppositional account of definite descriptions has proved more successful than the classically logical account advocated by Russell. Nevertheless, Grice's response remains a significant and original contribution to the other side of the dispute, and a striking example of Grice's own application of his notion of conversational implicature to an existing philosophical problem.

5 Pragmatics and "What is Said"

Grice may have been interested chiefly in the application of his theory of conversation to a range of long-standing philosophical problems, but the greatest impact of his theory has been felt in the neighbouring discipline of linguistics. The field of linguistic pragmatics, in particular, owes significant aspects of its focus and its terminology to Grice, and indeed arguably is indebted to Grice's pioneering work for its very existence.⁹ The theory of conversation has been applied, sometimes with and sometimes without modification, to many types of linguistic data, including those drawn from children's language (e.g. Perner and Leekam 1986), from language disorders (e.g. Ahlsén 1993) and from literary texts (e.g. Blake 2002). More theoretical pragmaticists can be characterised in terms of the extent to which they agree with, and where they differ from, Grice's theory.

⁸ Bezuidenhout offers a useful overview of that literature and of its relationship to Grice's observations.

⁹ See, for instance, Asa Kasher's claim that Grice opened a 'gap in the hedge', which first brought to the attention of linguists the idea that they might look for explanations of meaning that went beyond the linguistic system itself (Kasher, cited in Mey 1989: 825).

The Gricean coinage ‘implicature’ is fundamental to work in post-Gricean pragmatics, both in its neo-Gricean forms (e.g. Levinson 2000; Horn 1989, 2007) and, admittedly with some significant differences in application and scope, in relevance theory (Sperber and Wilson 1995; Carston 2002).

A belief common to all varieties of modern pragmatics, indeed a founding assumption of pragmatics as a branch of linguistics, is that not every aspect of what is communicated in context can be explained exclusively with reference to the language itself. The business of theoretical pragmatics could be defined as being that of describing and explaining the regularities and principles of human interaction that are not themselves linguistic in nature but that contribute, often in very significant ways, to how speakers put their meanings into words and how hearers interpret the words addressed to them. In Grice’s original framework, the distinction between those aspects of what is communicated that are contributed by the language and those that rely on other, non-linguistic factors, could be more or less comfortably discussed in terms of the distinction between ‘what is said’ and ‘what is implicated’. Recently, pragmaticists have voiced unease about the viability of an autonomous level of ‘what is said’. Such unease tends to focus on the apparent inability of the linguistic content of an utterance to yield an identifiable propositional meaning. Pragmatists of a variety of different theoretical outlooks have drawn attention to the apparent need for the intrusion of pragmatic principles into the specification of a Gricean notion of ‘what is said’. Pragmatic principles seem to be at work in the processes of reference assignment and disambiguation that Grice allowed occur in reaching ‘what is said’. They are at work also in various other ways in which linguistic content is filled out, including processes such as those which Recanati (2004) has described as ‘enrichment’ as in (23), ‘loosening’ as in (24) and ‘semantic transfer’ as in (25).

23. Mary took out her key and opened the door (where we assume that Mary opened the door with the key).
24. The ATM swallowed my credit card (where the usual requirement that the subject of the verb ‘swallowed’ be animate is relaxed).
25. The ham sandwich left without paying (where ‘the ham sandwich’ is taken to refer to the customer who ordered the ham sandwich).

Pragmaticists who have identified a need for pragmatic intrusion have responded to it in different ways. Levinson (2000), who has concentrated largely on pragmatic input into the resolution of anaphoric reference, argues that Grice seriously overestimated the power and coherence of conventional meaning in specifying ‘what is said’ and that generalised implicatures play a role in establishing even a basic truth-evaluable meaning. Relevance theorists such as Sperber and Wilson (1995) dispense with the notion of ‘what is said’ and argue instead that incomplete linguistic form needs to be enriched in context in line with the pragmatic principle of relevance before an explicit propositional content, or an ‘explicature’ can be identified. The contextualist Recanati (2004) retains a concept of ‘what is said’, but gives it a different definition such that it includes information dependent on context and determined by various pragmatic principles, in addition to linguistic meaning.

Grice's distinction between 'what is said' and 'what is implicated' has been challenged by experimental results, too. Some, but by no means all, experimental pragmaticists have argued that laboratory evidence does not support the Gricean picture in which 'what is said' is input into processing in relation to the conversational maxims in order to produce 'what is implicated'. The evidence in question is often based on some measure of processing time. In the case of supposed GCIs experimentalists have argued that there is no evidence that implicatures such as those from 'or' to 'not and', for instance, arises by default. If this were the case, then an interpretation that included the implicature should be quicker to reach, but 'reading time was significantly longer in the condition where the implicature is warranted by the context, indicating that implicatures are generated only in such cases' (Breheny et al. 2006: 445; see also Bezuidenhout and Morris 2004). In the case of figures of speech such as metaphors, it is argued that subjects appear to access the non-literal meaning straightaway, without having to go through a process of recognising a literal 'what is said' and then deriving the intended meaning from this on the basis of expectations of cooperation: 'People can read figurative utterances... as quickly, sometimes even more quickly, as literal uses of the same expressions in different contexts or equivalent nonfigurative expressions' (Gibbs 1999: 468; see also Glucksberg 2004).¹⁰

Pragmatics has not, however, entirely dispensed with Grice's conception of 'what is said'. Pragmaticists who defend some modified version of 'what is said' are those who are prepared to tolerate its having a minimal content. That is, they allow that 'what is said' may be radically underspecified with respect to what is communicated in context, but argue that this in itself does not make it unviable as an autonomous level of meaning. Levinson, Sperber and Wilson and Recanati all, in different ways and working within different frameworks, argue that a Gricean 'what is said' is unviable because it is not sufficient to yield an identifiable, truth-evaluable proposition. Those who argue that 'what is said' can survive being fragmentary, incomplete or not fully propositional include Borg (2004) who advocates 'minimal semantics', Cappelen and Lepore (2005) who argue for 'semantic minimalism', and Horn (2010) who argues for the adequacy of a 'minimally modified Gricean model'.¹¹

Pragmaticists who support a minimal but autonomous version of 'what is said' generally take issue with the idea that it must have some degree of psychological plausibility or interpretative viability. In this they are opposed to both the claims of theorists such as relevance theorists who argue that linguistic decoding alone gives a form that is too fragmentary to be recognised as what was explicitly communicated, and also the empirical claims of experimentalists that 'what is said' is untenable because it is not reflected in most likely or earliest interpretations. Supporters of a

¹⁰ Experimental work is predominantly, but not exclusively, opposed to Gricean analysis. For an example of experimental work that concludes in favour of Grice see for instance Winer et al. (2001).

¹¹ See also Martinich (2010).

basically Gricean ‘what is said’ point out that it was never intended to do the whole job of explaining what is communicated: ‘The most natural interpretation of an utterance need not be what is strictly and literally said’ (Bach 2001: 30). They also argue that Grice himself was aiming for a purely formal model of ‘what is said’: ‘one on which both speaker and audience may be wrong about what is said’ (Saul 2002: 347). It is certainly the case that Grice was aiming primarily for a theory of communication that was consistent with logic, and that he argued that, although an implicature must be calculable, it is not necessary for the process of calculation to be psychologically accessible to the interpreter. But on this last point he was equivocal, as the following extended worry about ‘what is said’ demonstrates:

We must of course give due (but not undue) weight to intuitions about the existence or nonexistence of putative senses of a word (how could we do without them?). Indeed if the scheme which I have been proposing is even proceeding in the right direction, at least some reliance must be placed on such intuitions. For in order that a nonconventional implicature should be present in a given case, my account requires that a speaker shall be able to utilize the conventional meaning of a sentence. If nonconventional implicature is built on what is said, if what is said is closely related to the conventional force of the words used, and if the presence of the implicature depends on the intentions of the speaker, or at least on his assumptions, with regard to the possibility of the nature of the implicature being worked out, then it would appear that the speaker must (in some sense or other of the word *know*) know what is the conventional force of the words which he is using. This indeed seems to lead to a sort of paradox: If we, as speakers, have the requisite knowledge of the conventional meaning of sentences we employ to implicate, when uttering them, something the implication of which depends on the conventional meaning in question, how can we, as theorists, have difficulty with respect to just those cases in deciding where conventional meaning ends and implicature begins? If it is true, for example, that one who says that *A or B* implicates the existence of non-truth-functional grounds for *A* or *B*, how can there be any doubt whether the word “or” has a strong or weak sense? I hope that I can provide the answer to this question, but I am not certain that I can. (Grice 1978: 49)

6 Conclusions

Grice admired philosophers who rather than simply identifying what should be done in relation to a particular philosophical question actually made an attempt at answering it. He once commented that, despite his many differences with Austin in terms of temperament, philosophical outlook and methodology, he approved of Austin’s response to the logical positivists that, in effect, if language was as important as they claimed to philosophy, then language should be given proper philosophical scrutiny. He paraphrased Austin as saying: ‘if you say that you are going to do something, then DO it’.¹² Grice was by no means the first philosopher

¹² Tape, 29 January 1983, Grice in conversation with Richard Warner and Judith Baker, H. P. Grice Papers, BANC MSS 90/135c, The Bancroft Library, University of California, Berkeley.

to notice that there may be some significant differences between what our words literally mean and when we mean, or what we communicate, when we use those words. But rather than simply observing that there were such differences, he waded in and had a go at explaining what they were and how they were effected. The range of applications in philosophy and also in linguistics to which Grice's theory of conversation has been put is testament to its success. As seems almost inevitable with any such ambitious project, it also has its problems. The various revisions, adjustments and alterations proposed in post-Gricean pragmatics indicate that even those who have shared Grice's basic conception of what is needed from a theory of communication have had reservations about the details of what Grice proposed. But perhaps the single biggest problem for Grice, the problem of the demarcation and indeed the viability of his distinction between 'what is said' and 'what is implicated', was one of which he was always aware. For Grice, the concept of 'what is said' was a necessary but an uncomfortable inclusion in his theory, which distanced his philosophy of language from the more austere account in his earlier work of meaning as being ultimately dependent on psychology. But this was a necessary concession in order to make possible the theory of conversation, and all that it could offer to philosophers and subsequently to linguists.

References

- Ahlsén, Elisabeth. 1993. Conversational principles and aphasic communication. *Journal of Pragmatics* 19: 57–70.
- Austin, J.L. 1940. The meaning of a word. In *Philosophical Papers*, Austin, J.L. 1961, 23–43. Oxford: Clarendon Press.
- Austin, J.L. 1950. Truth. *Proceedings of the Aristotelian Society*. Reprinted in *Philosophical papers*, Austin, J.L. 1961, 85–101. Oxford: Clarendon Press.
- Austin, J.L. 1956. A plea for excuses. In *Philosophical papers*, Austin, J.L. 1961, 123–152. Oxford: Clarendon Press.
- Austin, J.L. 1962a. *How to do things with words*. Oxford: Oxford University Press.
- Austin, J.L. 1962b. *Sense and sensibilia*. Oxford: Oxford University Press.
- Ayer, A.J. 1946. *Language, truth and logic*, 2nd edn. Harmondsworth: Pelican. (1st edn. London 1936).
- Bach, Kent. 2001. You don't say? *Syntheses* 128: 15–44.
- Bezuidenhout, Anne. 2010. Grice on presupposition. In *Meaning and analysis, new essays on Grice*, ed. Petrus Klaus, 75–102. Basingstoke: Palgrave Macmillan.
- Bezuidenhout, Anne, and Robin Morris. 2004. Implicature, relevance and default pragmatic inference. In *Experimental pragmatics*, eds. Noveck Ira, and Sperber Dan, 257–282. Basingstoke: Palgrave.
- Blake, Norman. 2002. *A grammar of Shakespeare's language*. Basingstoke: Palgrave.
- Borg, Emma. 2004. *Minimal semantics*. Oxford: Oxford University Press.
- Breheny, Richard, Napoleon Katos, and John Williams. 2006. Are generalized scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatics inferences. *Cognition* 100: 434–463.
- Burton-Roberts, Noel. 1989. *The limits to debate*. Cambridge: Cambridge University Press.
- Cappelen, Herman, and Ernie Lepore. 2005. *Insensitive semantics: a defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell.

- Carnap, Rudolf. 1937. *The logical syntax of language*. London: Routledge and Keegan Paul.
- Carston, Robyn. 2002. *Thoughts and utterances*. Oxford: Blackwell.
- Chapman, Siobhan. 2005. *Paul Grice, philosopher and linguist*. Basingstoke: Palgrave.
- Chapman, Siobhan. 2010. Paul Grice and the philosophy of ordinary language. In *Meaning and analysis, new essays on Grice*, ed. Petrus Klaus, 31–46. Basingstoke: Palgrave Macmillan.
- Frege, Gottlob. 1892. On sense and meaning. In *Translations from the philosophical writings of gottlob frege*, eds. Geach Peter and Black Max, 1980, 56–78. Oxford: Blackwell (1st edn. 1952).
- Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition and logical form*. New York: Academic Press.
- Gibbs, Raymond. 1999. Interpreting what speakers say and implicate. *Brain and Language* 68: 466–485.
- Glucksberg, Sam. 2004. On the automaticity of pragmatic processes: A modular proposal. In *Experimental pragmatics*, ed. Noveck Ira, and Sperber Dan, 72–93. Basingstoke: Palgrave.
- Grice, Paul. c. 1946–1950. Common sense and skepticism. In *Studies in the way of words*, Grice Paul, 1989, 147–153. Harvard: Harvard University Press.
- Grice, Paul. c. 1953–1958. G. E. Moore and the philosopher's paradoxes. In *Studies in the way of words*, Grice Paul, 1989, 154–170. Harvard: Harvard University Press.
- Grice, Paul. 1957. Meaning. In *Studies in the way of words*, Grice Paul, 1989, 213–223. Harvard: Harvard University Press.
- Grice, Paul. 1958. Postwar Oxford philosophy. In *Studies in the way of words*, Grice Paul, 1989, 171–180, Harvard: Harvard University Press.
- Grice, Paul. 1961. The causal theory of perception. In *Studies in the way of words*, Grice Paul, 1989, 224–247. Harvard: Harvard University Press.
- Grice, Paul. 1967a. Prolegomena. In *Studies in the way of words*, Grice Paul, 1989, 3–21. Harvard: Harvard University Press.
- Grice, Paul. 1967b. Indicative conditionals. In *Studies in the way of words*, Grice Paul, 1989, 58–85. Harvard: Harvard University Press.
- Grice, Paul. 1969. Utterer's meaning and intentions. In *Studies in the way of words*, Grice Paul, 1989, 86–116. Harvard: Harvard University Press.
- Grice, Paul. 1975. Logic and conversation. In *Studies in the way of words*, Grice Paul, 1989, 22–40. Harvard: Harvard University Press.
- Grice, Paul. 1978. Further notes on logic and conversation. In *Studies in the way of words*, Grice Paul, 1989, 41–57. Harvard: Harvard University Press.
- Grice, Paul. 1981. Presupposition and conversational implicature. In *Studies in the way of words*, Grice Paul, 1989, 269–282. Harvard: Harvard University Press.
- Grice, Paul. 1986. Reply to Richards. In *Philosophical grounds of rationality*, eds. Grandy Richard and Warner Richard, 45–106. Oxford: Clarendon Press.
- Grice, Paul. 1987. Retrospective epilogue. In *Studies in the way of words*, Grice Paul, 1989, 339–385. Harvard: Harvard University Press.
- Horn, Laurence. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61: 121–174.
- Horn, Laurence. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Horn, Laurence. 2007. Neo-Gricean pragmatics: a Manichaean manifesto. In *Pragmatics*, ed. Burton-Roberts Noel, 158–183. Basingstoke: Palgrave.
- Horn, Laurence. 2010. WJ-40: Issues in the investigation of implicature. In *Meaning and analysis, new essays on Grice*, ed. Petrus Klaus, 310–339. Basingstoke: Palgrave Macmillan.
- Karttunen, Lauri, and Stanley Peters. 1979. Conventional implicature. In *Syntax and semantics, vol 11: Presupposition*, eds. Oh C.-K., and Dineen D., 1–56. New York: Academic Press.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, Stephen. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: M.I.T. Press.
- Lüthi, David. 2006. How implicatures make Grice an unordinary ordinary language philosopher. *Pragmatics* 16: 247–274.

- Malcolm, Norman. 1942. Moore and ordinary language. In *The Philosophy of G. E. Moore*, ed. Schilpp Paul Arthur, 345–368. La Salle, Illinois: Open Court.
- Martinich, Al. 2010. The total content of what a speaker means. In *Meaning and analysis, new essays on Grice*, ed. Petrus Klaus, 252–267. Basingstoke: Palgrave Macmillan.
- McCawley, James. 1981. *Everything that linguists have always wanted to know about logic*. Oxford: Blackwell.
- Mey, Jacob. 1989. The end of the copper age, or: pragmatics 12½ years after. *Journal of Pragmatics* 13: 825–832.
- Moore, G.E. 1925. A defence of common sense. In *Philosophical papers*, Moore G.E., 1959, 32–59, London: George Allen and Unwin Ltd.
- Moore, G.E. 1939. Proof of an external world. In *Philosophical papers*, Moore G.E., 1959, 127–150. London: George Allen and Unwin Ltd.
- Perner, J., and S. Leekam. 1986. Belief and quantity: Three year olds' adaption to listener's knowledge. *Journal of Child Language* 13: 305–315.
- Petrus, Klaus. 2010. Introduction: Paul Grice, philosopher of language, but more than that. In *Meaning and analysis, new essays on Grice*, ed. Petrus Klaus, 1–30. Basingstoke: Palgrave Macmillan.
- Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.
- Recanati, François. 2004. *Literal meaning*. Cambridge: Cambridge University Press.
- Rorty, Richard (ed.). 1967. *The linguistic turn: Recent essays in philosophical method*. Chicago and London: The University of Chicago Press.
- Russell, Bertrand. 1905. On denoting. *Mind* 14: 479–499.
- Russell, Bertrand, and Alfred Whitehead. 1910. *Principia Mathematica*. Cambridge: Cambridge University Press.
- Saul, Jennifer. 2002. What is said and psychological reality: Grice's project and relevance theorists' criticisms. *Linguistics and Philosophy* 25: 347–372.
- Searle, J.R. 1969. *Speech acts*. Cambridge: Cambridge University Press.
- Sperber, Dan, and Deirdre Wilson. 1995. *Relevance*, 2nd ed. Oxford: Blackwell.
- Strawson, P.F. 1950. On referring. *Mind* 59: 320–344.
- Strawson, P.F. 1952. *Introduction to logical theory*. London: Methuen and Co.
- Winer, Gerald, Jane Cottrell, Tammy Mott, Matthew Cohen, and Jody Fournier. 2001. Are children more accurate than adults? Spontaneous use of metaphor by children and adults. *Journal of Psycholinguistic Research* 50: 485–496.

Writing Letters in the Age of Grice

Claudia Bianchi

Abstract This article aims to investigate the notion of implicature and its connections with speaker's intentions, communicative responsibility and normativity. Some scholars stress the normative character of conversational implicatures more than their psychological dimension. In a normative perspective, conversational implicatures don't correspond to what the speaker *intends* to implicate, but should be interpreted as enriching or correcting inferences licensed by the text. My chapter aims to show that the idea of an implicature that the speaker does not intend to convey is not persuasive. In Grice's theory conversational implicatures are speaker-meant: this means that inferences derived by the addressee but not intended by the speaker should not count as conversational implicatures. On the contrary, I will claim that propositions intended by the speaker and not recognised by the addressee should count as implicatures, if the speaker has made her communicative intention available to her audience.

1 Introduction

This article aims to investigate the Gricean notion of implicature and its connections with speaker's intentions, communicative responsibility and normativity. According to Grice, implicatures are part of what a speaker communicates, hence part of speaker meaning—and speaker meaning is a matter of speaker intentions. Some scholars stress the normative character of conversational implicatures more than their psychological dimension (Green 2002; Saul 2001, 2002a; Sbisà 2007). In a normative perspective, conversational implicatures don't correspond to what the speaker *intends* to implicate (or to what the addressee successfully infers):

C. Bianchi (✉)

Faculty of Philosophy, University Vita-Salute San Raffaele, via Olgettina 58,
20132 Milan, Italy

e-mail: claudia.bianchi@univr.it

conversational implicatures should be interpreted as enriching or correcting inferences licensed by the text. Implicatures have a normative status as integration or amendment of an utterance, justified by an appropriate argumentative path: a conversational implicature isn't necessarily a proposition believed by the speaker, but a proposition that *should* be accepted by the speaker.

My chapter aims to show that the idea of an implicature that the speaker does not intend to convey is not persuasive. In Grice's theory conversational implicatures are speaker-meant—conscious or even “designed”.¹ This means that inferences derived by the addressee but not intended by the speaker should not count as conversational implicatures. On the contrary, I will claim that propositions intended by S and not recognised by A should count as implicatures, if S has made her communicative intention available to her audience.

2 Grice: What Is Said and What Is Implicated

As is well known, Grice distinguishes between “what is said” and “what is implicated”—between the proposition expressed by an utterance (the truth-conditions of the sentence uttered) and the implicit meaning of the utterance, “what is implicated” by a speaker using a sentence in a given context—an inference licensed in context, and which cannot be identified with logical implication, logical consequence or entailment (inferences derived solely from semantic content).²

According to Grice, if someone asks S how Tom is getting on in his job, and S utters

1. Tom likes his colleagues and hasn't been to prison yet,

she is only implying—and not saying—that Tom is the sort of person likely to yield to the temptation provided by his occupation.³

Therefore, Grice's main task is to draw a distinction “within the total signification of a remark: a distinction between what the speaker has *said* (in a certain favored, and maybe in some degree artificial, sense of 'said'), and what he has *implicated* (e.g. implied, indicated, suggested), taking into account the fact that what he has implicated may be either conventionally implicated (implicated by virtue of the meaning of some word or phrase which he has used) or nonconventionally implicated (in which case the specification of the implicature falls outside the specification of the conventional meaning of the words used)”.⁴ In other words, Grice tries to fully characterise

¹ Grice (1975, 1989, p. 34).

² Cf. Grice (1989).

³ Grice (1975, 1989, p. 24).

⁴ Grice (1967, 1989, p. 118).

- A. what an expression E means;
- B. what a speaker S explicitly says (in Grice's "favored", technical sense) using E in a given occasion⁵;
- C. what S implicitly conveys (implicates, implies, indicates, suggests) using E in that given occasion.

Both B. and C. amount to *speaker meaning*. Hence, an implicature is a non-truth-conditional aspect of speaker meaning—part of what is meant when S utters E in context C, without being part of what is said by S with E.

The gap between expression meaning (A.) and speaker's meaning (B. + C.), and between saying (B.) and implying (C.) is filled by exploiting a set of expectations which both speaker S and addressee A share. Those expectations are based on an assumption: language use is a form of rational and cooperative behaviour, characterised by a high level of coordination.⁶ Conversation, then, is a rational, cooperative, goal-oriented activity—governed by a Cooperative Principle.⁷ The Cooperative Principle is specified by four categories, called Quantity, Quality, Relation and Manner, under which fall more specific maxims. In order to understand the speaker's meaning, A is guided by certain expectations concerning S's behaviour: namely the expectation that S's utterance will satisfy certain standards—being informative (Quantity Maxims), sincere (Quality Maxims), relevant (Relation Maxim) and clear (Manner Maxims).⁸

3 Implicatures

Sometimes what S says fails to be plausible, informative, relevant or appropriate. If the violation is overt, manifest and blatant, A, when she has reason to think that S is no longer cooperative, may infer additional propositions completing or revising what S says. According to Grice, then, in our linguistic interactions a speaker may often communicate much more than what she says explicitly. The proposition S communicates by uttering a particular sentence in a particular context without saying it Grice dubs *implicature*: this proposition is not part of the

⁵ According to Grice, what is said (i.e. B.) is "closely related to the conventional meaning of the... sentence... uttered" (i.e. A.) and must correspond to "the elements of the [sentence], their order, and their syntactic character". It is closely related but not identical to what the sentence means, because the sentence may contain ambiguities or indexicals: Grice (1969, 1989, p. 87).

⁶ Grice (1975, 1989, p. 26): "Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction".

⁷ Grice (1975, 1989, p. 26): "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged".

⁸ Cf. Grice (1975, 1989, pp. 26–27).

truth-conditional content of the sentence uttered—it does not contribute to its truth-conditions. The general idea is that A’s expectations may be exploited by S in order to generate further communicative effects. Uttering a sentence is an action that A takes as meaningful: she may then infer S’s communicative intention (what is implicated) by taking into account S’s utterance (what is said) and contextual factors that A (supposedly) shares with S.

In “Logic and conversation”, Grice offers a definition of conversational implicature:

A man who, by (in, when) saying (or making as if to say) that *p* has implicated that *q*, may be said to have conversationally implicated that *q*, provided that a) he is to be presumed to be observing the conversational maxims, or at least the Cooperative Principle; b) the supposition that he is aware that, or thinks that, *q* is required in order to make his saying or making as if to say *p* (or doing so in *those* terms) consistent with this presumption; and c) the speaker thinks (and would expect the hearer to think that the speaker thinks) that it is within the competence of the hearer to work out, or grasp intuitively, that the supposition mentioned in b) is required.⁹

Let’s summarize the definition.¹⁰ S says that *p* and implicates that *q* if¹¹:

- (a) S is to be presumed to be observing the Cooperative Principle;
- (b) the supposition that S is aware that, or thinks that, *q* is required in order to make her saying consistent with this presumption;
- (c) S thinks that it is within the competence of A to work out that the supposition mentioned in (b) is required.

An example will help clarify the definition: the reference letter in “Logic and Conversation”. Grice writes: “A is writing a testimonial about a pupil who is a candidate for a philosophy job, and his letter reads as follows: Dear Sir, Mr. X’s command of English is excellent, and his attendance at tutorials has been regular. Yours, etc.”¹² For sake of simplicity I will use the utterance

2. Mr. X’s command of English is excellent.

By uttering (2), S says that *Mr. X’s command of English is excellent* and implicates that *Mr. X is no good at philosophy* if:

- (a) S is to be presumed to be observing the Cooperative Principle: as Grice puts it: “if [S] wished to be uncooperative, why write at all?”;

⁹ Grice (1975, 1989, pp. 30–31).

¹⁰ In my chapter I will focus on the normative/psychological debate and will not address many interesting points raised by Grice’s definition; in particular I will not deal with the distinction between “saying” and “making as if to say”. On this and related points, see Bianchi (2009) and (2013).

¹¹ According to Saul “only if”: Saul (2002a, p. 231). For a different point of view, see Neale (1992, pp. 527–529).

¹² Grice (1975, 1989, p. 33).

- (b) the supposition that S is aware that, or thinks that, *Mr X is no good at philosophy* is required in order to make her saying consistent with this presumption: “[S] knows that more information than this is wanted. [she] must, therefore, be wishing to impart information that [she] is reluctant to write down”¹³;
- (c) S thinks that it is within the competence of the philosophy search committee to work out that the supposition mentioned in (b) is required 4 Contextualizing justification.

4 Near-Implicatures

Saul (2002a) argues against a common understanding of Grice according to which speaker meaning divides exhaustively into what is said and what is implicated: “there are many things which speakers mean that they neither say nor implicate”.¹⁴ She distinguishes between utterer-implicatures (intended by the speaker, but not recognized by the addressee), audience-implicatures (recognized by the addressee but not intended by the speaker) and conversational implicatures. In what follows I would like to examine the status of implicatures and “near-implicatures” (utterer and audience implicatures), with a more general aim in mind: investigating the normative dimension of language use. This is especially urgent in a Gricean perspective, where meaning is reduced to a complex array of audience-oriented intentions: in this perspective, restrictions and constraints are posited in order to prevent speakers from saying just anything, simply by intending it. In a similar vein constraints must be posited in order to prevent speakers from conversationally implicating just anything, simply by intending it.¹⁵

In order to show that speaker meaning doesn't divide exhaustively into what is said and what is implicated, Saul presents some variations of the Gricean reference letter.¹⁶

Case I. A student of mine, Fred, is a poor philosopher but a good typist. He is applying for a philosophy job and I am asked to write a reference letter. My letter reads as follows: “Dear Sir or Madam, Fred is a good typist, and his attendance at tutorials has been regular. Yours, CB”. For sake of simplicity I will use the utterance

3. Fred is a good typist.

By uttering (3), I say that *Fred is a good typist* and implicate that *Fred is no good at philosophy* if:

¹³ Grice (1975, 1989, p. 33).

¹⁴ Saul (2002a, p. 229).

¹⁵ Cf. Saul (2002a, p. 229).

¹⁶ The five cases are taken, with variations, from Saul (2002a).

- (a) I am to be presumed to be observing the Cooperative Principle;
- (b) the supposition that I think that *Fred is no good at philosophy* is required in order to make my saying consistent with the presumption indicated in (a): I know that more information than this is wanted. I must, therefore, wish to impart information that I am reluctant to write down;
- (c) I think that it is within the competence of the philosophy search committee to work out that the supposition mentioned in (b) is required.

Let's now examine a variation of Case I.

Case II. Suppose that, unbeknownst to me, Fred is applying for a job as a typist. In this scenario, by uttering (3), I say that *Fred is a good typist* but I do not implicate that *Fred is no good at philosophy*. Condition (b) isn't satisfied: *q* (namely that *Fred is no good at philosophy*) isn't required in order to make my saying consistent with the presumption that I am being cooperative. The audience (Fred's prospective employers) may legitimately keep the assumption that I am observing the Cooperative Principle without assuming that I think that *q*: according to Saul, "The implicature was blocked because a speaker cannot conversationally implicate something which the audience is not required to assume that she thinks".¹⁷

Let's now consider another example of near-implicature, involving a different condition.

Case III. A student of mine, Cedric, is a poor philosopher but a good typist. He is applying for a philosophy job and I am asked to write a reference letter. My letter reads as follows: "Dear Sir or Madam, Cedric is a good typist, and his attendance at tutorials has been regular. Yours, CB". For sake of simplicity I will use the utterance

4. Cedric is a good typist.

Suppose now that the philosophy search committee thinks (falsely) that I always write uncooperative reference letters (perhaps because I consider the practice of writing reference letters itself inappropriate). In this scenario by uttering (4), I say that *Cedric is a good typist* but I do not implicate that *Cedric is no good at philosophy*. Condition (a) isn't satisfied: the philosophy search committee doesn't presume that I am observing the Cooperative Principle, and therefore derives no proposition completing or revising what I said. They simply take (4) as an uncooperative remark.

Cases II and III show that speaker meaning doesn't divide exhaustively into what is said and what is implicated. There are propositions that S means but

¹⁷ Cf. Saul (2002a, p. 231): "The implicature was blocked because a speaker cannot conversationally implicate something which the audience is not required to assume that she thinks". In this example, as in other examples, there is a crucial distinction to be made: the one between the intended addressee (the addressee intended by the speaker) and the actual audience. In order to adhere to Saul's arguments I will not make use of the distinction in what follows.

doesn't implicate: Cases II and III are examples of utterer-implicatures, intended by the speaker, but not recognized by the addressee.

Moreover, Saul individuates another category of near-implicatures: audience-implicatures, recognized by the addressee but not intended by the speaker. Let's examine Saul's example.

Case IV. A student of mine, Felix, is a good philosopher and a good typist. I think that he is applying for a job as a typist; my letter reads as follows: "Dear Sir or Madam, Felix is a good typist, and his attendance at tutorials has been regular. Yours, CB". For sake of simplicity I will use the utterance

5. Felix is a good typist.

Suppose that, unbeknownst to me, Felix is applying for a philosophy job. In this scenario we may imagine that the philosophy search committee would derive from my utterance of (5) the proposition q (*Felix is no good at philosophy*): (5) isn't as informative as is required for the current purposes of the exchange, therefore in overt violation of the Quantity Maxim. However, according to Saul, in Case IV q doesn't qualify as conversational implicature: condition (c) isn't satisfied. I am indeed presumed to be observing the Cooperative Principle by the philosophy search committee [condition (a)]; also, the supposition that I think that Felix is no good at philosophy is required in order to make my saying consistent with this presumption [condition (b)]. But Saul argues that we cannot say that I think that it is within the competence of the philosophy search committee to work out that the proposition that *Felix is no good at philosophy* is required, because I do not intend to communicate that *Felix is no good at philosophy*, and I do not even believe that *Felix is no good at philosophy*: "The audience, which is actually a philosophy appointments committee, takes me to have conversationally implicated that Felix is a poor philosopher. They are, of course, wrong: clause (c) of Grice's characterisation was not satisfied as I had no idea that they would, or even could, work out from my utterance that I think that Felix is a poor philosopher, and I would not have made my utterance if I'd realised the situation. This claim, then, fails to be conversationally implicated".¹⁸

According to Saul, then, speaker meaning doesn't divide exhaustively into what is said and what is implicated: we must distinguish between utterer-implicatures (intended by the speaker, but not recognized by the addressee), audience-implicatures (recognized by the addressee but not intended by the speaker) and conversational implicatures.

¹⁸ Saul (2002a, p. 242).

5 The Normative Perspective

Speaker meaning is conceived by Grice as a complex array of audience-oriented intentions, whose distinctive feature is that their fulfilment is obtained by means of their recognition.¹⁹ According to Grice, then, being meant by S is a *necessary* condition of what is said. Now, what about implicatures? Is being meant by S a *necessary* condition of what is implicated?

Not for Saul. As I have said, Saul (along with other scholars such as Gauker, Green, and Sbisà) stresses the normative character of conversational implicatures more than their psychological dimension.²⁰ In this perspective, conversational implicatures correspond neither to what the speaker intends to implicate nor to what the addressee successfully infers: conversational implicatures should be interpreted as enriching or correcting inferences licensed by the text. This means not only that the addressee is capable of working out the implicature, but also that she *should* have worked it out—that she may rightfully attribute to the speaker the intention of conveying it.²¹ Implicatures therefore have a normative status as integration or correction of an utterance, justified by an appropriate argumentative path.²²

In my opinion, the idea of an implicature that the speaker does not intend to convey is not completely persuasive. In Grice's theory conversational implicatures are speaker-meant—conscious or even “designed”.²³ In this section, I examine some arguments against the idea that propositions derived by A but not meant by S should count as conversational implicatures.

1. Saul denies that being meant by S is a *necessary* condition of what is implicated, because it doesn't explicitly appear in conditions (a)–(c). We may concede that, and even that the meant-condition doesn't *follow* from conditions

¹⁹ Cf. Bach and Harnish (1979, p. 15): “its fulfilment consists in its recognition”. Bach and Harnish's theory is a development of Grice's, and of his intention-based and inferential view of communication. To many, however, their position is too strong.

²⁰ Cf. Green (2002, 2007), Saul (2002a, b) and Sbisà (2007).

²¹ Cf. Saul (2002a, p. 244): “There are, then, cases in which we can reasonably say that the audience *should* have worked out the conversational implicature, even if they failed to do so”; Sbisà (2007, pp. 122, 126 and 192).

²² According to Marina Sbisà, a conversational implicature isn't necessarily a proposition believed by the speaker, but a proposition that should be accepted by the speaker. This means that the speaker may be wrong about an implicature: even if she does not intend to convey a particular implicature, there are cases in which this should in any case be worked out by the addressee. In Sbisà's framework, implicatures are normative virtual objects. The alleged implicature does not count as conveyed meaning only if to attribute that communicative intention to the speaker would be absurd or contradictory: but if the text licenses it, the derivation of a particular implicature will be legitimate, even if S has no intention of conveying it.

²³ Grice (1975, 1989, p. 34).

(a)–(c).²⁴ However, the meant-condition appears elsewhere, especially in the Gricean description of the inferential calculus allowing A to derive a conversational implicature—where the supposition that S thinks that *q* is required:

He has said that *p*; there is no reason to suppose that he is not observing the maxims, or at least the Cooperative Principle; he could not be doing this unless he thought that *q*; he knows (and knows that I know that he knows) that I can see that the supposition that he thinks that *q* is required; he has done nothing to stop me thinking that *q*; he intends me to think, or is at least willing to allow me to think, that *q*; and so he has implicated that *q*.²⁵

2. Another philological point, underlined both by Neale (1992) and Davis (2007): the meant-condition appears in “The causal theory of perception”: “a speaker implies something only if it was something he ‘intended to get across’”.²⁶ True, the passage is omitted in Grice (1989), but only because the view presented is “substantially the same”; as Grice puts it, “This section [3] is here omitted, since the material which it presents is substantially the same as that discussed in Essay 2 [‘Logic and Conversation’]”.²⁷ In spite of this, Saul chooses to ignore the passage, because of its “odd” status.²⁸
3. However, the crucial point is another. The conditions (a)–(c) are meant to set apart conversational implicatures from other kinds of implicatures (i.e. conventional ones). The definition states that “A man who, by saying that *p* has implicated that *q*, may be said to have conversationally implicated that *q*, provided that [the three conditions are satisfied]”. We must look elsewhere for the general definition of implicature, in the opening of “Logic and Conversation” where Grice defines implicating as “mean without saying”.

Wayne Davis holds the same view, criticises Grice on this very point, and suggests an alternative definition: implicating must be conceived as “meaning something by saying something else”.²⁹ Davis’ proposal avoids Saul’s objections about lapsus, malapropisms, or poor translations—where S intends to say that *p* but accidentally utters a sentence which conventionally means that *q*. According to Saul, we cannot claim that S, uttering

6. We’re all cremated equal

has implicated that

7. We’re all created equal.

²⁴ Cf. Davis (2007, Sect. 2): “conditions [(a)–(c)] say nothing about what S intends”. For a different opinion, see Neale (1992, p. 528).

²⁵ Grice (1975, 1989, p. 31). Cf. Davis (2007, Sect. 2): “Grice does mention intention at the appropriate place when he is setting out the ‘Calculability Assumption’ of his theory”.

²⁶ Grice (1961, 1965 p. 448).

²⁷ Grice (1961, 1989, p. 229).

²⁸ Cf. Saul (2002a, pp. 238–239).

²⁹ Davis (1998, 2007).

But if we conceive implicating as “mean without saying”, then we are committed to hold that S, uttering (6), has implicated (7) (a proposition meant, but not said, by S). The alternative definition proposed by Davis underlines the *indirect* character of implicatures: in Case I, *by* saying that Fred is a good typist, S intends to communicate that Fred is a poor philosopher. Things work differently as far as malapropisms or lapsus are concerned: intuitively we wouldn’t claim that S, by saying that (6), intended to communicate that (7).

4. Last point: in the normative perspective advocated by Saul, Case IV should count as a conversational implicature, even if unmeant by S. In that context, A may seem right in ascribing the conversational implicature *q* (*Felix is no good at philosophy*) to S. In that particular communicative situation (the reference letter for a philosophy job written by a professor of philosophy for a student she knows very well), what is said isn’t as informative as required for the current purposes of the exchange, and is hence in overt violation of the Quantity Maxim: therefore the committee seems fully justified in deriving that particular implicature. However, in Case IV Saul describes *q* as a case of near-implicature—as if reluctant to ascribe to S unmeant implicatures.

6 Making Implicatures Available

In closing, I address an aspect symmetrical to the one dealt with in [Sect. 5](#). I said that being meant by S is a necessary condition of what is implicated; let’s now examine the question of whether being recognised by A is a necessary condition of what is implicated. More generally, are conversational implicatures propositions merely intended by the speaker and recognized by the addressee?

With Saul, I claim that this kind of restriction upon A is not necessary: some implicatures fail to be recognised by A. Let’s consider another of Saul’s variations of the reference letter.

Case V. A student of mine, Wesley, is a poor philosopher. He is applying for a philosophy job and I am asked to write a reference letter. My letter reads as follows: “Dear Sir or Madam, Wesley’s main virtues as a philosopher are punctuality, an attractive choice of fonts, and an encyclopaedic knowledge of illegal pharmaceuticals. Yours, CB”. For sake of simplicity I will use the utterance

8. Wesley has an encyclopaedic knowledge of illegal pharmaceuticals.

Suppose now that the philosophy search committee reads my letter too quickly, retains the information that Wesley has encyclopaedic knowledge and hires him. In this context, even if the committee didn’t recognise my implicature, we must acknowledge that I said *p* and implicated *q* (*Wesley is no good at philosophy*). The three conditions are indeed satisfied:

- (a) I am to be presumed to be observing the Cooperative Principle;
- (b) the supposition that I think that *Wesley is no good at philosophy* is required in order to make my saying consistent with the presumption indicated in (a);
- (c) I think that it is within the competence of the philosophy search committee to work out that the supposition mentioned in (b) is required.

In this scenario not only is the supposition that I think that *Wesley is no good at philosophy* required in order to make my utterance of (8) consistent with my being collaborative, but also I think that it is within the competence of the philosophy search committee to work out that this supposition is required.

This point introduces a normative element into my own intentional perspective. Being meant by S is a necessary condition of what is implicated; nonetheless, it isn't a *sufficient* condition. Being meant is a necessary condition of implicatures only if S's intention is non-arbitrary—that is connected with a behaviour that enables the addressee to identify it. S's intention must be "reasonable"—i.e. something that A in normal circumstances is able to work out using external facts (where, when and by whom the utterance is produced), linguistic co-text (what has been said so far), and background knowledge.³⁰ If, in Case V, I have made my implicature available to A—having thus fulfilled my communicative responsibilities—then I have implicated *q*.

In conclusion, conversational implicatures are more than merely intended by the speaker and recognized by the addressee: implicating (con conversationally) amounts to making available to the addressee the implicit message S wants to communicate. Of course, having fulfilled my communicative responsibilities isn't a guarantee of a successful communication: in Case V, I have implicated that *q* but I haven't *communicated* that *q*.³¹ This, however, comes as no surprise, as it holds also for what is said. To say something is to satisfy one's communicative responsibilities, even if this fails to guarantee communication. If I utter

9. George is Hellenic

I mean and say

10. George is Greek;

but if A doesn't know that "Hellenic" means "Greek", I fail to communicate it. S's having a complex array of audience-oriented intentions, and making it public and available to A, does not secure communication.³²

³⁰ Roberts (1997, p. 196). Cf. Donnellan (1968).

³¹ Saul (2002a, p. 245): "con conversationally implicating something... fails to guarantee audience uptake but does mean that the speaker has fulfilled her communicative responsibilities with regard to what she wants to communicate... she may not have communicated her intended message, but she has *made it available*".

³² Cf. Predelli (2002, pp. 315–316).

7 Conclusion

The task of this article was to critically examine the notion of implicature in a Gricean perspective, along with its connection with the speaker's intentions, communicative responsibility and normativity. In the normative view held by Saul, a conversational implicature isn't necessarily a proposition meant by the speaker, but a proposition that should be accepted by the speaker. My chapter aims to show that the idea of an implicature that the speaker does not intend to convey is not persuasive. In Grice's theory conversational implicatures are speaker-meant. This means that inferences derived by the addressee but not intended by the speaker should not count as conversational implicatures. On the contrary, I claim that propositions intended by S and not recognized by A should count as implicatures, if S has made her communicative intention available to her audience. In this latter case, we are facing a communicative failure: the implicature exists (an ontological matter) but is not recognized (an epistemological matter).³³

References

- Bach, Kent, and Robert M. Harnish. 1979. *Linguistic communication and speech acts*. Cambridge: MIT Press.
- Bianchi, Claudia. 2009. *Pragmatica cognitiva. I meccanismi della comunicazione*. Roma-Bari: Laterza.
- Bianchi, Claudia. 2013. Implicating. In *Pragmatics of speech actions*, eds. M. Sbisà, and K. Turner. Berlin: Mouton de Gruyter: 83–118.
- Davis, Wayne. 1998. *Implicature: Intention, convention, and principle in the failure of Gricean theory*. Cambridge: Cambridge University Press.
- Davis, Wayne. 2007. How normative is implicature? *Journal of Pragmatics* 39: 1655–1672.
- Davis, Wayne. 2010. Implicature. In *The Stanford encyclopedia of philosophy*, ed. Edward N. Zalta, (Winter 2010 Ed.), forthcoming <http://plato.stanford.edu/archives/win2010/entries/implicature/>.
- Donnellan, Keith. 1968. Putting humpty dumpty together again. *The Philosophical Review* 77: 203–215.
- Green, Mitchell. 2002. Review of Implicature: Intention, convention, and principle in the failure of Gricean theory, by Wayne Davis. *Philosophy and Phenomenological Research* 65: 241–244.
- Green, Mitchell. 2007. *Self-expression*. Oxford: Clarendon Press.
- Grice, Herbert Paul. 1961. The causal theory of perception. *The Aristotelian Society: Proceedings, Supplementary Volume*, 35, 121–52, reprinted in R. Swartz ed. *Perceiving, Sensing and Knowing*, Garden City, N. Y., Doubleday, Anchor, 1965, reprinted in Grice 1989, 224–247.
- Grice, Herbert Paul. 1967. Logic and conversation. In *Syntax and semantics 3: Speech Acts*, 41–58, eds. Cole, P. and Morgan, J., 1–143. New York: Academic Press, reprinted in Grice 1989.
- Grice, Herbert Paul. 1969. Utterer's meaning and intentions. *Philosophical Review* 68: 147–177, reprinted in Grice 1989, 86–116.

³³ Cf. Davis (2010).

- Grice, Herbert Paul. 1975. Logic and conversation (1967). In *Syntax and semantics 3: Speech acts*, eds. Cole, P. and Morgan, J., 41–58. New York: Academic Press, reprinted in Grice 1989, 22–40.
- Grice, Herbert Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.
- Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509–595.
- Predelli, Stefano. 2002. Intentions, indexicals and communication. *Analysis* 62(4): 310–316.
- Roberts, Lawrence D. 1997. How demonstrations connect with referential intentions. *Australian Journal of Philosophy* 75(2): 190–200.
- Saul, Jennifer. 2001. Wayne A. Davis, Implicature: Intention, convention, and principle in the failure of Gricean theory. *Nous* 35: 631–641.
- Saul, Jennifer. 2002a. Speaker meaning, what is said, and what is implicated. *Nous* 36: 228–248.
- Saul, Jennifer. 2002b. What is said and psychological reality: Grice's project and relevance theorists' criticisms. *Linguistics and Philosophy* 25: 347–372.
- Sbisà, Marina. 2007. *Detto, non detto. Le forme della comunicazione implicita*. Roma-Bari: Laterza.

Implicatures as Forms of Argument

Fabrizio Macagno and Douglas Walton

Abstract In this paper, we use concepts, structure and tools from argumentation theory to show how conversational implicatures are triggered by conflicts of presumptions. Presumptive implicatures are shown to be based on defeasible forms of inference used in conditions of lack of knowledge, including analogical reasoning, inference to the best explanation, practical reasoning, appeal to pity, and argument from cause. Such inferences are modelled as communicative strategies used to fill knowledge gaps by shifting the burden of proof to provide the missing contrary evidence to the other party in a dialogue. Through a series of illustrative examples, we show how such principles of inference are based on common knowledge about the ordinary course of events shared by participants in a structured dialogue setting in which they take turns putting forward and responding to speech acts.

Keywords Argumentation · Implicatures · Argumentation schemes · Analogy · Speech acts · Pragmatics · Implicit speech acts · Indirect speech acts

The notion of implicature was introduced by Grice in ‘The Causal Theory of Perception’ (Grice 1961) and developed in his further essays (Grice 1975)¹. The theory is meant to describe the relationship between what is said and what is meant in cases in which the conventional meaning commonly associated with the sentence (“what is said”) does not correspond with the speaker’s communicative intention, or his/her intended effect on the audience (“what is meant”) (Grice

¹ For an overview of the roots of Grice’s theory of implicature, see (Davis 2007).

F. Macagno (✉)
IFL, Faculdade de Ciências Sociais e Humanas-Universidade Nova de Lisboa,
Lisboa, Portugal
e-mail: fabrizio.macagno@fcsh.unl.pt

D. Walton
University of Windsor, Windsor, Canada
e-mail: dwalton@uwindsor.ca

1975, 1989, 220; Levinson 1983, 97). Implicatures are therefore consequences of a speaker's *saying* a specific sentence. On Grice's view, what a man says needs to be considered in the context of the expectations and presumptions of the community of speakers he belongs to (Grice 1975, 47). Grice collected such presumptions and expectations under general categories conceived as communicative norms (Sperber and Wilson 1986, 162) or presumptions. If an utterance superficially appears not to conform to the presumptions that the speaker is cooperative and is speaking truthfully, informatively, relevantly, perspicuously, and otherwise appropriately, the hearer tries to explain this failure by reconstructing a new unconventional meaning for the sentence (Bach 2003, 155). Grice conceived these presumptions as general and context-independent communicative principles, and they were reduced to principles of relevance by Sperber and Wilson (1986, 382; Wilson and Sperber 2004, 608) and applied to several forms of communication.

Grice's proposal for analyzing implicit meaning in terms of a process of reasoning triggered by a breach of dialectical expectations is based on three basic principles: the meeting of the parties' minds (or understanding of the intentions), the existence of shared rules of communication, and the ability to reconstruct "what is meant" from "what is said". These three principles have been studied from cognitive perspectives, transferring the concept of meaning to the field of intentions and beliefs (Grice 1975; Thomason 1990) and applying universal principles (cooperation and relevance) to explain inferential processes. However, such maxims and principles are not universal, but simply characterize some contexts of dialogue. In legal dialogues or negotiations such maxims cannot hold (Goodwin 2001). Moreover, the crucial step between what is said and what is meant has been never described.

The purpose of this paper is to investigate the relationship between reasoning and implicitness. What is the process of reasoning linking a sentence with an intention? Why do different people with different background knowledge draw the same inferences in different circumstances? Why does the process of retrieving the implicit meaning vary according to the contexts of dialogue? Our proposal is to interpret the grounds of Grice's theory using argumentation schemes in a dialectical framework, in order to show how implicatures work in this structure as implicit arguments.

1 Presumptions and Meaning

Grice's theory of implicature relies on the hearer's ability to understand the intended communicative purpose of an utterance, that is, its meaning. The mechanism the speaker activates to retrieve the communicative intention is triggered by a failure of communicative expectations that are crystallized in four maxims or in one super-maxim. Conceiving such presumptions as abstract norms, Grice reduced all communication to some unspecified but evidently particular type of dialogue in which such rules are taken to hold. This ideal dialogue, however, hides the

complexity of human conversation, whose regulating principles are context dependent. For instance, the most basic principle, the cooperative principle, does not hold in all contexts of dialogue. In legal cross-examination, the most important rule of the cross-examination is that the lawyer needs to avoid possible evasions (*Bronston v. US* 409, U.S. 359, 1973; Goodwin 2001). For instance, in this case, the lawyer asked a question, and retrieved the answer by resorting to implicature.

Case 1: Presumption of evasion

“Q. Do you have any bank accounts in Swiss banks, Mr. Bronston?”

“A. No, sir.”

“Q. Have you ever?”

“A. The company had an account there for about six months, in Zurich.”

The witness actually held a bank account in a Swiss bank, but was found to have testified truthfully, as he never stated the contrary. In this type of dialogue, the Gricean maxims, including the relevance maxims cannot hold.

What is needed to account for such conversational differences is a weaker requirement. Instead of considering maxims and super maxims as principles, we maintain that they are dialogical presumptions that depend on the type of dialogical purpose and dialogue type. In ordinary conversation the interlocutor is presumed to be cooperative and therefore willing to contribute to the dialogue. In other contexts of dialogue he is presumed to aim at winning the discussion, hiding some information, or trying to make the other party lose. In order to retrieve the intended meaning of “what is said”, or rather the purpose of a dialogical move, it is necessary first to understand the purpose of the dialogue the interlocutors are engaged in. Grice (1975, 45) underscored this first requirement of interpretation introducing the notion of “direction” of the dialogue, which imposes conditions on the possible conversational moves:

Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction. This purpose or direction may be fixed from the start (e.g., by an initial proposal of a question for discussion), or it may evolve during the exchange; it may be fairly definite, or it may be so indefinite as to leave very considerable latitude to the participants (as in a casual conversation). But at each stage, SOME possible conversational moves would be excluded as conversationally unsuitable

Aristotle first noticed that dialogical intentions can be broadly classified into types, depending on the institutional and conversational setting (Aristotle, *Topics*, 159a25–159a38).

In as much as no rules are laid down for those who argue for the sake of training and of examination—for the aim of those engaged in teaching or learning is quite different from that of those engaged in a competition; as is the latter from that of those who discuss things together in the spirit of inquiry; for a learner should always state what he thinks (for

no one tries to teach what is false); whereas in a competition the business of the questioner is to appear by all means to produce an effect upon the other, while that of the answerer is to appear unaffected by him; on the other hand, in dialectical meetings held in the spirit not of a competition but of an examination and inquiry, there are as yet no articulated rules about what the answerer should aim at, and what kind of things he should and should not grant for the correct or incorrect defence of his position—in as much, then, as we have no tradition bequeathed to us by others, let us try to say something upon the matter for ourselves.

On his view, the participants' institutional setting help us to understand their purposes and intentions in speaking, providing a blueprint for retrieving their specific dialectical goals. Walton (1989, 1990, further developed in Walton and Krabbe 1995 and Walton 1998) provided a typology of dialogue types (negotiation, information seeking, eristic, persuasion dialogue, deliberation and inquiry) in the form of a set of prototypical dialectical contexts defined by the interlocutors' roles, including their reciprocal knowledge based on the dialectical setting (such as work, family, etc.) and institutional roles (i.e. father, professor, businessman, etc.). A type of dialogue therefore can be considered as a normative framework in which there is an exchange of arguments between two speech partners reasoning together in turn-taking sequences aimed at moving toward the fulfillment of a collective conversational goal (Walton 1998, 30). On this perspective (Walton 2010, 14) the global goal of the dialogue determines the local goal of the interlocutors' specific moves made as speech acts, such as asking a question or putting forward an argument (Macagno 2008).

The argumentation that takes place in connected moves at this local level can be modeled by profiles of dialogue (Walton 1989, 65–71). For instance, in a decision-making dialogue about choosing a restaurant, the participants are presumed to bring forward arguments in favor of their preferences, express their opinions, and ask for more information. It would be taken to be unreasonable in such a setting to resort to threats, or to demand uselessly detailed information while putting off making a timely decision. Moreover, the interpretation of each move will be influenced and determined by the global purpose of the dialogue. The same discourse move “I like eating at Gino's” will be taken as a proposal and an implicit argument in a deliberation dialogue, as a report of personal tastes in an information-seeking dialogue, as a premise of a value judgment in a persuasion dialogue, as the denunciation of a personal offence in an eristic dialogue, or as the putting forth of an intention or goal in a negotiation.

If we want to analyze the structure of the reasoning underlying the mechanism of interpretation, we need to change perspective, and conceive interlocutors' communicative intentions using the theory of dialogical coherence relations. Grimes (1975, 209ff) interpreted abstract intentions as “rhetorical predicates”, later named “logical-semantic connectives” (Crothers 1979) or “coherence relations” (Hobbs 1979, 68, 1985; Rigotti 2005), such as explanation, alternative, support, etc. Such predicates, or relations, connect discourse sequences in two similar fashions, through subordination or coordination. In the first case the predicate is explicit and imposes a set of coherence conditions, or pragmatic presuppositions (Vanderveken 2002, 47;

Bach 2003, 163), on its arguments (Grimes 1975, 162). In the second case, such a predicate hides a deeper relationship (Lee et al. 1971) that needs to be reconstructed in order to understand the role of the discourse segments or sequences. For instance, coordination can express temporal, causal or explanation relations. From a pragmatic perspective, such relations can be considered as high-level speech acts (Grice 1989, 362; Carston 2002, 107–108), indicating the role of the first level speech acts, or rather, their felicity conditions (Vanderveken 2002, 28). If we structure the language used in terms of predicates, we can conceive the common goal as a high-level predicate (the discourse purpose), which assigns a role to each dialectical move, or discourse segments (Grosz and Sidnert 1986, 178; Walton 1989, 68). This relationship can be visualized in the structure shown below:

Case 2: Discourse coherence

- A: Where shall we go for dinner?
- B: We can go to Bob’s. It is good and really cheap.
- B*: I am going to an expensive restaurant. Go wherever you want.

The third move (Reply 1* in figure 1 below) represents a possible “unreasonable” reply to A’s question. It does not fulfill the role and the presuppositions of the global goal of the dialogue, as the purpose of advancing a proposal is not met by the discourse move. The pragmatic presuppositions of the dialogical predicate fail and the move can be described as an interruption of a dialogue, as opposed to a move that contributes to the dialogue’s properly moving along toward its goal (Fig. 1).

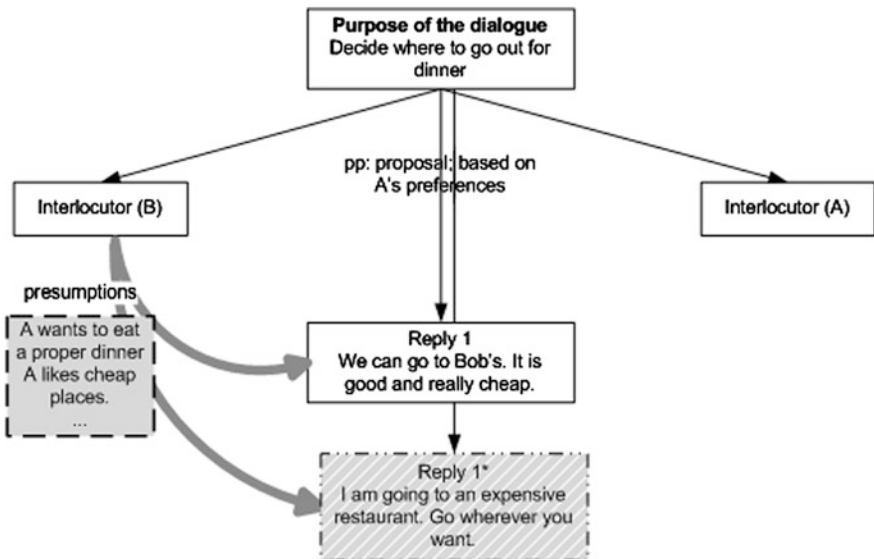


Fig. 1 Purpose of a dialogue as a high-level predicate

2 Interpreting Discourse Relations: Predicate Conditions and Presumptions

The collective purpose of a dialogue imposes some conditions on the possible dialogical moves. However, the interlocutors need to instantiate such generic conditions, and in order to do this they need to know what knowledge the interlocutor shares. For instance, if I want to make a proposal such as “Let’s go to Bob’s”, I take for granted that my interlocutor shares the meaning of the words uttered, the purpose usually conveyed by an imperative sentence, and knows what Bob’s is. However, we cannot know what our interlocutor knows now. All we can tell is that imperatives can convey proposals, that speakers usually know the ordinary meaning of the words, and that usually people remember where they have been before (see also Perelman and Olbrechts-Tyteca 1951). All we can do is to *presume* the existence of such information.

The concept of presumption is used to describe a particular type of inference based on generally accepted principles stating how things usually happen. They are defeasible generalizations, and hold as true until the contrary is proven (see Rescher 1977, 26). Presumptions are forms of inference used in conditions of lack of knowledge. For instance, if it is not known that a person is guilty, he is presumed to be innocent because a person is normally innocent, unless he or she commits some crime. If something has happened in a certain place the previous day, this place can be searched for evidence because it is presumed not have changed in the meantime. Such inferences are therefore strategies to fill a gap of incomplete knowledge by shifting the burden of providing the missing contrary evidence to the other party. Such principles of inference need to be shared and based on the ordinary course of events; in particular, the presumed fact needs to be more likely than not to flow from the proved fact supporting it.

Discourse is based on presumptions. The speaker adopts a conversational behaviour because he presumes that the interlocutor cooperates or does not cooperate. He conveys an intention through an action because such an action is presumed to be associated with an intention. He uses a certain word to convey a specific meaning because it is presumed to have that meaning. Speakers can also breach some presumptions, because they presume that the hearer will rely on a higher-level of presumptions. For instance, indirect speech acts or tautologies are based on the same mechanism of breaching presumptions related to the relationship between speech acts and intentions, and words and meanings. On this perspective, interpretation is an argumentative activity that is carried out based on presumptions and breaches, or rather clashes, of presumptions. While the first type of argumentative process is simply grounded on presumptive, or defeasible modus ponens (Lascarides and Asher 1991, 57) the second process, usually referred to as implicature, hides complex reasoning patterns aimed at explaining presumptive inconsistencies.

On our perspective, implicatures are indirect speech acts of a kind (Bach 1994, 13), whose presumptive meaning differs from the intended one. Such a

discrepancy, caused by a conflict of presumptions, need to be resolved through a process of explanation.

The first crucial aspect is the matter how an implicature is triggered by presumptions. Consider the following interrogative sentences.

Case 3

1. Can you pass the salt?
2. Have you the physical ability to pass the salt?

Such sentences can be used to inquire about a physical ability if uttered to a physically impaired person, or an interlocutor whose arm is in plaster. However, in a context in which it is not evident that such impairment is the case (1) would normally be considered as a polite request, and (2) as a rebuke. Both interrogative sentences are presumed to convey the speech act of asking a question, based on the intention of obtaining needed information. However, when asked in an ordinary context, such a presumption clashes with the stronger one that usually people can perform ordinary actions. The presupposition of the presumed speech act (Van-derveken and Searle 1985, 66–67) conflicts with this encyclopedic presumption and therefore the speech act cannot be felicitous. From such an interpretative failure a need for explanation arises, which can be provided by interpreting the sentence as uttered to perform a non-prototypical act. In this kind of case, the dialogical relation is reinterpreted. We can represent this process as (Fig. 2).

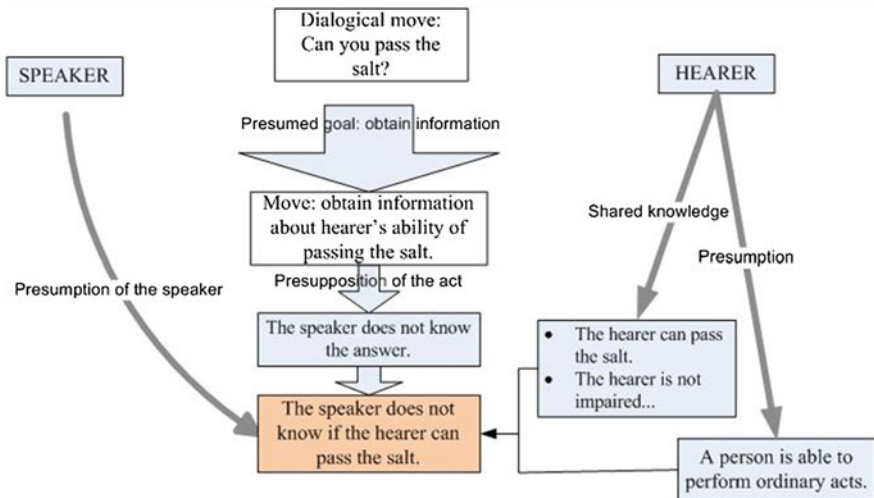


Fig. 2 Presuppositions of an indirect speech act

The speaker, by uttering the interrogative sentence, exploits a presumption which is commonly shared, namely that interrogative sentences are used to ask questions. However, such a move presupposes some conditions, and therefore advances some presumptions that are in conflict with other stronger presumptions (a person in normal conditions is able to perform ordinary acts). The explanation therefore corrects the weaker presumption (the speaker intended to obtain information) with the most reasonable possibility (advance a request). In (2) the explanation is more complex, as the precision of the requested information enters into conflict with a contextual factor (evidence of the interlocutor's ability). The explanation would be that the speaker intended to communicate that the only acceptable reason why the hearer does not pass the salt can be that he is impaired.

This process of reinterpretation of the discourse move (Asher and Lascarides 2006) can be guided by dialogical relations. For instance, we can consider the following cases of particularized conversational implicature:

Case 4

- A. Would you like to go to the theatre tonight?
- B. I have to study.

Case 5

- A. Do you want another piece of chocolate cake?
- B. Is the Pope Catholic?

These examples are forms of particularized conversational implicature, but at the same time are indirect speech acts of a kind. In the first case the hearer, B, uses an assertive, but not to inform the speaker that he has to study. The purpose of his move is to refuse a proposal. In the second case, B does not want to ask a question in order to acquire some missing information, but to accept an offer. In such cases the interpretation of the non-paradigmatic move is guided by the role imposed by the dialogical relation: in these two cases, a proposal or an offer require an acceptance or a refusal, and therefore B's move is presumed to be a reply to the specific proposal or offer.

Sometimes interpretation is guided by dialogical relations drawn from contextual information that further specify the function of the dialogical move. For instance, we can consider the following exchange between a sea-captain and his first mate (Fischer 1970, 272).

Case 6: Drunkard captain

The captain wrote in the ship's log: "The first-mate was drunk all day". When the first-mate read the log, he confronted the captain. The captain replied: "Well, it was true, wasn't it?". The following day the first-mate, whose normal duties include writing up the ship's log, got his revenge. He wrote in the ship's log: "The captain was sober all day".

In this case, the implicature is guided by the dialogical predicate drawn from the contextual information that the sentence has been written in a logbook, and the presumption is that in a logbook only exceptional events are recorded. We can represent the structure of the interpretation procedure as (Fig. 3):

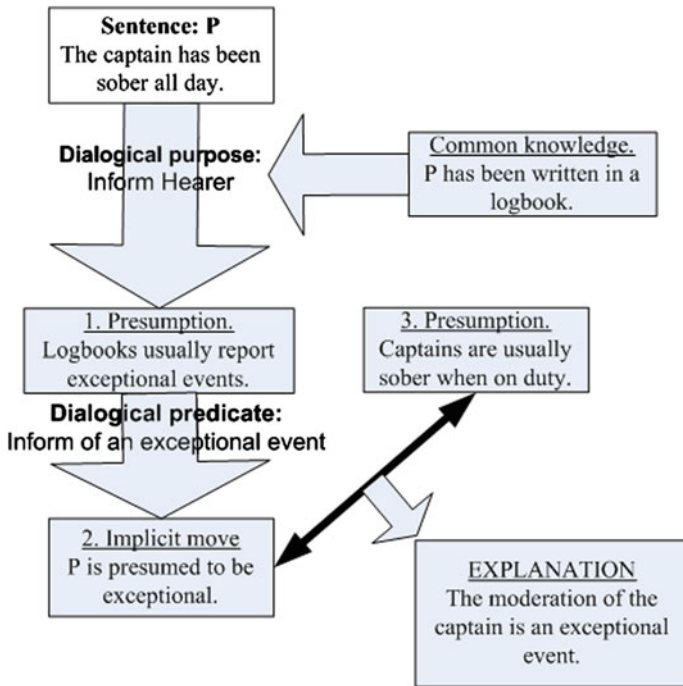


Fig. 3 Explaining conflicting presumptions

This account of implicature shows a crucial relationship between interpretation and dialogue theory in two key respects. First, implicatures need to be explained in terms of dialectical relevance. And second, they need to be analyzed as implicit arguments, involving a pattern of reasoning leading from a specific premise to a conclusion. Such pragmatic and linguistic phenomena can be therefore integrated and developed within dialectical argumentation theory, and can be starting points for developing argumentation theory into a theory of textual interpretation.

3 Conversational Implicatures as Implicit Arguments

On Grice’s view, two types of implicature are distinguished: conventional implicature, dependent on the conventional meaning that certain words have, and generalized and particularized conversational implicature, dependent on the use of

a word or a sentence. The mechanisms of implicature and interpretation set out above constitute the background for analyzing the reasoning structure of particularized implicature. As seen above, dialectical predicates impose precise requirements on the sequences of dialogue; when some inconsistencies arise between dialogical presumptions and encyclopedic presumptions, a process of reasoning is triggered to reconstruct the specific purpose of the move, providing a specific interpretation of the move. This process of reinterpretation is grounded on specific reasoning patterns.

3.1 Reasoning from Best Explanation

In case 6 above the conflict between the presumed exceptionality of the captain's soberness conflicts with the encyclopedic presumption that captains are usually sober on duty. This conflict is explained by retrieving the information that the captain is a drunkard; however, in other contexts other inferences could have been drawn. For instance, if a party was thrown on the ship, the captain would have been praised by that comment. If brought before a court, the statement would have been interpreted literally, without providing any further explanations. The reason for these possibilities can be found in reasoning from best explanation. Any conflict of presumptions needs to be explained, and possible solutions need to be found. Such solutions, however, can conflict with presumptions drawn from the circumstances of the case. On this perspective, the strongest, or best explanation is the explanation that in a given context is more hard to refute by a counter-presumption. The abstract scheme can be represented as an instance of abductive reasoning, or inference to the best explanation (Walton 2002, 44):

Argumentation Scheme 1: Reasoning from best explanation

- *F* is a finding or given set of facts.
- *E* is a satisfactory explanation of *F*.
- No alternative explanation *E'* given so far is as satisfactory as *E*.
- Therefore, *E* is a plausible hypothesis, based on what is known so far.

The argumentative structure underlying the particularized implicature in case 6 can be represented by the diagram in (Fig. 4).

This mechanism is based on the basic presumption that the speaker is using his discourse move for a specific purpose, which in some contexts is stronger than the explanation that he is simply mistaken. Depending on the context and the presumptions of dialogue, the process of explanation ends at the first explanatory step or proceeds further.

A similar example of this process can be found in the following famous case of particularized implicature (Grice 1989, 33).

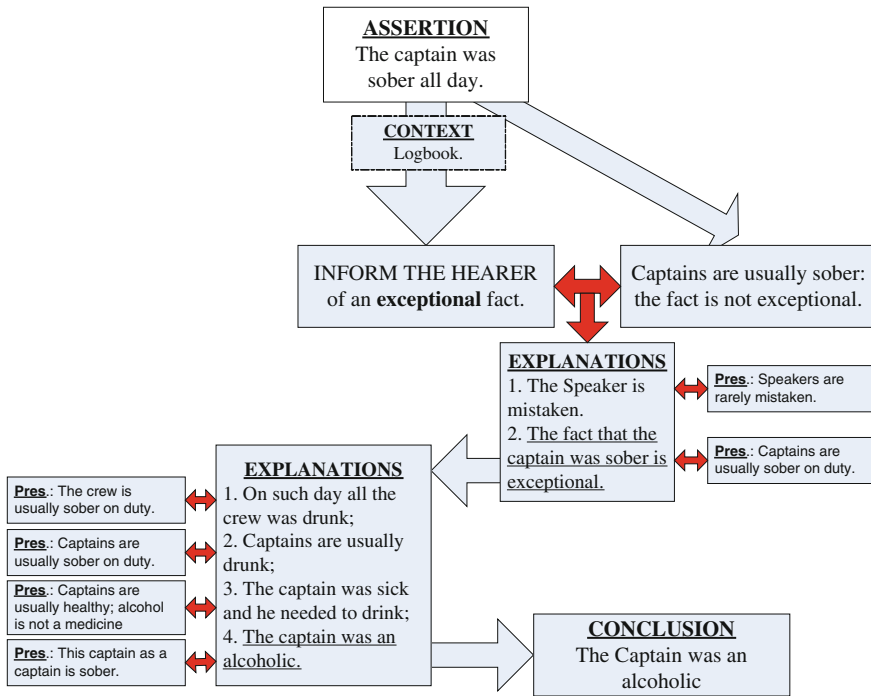


Fig. 4 Argumentative structure of the drunkard captain

Case 7: Recommendation Letter

A professor supposed to write a recommendation letter for his student, writes “Dear Sir, Mr. X’s command of English is excellent, and his attendance at tutorials has been regular. Yours”.

The described qualities cannot fulfill the role of “being an exceptional quality for the position”; therefore, this conflict between dialogical presumption and encyclopedic presumption needs to be resolved, and the best explanation is that no other qualities can be found in the student. The reasoning can be represented as (Fig. 5):

Reasoning from best explanation is not the only pattern of reasoning used in interpretation, but is the one that is directly triggered by the conflict of presumptions. Other types of reasoning however may intervene at the second level of meaning reconstruction.

3.2 Argument from Analogy

The second level of the reasoning underlying the reconstruction (or construction) of the meaning can be based on an argument from analogy. For instance, we can analyze the following examples (Yule 2008, 43–44):

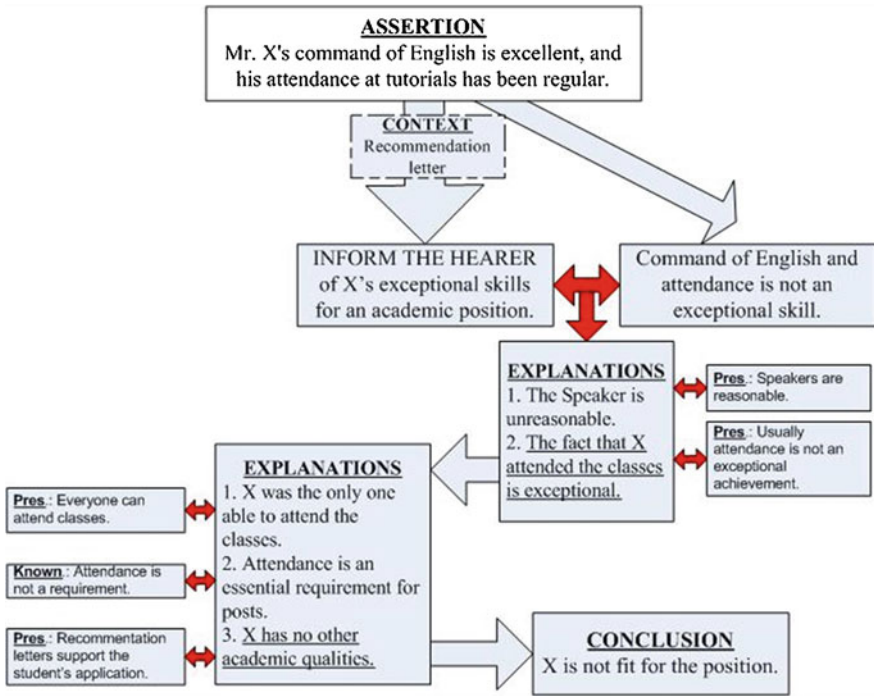


Fig. 5 Reasoning from best explanation—recommendation letter

Case 8

Bert: Do you like ice-cream?

Ernie: Is the Pope Catholic?

Case 9

Bert: Do vegetarians eat hamburger?

Ernie: Do chickens have lips?

In both cases, Ernie is asked obvious questions. However, he does not point out or attack the triviality of the answer; instead, he advances an implicit argument that we can represent as follows (Walton 1995, 135–136):

Argumentation Scheme 2: Argument from analogy

Major Premise: Generally, case C_1 is similar to case C_2 .

Minor Premise: Proposition A is true (false) in case C_1 .

Conclusion: Proposition A is true (false) in case C_2 .

As argued in (Macagno and Walton 2009), the proposition true in both cases is abstracted from the possible properties of the two compared cases. We can

consider it as a functional genus, an abstract category under which the two instances fall not absolutely, but only for the purpose of the argument. For instance we can consider the following example (Grimes 1975, 217):

Trying to do linguistics without any reference to meaning would be like going into battle with one hand tied behind your back.

Here the conclusion, corresponding to the implicit meaning, or conversational purpose of the utterance, is that doing linguistics without considering meaning is useless, or limiting. Such a conclusion is drawn from abstracting a general property that can be shared by the two parallel cases.

In both cases 8 and 9, the meanings of the answers (obviously yes/no) are abstracted from the analogy between the two questions; however, the implicit abstraction from the speech acts advanced as analogous leads to other possible reconstructions. For instance, in the first case the questions can be gathered under the same category of “Questions whose answer is known to everyone”, while in the second case under the genus “Stupid questions”. In both cases an implicit value judgment on the act of asking the question can be drawn. We can represent the implicit argumentation of case 8 as (Fig. 6).

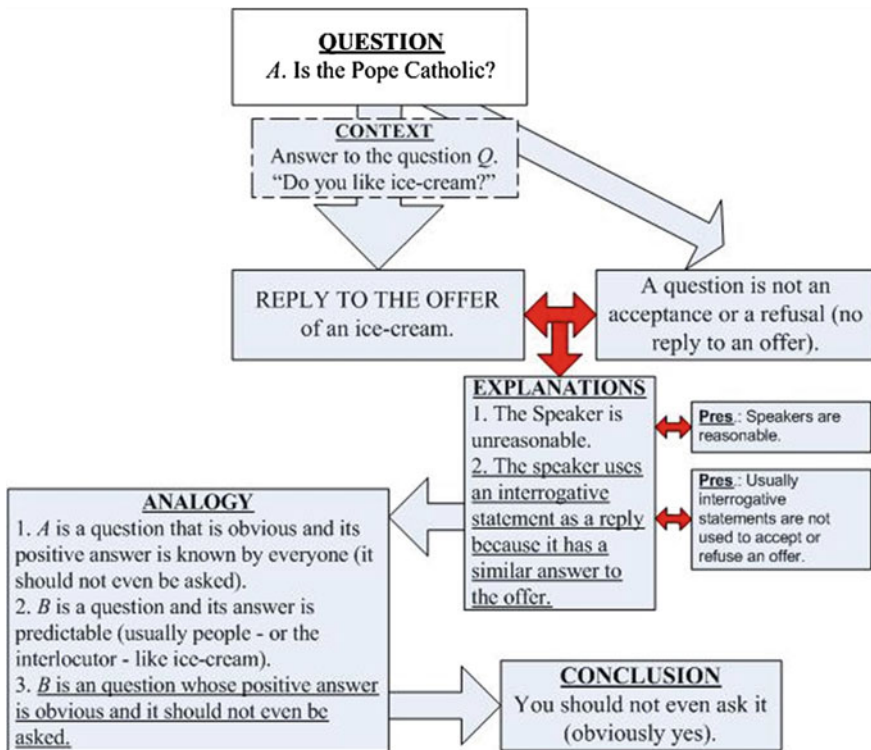


Fig. 6 Reasoning from analogy

Obviously this analysis is simply an approximation of a more complex mechanism. Behind every explanation there is abductive reasoning (see Lascarides and Asher 1991) leading from pieces of evidence (an interrogative sentence is used as a reply) to an explanation (the two speech acts are similar in some respects).

3.3 Practical Reasoning and Other Pragmatic Arguments

Sometimes assertive statements are used not to inform, but to lead the interlocutor to a specific action. In such cases, the inadequacy of the assertion to perform the prototypical action it is associated with can be explained in terms of implicit pragmatic arguments, in which the stated sentence is the only explicit element. For instance, consider the following case (Grice 1975, 51):

Case 10

A: I am out of petrol.

B: There is a garage around the corner.

The statement “I am out of petrol” contradicts a presupposition of the dialectical predicate ‘to inform the hearer’, the fact that the hearer may be interested in the stated information. The speaker cannot presume that the hearer is interested in his condition, or that the stated information is of any interest for him. Therefore, a mechanism of reinterpretation is triggered: lack of petrol is a condition of need (causal relationship), and as people are usually presumed to be interested in trying to help someone who is need when it is possible, the conclusion is that the hearer should be interested in helping the speaker find some petrol. This conclusion, however, does not fulfill the requirement of a proper act of informing, but only the requisite of an act of requesting. The pattern of reasoning on which the link between need and request for help (or appeal to pity) is grounded can be represented using the argument scheme for appeal to pity (Walton 1997, 105):

Argumentation Scheme 3: Appeal to pity

Individual x is in distress (is suffering).

If y brings about A , it will relieve or help to relieve this distress.

Therefore, y ought to bring about A .

In this case, the request for help varies according to the context: depending on whether the interlocutor can provide some petrol, drive the speaker to a gas station, or help in some other fashion, the precise interpretation of the relational predicate varies.

On the other hand, B’s reply is an invitation to bring about a specific action, based on A’s needs. His information is therefore an implicit argument from practical reasoning, in which the needed premise (“in garages it is possible to find

some petrol”) and the conclusion (“you should go to the garage”) are missing. The scheme is the following (Walton 1992, 89–90):

Argumentation Scheme 4: Practical reasoning

- My goal is to bring about *A* (Goal Premise).
- I reasonably consider on the given information that bringing about at least one of [*B*₀, *B*₁, ..., *B*_{*n*}] is necessary to bring about *A* (Alternatives Premise).
- I have selected one member *B*_{*i*} as an acceptable or as the most acceptable necessary condition for *A* (Selection Premise).
- Nothing unchangeable prevents me from bringing about *B*_{*i*} as far as I know (Practicality Premise).
- Bringing about *A* is more acceptable to me than not bringing about *B*_{*i*} (Side Effects Premise).
- Therefore, it is required that I bring about *B*_{*i*} (Conclusion).

These two schemes connect information with actions, and can therefore explain several forms of implicatures in which assertive sentences are interpreted as directives.

3.4 Causal Arguments

The relationship between the presumed purpose of a discourse move, as imposed by the dialogical predicate, and the role such a move is presumed to play can be sometimes based on causal arguments. We can consider the following implicature (Levinson 1983, 126):

Case 11

- A: What on earth has happened to the roast beef?
- B: The dog is looking very happy.

The assertion on the dog’s happiness is presumed not to fulfill the role of “providing information on the roast beef”. However, the relationship between “happiness of the dog” and “information on the roast beef” can be retrieved through an argument from sign (Walton 2002, 42).

Argumentation Scheme 5: Argument from Sign

- Generally, if this type of indicator is found in a given case, it means that such-and-such a type of event has occurred, or that the presence of such-and-such a property may be inferred.
- This type of indicator has been found in this case
- Such-and-such a type of event has occurred, or that the presence of such-and-such a property may be inferred, in this case

The expression of the dog is an indication of satisfaction; such satisfaction can be specified by the reference to a food, and therefore can be interpreted as the result of eating the roast beef. The same relation from sign can explain the following implicature.

Case 12

- A: What time is it?
 B: Some of the guests are already leaving

Here there are several possible reasons for guests to leave; however, the function of the discourse move is to provide a time indication. The leaving of the guests is therefore presented as an indication of time, based on the presumption that late time causes guests to leave.

In the aforementioned cases the information guiding the interpretation and specification of the argument scheme is provided by the dialogical information (context). Such a specification can be also provided by the shared information. We can consider the following case (Carston 2002, 109).

Case 13

Bill goes up to Scotland every weekend.

As Carston points out (Carston 2002, 110), “in different specific contexts this could implicate ‘Bill’s mother is ill’, ‘Bill has a girlfriend in Scotland’, ‘Bill gets as far away from London as he can when he can’, ‘Bill still hasn’t got over his obsession with the Loch Ness monster’, etc.” Here the sign is provided by information about Bill and specific presumptions related to specific behaviors. For instance, the information that Bill has a mother in Scotland, or that he does not have any girlfriend in London, or that he is obsessed with the Scottish monster triggers the presumption that “people tend to stay closer to what they love or care (or love or necessity cause people to desire to see each other)”, while the fact that Bill hates London elicits the presumption that “people tend to escape what they hate (or bad conditions cause people to leave)”.

Argument from sign can be conceived as an instance of causal argumentation, linking an event to its effects. The general scheme of argument from cause can be represented as follows (Walton 1995, 140):

Argumentation Scheme 6: Argument from Cause

- Major Premise: Generally, if *A* occurs, then *B* will (might) occur.
 Minor Premise: In this case, *A* occurs (might occur).
 Conclusion: Therefore in this case, *B* will (might) occur.

A clear example would be case 4 above, in which the interlocutor instead of answering negatively to the proposal of going to the theatre replies that he “has to

study”. Here, the goal of the utterance is not to inform the interlocutor about the speaker’s duties, but to lead him to draw a conclusion from cause to effect. Study is presumed to be incompatible with going to the theatre: if someone studies, he needs to stay at home, and if someone stays at home, he cannot at the same time see a movie. The latter causal relationship is presented as an alternative: either *A* or *B*; not *A*; therefore *B*.

This representation of implicatures can explain both the reasoning mechanism underlying reinterpretation of the discourse relation (such as in case 6, 7 and 10) or the reinterpretation of the role of the discourse move (or sequence). In all the aforementioned cases the reasoning depends on specific information provided by the context and the co-text. Sometimes, however, the argumentative procedures are context-independent and are simply related to the linguistic structure.

4 Conventional Implicatures as Implicit Arguments

If we analyze Grice’s generalized and conventional implicatures in this framework, we can notice that they represent two types of implicitness, which we will refer to as implicature within predicates, or presuppositions, and implicatures within paradigms, or implicit arguments from alternatives.

Grice notices that connectors like ‘therefore’ convey an implicit meaning which is within the word itself (Bach 1999). For instance consider the following sentence (Grice 1989, 25).

1. He is an Englishman; he is, therefore, brave.

It would be infelicitous to say “He is an Englishman; it is, therefore, sunny outside”. Grice realizes that ‘therefore’ conveys the implicit information that there is a relationship between ‘being an Englishman’ and ‘being brave’. Consider the following case.

2. Ben is nice but he drives a Ford Capri.

In 2 there is a contrast between “Ben is nice” and “Ben drives a Ford Capri” that is taken for granted (Carston 2002, 108). Such conventional implicature can be explained as a presupposition of the predicate (see Ducrot 1968, 1972; Barker 2003, 5), that is, a condition or requirement on the predicate affecting the felicity of the speech act.

Grice (1989, 37) pointed out another type of implicit content conveyed by some utterances. He analyzed sentences of the following kind.

- 3. The flag is white
- 4. X is meeting a woman this evening
- 5. X went into a house yesterday and found a tortoise inside the front door
- 6. Some guests are already leaving

Grice (1989, 51–52) noted that these sentences, when uttered in any type of context, always convey the same implicit information, namely:

- 3a. The flag is not white and other colors/the flag is white even though stained with blood.
- 4a. X is meeting a woman that the interlocutor does not know.
- 5a. X went into a house that is not specific (not his house).
- 6a. Not all the guests are already leaving.

These inferences not only depend on the semantic features of the quantifiers, but also on the paradigm of possible alternatives. The inference is not exclusively drawn from an implicit content of the word, but from the structure of the word selection. If we consider language as a matter of choices between linguistic items used to convey a meaning, depending on the type of information we need to convey we need to select one possibility while excluding others. For instance, consider the following sentence.

The computer is heavy.

We have made a choice among other predicates that can be attributed to ‘computer’, and convey a specific type of information, namely information about its weight. We can represent such a choice as shown in (Fig. 7).

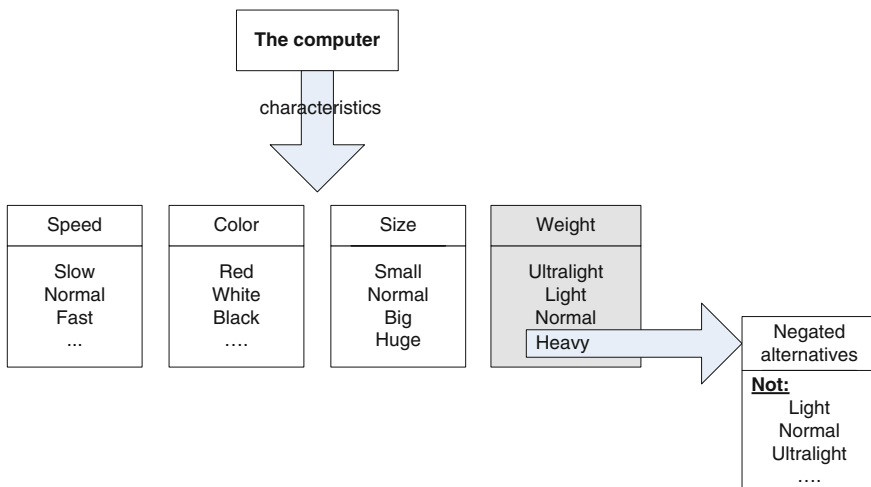


Fig. 7 System of choices in language

The affirmation of a predicate, on this perspective, is an implicit negation of the other possible alternatives (see Pap 1960, 53) falling within the same genus, namely the same semantic characteristic which is contextually identified. Conventional implicatures can be therefore described as a particular type of disjunctive syllogism applied to natural language, in which an element is presented as an alternative to a set of possible choices falling within the same semantic category, or genus. This type of reasoning is grounded on Aristotle's semantic principle expressed in his (*Topics*, 121a27–121a38), stating that the genus cannot be predicated of an entity unless one of its species is predicated of it too. We can represent the reasoning as follows:

Argumentation Scheme 7: Argument from alternatives

- x is A
- A belongs to the genus β
- If something is β , it is either A , or B , or C ...
- Therefore x is not B (not C , etc.)

Identifying the semantic genus is crucial for the process of interpretation, because it also determines and specifies the exact semantic feature of the noun to which the predicate is referred. For instance, if I say that a computer is fast, I am considering the computer as an electronic device, whereas if I describe it as heavy I am considering it merely as an object.

This structured framework can account for some other kinds of implicatures. For instance, if we consider the sentence “the flag is white” we implicate that the flag is *only* white, and not, for instance, white and red, even though it does not exclude that it may be dirty (stained with blood) (Levinson 1995, 97). The semantic paradigm is the symbolic color, and therefore the sentence specifies the ‘flag’ as the most evident part of the emblem, and not as fabrics, nor as the frame. The semantic paradigm ‘symbolic color’ not only specifies the paradigm of predicates that can be attributed to the flag, but also the concept of ‘flag’. A flag cannot be transparent, nor can the colors be mixed in order to be identified as such. It is as an emblem that the flag is white, but it is as a piece of fabric that it is red with blood (Fig. 8).

Semantic paradigms therefore work by specifying the semantic features of the object of the predication, and the possible choices. This structured process is a form of interpretation, since establishing the semantic predicate governing the paradigm means retrieving the meaning of the sentence. Depending on the type of paradigm, and therefore the governing semantic predicate, or paradigmatic genus, the meaning varies.

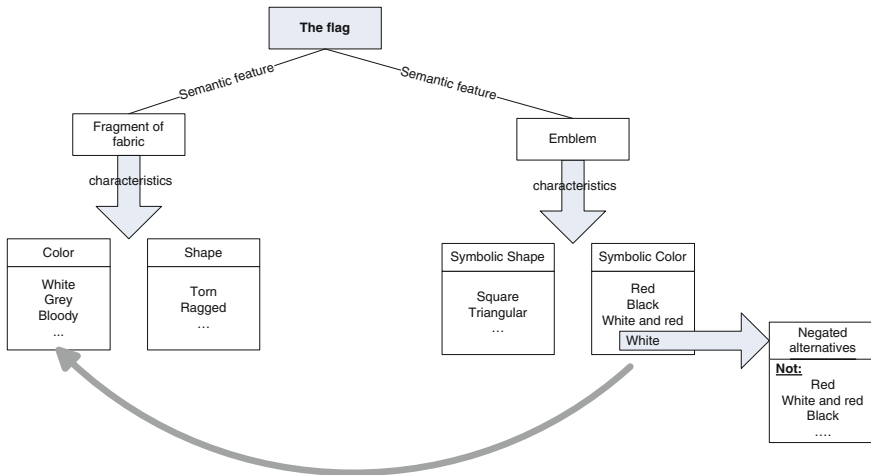


Fig. 8 Semantic paradigms and the specification of aspects of meaning

Not only can the meaning of predicates be determined in this fashion, but also the interpretation of quantifiers. Let’s reconsider cases 5 and 6. The indefinite article conveys the information that the noun which is determined is not known by the other party, or is not specific for the speaker himself [for a treatment of articles and their paradigm of specificity see (Haspelmath 1997)]. It contrasts with the definite article, which is used to represent the assumption that the entity is part of the domain of the hearer’s knowledge or can be retrieved by him. In the semantic paradigm characterized by the semantic trait of “uniqueness”, only two alternatives are possible, namely the specific and non-specific article. The choice of the second excludes that the entity is known, or knowable by the hearer. Similarly, considering the more generic paradigm of quantifiers, in 7 ‘some’ is contrasted with ‘all’, ‘none’ and ‘one’. Therefore, the speaker, by making such choice, excludes that only one guest is leaving, that it is not true that no guests are leaving, and that not all the guests are leaving.

Such mechanisms of specification allow one to explain several types of processes of interpretation, from prepositions to quantifiers or predicates.

5 Conclusion

Grice's theory of implicatures has been used in this paper to bring out aspects of the previously unclear relationship between reasoning and meaning. We have shown how conversational implicatures represent implicit meaning triggered by the use of a sentence and how they can be considered interpretations of the meaning of a word or a speech act. On this perspective, conversational implicatures are triggered by conflicts of dialogical and epistemic presumptions that are resolved by a process of best explanation, which in turn are based on argumentation schemes such as inference to the best explanation, practical reasoning, argument from sign, appeal to pity and analogy. Depending on the context, the presumptions on which the process of explanation is based on vary, and therefore the conclusions of the implicit arguments can be different. The argumentative structure underlying conventional implicatures stems from a different reasoning process, based on a linguistic structure. In this case, the interpretation of a word's meaning is carried out through the exclusion of its alternatives within its semantic paradigm, or genus.

References

- Aristotle. 1984. *Topics*. Translated by W. A. Pickard-Cambridge. In *The complete works of Aristotle*, vol. I, ed. Jonathan Barnes. Princeton, NJ: Princeton University Press.
- Asher, Nicholas, and Alex Lascarides. 2006. Indirect speech acts. *Synthese* 128: 183–228.
- Bach, Kent. 1994. Meaning, speech acts, and communication. In *Basic topics in the philosophy of language*, ed. Robert Harnish, 3–21. Englewood Cliffs: Prentice-Hall.
- Bach, Kent. 1999. The myth of conventional implicature. *Linguistics and philosophy* 22: 237–366.
- Bach, Kent. 2003. Speech acts and pragmatics. In *Blackwell. Guide to the philosophy of language*, eds. Michael Devitt and Richard Hanley, 147–167. Oxford: Blackwell.
- Ballard, Lee, Robert Conrad, and Robert Longacre. 1971. The deep and surface grammar of interclausal relations. *Foundations of language* 7(1): 70–118.
- Barker, Stephen. 2003. Truth and conventional implicature. *Mind* 112(445): 1–33.
- Carston, Robyn. 2002. *Thoughts and utterances*. Malden: Blackwell.
- Crothers, Edward. 1979. *Paragraph structure inference*. Norwood: Ablex.
- Davis, Wayne. 2007. How normative is implicature? *Journal of pragmatics* 39: 1655–1672.
- Ducrot, Oswald. 1968. Le structuralisme en linguistique. In *Qu'est-ce que le structuralisme?*, eds. Oswald Ducrot, Todorov Tzvetan, Dan Sperber, Moustafa Safouan, and Wahl François, 13–96. Paris: Seuil.
- Ducrot, Oswald. 1972. *De Saussure à la philosophie du langage, preface to John Searle, Les actes de langage, 7–34*. Paris: Hermann.
- Fischer, David. 1970. *Historians' fallacies*. New York: Harper & Row.
- Goodwin, Jean. 2001. The noncooperative pragmatics of arguing. In *Pragmatics in 2000: Selected papers from the 7th International Pragmatics Conference*, vol. II, ed. Enikő Nemeth, 263–277. Antwerp: International Pragmatics Association.
- Grice, Paul. 1961. The causal theory of perception. *Proceedings of the Aristotelian society, Supplementary* 35: 121–152.

- Grice, Paul. 1975. Logic and conversation. In *The logic of grammar*, eds. Donald Davidson and Gilbert Harman, 64–75. Encino, California: Dickenson.
- Grice, Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.
- Grimes, Joseph. 1975. *The thread of discourse*. The Hague: Mouton.
- Grosz, Barbara, and Candice Sidnert. 1986. Attention, intentions, and the structure of discourse. *Computational linguistics* 12(3): 175–204.
- Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.
- Hobbs, Jerry. 1979. Coherence and coreference. *Cognitive science* 3: 67–90.
- Hobbs, Jerry. 1985. *On the coherence and structure of discourse*. Report No. CSLI-85-37, Center for the study of language and information, Stanford University.
- Lascarides, Alex, and Nicholas Asher. 1991. Discourse relations and defeasible knowledge. In *Proceedings to the 29th annual meeting of the association of computational linguistics (ACL91), Berkeley USA*, 55–63. Stroudsburg, PA: Association for computational linguistics.
- Levinson, Stephen. 1995. Three levels of meaning. In *Grammar and meaning*, ed. Frank Palmer, 90–115. Cambridge: Cambridge University Press.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Macagno, Fabrizio. 2008. Dialectical relevance and dialectical context in Walton's pragmatic theory. *Informal logic* 28(2): 102–128.
- Macagno, Fabrizio, and Douglas Walton. 2009. Argument from analogy in law, the classical tradition, and recent theories. *Philosophy and rhetoric* 42(2): 154–182.
- Pap, Arthur. 1960. Types and meaninglessness. *Mind* 69(273): 41–54.
- Perelman, Chaim, and Lucie Olbrechts-Tyteca. 1951. Act and person in argument. *Ethics* 61(4): 251–269.
- Rescher, Nicholas. 1977. *Dialectics: a controversy-oriented approach to the theory of knowledge*. Albany: State University of New York Press.
- Rigotti, Eddo. 2005. Congruity theory and argumentation. *Studies in communication sciences*, Special Issue: 75–96.
- Sperber, Dan, and Deidre Wilson. 1986. *Relevance: Communication and cognition*. Oxford: Blackwell.
- Thomason, Richmond. 1990. Accommodation, meaning, and implicature: Interdisciplinary foundations for pragmatics. In *Intentions in communication*, ed. Philip Cohen, Jerry Morgan, and Martha Pollack, 325–363. Cambridge: MIT Press.
- Vanderveken, Daniel, and John Searle. 1985. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press.
- Vanderveken, Daniel. 2002. Universal grammar and speech act theory. In *Essays in speech act theory*, ed. Daniel Vanderveken, and Susumu Kubo, 25–62. Amsterdam: John Benjamins.
- Walton, Douglas, and Erik Krabbe. 1995. *Commitment in dialogue*. Albany: State University of New York Press.
- Walton, Douglas. 1989. *Informal logic*. New York: Cambridge University Press.
- Walton, Douglas. 1990. What is reasoning? What is an argument? *Journal of philosophy* 87: 399–419.
- Walton, Douglas. 1992. *Slippery slope arguments*. Newport News: Vale Press.
- Walton, Douglas. 1995. *A pragmatic theory of fallacy*. Tuscaloosa and London: The University of Alabama Press.
- Walton, Douglas. 1997. *Appeal to pity*. Albany: State University of New York Press.
- Walton, Douglas. 1998. *The new dialectic, conversational contexts of argument*. Toronto: University of Toronto Press.
- Walton, Douglas. 2002. *Legal argumentation and evidence*. University Park: The Pennsylvania State University Press.

- Walton, Douglas. 2010. Types of dialogue and burden of proof. In *Computational models of argument: Proceedings of COMMA 2010*, eds. Pietro Baroni, Federico Cerutti, Massimiliano Giacomin, and Guillermo Simari, 13–24. Amsterdam: IOS Press.
- Wilson, Deirdre, and Dan Sperber. 2004. Relevance theory. In *Handbook of pragmatics*, ed. Laurence Horn, and Gregory Ward, 607–632. Oxford: Blackwell.
- Yule, George. 2008. *Pragmatics*. Oxford: Oxford University Press.

Some Remarks About Speech Act Pluralism

Marina Sbisà

Abstract This paper moves from a well known pragmatic theme in the philosophy of language—speech acts—to the philosophy of action. It focuses on the problem whether an utterance token can carry out more than one speech act, to which the attention of philosophers of language was drawn by Cappelen and Lepore's (2005) defense of speech act pluralism. It is argued that Austin's notions of the locutionary, illocutionary and perlocutionary acts already presuppose speech act pluralism. In particular, it is shown that the notion of perlocutionary act, as defined by Austin, makes sense only within a pluralist framework, where no additional behavior on the part of the speaker is needed in order for her performance of a perlocutionary act to take place, besides those of the locutionary and illocutionary acts. The central section of the paper discusses speech act pluralism as it applies to illocutionary acts, in the light of an Austin-inspired view of illocution developed by the author in her previous work and of the analysis of the conversational practices of multi-action ascription recently outlined by Levinson. As to the locutionary act, it is argued that plural ascriptions are by far more limited than is claimed by Cappelen and Lepore, if they exist at all. In conclusion, it is shown that all the views admitting of speech act pluralism that are examined in the paper (those of Cappelen and Lepore, Levinson, and Sbisà) require an ascription-centered conception of action, such as that outlined by Austin in his papers in the philosophy of action.

M. Sbisà (✉)
University of Trieste, Trieste, Italy
e-mail: sbisama@units.it

1 What is Speech Act Pluralism?

I call here “speech act pluralism” the view according to which an utterance token can carry out more than one speech act.

This view, albeit concerning speech acts (actions we do with words), is not compatible with most received versions of speech act theory, since in these, the speech act is generally considered as the expression of a single (albeit complex) intention of the speaker. Cases in which two speech acts are recognized to be somehow performed in issuing one and the same utterance, such as that of a question that counts as an indirect request, are taken to be cases in which the “secondary” speech act (the question) is merely the means by which the “primary” one (the request) is performed (Searle 1979: 33), which is the only act that the speaker had the actual communicative intention to perform.

Herman Cappelen and Ernie Lepore argue for speech act pluralism in the last chapter of their (2005), after defending semantic minimalism against contextualism. Their version of speech act pluralism is, as they themselves admit, more a collection of observations than a full-fledged theory. They note that an excerpt from a conversation typically includes few (if any) grammatically complete sentences, that in order to react properly to what the speaker said (or to report it, etc.) we have “to reconstruct utterances up to a point where they express thoughts”, and that no one way of doing so is uniquely correct (2005: 192): many different pieces of information concerning the speech event, its participants, its context and in general the world might be pertinent to the task, so that the consideration of each possible subset of them yields a lot of different ascriptions of speech act content (and therefore, of speech acts). They state this idea most explicitly in their First Thesis of Radical Speech Act Pluralism: “No one thing is said (or asserted, or claimed, or...) by any utterance: rather, indefinitely many propositions are said, asserted, claimed, stated, etc.” (Cappelen and Lepore 2005: 199).

Cappelen and Lepore are willing to affirm and defend various consequences of their version of speech act pluralism. For example, they reject the received assumption that speech act content is fixed by facts about the speaker, her audience, and their common context (an assumption they call “Original Utterance Centrism”) and argue that assumptions made by people who have nothing at all to do with the original context of an utterance can also contribute to fixing what the speaker said (2005: 201). This may sound counter-intuitive, but it is consistent with their First Thesis. They also claim (2005: 202) that speakers do not have privileged access to the content of their speech acts: indeed, if the audience’s beliefs or knowledge legitimately influence the ascription of speech act content, such a content may well comprise ideas that are completely foreign to the speaker.¹

¹ Cappelen and Lepore (2005) also connect their claims about speech acts and their contents to their own semantic minimalism, but here I do not address this aspect of their speech act pluralism.

Cappelen and Lepore's speech act pluralism appears as an isolated challenge to philosophers interested in speech acts, context, and pragmatics in general. However, it does not come in absolute isolation. In the neighboring field of Conversation Analysis, a version of speech act pluralism has recently been taken into serious consideration (though not by that name).

Although in CA-inspired research it is possible to distinguish a received tendency to speech act monism, that is, an inclination to ascribe to each turn no more than one action, CA has always been more sensitive than speech act theory to the many, multifarious activities we carry out when we speak and to the complexities of the organization of verbal interaction. It has thus recognized that the relationship between turns and actions is not simple: on the one hand, a turn can perform more than one action, if it comprises distinct turn-constructive units onto which the actions can be mapped; on the other hand, the words uttered in a turn are often multi-action compatible, that is, liable to have different actions mapped onto them. These complications do not amount *per se* to genuine speech act pluralism, since with the former, a one-to-one correspondence at least between turn-constructive units and actions is preserved, while with the latter, multi-action compatibility may be attributed to the words (the utterance type) as opposed to the utterance token itself. But they go some way towards it. If the same utterance token is part of two action sequences belonging to different interactional "tracks" (Goffman 1974) or "projects" (Levinson 2011), multi-action compatibility may enable it to play a distinct role in each. Moreover, in multi-action turns the different actions performed are not always distributed over turn-constructive units, but at least sometimes may be attributed to the turn as a whole (as Levinson 2011 notes), for example in the presence of a response that recognizes the turn's double job.

The theme of speech act pluralism has been developed also in the sociological and socio-semiotic study of agency. Recently, François Cooren (2010) has claimed that one and the same event or stretch of behavior can not only be ascribed more than one action, but also be viewed as constitutive (under different respects) of actions of different agents. This implies a vertiginous plurification not only of actions (and speech acts), but of agencies too.

In this context of multi-disciplinary interest for speech act pluralism, it seems to me interesting to explore where pluralist hypotheses about speech acts come from, what aspects of the speech act they can consistently be applied to, and which facets of speech or of verbal interaction they shed light upon. In this paper, I do three things in this direction. First, I go back to Austin's notions of the locutionary, illocutionary and perlocutionary acts and argue that these distinctions already presuppose speech act pluralism. Then, I discuss speech act pluralism as it applies to illocutionary acts, in the light of the Austin-inspired view of speech acts I have developed in my own research and reflection (Sbisà 1984, 1989, 2001, 2002a, 2002b). Finally, I come back to Cappelen and Lepore's remarks about speech act pluralism and briefly discuss some of them.

2 Roots of Speech Act Pluralism: The Locutionary-Ilocutionary-Perlocutionary Distinction

When Austin introduces his notions of locutionary, illocutionary and perlocutionary act, and with them those of phatic, phonetic and rhetic act (1975: 91–98), at first sight he appears to be analyzing the speech act into parts. The phonetic, phatic and rhetic act (one might say) add up to the entire locutionary act; the locutionary act (providing meaning) and the illocutionary act (providing force) add up to the entire speech act. The perlocutionary act appears to be an additional and perhaps optional component, which may come into what Austin emphatically referred to as the “total speech act in the total speech situation” (1975: 52; 148).

This quite natural reading of Austin’s distinctions is wrong, however. Let us first consider the phonetic, the phatic and the rhetic act. Two models representing them as parts or components of a whole are conceivable: an additive model in which the phatic act is added to the phonetic act and the rhetic act is added to the result of the former addition to make up the entire locutionary act, and an embedding model, in which the phonetic act is part of the phatic act and the phatic act is a part of the rhetic act, which then coincides with the entire locutionary act. But neither applies to Austin’s definitions. The phonetic act is “the act of uttering certain noises”, while the phatic act consists of “uttering noises of certain types belonging to and as belonging to a certain vocabulary, [...] conforming to and as conforming to a certain grammar, with a certain intonation, etc.”, which seems to comprise the uttering of noises as a part (1975: 92): this might lead us to prefer the embedding model. But the phatic act does not always include the phonetic one. For example, there are aspects of the phonetic act that are not relevant to the identification of the phatic act that the speaker has performed in issuing her utterance. Pitch distinguishes one phonetic act from another, but the pitch of the voice uttering a sentence is (in Western culture) largely irrelevant to the identification of the phatic act that has been so performed. The phatic act, in turn, is not a part of the rhetic act, since the rhetic act can and often must be identified without reporting the precise words uttered by the speaker (cf. Austin 1975: 95). For example, in the cases of tense and of indexicals, the rhetic act performed by A in saying “*I am hungry*” at noon on 01/01/2001 can be correctly reported by saying “On 01/01/2001, noon, A said that *she was hungry*”. So it seems that the words “I” and “am” are not part of what identifies the rhetic act performed by A, while we must report them if we want to identify the phatic act. Finally, the rhetic act is not to be identified with the locutionary act (as it should, as a consequence of the embedding model): to report a complete locutionary act, Austin (1975: 101–102) uses a complex formula bringing together phatic and rhetic aspects, so that it seems that, to this aim, the mere indirect report (which identifies the rhetic act) is not enough. So, the embedding model does not work.

One may attempt to adopt the additive model and say that the phonetic, phatic and rhetic act are non-overlapping parts of the locutionary act. This model allows for a distinction between the rhetic and the locutionary acts: the former is that part

of the latter which is missing when we have only the phonetic and phatic acts. But while the rhetic act can perhaps be taken as a non-overlapping addition to the phatic and phonetic acts, the phatic act, as defined by Austin, is not a real addition to the phonetic act: as Searle (1968) observed, the fact that the sounds uttered conform to certain rules, and are uttered as conforming to them, is not a new act to be added to the uttering of the sounds, but a manner in which they are uttered. Thus, the additive model too fails to account for Austin's distinctions concerning the locutionary act. Neither model works: there is no way of construing the phonetic, phatic and rhetic acts, as defined by Austin, as parts or components of the locutionary act.

If phonetic, phatic and rhetic acts do not add up to the locutionary act, there is no act encompassing them as parts or components. So, how many "acts" does the speaker perform when she says something? The answer is simple: in the Austinian perspective the phonetic, phatic, and rhetic acts, together with the locutionary act, are aspects of the total speech act, and not parts or components; moreover, each can be considered as an act, or even an "action", to be identified by reference to what it does or what it brings about, that is, by reference to its effects, for which the speaker can be ascribed responsibility. So, the locutionary act (the lowest step in the locutionary-illocutionary-perlocutionary distinction) is already one and many.

Coming to the locutionary-illocutionary distinction, here too one might want to adopt either an additive or an embedding model. The additive model is definitely unattractive: it represents meaning and force as produced independently of one another and suggests that they might be expressed by separable parts of the utterance. This runs contrary to Austin's characterization of illocutionary acts as those acts that are performed "in saying" something, which assumes, so to speak, a co-presence of two actions (one of which is the vehicle of the other) in one and the same stretch of verbal behavior. So, we should not use the additive model to explain the Austinian relationship between locutionary and illocutionary act. But we cannot use the embedding model either. In an embedding model, the outmost level of the model, which includes the inner levels, turns out to be identical to the whole. In such a model, the illocutionary act turns out to be identical to the entire speech act, as assumed by Searle (1969: 23) (and not by Austin). In Searle's theory, as is known, the locutionary act disappears and is replaced by an "utterance act" and a "propositional act", which are more suitable for combination according to the embedding model. But in Austin, again, the locutionary and illocutionary acts are not characterized as parts or components of the whole speech act either according to either an additive or an embedding model. Rather, like the phonetic, phatic, and rhetic acts within the locutionary act, they are two aspects of the total speech act, each of which can be considered as an act, or (as Austin insists, cf. 1975: 106) an "action". It seems inescapable to conclude that, in the Austinian perspective, one and the same stretch of verbal behavior does perform at least two different actions—a locutionary act and an illocutionary one.

To throw more light on this conclusion I would like to recall a quite mysterious passage of Austin's where he claims that the distinction among acts such as the

locutionary, the illocutionary and the perlocutionary acts is made apparent when something goes wrong with one of them: “typically we distinguish different abstracted ‘acts’ by means of the possible slips between cup and lip, that is, in this case, the different types of nonsense which may be engendered in performing them” (1975: 147). In this passage, Austin uncovers his conviction that conceptual investigations are often best carried out by considering marginal cases and failures, which illuminate structures and distinctions that one would not think of while contemplating “normal”, successful cases. As regards actions in particular, his writings clearly show not only that he considers the viability of action ascription as indispensable for a stretch of behavior to be the performance of an action, but also that he takes the ways in which we describe failures as a cue to the action’s own structure (Austin 1979: 193–194).² One might say that remarks on locutionary and illocutionary failures illuminate the structure of the “total speech act”. But, considering that the separate aspects of the total speech act are taken by Austin to be themselves *actions*, such remarks should rather be understood as illuminating the structures of the locutionary and the illocutionary act themselves. The “total speech act” remains, in Austin’s thought, something like a theoretical target, warning us that *all* aspects of a speech act are relevant to the philosophical task of elucidating what we do when we speak. Austin’s use of that term should not be taken to imply that there is an all-encompassing way to deal with speech acts as units of communication or interaction, over and above the analysis of speech and conversation at the locutionary, illocutionary, and possibly perlocutionary level.

There is, finally, something specific to say with regard to the distinction between the illocutionary act and the perlocutionary act. The perlocutionary act is defined by Austin as consisting of the production of “consequential effects upon the feelings, thoughts, or action of the audience, or of the speaker, or of other persons” by one’s locutionary and illocutionary acts (1975: 101). It is clear from this definition that no performance of the speaker is needed in order for it to be accomplished, apart from the carrying out of that very stretch of behavior which is constitutive of the locutionary and illocutionary acts. Indeed, the perlocutionary act involves something that goes beyond that stretch of behavior, namely, the addressee’s psychological or behavioral response to it: so, as claimed by Gu (1993), it may seem that it is more an accomplishment of the addressee than one of the speaker. This makes the notion of perlocutionary act appear mysterious and even paradoxical. Mystery and paradox fade away, though, once speech act pluralism is introduced and it is admitted that we do not need additional verbal behavior on the part of the speaker in order to recognize her performance of a perlocutionary act (in addition to those of the locutionary and illocutionary acts).

² It is ironic that Jacques Derrida (1977) should have accused Austin of lacking appreciation for failures, when most of Austin’s philosophy was actually guided by attention to failures and an awareness of their importance.

It is interesting to observe, moreover, that the perlocutionary act performed by a certain utterance token need not necessarily be only one. Even in one and the same context, the speaker's utterance can produce more than one contextually relevant consequence on the addressee (not to mention other participants). If an order is obeyed, the speaker has succeeded in making the addressee do something; but if, while obeying, the addressee gets angry at the speaker for imposing on him a task he dislikes, the speaker has also produced the additional consequence of making the addressee get angry, which is a second perlocutionary act. This confirms the indispensability of speech act pluralism, or perhaps, more generally, of some form of action pluralism.

3 Illocutionary Pluralism

I would now like to focus on the illocutionary act and consider whether, in issuing an utterance, we can perform more than one illocutionary act.

The reply to this question obviously depends on the definition or characterization of the illocutionary act one assumes to hold. Within a view that identifies the illocutionary act with the whole speech act and takes it to be the expression of a single (albeit complex) intention of the speaker, speech act monism holds and, *a fortiori*, there is no room for illocutionary pluralism. At most, it can be conceded that the same sentence or string of words, uttered in different contexts (to different addressees, for different aims, etc.), may express different communicative intentions. But this is not yet speech act pluralism: in each context and occasion of use, the sentence or string of words uttered will in fact embody no more than one illocutionary act. Our question is, rather, whether there are cases in which one and the same utterance token (which already comes with a context) is the vehicle of more than one illocutionary act.

3.1 A View of Illocutionary Acts

In my research on speech acts, I have proposed, developed and applied a view of illocutionary acts as bringing about changes in interpersonal relationships (see e.g. Sbisà 1984, 2001), which is compatible with speech act pluralism and with illocutionary pluralism in particular.

I shall introduce the main lines of this view concisely, starting from some examples:

- a command changes the interpersonal relationship between speaker and addressee, so that the addressee has a new obligation;
- a promise changes the interpersonal relationship between speaker and addressee, so that the speaker has a new obligation;

- an apology changes the interpersonal relationship between speaker and addressee so that the speaker is freed from a debt she had with respect to the addressee;
- a warning changes the interpersonal relationship between speaker and addressee so that the addressee can no longer hold the speaker responsible for the situation addressed by the warning and its possible consequences;
- a statement changes the interpersonal relationship between speaker and addressee so that in subsequent speech or behavior, the addressee is licensed to rely upon the results of the speaker's judgement.³

These examples are limited in number and incomplete. They are limited in number, not only because there are only five of them, but because the number of illocutionary acts (types, sub-types or even tokens) we can describe this way is enormous if not endless. Following Austin (1975: 150), one may want to specify the meaning of all the verbs that have a possible performative use in terms of effects on the interpersonal relationship. But one may also want to add to such descriptions a certain number of generalizations, specifying the prototypical effects of groups or classes of illocutions, such as “verdictives”, “exercitives”, “commissives”, “behabitives” (Austin’s terminology), and perhaps also “directives”, “expressives”, “permissives”, “interrogatives”, and so on.⁴ Finally, one may want to specify what change in the interpersonal relationship of the participants has been brought about by a certain utterance on a certain occasion. This can be done, I believe, with any utterance token whatsoever, and the resulting descriptions will very often include *ad hoc*, idiosyncratic elements, thus differing from one another endlessly.

Our examples are incomplete, because one may want (or need) to describe the effect of the speech act on the interpersonal relationship in greater detail, considering, for example, not only those aspects of it which concern one of the participants (say, in commands, the addressee, who is given the obligation to do something), but also those which concern the other participant (the speaker, who after all is obliged at least not to hinder the addressee from doing what he has been told to do, e.g. to abstain from doing it in the addressee’s place). One may want to describe the (prototypical) initial situation that is required to be the case for a certain utterance to bring about a change of a certain kind in it: a situation in which the speaker has authority, or one in which she is or counts as capable of certain

³ There are controversial elements in these examples, particularly in the last two, but I cannot discuss them here.

⁴ I do not include in this list Searle’s “declarations” (1979: 16–20) or “declaratives”, since such a generalized description is not possible for declaratives as a class. Indeed, each declarative (type) brings about its own (type of) effect, which regularly involves an effect on the interpersonal relationship, at least as an element of a broader, social effect. A class of declaratives is possible only in a theory that does not view (other types of) illocutionary acts as the bringing about of changes in any kind of state of the world. It is therefore incompatible with the perspective adopted here.

performances, one in which she is in a position to know about something potentially dangerous for the addressee, one in which she has damaged the addressee or the addressee's face and owes him a remedy; etc. One might want to describe the roles the addressee is expected to play in each of these kinds of initial situations, or specify what kinds of agents may count as speaker and addressee in different cases, or say what kinds of contents the change in the interpersonal relationship is designed to touch upon, and so on.

Moreover, our examples are incomplete because the idea they exemplify—that of bringing about a change in an interpersonal relationship—is not precise enough to discriminate between illocution and perlocution. Perlocution too can well be said to affect the interpersonal relationship between speaker and addressee: for example, the speaker may reassure her interlocutor, or alarm him, or surprise him by her utterance. Getting closer to the gamut of attitudes and performances that are usually associated with the various types of illocutionary acts, we can say that, by her utterance, the speaker may get the addressee to believe something or disbelieve it, or to do something or abstain from doing it. Among speech act philosophers, there is a recurrent tendency to consider the effects mentioned in our initial set of examples and the psychological or behavioral effects referred to here as equally pertaining to perlocution, because they appear to “go beyond” the speech act itself. But this widespread feeling depends on the presupposed acceptance of a conception of action which, following (by and large) Davidson (1980), identifies actions with bodily movements animated by an intention: according to such a view, any effect not coinciding with the production of the bodily movement itself is external to the “real” action (even if it is part of the same causal chain and can be traced back to the responsibility of the same agent). Already Austin was aware that a view of action centered on basic action, if applied to speech acts, would obscure and even hide from sight the intimate connection between an utterance's being an action of a certain kind and its bringing about certain effects (not coinciding with the mere fact that certain articulated sounds have been uttered).⁵

So, albeit in both the groups of effects distinguished above we find effects related with interpersonal relationships, the two groups differ greatly. Their difference was characterized by Austin in terms of the “conventional” nature of the effects of the illocutionary act, and therefore of the illocutionary act itself (Austin 1975: 103; 107; 117; Sbisà 2007). In speaking of the illocutionary act as bringing about “conventional” effects, I mean that its effects could not come into being without some kind of coming together or agreement, overt or (more often) tacit and by-default, of the relevant members of the relevant social group. Since such an agreement cannot be completely *ad hoc*, starting from scratch again every time, repeatable patterns and routines are needed, including possible standardized or even conventional *means* for performing the illocutionary act. So, there is some link between the conventionality of the means for performing an illocutionary act,

⁵ This is why he warns the reader against reducing action to the “minimum physical act” (1975: 107).

which has been so often discussed (and sometimes denied) in speech act theory, and the conventionality of the effects that the illocutionary act brings about, so often neglected. But the latter is primary, while the former is derivative. In this perspective, the idea that illocutionary acts bring about changes in interpersonal relationships is to be qualified by adding that these changes should affect the conventional aspects of interpersonal relationships. The examples of illocutionary effects provided at the beginning of this paragraph are all conventional effects in this sense, involving obligations, debts and their ritual remedies, entitlements, or the legitimation of expectations. All these states have to do with conditions of the relevant agents that are represented by means of modal verbs such as *can* and *ought* and other words associated with them, can be further acted upon and even cancelled by other illocutionary acts, and are vulnerable to the discovery of initially unknown non-default conditions (Austin's infelicities).

3.2 *Plural Illocutionary Effects*

Within this framework, the question about illocutionary pluralism can be reformulated as follows: can one and the same utterance token bring about more than one change in the conventional aspects of the interpersonal relationship between speaker and addressee?

First of all, a warning. Changes in the interpersonal relationship between speaker and addressee are never simple: there must be a balance between the status or condition assigned to the addressee and the one assigned to the speaker herself. This may make it difficult to tell whether, on a given occasion, we face one change or more than one. Let us assume that there are patterns of balanced effects concerning the respective statuses of speaker and addressee and that the effects belonging to the same pattern count as one. The question we raise is, then, whether one utterance token can activate more than one of these patterns. In order to answer this, I shall illustrate various kinds of cases about which the issue of illocutionary pluralism can be raised.

(1) Utterances ambiguous between two illocutionary patterns.

An utterance token may contain words and display paralinguistic features that may be taken to indicate one illocutionary force or another. It is up to interactional negotiation, in this case, to make the utterance count as one of the two illocutionary acts it is compatible with. Consider an utterance in which the speaker suggests to the addressee some joint action, which is alleged to be in the addressee's own interest. Is it a proposal or an offer? The way in which the utterance is taken will be revealed by the addressee's reply. If the reply comprises, for example, an act of thanking, the pattern of the offer will emerge and the speaker will count as having made an offer. If the reply comprises, for example, consideration of pros and cons, the emerging pattern will be that of a proposal. Consider also praise and compliments. There is an element of praise in

compliments, but not all acts of praising are compliments and not all compliments are competent evaluative judgements about the addressee's merits. One and the same utterance in the same context may display an ambiguous profile, which will be disambiguated by the response it gets, which may acknowledge the speaker's competence or authority in evaluating, making the pattern of praise emerge, or express the addressee's gratitude but also to some extent his polite partial disagreement, making the pattern of compliment emerge.

(2) Utterances whose designed effect conflates two illocutionary patterns (or more).

An utterance may be designed (often with some awareness, but not deliberately) to bring about an effect that conflates two illocutionary patterns. Consider the acts of promising and of apologizing. Their effects are very different, but not unrelated, so that they can merge with one another. If a speaker says "I won't do that to you ever again", an effect of the commissive kind can be expected (the speaker assigns to herself the obligation not to do something that is bad for the addressee ever again in the future, the addressee is granted a legitimate expectation that the speaker will so behave), but there can also be expected an effect of the behabitive kind (the speaker reacts to a situation in which she has a debt as regards the addressee and remedies it by displaying her willingness to avoid certain behavior in the future). In similar cases, participants may be perfectly at ease with the fuzzy aspects of their relationship and its changing shades, but describing what they are really doing might be difficult for the analyst, because it does not fit a standard pattern. An analogous difficulty is experienced when we attempt to describe the effects of mitigated speech acts, especially when they are finely tuned to fit the needs of the particular situation in which they occur (Sbisà 2001).

(3) Utterances that count as so-called indirect speech acts.

As is known, an utterance token can be both a question and a request. One may say that the "real" illocutionary act performed by the speaker is the request, but undeniably, there are cases in which the utterance is responded to as a question, rather than as a request: "Can you reach that book over there?" "No (I am not tall enough)". So it seems more sensible to deal with the utterances used to make indirect requests as multi-action compatible, and leave to the audience's uptake the task of selecting their illocutionary effect. But the audience's response may contain elements recognizing the performance of both illocutionary acts. Consider: "Can you reach that book over there?" "No. Sorry!", where "No" is a reply to the question that has been asked, while "Sorry", aimed at remedying the debt deriving from the unfulfilled request, recognizes that a request was made. Indeed, when a request is made by means of an interrogative sentence, both question and request may be effective, each in its own way. The task of the question is to elicit an answer (by assigning to the addressee the obligation to provide it, or perhaps, by creating a legitimate expectation that the next turn will be an answer). The task of the request is to direct the addressee's behavior: it elicits non-verbal action (by assigning to the addressee an obligation, which may be recognized, before or independently of its fulfillment, by an act of acceptance or of refusal to perform). It

is as if the utterance token were a step not in one conversational sequence only, but in two sequences, one of which made of verbal acts, the other of both verbal and non-verbal ones.

In the cases of the kind described in (1), insofar as the addressee or other participants, in manifesting their uptake of the speech act in their responses, select one illocutionary act as actually performed, the illocutionary act that is performed is only one. So these cases are not cases of illocutionary plurality. The “plural” element in them only concerns the speaker’s openness to accept either understanding of her utterance. Also the kind of cases described in (2) does not *per se* involve illocutionary plurality: a composite illocutionary effect is not therefore plural. But cases of this kind may come very close to a plural illocutionary performance, because it is easy for the audience’s response to shift from recognizing the performance of one composite illocutionary act to recognizing the performance of two illocutionary acts. Finally, utterances of the kind described in (3) can be taken to display illocutionary plurality at least every time they receive responses recognizing both the so-called “primary” and “secondary” illocutionary acts. Something similar might conceivably happen also with some ambiguous utterances of the kind described in (1).

Considering these cases in the light of the view of illocutionary acts presented in Sect. 3.1, illocutionary pluralism appears plausible. Indeed, if we take the illocutionary act to be the bringing about of a conventional effect, and view the conventionality of effects as their being (roughly speaking) dependent on the audience’s uptake, there is no reason to suppose that for each utterance token, only one conventional effect will be recognized (and brought about). So, the audience’s uptake (against a background of a multiplicity of illocutionary indicators, including textual strategies) seems to play a central role in allowing for illocutionary plurality.

It could be said, though, that while this may explain illocutionary plurality locally (how it happens that a certain utterance token brings about more than one illocutionary effect), it does not account for the general possibility of illocutionary plurality. This other kind of explanation is to be looked for not in the dynamics of the illocutionary act, but in that of social interaction. The structure of interactional events resembles a bunch of intertwined threads, rather than a single linear sequence: there may be several different tracks present in an interactional event, each one corresponding to a course of action in which one or more steps may consist of one or more verbal utterances. The same utterance token may be a step or part of a step along one or more of these tracks, doing different jobs in each and eliciting responses to each of these jobs. Following Levinson (2011), we could say that an utterance token may do one job as a step in a conversational sequence and another as a step in a sequence of actions aimed at a project that the agent/speaker is pursuing. I would like to add that some, or even all, of the possible different jobs of an utterance token may consist of the bringing about of conventional effects on the interpersonal relationship, and therefore of illocutionary acts.

Moreover, the multiplicity of interactional tracks is perhaps not the only feature of interactional events that can generate illocutionary plurality. An interactional track is

liable to be given alternative internal organizations. When a conversation begins, we easily identify its first move (leaving aside here the fact that the first verbal move might also be viewed as a response of the speaker to the interlocutor's arrival or presence, and therefore as a second move). But it may become less and less clear in the course of a conversation when sub-sequences begin and end. Most utterance tokens are prompted by something that has been said or done or has occurred prior to them and at the same time call for responses recognizing at least their illocutionary effects. So each utterance token might be viewed as the central element of a triplet (roughly corresponding to the "narrative scheme", manipulation-action-sanction, elaborated upon by semioticians: see Greimas and Courtés 1979; Sbisà 1986, 2002b) and the interactional sequence as a set of partly overlapping triplets. The plural potentialities of sequential positioning do not always emerge, but neither are they ever completely neutralized in favor of a univocal sequential organization mandated by linguistic markers. Even if we leave aside the multiplicity of interactional tracks, therefore, the various possible segmentations of one and the same track into sub-sequences may suffice to assign multiple jobs to an utterance token, thereby endowing it with illocutionary plurality. This last source of illocutionary plurality may go deeper than the others discussed above, since narrative contextualization might turn out to be an indispensable component of our notion of action.

4 Locutionary Pluralism or Speech Act Content Pluralism?

In concentrating on illocutionary acts, I nearly lost track of that aspect of the speech act which is the main target of Cappelen and Lepore's speech act pluralism, namely, speech act content. Cappelen and Lepore are not concerned with illocutionary pluralism: in exemplifying speech act plurality, they mention such illocutionary forces as assertion, claim and statement merely in order to say that an utterance token may do anyone of these things with many possible contents. Moreover, as I reported above, they believe that when we take into account actual speech production, our first task is to reconstruct utterances "up to a point where they express thoughts". This shows that they are mainly interested in the cognitive or informational content of speech as opposed to its interpersonal and relational function. They seem insensitive even to the difference, often held to be a basic illocutionary difference, between "saying that" and "telling to". Thus, we may conclude that the observations made by Cappelen and Lepore (2005) outline not speech act pluralism in general (let alone illocutionary pluralism), but some kind of locutionary pluralism.

But does speech act pluralism really apply to the locutionary act? A reply in the affirmative would account for many of the issues raised by so-called contextualism. If it is possible for an utterance to express both the "minimal proposition" (determined by linguistic rules plus processes of disambiguation and indexical

saturation) and some other contextually enriched proposition, thus performing more than one speech act at once, this would undermine the contextualist claim that the proposition expressed by each utterance of any sentence whatsoever is partly determined by the context, since it is not the case that only one proposition is actually expressed and moreover, not all propositions expressed are context dependent. This is indeed what Cappelen and Lepore wish to claim with their speech act pluralism.

However, I am doubtful about applying speech act pluralism to locutionary acts. This would amount to claiming that, in an utterance token, words may be used simultaneously with more than one sense and more than one reference. Perhaps, in some cases of vague or ambiguous reference or predication, or in some cases of loose use or figurative speech, we get close to a situation in which this happens (and in which it can therefore be said that more than one proposition is expressed). However, underdetermination should be kept distinct from plurality. Underdetermination invites interactional negotiation and selection by the audience, whereas plurality aims at multiple recognition and is confirmed by it. Now, most of the examples of words conveying more than one sense and more than one reference which I can think up are such that the previous or the following part of the conversational sequence in which they occur would either provide disambiguation or yield a negotiated meaning. Conveyed meaning may remain underdetermined only in the case of implicature, in which as all authors recognize (albeit with differing explanations), the series of contents made available by the speaker is often left open.⁶

But I suspect that, if we come to the reference of names, indexicals or descriptions (as used in a certain utterance token), or to what is said of their referents by the predicates that compose with them in the uttered sentence, it is all too important for speaker and addressee to converge on one reading for there being any room left for locutionary (and particularly rhetic) plurality. In their attempt to save speech act pluralism at this level, Cappelen and Lepore have to reject the difference between saying and implicating altogether. They do so rather hastily, without pausing to reflect on whether there is anything to be saved in it, or learnt from it.

Even if I find Cappelen and Lepore's criticism of "Original Utterance Centrism" (2005: 201) innovative and to some extent correct, I believe it ceases to be such if the consequent openness of the series of contextual interpretations of an utterance token is left limitless. It may be wise to decentralize utterance interpretation, admitting that it may be influenced by factors going beyond the conscious control of the participants (see, for example, Gauker's arguments in favour of an objective and mind-transcendent conception of context: Gauker 1998, cf. Sbisà 2002a). But it is definitely unwise to relax the requirements for a correct interpretation (or a good speech act report) too much, forgetting about those

⁶ In a Gricean perspective, this series is left open under the condition that any of its members, added to the content of the utterance or replacing it, should be capable of making the utterance fit the relevant standards of conversational cooperation.

argumentative routes for the retrieval of said or implicated content that make pragmatic meaning rational.

I would like to suggest, though, that within the framework of certain cases of illocutionary plurality, some kind of speech act plurality connected with the locutionary level can actually be found. From a certain perspective and as a second step in a sequence “It’s cold here” (as a reaction to the cold air coming from the window), is a complaint and expects some recognition (in terms of agreement or disagreement, of expression of sympathy or disapproval). Its content (to which the reference of “here” and the meaning of the predicate “cold” give their contributions) concerns the temperature in the room the speaker has just entered. But as a first, initiative step, it may count as a suggestion to close the window. From the former perspective to the latter, there occurs a shift concerning not only illocutionary force, but also the content of the change in the interpersonal relationship that is brought about. The complaint is about temperature in the room and reveals to the audience the speaker’s attitude towards that state of affairs. The suggestion is about doing something (closing the open window). In the utterance token under consideration, no word is issued related to the act of closing a window, and no word is issued referring to windows (or to a specific window in the context). So, we cannot say that the window’s being closed is part of the locutionary meaning of the utterance token, even if we focus on the rhetic act alone. The new content, which is however different from the content of the attitude displayed by the act of complaining, is not a propositional content “expressed” by the utterance: it is the implicit content to which we apply the modal predicate “ought”, which is assigned to the addressee by the illocutionary act of making a suggestion. So, there are two “contents” here, both of which may be called “speech act contents”, but only insofar as there are two illocutionary acts. The relationship of these two contents with the Austinian notions of sense and reference, that come into the rhetic act, and with the philosophical notion of expressed proposition, is not completely clear, but I will not discuss it further here. I would only like to remark that asserting, claiming and stating, that is, the examples of illocutionary acts mentioned by Cappelen and Lepore (2005: 199), are a handful of illocutionary acts strictly associated with one another. They could even be considered as shades of one and the same procedure. There might be situations or speakers to which it makes a difference whether an utterance token is an assertion or a claim or a statement.⁷ Moreover, the locutionary act (in its rhetic aspect, that is as reference and predication, or as the expression of a proposition) performed by an utterance token when it is taken to be an assertion is very likely the same as that which is performed when the same utterance token is taken to be a claim or a statement. So,

⁷ To my ear, for example, asserting has a commissive shade (its most salient effect is the speaker’s commitment), while claiming has an exercitive shade (its most salient effect is some kind of obligation on the part of the audience to take the speaker’s utterance as true, unless they explicitly question it), and stating has a verdictive shade (it is a judgement, licensing analogous judgements on the part of the audience).

conditions allowing for plurality of speech act content do not seem to hold in the cases taken into consideration by Cappelen and Lepore.

5 Some Conclusions

In conclusion, I would like to underscore a coincidence. Cappelen and Lepore, Levinson, and myself have all come to deal with actions done in speech as with something the identity of which depends on some kind of ascription. They are not a matter of the speaker alone, but involve her audience; they are not a matter of the speaker's behavior, but also of the audience's uptake or acknowledgement, often coming after some negotiation.

Cappelen and Lepore reveal their commitment to an ascription-centered view of action when, in arguing against Original Utterance Centrism and the alleged privileged access of speakers to the content of their speech acts (2005: 201–202), they observe that in indirect reports we can use expressions that would and could not be used by the original speaker. On their view, the fact that the speaker's utterance can be reported in a certain way implies that the speaker has performed a speech act with that content. To illustrate the idea: if Lois says "Superman flies", without knowing that Superman is identical with Clark Kent, I, knowing (and addressing to an audience that knows) that Superman is identical with Clark Kent, can report her rhetic act by saying that she said that Clark Kent flies. Thus I ascribe her an assertive speech act with the content that Clark Kent flies, and even the act of referring to Clark Kent. I am not sure such reports are always acceptable (I would specify contextual conditions), but since they are acceptable at least sometimes, we should admit that the ascription of a certain speech act (particularly, a speech act with a certain rhetic component or propositional content) to a speaker may make her have performed that speech act. So Cappelen and Lepore, in making speech act content depend on speech act content ascription, make speech acts, as actions, depend on action ascription.

Also Levinson, in his discussion of action formation and ascription in conversation (2011), seems to endorse an ascription-centered conception of action. He finds the received notion of "action recognition" misleading, because it presupposes that actions have a correct identity prior to and independent of the negotiated process of attributing an action to a turn. He does not want to accept this assumption. But if actions done with words have no correct identity prior to the response they receive by an audience and to the acceptance of that response by the speaking agent herself, then it is action ascription and the acceptance of it that play the major role in the identification of actions.

As to myself, I have concentrated on the effort of understanding what exactly is ascribed to the speaker when she is attributed an action done with words. On my account, an action is the bringing about of a change in the world, a speech act (broadly intended) is the production of any such change on the basis of verbal behavior, and an illocutionary act is the production of a change in the conventional

aspects of the interpersonal relationship among the participants. What is ascribed to the speaker/agent is responsibility for the change brought about: the new situation is recognized to hold and to be caused at least in part by her behavior. In the case of illocutionary acts, this ascription involves distinguishing, on the basis of multiple verbal and non-verbal cues, the change or pattern of changes that an utterance token is designed to bring about, but at the same time, manifesting the audience's uptake (the "securing" of which is essential to the illocutionary act) actively contributes to the success of the illocutionary act and, therefore, to the real occurrence of that change.

This convergence suggests that speech act pluralism (of any kind) makes sense only in the framework of an ascription-centered conception of action. Such was, not by chance, the view of action outlined by J.L. Austin in his papers in the philosophy of action (1979: 173–204, 272–287) as well as in some passages of *How to do things with words*, a view compatible with pluralist ontology, not only as regards speech acts, but actions in general.⁸

As we have seen, there are differences beyond this convergence, particularly as to the motivations for endorsing speech act pluralism and the limits of its application. I share with conversation analysts, and with Levinson in particular, the same preoccupation to describe as precisely as possible the dynamics that make multiple action ascription conceivable and, in certain circumstances, correct. This preoccupation does not seem to be shared by Cappelen and Lepore, whose speech act pluralism does not admit of procedures for distinguishing reliable ways of ascribing actions (be they single or plural) from unreliable ones.⁹

References

- Austin, John L. 1962. *Sense and sensibilia*. Oxford: Oxford University Press.
- Austin, John L. 1970. Intelligent behaviour. A critical review of *The concept of mind*. In: Ryle. *A collection of critical essays*, eds. Oscar P. Wood and George Pitcher, 45–51. Garden City, NY: Doubleday.
- Austin, John L. 1975. *How to do things with words*, 2nd ed. Oxford: Oxford University Press.
- Austin, John L. 1979. *Philosophical papers*, 3rd ed. Oxford: Oxford University Press.
- Cappelen, Herman, and Ernie Lepore. 2005. *Insensitive semantics*. A defense of semantic minimalism and speech act pluralism. Oxford: Blackwell.
- Cooren, François. 2010. *Action and agency in dialogue*. Amsterdam: John Benjamins.
- Davidson, Donald. 1980. *Essays on actions and events*. Oxford: Oxford University Press.
- Derrida, Jacques. 1977. Signature event context. *Glyph* 1: 172–197.
- Gauker, Christopher. 1998. What is a context of utterance? *Philosophical Studies* 9: 149–172.

⁸ That Austin's linguistic-phenomenological ontology was pluralist is testified, for example, by some of his remarks about the things we perceive in his (1962) and by the criticism he leveled at Gilbert Ryle in his review of Ryle (1949) (Austin 1970).

⁹ I am grateful to Paolo Leonardi and Ken Turner for their comments on a previous draft of this paper.

- Goffman, Erving. 1974. *Frame analysis: An essay on the organization of experience*. New York: Harper and Row.
- Greimas, Algirdas J., and Joseph Courtés. 1979. *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Paris: Hachette.
- Gu, Yueguo. 1993. The impasse of perlocution. *Journal of pragmatics* 20: 405–32.
- Levinson, Stephen C. 2011. Action formation and ascription. In: *Handbook of conversation analysis*, eds. Tanya Stivers and Jack Sidnell. Malden, MA: Wiley-Blackwell (forthcoming).
- Ryle, Gilbert. 1949. *The concept of mind*. London: Hutchinson.
- Sbisà, Marina. 1984. On illocutionary types. *Journal of pragmatics* 8: 93–112.
- Sbisà, Marina. 1986. Manipulation et sanction dans la dynamique des actes de langage. In: *Exigences et perspectives de la sémiotique. Recueil d'hommages pour A.J. Greimas*, eds. Herman Parret and H.G. Ruprecht, 529–538. Amsterdam: John Benjamins.
- Sbisà, Marina. 1989. *Linguaggio, ragione, interazione*. Bologna: Il Mulino.
- Sbisà, Marina. 2001. Illocutionary force and degrees of strength in language use. *Journal of pragmatics* 33: 1791–1814.
- Sbisà, Marina. 2002a. Speech acts in context. *Language & Communication* 22: 421–436.
- Sbisà, Marina. 2002b. Cognition and narrativity in speech act sequences. In *Rethinking sequentiality*, eds. Anita Fetzer, and Christiane Meierkord, 71–97. Amsterdam: John Benjamins.
- Sbisà, Marina. 2007. How to read Austin. *Pragmatics* 17: 461–473.
- Searle, John R. 1968. Austin on locutionary and illocutionary acts. *The Philosophical Review* 77:405–424.
- Searle, John R. 1969. *Speech acts*. Cambridge: Cambridge University Press.
- Searle, John R. 1979. *Expression and meaning*. Cambridge: Cambridge University Press.

Speech Act Pluralism, Minimal Content and Pragmemes

Michel Seymour

Abstract As speech acts in contexts, pragmemes serve to illustrate speech act pluralism. What is less clear is whether they play an important role in determining the primary meanings of sentences. Semantic contextualism is the view according to which word meaning or sentence meaning cannot be detached from some extralinguistic features of their utterance. Semantic minimalism suggests another way of conceiving the relationship between sentence meaning and pragmemes. Some sentence-types may express only “proposition-radicals”, as suggested by Kent Bach. Are there however pragmemes that determine primary sentence meanings and that are not prescribed by the very semantic features of the sentence? Carston and Recanati both argue that there are. However, cancel ability reveals the presence of a minimal accessible content that could be expressed without these additional features. Are there pragmemes determining primary sentence meanings that are not prescribed by semantic features and that are not cancelable? In this paper, I argue that there are no such examples. Pragmemes may contribute to the determination of the content of certain *assertions*, but they do not contribute to the determination of minimal content of the sentence-types used in these utterances. I conclude that a proper appreciation of the role of pragmemes forces us to accept speech act pluralism and bifurcationism, the idea that there are two levels of content: minimal and maximal. That is, different pragmemes produce different inferential augmentations of a minimal level of linguistic meaning. But this is precisely what semantic minimalism is all about.

I wish to thank Jessica Moore for her translation of some of the parts of this work.

M. Seymour (✉)

Department of Philosophy, Université de Montréal, Montreal, Canada
e-mail: seymour@videotron.ca

1 Introduction

As speech acts in contexts, pragmemes play a major role in the pragmatics of language and they serve to illustrate what Hermann Cappellen and Ernest Lepore call speech act pluralism (Lepore and Cappellen 2005). That is, a speaker may assert many different things by using the very same sentence in different contexts. What is less clear is whether pragmemes play an important role in determining the primary semantic content of sentences.

The fundamental debate is whether or not the literal meaning of type-statements that are devoid of indexical expressions is partly determined by the context of utterance. Those who recognize a minimal semantic content conveyed by sentences not containing indexicals and who claim that their literal meaning is independent of the context of utterance can be described as minimalists. Those who maintain that the literal meaning of these type-statements is most often determined by the context of utterance are contextualists. Whether it is cast as an opposition between formal semanticists, institutionalists or literalists, on the one side, and interpretationists, intentionalists or pragmaticians, on the other side, the debate echoes the one that took place between formal semantics and ordinary language philosophy early in the second half of the 20th century. We could say that the debate has started up once again with renewed vigor, and that it is, in fact, the most important controversy to arise in the past thirty years in the analytic philosophy of language.

There are of course different sorts of contextualist debates: in epistemology, political philosophy, etc. Here I concentrate on the semantic debate. On the contextualist side, we must mention for instance Carston (2002), Levinson (2000), Recanati (2004), Searle (1983), Sperber and Wilson (1995), Stainton (2006) and Travis (2001). In the minimalists' side, we could name Borg (2004), Lepore and Cappellen (2005), Predelli (2005), Salmon (2005) and Stanley (2005).

What does semantic contextualism consist of? To better understand this philosophical position, we must take note of the following fact. There was a time when it was believed that the meaning of an *utterance* was essentially a function of the meaning of the uttered *sentence*. Performing an assertion was nothing but the tokening of a linguistic type. But we now know that the meaning of an utterance almost always depends upon other factors that differ from those determined by the conventional meanings of words, and this is why speech act pluralism must now fully be acknowledged. But the fundamental question that remains is whether conventional meaning offers at least a nucleus onto which the far richer intended meanings of the speaker are grafted, or if pragmatic features intrude in the very determination of what is literally expressed. It is with this last claim that contextualist arguments come into play.

Semantic contextualism must be distinguished from semantic intentionalism. This last doctrine stipulates that the meaning of a sentence is determined by what the speaker has in mind. It is slightly different from the contextualist claim according to which literal meaning of sentences not containing indexicals are often

determined by the context of utterance. The intentionalist and contextualist theses are of course compatible in principle, since they may up to a certain point reinforce each other in establishing the existence of a correlation between literal meaning and facts occurring when speakers utter a sentence. But these facts need not be psychological facts and affirming that two views are mutually reinforcing is one thing, and demonstrating that the two are identical is quite another. First, there are variants of semantic intentionalism that reject semantic contextualism. Grice (1975), for example, believes that in a given context of utterance, the conventional meaning of words constitutes the most important part of what is uttered, and he maintains that optional pragmatic factors only rarely disturb the delivered semantic content. Thus, he tends to minimize the importance of pragmatic factors in the determination of intuitive semantic content attached to a given utterance. The disambiguation and saturation of statements together with conversational principles are sufficient to derive, from the semantic potential of sentences, minimal truth-conditions that are intuitively associated by speakers. There can of course also be conversational implicatures, but these are an additional layer on top of minimal content. But he also argues for such a minimalist position while simultaneously adopting a conceptual reductionist approach to the subject of conventions governing the rules that give the invariant meaning of words. So in this sense, the Gricean account is all at once minimalist and intentionalist. A similar line of argument is developed by Schiffer (1972), Bennett (1976) and Lewis (1969). According to this view, the very concept of linguistic meaning is one that should be analysed in terms of intentional meaning. This conceptual reduction of conventional meaning can be described as “conceptual intentionalism” or “analytical intentionalism.” It is a case of conceptual analysis while the contextualist claim comes from the observation of empirical facts concerning meaning. Word meaning is for Griceans reducible to speaker’s meaning, but not in the contextualist sense.

Conversely, it is possible to argue for an account of “meaning as use,” interpreted as suggesting the existence of a close connection between meaning and verbal events, and thus agree with the main tenet of contextualism while rejecting psychologism. Baker and Hacker (1984) offer an excellent illustration of this particular philosophical posture. We must also take note that interpretationism, as defended by Davidson (1984, 1989, 1999), is a form of contextualism, but it entails that the interpretation of the hearer, and not the intention of the speaker, is the relevant factor in the determination of the meaning of most utterances. It is therefore also not to be confused with intentionalism. Of course, as I suggested above, it is possible to argue for a particular intentionalist variant of the contextualist thesis, but there are other instances such as ordinary language philosophy and interpretationism.

So semantic intentionalism must be distinguished from semantic contextualism. That being said, I am mostly concerned in this paper with the intentionalist version of semantic contextualism.

It is also important not to confuse semantic contextualism with the empirical facts upon which this doctrine is based. We could perhaps recognize in many cases

the existence of a positive correlation between the *intuitive* truth-conditions (or assert ability conditions) of *utterances* and the intended meaning of speakers—without coming to the conclusion that intentional meaning intrudes in the *literal* meaning of a sentence. The empirical evidence which corroborates the contextualist hypothesis can hardly be denied, whereas the explanation in terms of the role of context in the determination of the *literal* meaning of the sentence is the more controversial aspect of contextualism. One can grant that the intended meaning of a speaker and/or the normal interpretation of hearers play a major role in determining the meaning of what is *asserted*. But does it determine what is literally *said*? As such the existence of a determination of context on the content of an utterance can be accounted for just as an instance of speech act pluralism.

This point illustrates that of the main questions at stake is whether one can accept the distinction between illocutionary acts of assertions and locutionary acts of saying. The empirical existence of a close link between intended meanings and assertions does have a bearing on the literal content of a sentence only if one denies the existence of locutionary acts of saying. For if what is asserted is only something done in the context of saying something, then the fact that my intention determines what I have asserted does not yield the conclusion warranted by the contextualist; that is, it does not disturb what has minimally been said.

An author such as Wittgenstein (1953) recognizes the importance of pragmatic factors in the determination of satisfaction-conditions of assertions, and would thus be inclined to acknowledge the empirical observation that serves as a basis for contextualism. Moreover, he was no doubt eager to show how the conventional meanings of sentences were community relative. But as an institutionalist philosopher, he was defending the autonomy of the institution of language, and he did therefore reject the suggestion that word meaning took place essentially in the intended meaning of a particularized context of utterance. Meaning is not what occurs in an eventful situation. Language is a practice governed by rules, understood as a system of ordinary and ostensive definitions, together with an agreement on the application of words in some paradigmatic instances. Even if the semantic rules do not anticipate all their applications, the decision made in a specific context of utterance concerning what has been said is an institutionalist decision that serves only to refine the rules (Seymour 2005). For Wittgenstein, the context of utterance plays a role in relation to the initial rules of language that is similar to the role played by jurisprudence in relation to a system of laws. Even if the laws do not have a determinate meaning, it does not mean that there are no laws. Also the decisions made by the judges concerning the application of a law are not *sui generis* and spontaneous, for they are made in accordance with the rules contained in the existing laws. And finally, these decisions can be interpreted as entailing nothing but a refinement of the meaning of the already existing laws. So if context play a role, it is only the institutional context, for the variations of meaning are always to be understood in relation with the institutional decisions that are made on the basis of already existing rules. Wittgenstein is not arguing for a close connection between meaning and events. As a staunch anti-psychological philosopher, Wittgenstein would certainly challenge the intentionalist version of

semantic contextualism. I also believe that he would challenge the social version as well, because he rejects the connection that purportedly holds between literal meaning and events. An utterance does not all by itself *determine* meaning, not even an utterance made using a term or a sentence having conventional meaning. The variations created by the use of an expression in a context of utterance becomes a semantic variations only when the linguistic community, just like a judge, stipulates that this variation is semantical. Wittgenstein would consider semantic contextualists to be largely correct in the *observations* that they are making, since, very often, in many different contexts of utterance, the intuitive satisfaction-conditions of utterances are considerably different from the semantic content of sentences. But once again, this amounts to speech act pluralism and not to the claim that sentence-meaning is determined by the context of utterance. If 'context' becomes relevant for conventional meaning according to the Wittgensteinian philosopher, it is only in the sense that an institutional decision has to be made on how to apply the rules, granted that they do not anticipate all their applications. This institutional decision modifies the meaning of the rules. But it is a variation occurring within the institution of language and at no time utterance-events do intrude in the literal meanings of words and sentences. It is rather the rules and the jurisprudence involved in their application that tells us what is meant.

Is it possible to separate natural language from its social use in context? The institutionalist account already in some sense recognizes the influence of the community in the determination of meaning. Meaning is social and community relative. But to say that meaning is use under this account is to say that the semantic rules are those that a community accepts in a dictionary for a certain time on a given territory, and it is also to say that they are always changing through the "jurisprudence" of institutional decisions made either by the experts, or by a critical mass of people within the community as a whole. Now this account is compatible with a fairly strong distinction between literal meaning and pragmatic meaning. So it has nothing to do with the contextualist claim according to which the background assumptions made by the members of the community influence literal meaning of a sentence in that community.

The contextualist philosopher asks: how are we to characterize the semantic rules associated with sentence type and how could we describe the semantic rules in reference to new situations? The problem is apparently that we come across new situations all the time. This amounts to acknowledge the indeterminacy of the semantic rules attached to sentences. I wholeheartedly agree with the indeterminacy thesis, but my own account is one in which the institutional rules must be accompanied by the jurisprudence of institutional decisions made concerning the application of the rules. This institutionalist model is quite different from the contextualist model in which it is claimed that speakers intentions, occurrences or interpreters determine content. This is why the institutional model is compatible with a strong distinction between semantics and pragmatics. And the institutional model allows us to cope with novelty.

Lepore and Cappelen (2005) are also eager to recognize that a very large quantity of propositions may be asserted with a sentence in the context of

particular utterances. This is because that which is asserted depends upon factors that vary greatly and that may differ from the semantically expressed proposition. They accept that speech acts potentially express a very large and indeterminate quantity of propositions with a single sentence. Furthermore, they agree that in many cases we can express a full proposition only if the context of use comes into play. This happens when the sentence used contain expressions belonging to a very basic set of indexical expressions. Their semantic minimalism must be understood as the claims that (1) the meanings of words are in general not determined by the context of utterance, (2) the sentences that are determined by context are grammatically sensitive to it and (3) this sensitivity is explained by the presence of words belonging to a basic set of indexical expressions. (4) In these cases, context is understood in the limited sense implying nothing more than time, place, speaker, and proximal or distal features.

2 Secondary Sentence Meaning

Semantic contextualism stipulates that nothing or almost nothing can be meant with language independently of a conversational practice. The presence of the context of utterance is a necessary condition for sentences to express their full semantic potential. The true meanings are those that are expressed by verbal or mental *occurrences*. There is a moderate version held by Carston (2002) and Recanati (2004) according to which these contextualist claims apply only to sentences but not to words, since words do in general have conventional meanings apart from the context of utterance. But there are also radical versions such as those of Sperber and Wilson (1995) and Travis (2001) in which the very same claims are said to apply to words as well. Words have, at best, a “semantic potential,” but their full and complete meaning depends upon the context of utterance. Despite the important differences between moderate versions and more radical versions, proponents of these two views agree that pragmemes play a role in determining the primary meanings of sentences. In this paper, I consider only the moderate version and I ignore the issue of pragmatic intrusion into word meaning. I concentrate on sentence meaning only.

So the question we want to ask concerns the argument for the suggestion that the literal meaning of sentences is determined by the intentions of speakers in a context of utterance. Semantic minimalism sees literal sentence meaning and pragmemes as involving two different layers of meaning. Word-types and sentence-types enjoy a certain semantic autonomy relative to their occurrences in particular conversational contexts. According to this picture, a very large class of sentence-types express minimal propositions and do not require a verbal event in order to express minimal content. The only exception to this general rule is provided by sentences containing expressions belonging to the basic set of indexical expressions. According to this account, when the sentence expresses a minimal proposition, pragmemes may serve to determine a secondary sentence meaning,

but not a primary one. Pragmemes behave in this case like conversational implicatures, that is, as intended meanings that add up to the minimal sentence meanings. So they do not determine primary meanings. Why do we say that the implicated content is added up to a first layer of content? The reason is that the very same act of saying could be associated with a quite different conversational implicature. To put it differently, conversational implicatures are cancelable.

Capone (2009) has argued that some particularized conversational implicatures were not cancelable, but he reached that conclusion while considering very specific conversational situations. However, if he is right this only means that conversational implicatures cannot be cancelled from a specific conversational context, and it does not imply that they could not be cancelled from a specific act of saying. So for instance, in the context of writing a letter of recommendation for a candidate to become professor in a university department, it is impossible not to infer a particular negative implicature if I merely write that the candidate has a good handwriting. There seems to be no way of suggesting anything else. So in such a case, it seems that the content expressed by the utterance of the sentence is determined by pragmemes. But in the context where the same person would be applying for a job involving essentially writing abilities, the very same act of saying could in that context of utterance become quite positive. What does this show? It shows only that context of utterance plays a role in the determination of utterance meaning. Now this can have a bearing on literal meaning only if we assume at the outset that utterance meaning determines sentence meaning. But this is precisely what is at issue. The argument purporting to show that the context of utterance determines sentence meaning cannot work if it rests on the claim that some conversational implicatures are not cancelable in some contexts of utterance. For then the argument is circular, since it has a bearing on sentence meaning only if we are presupposing the claim that we are trying to make. The fact that an implicature cannot be cancelled from a particular context of utterance does not imply that it is not cancelable *per se*. Showing that a particular conversational implicature cannot be cancelled from a context of utterance only shows that different contexts of utterance determine different utterance meanings. It illustrates speech act pluralism and does not necessarily have an impact on literal meaning. A conversational implicature may play a role in the primary sentence meaning only if it remains present in different contexts of utterances. The fact that a particular implicature cannot be cancelled from a particular context of use is compatible with its cancel ability within different contexts of use. Particularized conversational implicatures may be difficult to avoid in a particular context of utterance, but the very same act of saying involved in them could have been made in quite a different particularized context of utterance, and this is all we need to argue that a conversational implicature is cancelable.

3 Are There Propositional Radicals?

Pragmemes very often determine secondary meanings. These secondary meanings are on top of the ones that are expressed by the sentence itself, or by the sentence and the context if the sentence contains indexical expressions. But could they determine primary meanings of sentences that are devoid of indexical expressions? It seems that there are cases that can illustrate this thesis. Some sentence-types may perhaps express only “proposition-radicals,” as suggested by Bach (2005), although they do not contain indexical expressions, on the surface at least. Pragmemes, according to this view, serve to determine the missing ingredients in the primary meaning. However, these cases are not clear counterexamples for minimalism, and not only because of the existence of a constant propositional radical. The reason is that the recourse to contextual features may be prescribed by semantic rules attached to the sentence itself. Minimal content is indeed not clearly threatened if the requirement of context must be imposed by some semantic conditions associated with the sentence type. The requirement of context in determining content is taking place only because semantic rules tell us that we must look at the context of utterance in order to determine content.

This is perhaps the main motivation behind Lepore and Cappellen’s wish to restrict context sensitive expressions to the basic set of indexical expressions (e.g. ‘I’, ‘you’, ‘we’, ‘they’, ‘that’, ‘this’, now, tomorrow, here, there, etc.), for these have an autonomous linguistic meaning. Indeed, the semantic character of an indexical such as ‘I’ is expressed by “the utterer of the this token” or, perhaps more conveniently, “the utterer of the token that is here and now.” This does not force us to draw a contextualist conclusion, since it is an autonomous semantic rule associated with the word-type ‘I’ that prescribes recourse to the context of utterance. In Kaplan’s (1989) sense, the ‘characters’ that are attached to words themselves are not determined by context. As function from context to content, characters are autonomous and are not themselves determined by context. Since the indexical expressions belonging to the basic set do have characters, it seems that by restricting context sensitive sentences to those containing expressions belonging to the basic set, we avoid the pitfalls of contextualism.

The problem, though, is that there are sentences that do not contain expressions belonging to the basic set but that are still incomplete in some sense. A sentence like “Rowan is Mr. Bean” may itself be context sensitive although it does not contain indexical expressions. But this may simply be because of an institutionalized use of proper names according to which they are meant to vary in different contexts. Like indexicals, proper names may have a ‘character’ expressing a function from context to content. This function may often be constant but it can also determine a different content in different contexts. Proper names may partly be assimilated to the set of indexical expressions. A proper name such as ‘Rowan’ is a shorthand for the definite description “The individual named ‘Rowan’,” and the definite article may refer to a unique individual in the context. Similarly for “It is raining.” This sentence contains a covert location variable that can be saturated

by means of an overt locative phrase. In all these different cases, sentences would be context sensitive despite the fact that they do not explicitly contain expressions belonging to the basic set, but it is just because they do contain implicitly such expressions.

For more controversial cases like “Mary has had enough” and “John is ready,” Montminy (2006) has argued that the best minimalist move would have been to treat them as saying something like “Mary has had enough of something or other” and “John is ready for something or other.” If that were the correct account, then “A is not ready” would be a shorthand for “A is ready for nothing at all,” which is absurd in most if not all contexts. But the correct minimalist analysis perhaps compels us to say that the sentence is to be completed by a reference to some specific element (an event, action or another kind of specific thing). We must perhaps acknowledge the fact that “A is ready” behaves almost like a proposition-radical in Kent Bach sense. But at the same time, it is available for logical inferences like

All those who are ready won't be surprised
 A is ready
 —
 A won't be surprised

Does that mean that there are minimal truth conditions for “A is ready”? Or is the sentence a proposition radical? I for one would tend to argue that the sentence implicitly contains an indexical, or a demonstrative. This seems to follow from the fact that it is all at once in need of completion by a reference to something specific, and available for logical inferences. Cases like “A is ready” seem to behave like ‘proposition radicals,’ but proposition radicals are as a matter of fact sentences that implicitly contain a demonstrative expression. This is perhaps a controversial move that Bach would not approve, but it is perhaps one that minimalists must approve. The utterance of “A is ready” with the intention to mean ‘for supper’ is semantically equivalent to an utterance of “A is ready for this”, where ‘this’ is used as a demonstrative to be completed by a demonstration in which the utterer is pointing to a plate full of food served in the evening. Now in “A is ready for this,” we have a semantically constant sentence, for the demonstrative itself expresses a semantical rule that prescribes saturation. It is of course in some sense still incomplete, but just as any old sentence containing context sensitive expressions belonging to the basic set.

If someone asks you out of the blue if you are ready, you will be inclined to answer: “for what?” This shows how incomplete the sentence is. But its incomplete character is explained by the elliptical and therefore implicit presence of a demonstrative expression like ‘this’, so that the sentence should read “John is ready for *this*” (or “Mary has had enough of *this*”). Here I follow Capone’s minimalist explanation of the nature of such incomplete propositions (Capone 2008). If they could be interpreted as implicitly containing empty slots that can be interpreted as demonstratives or discourse-deictic anaphoric expressions, sentences expressing incomplete propositions would indeed be harmless for

minimalism, for they would be analysed as implicitly involving expressions belonging to the basic set.

Similarly, “John is tall” and “Mary is rich” would also contain implicit semantic empty slots calling for completion by a particular reference class. These sentences should perhaps be analyzed as “John is tall (relatively to *this* reference class),” and “Mary is rich (relatively to *this* reference class).” Quantified statements like “All came for breakfast” would be an elliptical form for “All of *them* came for breakfast” (or “*They* all came for breakfast”), and would thus also be implicitly containing expressions belonging to the basic set of indexical expressions.

So it may be necessary to enlarge the set of context sensitive sentences beyond those explicitly containing words belonging to the basic set. But it would by no means ruin the main claims made by minimalists, for these sentences could be in need of saturation, not modulation. We may indeed only superficially be compelled to enlarge the basic set of indexical expressions, for the involvement of context is prescribed by semantic features such as the implicit (elliptical) presence of demonstratives. Minimalists need not be claiming that there are minimal sentence meanings associated with each and every indicative sentence of the language not containing indexical expressions. In many cases, they only have to postulate minimal ‘proposition radicals’ understood as expressions associated with the implicit semantic constraints of expressions contained in the basic set. There are empty slots involved in proposition radicals at the surface level, but they would always be counted as elliptical for sentences containing expressions belonging to the basic set of indexical expressions. So in addition to sentences expressing minimal propositions, there may perhaps also be sentences expressing proposition radicals that implicitly contain expressions requiring the presence of context. All of this is perfectly compatible with semantic minimalism.

4 Can Pragmemes Determine Sentence Meanings?

Contextualists argue that an innumerable amount of sentences not containing indexicals are dependent on context. They claim that there is an unlimited dependence of the meaning of different linguistic items on other sentential elements showing that the scope of items needing indexes is much larger than commonly accepted. A classic situation serves to illustrate the point. It concerns the use of color words. I stated at the outset that I would not discuss pragmatic intrusion concerning word meaning, but let us consider for a moment the use of color words. If someone says that red grapefruits are on sale at the department store, it seems that the truth conditions of the sentence cannot be determined without knowledge of speaker’s intentions. Is the speaker referring to fruits that are red on their surface or red inside their surface? The minimalist answers that the truth conditions of the sentence make reference to grapefruits that are red ‘period’, whether on their surface or inside their surface. Of course, the speaker may with

her use of the word 'red', intend to refer to those grapefruits that are red inside as opposed to those that are yellowish inside and outside. But once again, this is nothing more than a determination of speaker's meaning on the meaning of her utterance, and not a determination of speaker's meaning on the literal meaning of the sentence itself. In the context of utterance, the sentence becomes an elliptical way of saying 'red color inside the surface' by using the word 'red'. The additional information is relevant only for determining what the speaker means in the course of saying what she is as a matter of fact saying, not for determining the content of the sentence itself. A similar situation occurs in the case of the sentence "Pierre went to the gym." For the minimalist philosopher, the minimal truth conditions suppose that Pierre went to the gym "period", no matter whether he went inside the gym or in the vicinity of the gym. Most of the time, I may intend to refer to the fact that he went inside the gym, but this is what happens in particular utterances. My diagnosis is that contextualists very often confuse what is expressed by the sentence with what is expressed by an utterance of the sentence in the context of an illocutionary act. They take for granted that we must be looking for contextual features in order to determine the literal content of sentences, while context is as a matter of fact relevant only for determining the content of her illocutionary act.

We have seen that pragmemes do not play a role in determining primary sentence meaning. Of course, but this is perfectly compatible with the fact that pragmemes involve conversational implicatures and determine secondary meanings. We have also seen that pragmemes do determine primary meanings for very specific sorts of sentences, as long as these are interpreted as grammatically sensitive to context and as long as they do so because of the explicit or implicit presence of expressions belonging to the basic set. But in these examples, pragmatic features are not optional, since they are semantically called for. So we still do not have a strong case for semantic contextualism.

We must now ask whether pragmemes can sometimes also optionally determine primary meanings for sentences. It would then be a case of modulation and not of saturation. We are wondering whether pragmemes can sometimes involve explicatures, that is, speech acts involving a pragmatic intrusion in the determination of literal content that determine optional primary sentence meanings. Carston (2002) and Recanati (2004) both argue that there are many cases like this. Recanati's contextualism stipulates the existence of a relation of dependence between literal meaning and pragmemes such that the literal truth conditions would all at once be (1) primary, (2) optional and (3) intentional.

(1) The pragmatic factors concerned are *primary* in the sense that they play a part in the determination of the literal truth-conditions of what is said with a sentence. They can be distinguished from secondary pragmatic factors that add an additional layer of (pragmatic) meaning to the literal meaning. As primary factors they are not to be confused with irony, metaphor, conversational implicatures and indirect speech acts. All of these presuppose statements that already have a literal meaning. But we are more interested in phenomena that play a part in the specification of literal or primary truth-conditions. (2) The relevant pragmatic factors are also *optional* in the sense that they do not stem from semantic rules associated

with expressions. Indexical and demonstrative expressions, for example, have functioning rules that semantically enforce recourse to the context of utterance. The context allows us to complete the statement, which then expresses full and complete truth-conditions. This would be the phenomenon of saturation that minimalists are in a position to accept. It should not be confused with the phenomenon of modulation that contextualists are postulating, in which facultative pragmatic factors play a useful role for determining the literal truth-conditions of a given statement. In order to show that the primary truth conditions that pragmemes determine are not prescribed by the literal meaning of the sentence, Recanati argues that they are optional, which is another way of saying that they are cancelable. If indeed the intended meanings are optional, then they are not imposed by the very semantics of the sentence and they seem to offer a clear case of intrusion of pragmemes in the very content of the sentence. (3) Finally, the pragmatic factors concerned are also described as *intentional* in the sense that it does not suffice to refer to a limited notion of context that implies nothing more than time, place, speaker, and proximal or distal features. We must also bring in the intended meanings and the beliefs of the speakers.

The essential idea of Carston and Recanati is that primary, optional and intentional pragmatic factors will at times intrude *between* the conventional meanings of words and the actual truth conditions of sentences. The literal truth-conditions of many utterances are determined in part by the conventional meaning of constitutive expressions, but also by these sorts of pragmatic factors. According to them, the truth-conditions that stem merely from the conventional meaning of words are not always relevant. The truth-conditions that are meant by the speaker—and which are accessible to hearers—are also at times the relevant determining factor. Recanati maintains that diverse pragmatic phenomena serve to illustrate this point of view. There are cases of enrichment like for example, “he took his key and opened the door,” in which we understand that he opened the door *with* his key; cases of loosening like for example, “the bank machine swallowed my credit card,” in which we directly grasp what is meant without ever considering real swallowing; and cases of transfer like, for example, “the ham sandwich left without paying,” in which we directly grasp that it is the eater of the sandwich who took off without paying.

Recanati imposes an accessibility condition on literal meaning. It is the meaning that we all directly have access to in any context of utterance. So he also maintains that the intended meaning goes hand in hand with what is recognized by the addressee. In other words, he maintains that uptake must be secured. The intended meaning is also the meaning to which a normal interpreter would have access. He therefore goes on arguing that sometimes, speakers and hearers have direct access to intuitive truth conditions that do not correspond to those that seem to be expressed by the sentence itself. He also claims that speakers and hearers do not even compute the available meaning on the basis of the truth conditions supposedly determined by the sentences themselves. This is why the intended truth conditions are said to be primary. They are not parasitic upon a minimal content that would be asserted by the speaker and grasped by the hearer.

How can we reconstruct the argument for contextualism? Here is a first attempt. I shall formulate it assuming for the sake of argument that the literal meaning of a sentence is to be parsed in terms of truth conditions. It is assumed that truth conditions constitute the most important semantic ingredient of a sentence in the indicative mood. Now very often the truth conditions of *utterances* require for their specification the presence of pragmemes. There is therefore a close link between the literal meaning of sentences and pragmemes, and this suggests that pragmemes are the primary vehicles of meaning. We should by now know that his version of the argument is of course not valid. The fact that the truth conditions of an *utterance* requires context does not prove that there is a close link between the literal meaning of the *sentence* uttered and the context of utterance. It establishes a connection between an intention and the utterance of the sentence and not with the sentence itself. There is however a variant of the same argument that might look more sound. It is this variant that seems to be defended by Recanati. A missing premise in the first version of the argument helps establishing a close connection between the primary meaning of sentences and the utterances of those sentences. The availability principle provides the missing premise in the argument. The availability principle stipulates that the truth conditions of a sentence must be those that are available to both speaker and hearer (Recanati 2004, 20). Now the primary truth conditions that people have access to are very often the ones that are determined by pragmemes. Furthermore, this access is not mediated by the literal meaning of these sentences. That is, speaker and hearer have a direct access to these truth conditions. From these premises, it seems we can draw the appropriate contextualist conclusion. It should be stressed that this view is compatible with the suggestion that there are truth conditions expressed by the sentence itself. It is just that the semantic content expressed by the sentence itself is not relevant in determining the content that is directly available by speaker and hearer. The meaning of the sentence itself is not be directly available to those who are parsing the sentence in that context, and so it is not relevant in these contexts.

It is important to note the analogy between the argument thus reconstructed and the thesis argued for by Dummett (1993) that a theory of meaning must also be a theory of understanding. This claim played a major role in Dummett's argument that the meaning of a sentence could not simply be correlated with a set of truth conditions. Speakers and hearers very often do not have a direct access to the truth conditions of sentences. So a theory of meaning as truth conditions falls short of determining what is understood by speakers and hearers. If truth conditions fall on the side of 'denotation,' there has to be a corresponding side of 'sense' that captures what is grasped both by speakers and hearers. Dummett thought that what is grasped is a verification procedure. Recanati's argument is somewhat similar, for he imposes a normative constraint on what is to count as literally expressed by the sentence. This cannot merely be the proposition expressed by the conventional rules associated with the sentence. Literal meaning has to include what is directly accessible to both speakers and hearers. And then Recanati claims that the intuitive truth conditions that are associated to a sentence often do not coincide with the official ones prescribed by the sentence itself.

There is however an important difference between Dummett's argument and Recanati's argument. Dummett's verification procedures are themselves associated with sentence types and not with sentence tokens. Dummett is not a contextualist philosopher. Recanati's intuitive truth conditions are accessible by speakers and hearers in particular contexts of utterance. The accessibility constraint applies to what takes place in a conversational context. The question should then be asked: why should we map the intuitive conversational truth conditions onto the sentence itself and draw conclusions concerning its literal meaning?

Dummett's claim that a theory of meaning must be a theory of understanding enables us to see a little more clearly what is involved in Recanati's own theory. But it does not do all the work needed for concluding that the intuitive truth conditions are part of the primary meaning of the sentence. In order to see this, it is important to note that the accessibility constraint may in principle be satisfied in a given context by the truth conditions expressed by the sentence itself, even if these are not intuitively those associated by hearer and speaker in another context. This is a consequence that follows from the fact that the pragmatic features are optional. If they are truly optional, then they may fail to occur in a given context. In that context, speaker and hearer may directly have access to the truth conditions of the sentence itself. So if Recanati is truly committed to provide optional pragmatic features, he should consistently treat the literal meaning of the sentence as accessible both to speaker and hearer. And so according to his own account, the literal meaning of the sentence is in these contexts the primary meaning of the sentence. If pragmemes are really optional, and if there can be contexts in which the semantic content of the sentence, that is, the minimal proposition (or its minimal truth conditions), is directly accessible to all the participants in the conversation, the literal content of the sentence is in those contexts part of the normal interpretation of the sentence. So it appears that the minimal content is accessible after all, and thus could be treated as the primary content expressed by the sentence, allowing for the presence of a secondary layer of intended meaning in other contexts. But this is not what Recanati is suggesting. He is rather arguing that the primary meaning of a sentence in a given context is determined by what is accessible *in that context*. So if in a specific context, the intuitive truth conditions associated by the speaker and hearer are different from the ones expressed by the sentence itself, then in that context, the literal meaning of the sentence expresses a content that is determined by pragmemes. But it now appears that the notion of accessibility is context relative. The fact that the primary truth conditions could coincide in some context with those that are expressed by the sentence could be used to show that these minimal truth conditions are accessible in that other context, but this according to Recanati has no implications for determining what is accessible to both speakers in a different context. The determination of primary truth conditions should always be a function from what is accessible in a context.

Here one senses the danger of a circular argument. For we are invited to accept that primary meaning is determined by what is accessible in a context, in the course of an argument purporting to show that the primary meaning of a sentence is context relative. Why should the accessibility constraint be construed as context

relative? Is Recanati not cheating here a little bit? We want to determine whether literal truth conditions are context relative, and we are told that they must be accessible. But if it is implicitly stipulated that accessibility must be context relative, then the accessibility premise in the argument is doing all the work.

I shall soon return to the circular character of the contextualist argument. For the moment, let us simply conclude that there appears at best to be contexts in which content is determined by the minimal proposition expressed by the sentence, and contexts in which the content is determined by pragmemes. But which is the content that is so determined? Is it the content of utterances of the sentence or the content of the sentence itself? If accessibility is context relative, it seems that we can only conclude that it is the content of utterances. So one may wonder whether pragmemes do indeed determine the literal meaning of the sentence. Are they not instead only determining the truth conditions of a sentence-in-a-context-of-utterance, and therefore determining the truth conditions of an utterance of the sentence, and not of the sentence itself? And if so, then where is the intrusion?

5 One Last Refuge

Recanati ran into trouble because he correctly realized that pragmatic features had to be optional. They had to be optional because they should not be prescribed by the grammatical sensitivity of the sentence to context. The problem with that solution, as we have seen, is that it implies that the truth conditions expressed by the sentences themselves will also be accessible in some contexts. It is a problem because these truth conditions look very much like the minimal truth conditions that contextualists are trying to avoid. The only way out is to contextualize the accessibility constraint. In other words, it is assumed that the meaning of a sentence is what is accessible in a certain context. From such a premise, it is then easy to draw the conclusion that meaning is context relative. But apart from the fact that the conclusion of the contextualist argument seems to be contained in such a premise, there is also the problem that if availability is context relative, contextualized intuitive truth conditions will have to be associated with utterances, not with the sentences themselves. We are back to speech act pluralism which is not an issue opposing minimalists and contextualists, and contextualists are still unable to show how to go from there to the conclusion that sentence meaning is context relative.

In order to avoid circularity and an argument that has a bearing only for utterances and no impact on the primary meaning of sentences, it is perhaps important to see whether there are pragmatic features that have an impact on the primary meaning of sentences without being optional and without being prescribed by the semantic constraints of sentences. Are there pragmemes determining primary truth conditions that are not prescribed by semantic features and that are not cancelable? Burton-Roberts (2006) and Capone (2006) argue that there are. For example, “Pierre shrugged and left” means (via explicature) “Pierre shrugged and

then left,” and cannot be interpreted otherwise although nothing in the sentence is grammatically sensitive to context. But it is not clear why we could not read the sentence as describing two simultaneous events. And if so, interpreting it as describing a sequence of events can be cancelled (Pierre shrugged and left *all at once*). Similar remarks apply concerning the example discussed by Carston (2002, 138). The explicature of “Pierre ran to the edge of the cliff and jumped” is something like “Pierre ran to the edge of the cliff and jumped over the edge of the cliff.” For Carston, this explicature can be cancelled by saying that Pierre ran to the edge of the cliff and jumped but stayed on the top of the cliff. It is simply false to suggest that Carston is playing on two different meanings attached to the word ‘jump’: a directional meaning and a transitive meaning. The fact of the matter is that the sentence initially contains an intransitive occurrence of the verb. But when the sentence is *uttered*, the speaker and hearer most often parse truth conditions that go beyond the literal meaning of the sentence, in accordance with a transitive use of the verb ‘jump’. But since someone may also in other contexts use the same sentence in accordance with the literal intransitive word ‘jump’, the usual explicature can be cancelled. Even if the utterance of the sentence usually make use of a transitive interpretation, the tragic interpretation need not be the good one.

A more interesting case is the following:

1. If the king of France died and France became a republic I would be happy, but if France became a republic and the king of France died, I would be unhappy,

This sentence ‘*prima facie*’ appears to be contradictory, but the contextualist philosopher argues that it is not really so. Since one cannot imagine someone intentionally accepting a contradiction, the contradictory reading, if Capone is right, is simply not in the cards. There is only one possible interpretation that avoids the contradiction and it is one in which the conjunction is read as involving an ordered sequence of events. We should parse (1) as saying that if the king of France died and *then* France became a republic I would be happy, but if France became a republic and *then* the king of France died, I would be unhappy. But consider a context in which the speaker says: although I am a republican, I like the King of France, and it is not good news if I hear that the king has died, no matter how and when he died. But I always prefer to hear the bad news first and then the good news, no matter what is the actual sequence of events. So for instance, if I am told that the king of France died and told immediately after that France became a republic I would be happy, but if I am told that France became a republic and then told that the king of France died, I would be unhappy. This might be the explicature for the initial sentence. So it is not true that the first explicature is not cancelable.

Here is a third possible context of utterance. The speaker has mixed feelings concerning the complex state of affairs involving a revolution and the death of the king. As a matter of fact, she has contradictory emotions of happiness and unhappiness. In her assertion of (1), she is expressing these contradictory feelings by referring twice to the complex state of affairs, and the order in which the facts

are mentioned is not relevant. The point here is that when it comes to emotions, contradictions are quite possible. We live our lives full of contradictory emotions and mixed feelings about our personal and social environment.

A fourth possible context of utterance is the situation where the teacher in a classroom intends to show that from a false premise one could infer a true proposition as well as a false proposition. Let us assume that the conjunctive proposition “The king of France died and France became a republic” is false, and that the sentences “I am happy” and “I am unhappy” are respectively true and false. She then proceeds in uttering (1) to illustrate the point.

But no matter what is the specific context of utterance, we are not even condemned to fix the truth conditions of the sentence by appealing to contextual features in order to avoid the disturbing consequence of having to deal with the intentional acceptance of a contradiction. The logical form of the sentence expressed is:

2. (If P & Q then R) AND (if Q & P then –R)

or, in a more condensed form,

3. (If P& Q) then (R & –R)

If we construe the ‘if-then’ conditional form as a material implication, then there is a case where (3) could be true. The only case where (3) is true is the one in which the antecedent is false, since from a false premise, one can infer true and false conclusions.

But (1) raises general questions concerning the possibility that someone would use the sentence to utter what looks like a blatant contradiction. So let us rephrase it to illustrate the difficulty. Let us avoid the formulation of a material conditional and avoid referring to contradictory *feelings* granted that it is possible to entertain contradictory emotions toward the same state of affairs. We get something like

4. I am informed that the king of France has died *and* France became a Republic, and I believe that it is a good thing; I am **then** informed that France is a republic *and* the king of France has died, and I believe that it is not a good thing.

Here it seems hard to avoid the annoying impression that the speaker is involved in an intentional contradiction, unless we explicate the italicized conjunctives ‘and’ as directly meaning ‘and then’. In other words, we apparently here have no choice but to articulate the italicized conjunction in such a way as to reveal a sequence of events, if we are to avoid an intentional contradiction. We can then reinstate Capone’s interpretation, as in the initial interpretation and this time, we apparently have no way out. We no longer are in a position to read (4) as invoking a preference to hear the bad news first, or as involving contradictory emotions. And we do not have an explicit reference to a conditional sentence. But (4) need not be construed as involving an explicated ‘and then’ construction for the italicized ‘and’, since as the bold occurrence of ‘then’ shows that there are two occasions in which one is being presented with information concerning France and

the King. So the speaker may simply be telling the hearer that he changed his mind. So he need not be ordering in a sequence of events the death of the king and the advent of a French republic. She rather means to suggest that the two events were for her initially seen as a good thing, and they no longer are. So once again, Capone's explicatures are cancelable.

Here's now the ultimate attempt to impose a contextual reading on the conjunctive 'and' and force into it an explicature that apparently cannot be cancelled:

5. I believe that London is pretty *and* I believe that London is not pretty

Granted that it is impossible to entertain contradictory intentional beliefs, the italicized 'and' can only be read as implying something like 'and then', suggesting that the speaker has changed his mind. But the very same sentence could be used in the classroom to give an example of the impossibility of contradictory intentional beliefs. In other words, the whole sentence (5) is mentioned and not used by the teacher. Nevertheless, it is a case where it is important not to read the conjunction as involving a reference to a sequence occurring in time, for that would not necessarily be intentionally believing two contradictory propositions. If this is correct, then it follows that the 'and then' reading of the conjunction is cancelable.

As I have shown in the case of conversational implicatures, cancelability is a notion that should be applied with reference to all contexts of utterance and it should not be construed as indexed to a context. That is, if it's cancelable in one context, then it's cancelable period. I agree that some utterances performed in certain contexts are to be explained by a specific explicature and that all others should be excluded in that context. So some alternative explicatures do not even arise in these contexts. But explicatures should not be indexed to a context, unless of course we stipulate at the outset that they must be context relative. So we only need to show that an explicature is cancelable in one context in order to show that it's cancelable at all. Otherwise, we would once again build contextualism in the very notions we are using. The substantial conclusions would already be contained in these facts.

6 Saying and Asserting

Let me now move to considerations that diagnose the confusions leading one to embrace semantic contextualism. The first problem stems from the usual failure to distinguish between locutionary acts of saying and illocutionary acts of asserting. I believe that Searle was wrong to abandon the distinction between locutionary acts and illocutionary acts. The locutionary act is an act of saying something meaningful while uttering something. The speaker says the content expressed by the sentence uttered. She refers to objects, expresses senses, predicates properties, implies semantic consequences and presupposes semantic presuppositions. This is what is involved in her locutionary act of saying something. In particular, there is no commitment to truth involved in an act of saying (or locutionary act). Whether I

assert, promise, order, or declare that *p*, I am in each case saying that *p*, so saying is not just another illocutionary act. I can consider whether *p*, ask whether *p* is true, I can wonder what would happen if *p* and I can ascribe to someone the belief that *p*. In all these cases, I am saying that *p* without asserting it.

What is the difference between an illocutionary act of assertion and a locutionary act of saying? There are two separate issues involved: the locutionary / illocutionary distinction and the distinction between saying and asserting. In indirect speech or in the utterance of propositional attitude sentences, the subordinate clause is said but not asserted. Consider the sentence:

6. Graham believes that we can save the world

If that sentence is uttered, then the subordinate clause is also uttered. So there is a phonetic act involved. And since the subordinate clause is a well formed grammatical sentence, then the phonetic act involved is also a phatic act. And since the sentence is meaningful, it is also a rhatic act. So the utterance of the subordinate clause is a locutionary act. But if I utter (6) I perform no assertion of the subordinate clause. Of course, very often, when we perform a locutionary act on a sentence, we do also simultaneously perform an illocutionary act on that sentence. The illocutionary act is performed in the course of performing the locutionary act. So if you utter

7. We can save the world

The act of saying goes along with a commitment to the truth conditions, but it is because in the course of saying it, you performed an assertion. If you had uttered "He believes that we can save the world," you would also have performed a locutionary act of saying the same minimal proposition, but without commitment to the truth of the proposition expressed, because it is embedded in a larger sentential context. The locutionary act is the act of expressing a proposition and expressing minimal truth conditions, but it does not involve a commitment to the truth of the sentence.

The other issue concerns the distinction between saying and asserting. There may be a use of 'saying' that amounts to 'asserting'. But here I am using 'saying' in a somewhat technical sense synonymous with what I take to be the locutionary act. We must coin an expression to refer to what is happening when we are using (as opposed to mentioning) a sentence without committing ourselves to the truth of that sentence. Is there a sense of 'saying' in ordinary parlance that captures this technical notion of a locutionary act? The answer is 'yes'. There are at least two institutionalized uses of the word 'say', and I am going to use them in the next sentence. Let me just say-1 that when we do not make an assertion, but we order, promise, declare something, we are always saying-2 something. In the sentence just uttered, 'say-1' is the illocutionary use of the ordinary word 'say', and 'say-2' is the other use, the one that corresponds with the locutionary sense. If I utter "Alice believes that school is out for summer" and "Al believes that school is out for summer," I say-1 two different things about Alice and Al. That is, I make two

different assertions. But in both cases, I say-2 the same thing about them. To make use of Davidson's vocabulary, there is a same saying relation between the two subordinate clauses.

We can illustrate the use of 'say-2' in ordinary parlance with another example. If you utter

8. Are you experienced?

and someone does not hear you clearly, she could ask, "what did you say-2"? If you wrongly stick with a concept of saying understood as always involving an illocutionary act committing one to truth, then you should answer:

9. I was not saying anything. I was just asking a question.

This sounds odd. In my sense of say (say-2), you should simply repeat your question.

Why is it so important to distinguish between locutionary acts of saying and illocutionary acts of assertions? There is a philosophical argument that can be made on the basis of the distinction that has a bearing on the issue of contextualism. Locutionary acts express minimal propositions or minimal truth conditions. Full blown assertions may come equipped with loads of presuppositions and background beliefs, and therefore often determine maximal truth conditions. We may use sentences that express minimal propositions or minimal truth conditions in order to capture, express or describe what someone is saying-2, but we also very often use them while presupposing the complex cognitive architecture of each other's mental framework. The sentence "Pierre is cutting the grass" expresses a minimal proposition or minimal truth conditions and it is true in the minimal sense if Pierre cuts the grass, whether he is using a lawnmower or a razor blade. But in the thick, robust sense, it may be asserted with certain expectations and background presuppositions. It is in this latter sense that we are entitled to claim that the 'truth conditions' have not been satisfied if Pierre only used a razor blade.

Illocutionary acts are done in the course of saying something. What am I doing if I perform an assertion while saying something? In an assertion, I express a belief, I presuppose the existence of a justification for the content of my speech act, and I imply that the content of my speech act is true at least in part because of the presupposed justification. In a promise, I express an intention, I presuppose that you expect me to do the thing specified by the content of my speech act, and I imply that the content of my speech act will be fulfilled at least in part because of your expectation. In an order, I express a desire, I presuppose the existence of an expectation on my part that you do the thing specified in the content of the speech act, and I imply that you should bring about the situation described by the content of my speech act at least in part because of my expectation. In an expressive illocutionary act, I express an emotion, I presuppose the existence of a state of affairs and I imply that the content of my speech act is justified partly because of the presupposed state of affairs. Finally, in a declarative illocutionary act, I express my decision to bring it about that p by saying 'p', I presuppose that I have the

authority to bring it about that *p* by saying it, and I imply that *p* is brought about by my utterance because of my presupposed authority.

Now all of this is done in the course of saying something. So saying *p* is one thing, and asserting *q* while we are saying *p* (whether or not $p = q$) is another thing. There are alternative ways of construing the taxonomy of illocutionary acts (viz. Searle's notion of direction of fit), but the important point is that it does not have much to do with locutionary acts of saying things.

7 Intentional and Material Reports

I have distinguished two uses of saying: one that corresponds with asserting, that is, an illocutionary act, and one that corresponds to the locutionary act. In other words, I grant that sometimes we use 'saying' as synonymous with asserting, but there is still a distinction to be made because saying is also sometimes used in the sense of a locutionary act. But we should also introduce a distinction between 'intentional' reports and 'material' reports in order to capture another distinction between two different uses of saying in its locutionary sense (Seymour 1999, 1992). As we shall see, this distinction can also be used to refute contextualist philosophers.

Let me first discuss very briefly the distinction between these two kinds of report. An intentional report describes an intentional state, act or action. In the full blooded sense of intentionality, the state, act or action will have all the usual features of intentionality: directedness (intentional object), intentionality (with an *s*), reflexivity and first person authority. According to that last feature, if someone is in an intentional state of belief that *p*, she knows that she believes that *p*. If someone performs an intentional act of saying, the person knows what she is saying.

A material report describes a state, an act or an action by supposing much less than full blooded intentionality. It assumes the existence of a functional state of the agent. The properties of directedness and perhaps to a certain extent also intentionality are present, but not necessarily reflexivity and first person authority. In the material sense, Fido might be described as believing that there is a cat in tree because he behaves in a way that seems to take for granted the existence of a certain state of affairs. Unconscious beliefs provide another example. Oedipus believed he wanted to marry Jocasta, but he did not realize that Jocasta was his mother. So he did not intentionally believe that he wanted to marry his mother. But at the level of his unconscious states, he might have believed it.

Of course, since intentional states, acts and actions are themselves types of functional phenomena, material reports can also apply to them. If I intentionally believe that *p*, I also as a matter of fact believe that *p*. Intentional states are types of functional states. A material report is one that describes a state, an act or an action as functional without assuming full blooded intentionality. But it does not deny the presence of full blooded intentionality in what it is describing. This is why it can also apply to full blooded intentional states, acts or actions. But since it does not

assume full blooded intentionality, it can be used to describe functional states, acts or actions that do not exhibit all the properties generally associated with full blooded intentionality.

Another instance of application of material reports would be concerning certain kinds of locutionary acts. When I utter some sentence, there are things that I am saying that I do not necessarily fully comprehend or entertain. I might of course know what I am saying while saying it, but I might also fail to attend or grasp all the elements involved in my act of saying. This might be because of my ignorance (not fully grasping the meaning of a word, for instance) or simply because I did not fully attend to what I was saying.

This can happen when I intend to mean something in the course of saying some other thing. There is something I intend to mean in the course of my act of saying and, precisely for that reason, I am not entirely vigilant about the actual meaning of what I am in fact saying. So when someone utters 'she took her key and opened the door' or 'the cash machine swallowed my credit card' or 'the ham sandwich left without paying,' she does not necessarily realize what she is actually saying. There are true material reports that could describe what she is actually saying even if they would not describe what is taking place in her mind in the course of her actual intentional assertion.

So most cases that seem to serve the cause of contextualist philosophers can perhaps be explained by using the distinction between intentional and material reports and applying it to the locutionary act of saying. Although the speaker intentionally meant that the key was used in opening the door, that the cash machine did less than swallow something and that it is a person that left without paying, she did as a matter of fact say in the material sense what is literally expressed by the sentences she uttered.

Now in their use of language, speakers also defer to others. So if someone attracts her attention to what she actually said in the material sense, she will now entertain or realize what she earlier failed to entertain. Since she defers to others, she will recognize that what she did not consider or apprehend was actually said.

To conclude on this, we must accept the distinction between what the speaker actually says and what the speaker intentionally asserts while saying it, even in contexts in which the speaker does not intentionally entertain the content of what is said. These contexts do not prove that the intended meaning of the speaker intrudes in the literal meaning of the sentence, because the speaker can be described as having said (in the material or functional sense) what is expressed by the sentence. I earlier argued for the existence of minimal content even in the case where both speaker and hearer have access only to the content of the illocutionary act. My point was that the content of the sentence used was also accessible both to the speaker and the hearer. My argument for this was that since the intended meaning was optional, there should be other contexts in which the meaning of the sentence itself was fully accessible to them; and I argued that this was all we needed in order to claim that sentence meaning met the accessibility condition, even in the hard cases discussed by contextualists. Now I claim that there are other reasons for suggesting that what is expressed by the sentence meets the

accessibility condition in the three examples mentioned. This condition is met even in the context in which the speaker only has in mind what she means without knowing, or without fully attending to, what she is actually saying. In this kind of situation, even if the speaker may perhaps fail to perform an intentional act of saying, she can be described as actually saying (in the material sense) what is expressed by the sentence she is using. For if someone informs her or attracts her attention to what she is actually saying, she will then intentionally be attending to what was initially expressed by the sentence itself. As a deferring member of a linguistic community, she will herself acknowledge that she was saying what the sentence was expressing. Her disposition to defer to people will induce her to recognize that what she was in fact saying was different from what she intended, but that it was nevertheless what she said. What she said was the minimal content postulated by minimalist philosophers.

8 Circular Contextualist Arguments

I am concerned about the implications of ignoring the locutionary act of saying in the argument that leads to contextualism. If we have methodologically decided to consider only illocutionary acts of assertions, we will then perhaps inevitably be led to think that sentences can only be meaningful in contexts. Cappellen and Lepore admit speech act pluralism, and this relates to what takes place at the illocutionary level. They agree that we can use a sentence in many different ways in order to perform different kinds of illocutionary acts: that is, acts with the same illocutionary force (assertion) but with different contents. We should not be surprised about this since, as we saw, illocutionary acts of assertions come with expressed beliefs, pragmatic presuppositions and pragmatic implications. If we begin by ignoring at the very outset the legitimacy of the distinction between saying and asserting, this can play a major role in the argument for contextualism. I am wondering whether we are not ruling out from the very beginning the alternative conclusion, that is minimalism.

This would be a somewhat circular argument in favor of contextualism. First, we reject the distinction between locutionary acts of saying and illocutionary acts. We then proceed to consider illocutionary acts only. We then note that many different intentional illocutionary acts of assertions can be performed on a single sentence even if it is devoid of indexicals and demonstratives. Primary pragmatic features in the context seem to explain this pluralism. Therefore, literal meaning is strongly dependent on context, and pragmemes become the primary vehicle of meaning. But in this argument, the rejection of the distinction between locutionary and illocutionary acts of assertions in favor of intentional illocutionary acts is an indication of a possible circular argument. If our objects of study are intentional acts of asserting many different things in different contexts, then obviously, meaning is context relative and pragmemes are the primary vehicles of meaning.

Of course, the contextualist philosopher may be willing to accept the distinction between locutionary acts of saying and illocutionary acts of asserting. At first, granting this point might be seen as granting the truth of minimalism, because it seems like she is willing to accept the existence of minimal propositions or minimal truth conditions. But just like the radical contextualist philosopher might be willing to accept that words have a ‘semantic potential’, the moderate contextualist philosopher might be willing to accept that sentences only have a semantic potential, but it is one that is not relevant for determining the intuitive truth conditions that both speaker and hearer associate to the sentence. But I still have worries about the circularity of the argument, even when the distinction between saying and asserting is granted. For in the argument that purport to establish the importance of the intuitive content accessible to both speaker and hearer, the contextualist philosopher is assuming that what is important is located at the level of occurrences, and this also influences the conclusion of the argument which is precisely that meaning is relative to contextual features in the context of utterance. I have expressed this worry in my critical study of François Recanati’s *Literal Meaning* (Seymour 2006). Recanati replied that he is perfectly willing to admit that, in some sense, we may be saying the same thing when we assert two different things. And he thinks for this reason that he is not vulnerable to a criticism of circularity in his argument. There may be something minimal that we say even when we assert many different things, but that does not prove minimalism. On the contrary, Recanati insists that the content of what is said, even if it exists, plays no important role in the process of a normal interpretation, which is after all the only game in town. But it is here that I locate the most important danger of a circular argument. The real problem concerns the emphasis on interpretation or actual processing. If we are all interpreters struggling to decipher particular inscriptions, and if what is crucial concerning primary meaning is to capture what is going on in utterance events, well then of course, we must concede immediately victory to the contextualist philosopher. Linguistic inscriptions are events, or tokens, and it has already been granted that we could make very different illocutionary acts of assertions with the same sentence, even if it is devoid of indexicals and demonstratives. Now if doing this is the only game in town, then of course, pragmemes are at center stage and they have won the day.

There may of course be many different contexts of utterance in which the literal, minimal, primary content does not appear to coincide with the content of the interpretation. But this has a bearing on literal meaning only if, from the start, we assume that literal meaning is what is taking place at the illocutionary level.

There is an important discrepancy between what is actually said and what is interpreted in cases of enrichment (she took her key and then unlocked the door), loosening (the cash machine swallowed my credit card) and transfer (the ham sandwich has left without paying). What is ‘literally’ said may be interesting, but according to the contextualist, it is not relevant to what is actually asserted and interpreted. But the minimalist philosopher is eager to reply: so what? What is the bearing of these observations on the issue of literal meaning? It is here that the contextualist philosopher is forced right from the start to answer that what is

important in meaning is what goes on at the level of speech acts. But wasn't it precisely the conclusion the contextualist was looking for?

So I am afraid that contextualist conclusions are very often implicitly contained either in the accessibility constraint, if it is interpreted as accessibility-in-a-certain-context, in the rejection of the locutionary/illocutionary distinction, or as suggested above, in the methodological principle that interpretation or information processing is the only game in town. We saw also that another instance of circularity is bound to occur in arguments that assume that one could diagnose pragmatic intrusion if some pragmatic features cannot be canceled-in-the-context-of-utterance. If what is important occurs at the level of assertions, interpretations and accessible truth conditions in a context, well then of course, it is hard to resist the contextualist conclusion. But these methodological assumptions make almost all the work in the argument, and they already presuppose the truth of contextualism. If we insist that what is important concerning meaning is what takes place at the level of intended meaning, interpretation or illocutionary acts, then what is important is automatically related to occurrences, tokens and events along with their contextual features. We should not then be surprised to be in a position to conclude, along with Recanati, that meaning is relative to illocutionary acts and that illocutionary acts are the primary vehicles of meaning. Recanati's argument is that even if there is a minimal proposition expressed by a locutionary act, it doesn't necessarily play a part in occurrent on-line processing, but he is assuming from the start that what is important to meaning theory is "occurrent on-line processing," and this is why he is in a position to conclude that meaning is context relative.

A similar debate has been raised with Rob Stainton in a private exchange. He asks: "Are we trying to model natural languages, understood as systems of expressions, or are we modeling human psychological processing of language in context?" If we are assuming from the start that meaning theory is "modeling human psychological processing of language in context," well who will be surprised about the conclusion that meaning is relative to context?

9 A Coherent Theory?

One could also question the very coherence of the contextualist argument if the moderate contextualist philosopher is willing to grant the locutionary/illocutionary distinction. Can contextualism be coherently defended with the locutionary/illocutionary distinction if primary pragmatic features are optional? For if they are truly optional, then there may be contexts of utterance in which they are not taking place. That is, the options are cancelable. In these other cases, the speaker does not perform the same pragmemes and her intended meaning might in these contexts coincide with the content of what is said. In accordance with the accessibility principle, this minimal content should therefore also be accessible to the hearer. The conclusion would then be that the truth conditions that follow from the sentence itself are accessible and should be primary truth conditions. Now since for

any given meaningful sentence of the language, there are minimal truth conditions or minimal propositions expressed by the sentence itself and accessible to the speaker and hearer. But this is exactly what the contextualist philosopher is supposedly denying.

According to Recanati's theory itself, optionality entails that for each sentence there must be at least one use made in accordance with its locutionary meaning, and so the theory itself is committed to acknowledging the existence of two relevant levels of meaning: a sentence (or literal or locutionary) meaning (that is, minimal propositions or minimal truth conditions) and a pragmatic meaning. Both are accessible if primary pragmatic features are truly optional. So aren't we stuck with two levels of meaning, as the syncretic view suggests? But the syncretic view is a particular version of minimalism, not of contextualism. The cancel ability or optional character of the pragmatic features signals the presence of a minimal content that could have been expressed without these additional features. This minimal content must therefore also be somehow always present, even in the context of utterance in which pragmemes are said to intrude in the determination of primary truth conditions. Now if minimal proposition are acknowledged, they must determine minimal truth conditions. And so the intuitive truth conditions associated by speaker and hearer in a context become secondary. So it seems that Carston and Recanati cannot have it both ways. Pragmemes cannot all at once determine *primary* truth conditions and be cancelable. The only way out of this dilemma is to construe cancel ability and accessibility as context relative notions.

The crucial issue is that of accessibility. For if the potential truth conditions are accessible in certain contexts, then they are accessible period, unless of course, we arbitrarily decide to constrain the accessibility principle to the context of utterance. This, I believe, is what is taking place in most if not all contextualist arguments. Since optionality implies that the potential minimal content is accessible in certain contexts, accessibility has to be construed as a contextual feature in order to avoid the conclusion that there are minimal literal contents in each context of utterance. Minimal content will be accessible in certain contexts but it won't be accessible in other contexts. So it appears that the only way out of incoherence for the contextualist is to say that there is no such thing as accessibility divorced from context. Literal meaning must intimately be related to what is accessible in a context. The contextualist philosopher must assume from the start that the medium in which literal meaning is determined is the context of utterance. So a circular argument seems to be the only way out not only for proving contextualism but also for avoiding an internal incoherence in the theory.

10 The Final Blow

For the sake of argument, let us avoid the problems of circularity and internal coherence of the theory. Let us suppose that the contextualist philosopher can make the controversial and circular methodological claims according to which the

most important level of meaning is language processing. Let us ignore the fact that this claim contains in a way the essential ingredients involved in the conclusion of the argument. Let us instead consider the debate concerning the claim that what is said, the minimal proposition (or minimal truth conditions), is not relevant in this process.

The minimalist philosopher could still argue against contextualism that even if language processing were crucial, minimal propositions or minimal truth conditions would still be essential to the language processing of sentences involved in cases of enrichment, loosening and transfer. Without them, interpretation would never get off the ground. This is especially so in the case of Ms Malaprop. When Ms Malaprop is interpreted as asserting that this is a nice arrangement of epithets, her actual act of saying that this is a nice derangement of epitaphs does not figure in the net result of the process of interpretation. But the interpreter must first consider the literal meaning of the sentence uttered, that is (“this is a nice derangement of epitaphs”), and find it odd in the context of a poetry class, for instance. But then the interpreter notices that another sentence, phonetically similar to the one uttered, has a literal meaning that is relevant in the context (“this is a nice arrangement of epithets”). So the interpreter is led to conclude that Ms Malaprop asserted the latter. Now even if the net result of the interpretation does not include what was said, it is hard to claim that what was said did not play an important role in the process of interpretation. As Wittgenstein would put it, capturing what Ms Malaprop wanted to say is moving from a form of expression to another form of expression (Wittgenstein 1953 § 334).

I acknowledge that Recanati’s arguments are compatible with admitting that there are locutionary acts. He simply argues that, very often, speaker and interpreter do not process the content expressed by the sentence, and that it is useless precisely for that reason. But there are problems with his examples, similar to the problems I just raised concerning Ms Malaprop. To give an obvious illustration, if someone utters “she took her key and unlocked the door,” who will deny that the minimal proposition expressed by the sentence was of some use? The information that she took her key and unlocked the door is surely contained in the information that she took her key and unlocked the door with her key! I would argue the same thing concerning many (all?) cases of loosening (“the cash machine swallowed my card”). The first time that we hear such a sentence we imagine a strange deglutition process. We then very rapidly get used to the secondary meaning, but the fact that we do get used to it does not turn it into a primary meaning. I would argue in a similar vein against Recanati concerning transfer, for when we say that the ham sandwich left without paying, we are first processing the proposition literally expressed by the sentence, in order to arrive at the conclusion that it is the client that ate the ham sandwich that left without paying. You first hear what has been asserted as very strange indeed, and then you get the point. According to Recanati, we never reach the stage where the minimal proposition is processed. We merely feel the incompatibility between the subject and predicate, and then shift the process of interpretation by taking into consideration primary pragmatic features. The first proposition that we grasp is the one determined by these pragmatic features, and not

the one literally expressed by the sentence. For the minimalist, on the contrary, we do reach the stage of processing the minimal proposition, for feeling the incompatibility between the subject and the predicate amounts to an attitude toward the sentence containing these subject and predicate. In order to arrive at a final interpretation, we must first consider the strange character of what is being said, and then consider “secondary pragmatic features.” The minimal propositions or truth conditions may not end up in the final interpretation, but they are occurring at one point in the process of interpretation.

One may think that the debate might be settled by considering “reaction experiments.” Perhaps these can show that the speaker who hears about the ham sandwich leaving the restaurant automatically computes that it is a person that left. Metaphors and idioms investigated in those reaction experiments may perhaps show that interpretations are made without considering literal meaning first. But if the experiment has been performed on someone who is accustomed to a secondary meaning, the fact that she no longer computes the primary meaning does not turn secondary meaning into primary meaning. Of course, one could reply that reaction experiment in a particularized situation allow us to draw conclusions concerning literal meaning, but this is problematic for reasons that were already stated. We should not draw any conclusion from such a result unless of course we stipulate at the outset that literal meaning is determined by pragmemes in a context, for this would be precisely asserting the conclusion of the argument, and not providing an argument.

11 Conclusion

In this paper, I have been looking for an argument in favor of semantic contextualism that would show how sentence meaning is determined by the intended meaning of speakers in contexts. As we now know, thanks to the contributions of those working in the fields of pragmatics, an intentional act of assertion may give rise to all sorts of things: pragmatic presuppositions, conversational implicatures, metaphors, irony, indirect speech acts and so on and so forth. There may be some minimalists like Stanley and Szabo (2000) who would deny that a wide variety of speech acts can be performed in intentional acts of assertion. But I tend to favor an enlightened version of minimalism that acknowledges speech act pluralism.

Nevertheless, while we may be asserting all sorts of things in the context of a particular utterance, we may at the same time be saying the very same thing if we use the same sentence and if the sentence contains no indexicals and is not ambiguous. Our locutionary act remains the same unless it contains indexicals and it is ambiguous. What we are saying is the minimal proposition (or minimal truth conditions). I have argued that conversational implicatures were a secondary level of (pragmatic) meaning and that they were cancelable. I have also suggested that we could harmlessly increase the number of context sensitive sentences beyond those that explicitly contain expressions belonging to the basic set, as long as we construe them as implicitly containing these indexicals. I then investigated the

possibility of primary, optional and intentional pragmatic features determining literal truth conditions. It was shown that if one accepts the distinction between locutionary acts and illocutionary acts, then the notion of accessibility appealed to by Recanati could be used against him. For if pragmatic features were really meant to be optional, then there is at least one context of utterance in which what is said is accessible and this is all that we need in order to defend the syncretic view.

Instead of referring to what Recanati calls the ‘syncretic’ view, I would use ‘bifurcationism’, because there are two levels of meaning: the minimal proposition expressed by the sentence (what is said by the speaker) and the additional intentional pragmatic meaning conveyed by the full illocutionary act (what is intentionally asserted by the speaker). Recanati does not deny in principle the existence of such a double level of meaning, but he argues that in the case of enrichment, loosening and transfer, the so called literal meaning of the sentence is not cognitively relevant for the speaker and normal interpreter. So the literal truth conditions are in these cases determined by speakers’ intentions and normal interpreters in the context of an intentional act of assertion. Recanati rejects the syncretic view (or bifurcationism) not because it is not possible to distinguish in principle different levels of meaning, but because there are apparently cases where no one computes (parses, cognitively entertains) the minimal proposition or the minimal truth conditions. So even if we admit the existence of minimal propositions or minimal truth conditions, they should according to Recanati be ignored in many cases and replaced by propositions or truth conditions that are determined by what the speaker meant in the context, which also happens to coincide with what the normal interpreter understands. Therefore, since pragmatic features like enrichment, loosening and transfer determine primary propositions or truth conditions, they are primary pragmatic features, and this refutes bifurcationism.

It is possible for Recanati to accept the distinction between saying and asserting and still be arguing for his idea that the so called minimal proposition or minimal truth condition is irrelevant. Recanati could perhaps claim that in certain contexts of utterances, the speaker is not intentionally saying the thing that is literally expressed by the sentence. Simultaneously, the normal interpreter is not processing the minimal content of the sentence in order to achieve his interpretation. So this is why the literal, minimal proposition (or truth conditions) is irrelevant. And this is why Recanati believes that the syncretic view is false.

I have argued that this argument presupposed methodological assumptions that turn contextualism into a circular argument. I also suggested that this circularity was needed in order to avoid incoherence. I then challenged Recanati on each of the examples he has provided in his attempt to refute bifurcationism. I would want to say that in all the examples discussed, the minimal proposition is in different ways relevant for the normal interpreter and even for the speaker. It seems that I am cognitively entertaining the minimal proposition when I am intentionally saying a sentence that expresses such a minimal proposition. But even if there were good examples where it appears that neither the speaker nor the hearer really considers the minimal proposition in their language processing, it would not affect the general criticism of circularity.

I suggest instead a general twofold approach in which bifurcationism and speech act pluralism form a sophisticated and enlightened version of semantic minimalism.

References

- Bach, Kent. 2005. Context ex Machina. In *Semantic versus pragmatics*, ed. Zoltan Szabo, 15–44. Oxford: Oxford University Press.
- Baker, Gordon, and Peter Hacker. 1984. *Language, sense and nonsense*. Oxford: Basil Blackwell.
- Bennett, Jonathan. 1976. *Linguistic behaviour*. Cambridge: Hackett.
- Borg, Emma. 2004. *Minimal semantics*. Oxford: Oxford University Press.
- Burton-Roberts, Noel. 2006. Cancellation and intention. *Newcastle Working Papers in Linguistics* 12–13: 1–12.
- Capone, Alessandro. 2006. On Grice's circle (a theory-internal problem in linguistic theories of the Gricean type). *Journal of Pragmatics* 38: 645–669.
- Capone, Alessandro. 2009. Are explicatures cancelable? *Journal of Intercultural Pragmatics* 6: 55–83.
- Capone, Alessandro. 2008. Book review of *Insensitive semantics. A defence of semantic minimalism and speech act pluralism*. by Herman Cappelen, Ernest Lepore. Oxford: Basil Blackwell, Oxford, 2005. *Journal of Pragmatics* 40/9: 1626–1635.
- Carston, Robyn. 2002. *Thoughts and utterances: The pragmatics of verbal communication*. Oxford: Blackwell.
- Davidson, Donald. 1989. The Myth of the subjective. In *Relativism: Interpretation and confrontation*, ed. M. Krausz, 159–172. Notre Dame: University of Notre Dame Press.
- Davidson, Donald. 1999. The emergence of thought. *Erkenntnis* 51: 7–17.
- Davidson, Donald. 1984. Thought and Talk. In *Inquiries into truth and interpretation*, 155–170. Oxford: Oxford University Press.
- Dummett, Michael. 1993. What is a theory of meaning? (I). *The seas of language*, 1–33. Oxford: Oxford University Press.
- Grice, Paul H. 1975. Logic and conversation. In *Syntax and semantics*, vol 3., ed. P. Cole, J. Morgan, 41–58. New York: Academic. Reprinted in *Studies in the Way of Words*, 22–40. Cambridge: Harvard University Press. 1989.
- Kaplan, David. 1989. Demonstratives. In *Themes from Kaplan*, ed. J. Almog, 481–563. Oxford: Oxford University Press.
- Lepore, Ernest, and Hermann Cappelen. 2005. *Insensitive semantics*. Oxford: Basil Blackwell.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: MIT Press.
- Lewis, David. 1969. *Convention: A philosophical study*. Cambridge: Harvard University Press.
- Montminy, Martin. 2006. Semantic content, truth-conditions and context. *Linguistics and Philosophy* 29: 1–26.
- Predelli, Stefano. 2005. *Contexts: meaning, truth, and the use of language*. Oxford: Oxford University Press.
- Recanati, François. 2004. *Literal meaning*. Cambridge: Cambridge University Press.
- Salmon, Nathan. 2005. Two conceptions of semantics. In *Semantic versus pragmatics*, ed. Zoltan Szabo, 317–328. Oxford: Oxford University Press.
- Schiffer, Stephen. 1972. *Meaning*. Oxford: Oxford University Press.
- Searle, John. 1983. *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Seymour, Michel. 1992. A sentential theory of propositional attitudes. *Journal of Philosophy* LXXXIX: 181–201.

- Seymour, Michel. 1999. Two concepts of beliefs. In *Consciousness and intentionality: Models modalities of attribution*, ed. Denis Fisette, 311–344. Boston: Kluwer.
- Seymour, Michel. 2005. *L'institution du langage*. Montréal: Presses de l'Université de Montréal.
- Seymour, Michel. 2006. Le contextualisme sémantique en perspective. Au sujet de *Literal Meaning*, de François Recanati. *Philosophiques* 33(1): 249–262.
- Sperber, Dan, and Deirdre Wilson. 1995. *Relevance*. Oxford: Basil Blackwell.
- Stainton, Robert. 2006. *Words and thoughts: Subsentences, ellipsis, and the philosophy of language*. Oxford: Oxford University Press.
- Stanley, Jason. 2005. Semantics in context. In *Contextualism in philosophy: Knowledge, meaning, and truth*, ed. Gerhard Preyer, and Peter Georg, 221–254. Oxford: Oxford University Press.
- Stanley, Jason, and Zoltan Gendler Szabo. 2000. On quantifier domain restriction. *Mind and Language* 15: 219–261.
- Travis, Charles. 2001. *Unshadowed thought: Representation in thought and language*. Cambridge: Harvard University Press.
- Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford: Basil Blackwell.

Language Adds to Context

Paolo Leonardi

Abstract Contextualism deems that the meaning of most utterances, if not any, varies from context to context. Here, I argue, instead, that utterances added to a context change the context by shifting salience. Considering truth conditions together with appropriateness, I try to show that utterances claimed to make the same assertion—such as ‘Rain is over’, ‘Here, rain is over’, ‘In Florence, rain is over’ and ‘Here in Florence, rain is over’—differently affect salience, and hence in an important sense do not assert the same. The idea is rather that ‘Rain is over’, for instance, says the same in any context, but that it differently shifts salience in different contexts, and the various assertions listed may get at the same salience only in different contexts.

1 The Model of Context Sensitive Expressions

The model of a context-sensitive expression is indexicals and demonstratives. The meaning of ‘I’ and that of ‘this’ are a function of an utterance role. When an individual, an object, an aspect, a moment, a place, referred to by an expression can be individuated quoting the very utterance of that expression I would say that the individual, the object, etc, play an utterance role. ‘I’ refers the agent of the utterance, ‘this’ to the thing demonstrated by the utterance, ‘now’ to the moment of the utterance, ‘here’ to the location of the utterance, etc. This seems true of no expression beyond indexicals and demonstratives.¹ ‘Tall’, ‘ready’, and indefinitely

¹ Kaplan lists the following indexicals and demonstratives: «...the pronouns ‘I’, ‘my’, ‘you’, ‘he’, ‘his’, ‘she’, ‘it’, the demonstrative pronouns ‘that’, ‘this’, the adverbs ‘here’, ‘now’, ‘tomorrow’, ‘yesterday’, the adjectives ‘actual’, ‘present’, and others.» (1989: 489).

P. Leonardi (✉)
Università di Bologna, Bologna, Italy
e-mail: paolo.leonardi@unibo.it

many other words are said to be context-sensitive. Speaking of her child, Adelaide says «Marco is tall». Understanding Adelaide's *tall* as told of children nine-years old doesn't seem dependent on the utterance role of 'tall' (and its expansion, 'tall for a child nine-years old', doesn't make explicit an utterance role of 'tall'). Does 'tall' point anyway to Marco's age? As it would point to President De Gaulle's age, if Adelaide had instead said «Charles De Gaulle was tall». If we were asked, we would probably agree that Marco is a 'tall child', as we would agree that President De Gaulle was a 'tall man'. If this is a kind of a context-sensitivity, which kind is it? Alternatively, we might think that 'tall' means 'of a height greater than average of the same sex and age cohort,'² something which we fully understand even knowing neither the sex nor the age of the person involved, and, from that and from the evident matter of fact that Marco is a child and President De Gaulle was a man, reason to 'tall child' and 'tall man', respectively.³ Analogous remarks apply to any other most famous contextualist example. 'Ready', for instance, attributes the property of being in a suitable state for an action or an activity, or for facing a situation.⁴ The context possibly supplies the action, the activity or the situation one is ready for, say, 'ready for dinner'. But neither 'for dinner' nor 'for the test' express an utterance role. How is 'ready' differently context-sensitive?⁵ Or is it that again we reason from the meaning of the word and from the evidence the circumstance affords to yield what one is ready for?⁶

² There are other relevant parameters, over which I skip for the sake of simplicity.

³ We could also imagine to expand «Marco is tall» to «Marco is tall for his age», in which no real contextual parameter surfaces.

Fara remarks that not all contextual dependency can be reduced to a comparison class, and offers as an argument the case in which by a dreadful coincidence all tall children die, a case which could not yet induce us to say tall the tallest of the surviving children. I think that this depends on our memory surviving the tragic accident. (2000: 54-9).

⁴ Indexical and demonstratives are token reflexives. 'Who utters *I*' can't therefore be substituted for *I*, because then *I* would have been quoted but not really uttered, whereas Adelaide can say 'Marco is a tall child' instead of 'Marco is tall', even if, as we will see, the two expressions aren't synonymous.

⁵ An argument to this effect can be found, for instance in (Travis 1985 (2008) and in Bezuidenhout (2002). Most authors are well aware of the difference between context-sensitivities.

⁶ There are two other problems, which I believe convergent with the contextualist's.

- (a) A sentence might be true in indefinitely many different situations. 'The President gives a party for Independence Day' might be true of any U.S. President of any year of Presidency, and if other countries have their Independence Day, it might be true of any President of theirs and of any year of their Presidency too. Again, the alternative possibilities do not depend on any utterance role of 'President' or of 'Independence Day'. Indeed, the possibilities point to different circumstances of evaluation, among which the sentence doesn't discriminate—which it could do if there occurred an indexical. The sentence 'The President gives a party for Independence Day' in almost any of its uses speaks of a particular President and of a particular Independence Day. Of course, 'President Barack Obama gives a party for Independence Day 2011' seems to tell which is the relevant situation. But we could imagine another year counting with another 2011 year and another President whose name be 'Barack Obama', not impossible if you remind yourselves that the U.S. have already had two

I'll first investigate the supposedly different context sensitivity, and later suggest a solution that doesn't make 'tall', 'ready' and any other expression context sensitive, but yet a solution which is context sensitive: rather than looking at context as something which makes precise what we say, I'll take things the other way around and look at language as something which is added to context for modifying it in a particular way. We speak in a situation cognitively structured, and what we say further structure it. In a circumstance in which we are discussing Anne's next exam, it is appropriate saying 'Anne is ready' for claiming she is ready for the exam; in a circumstance in which we discuss whether inviting Anne to go to the movie, it is appropriate saying 'Anne is ready for the exam' for suggesting that she can be invited. What is said is placed, and if it does not fit it hits the wrong place.⁷

2 Context Sensitive Expressions Nonetheless?

Here are some possible ways we take non-indexical and non-demonstrative expressions to be context-sensitive nonetheless.

- a. Perhaps, Adelaide's utterance—«Marco is tall»—contains a covert element referring to Marco's age. The constituent is unpronounced, and its general form is, say, *for an X-years old*. Any comparative or grading term compares and grades in relation to a comparison class or a scale, or not just comparative and grading terms but possibly any term. «Please, buy some lemons, four red apples and a red watermelon!» The red apples have red skin, the red watermelon has

(Footnote 6 continued)

President John Adams and two President George Bush! Besides, there are indefinitely many aspects under which the sentence 'President Barack Obama gives a party for Independence Day 2011' allows for alternative possibilities. The time and the place the party is at, the people invited—relatives, relatives and friends, officials, US officials, foreign ambassadors, etc. Turn the sentence to the past, 'President Barack Obama gave a party for Independence Day 2011'. All the data could presumably be filled in, but any filling will leave us a choice among indefinitely many alternative possibilities.

- (b) Speaking of truth, John L. Austin, claimed that the truth of a statement depends on Descriptive and on Demonstrative conventions. The *Demonstrative* convention correlates the words with the *historic* situations in the world. (Austin 1950 (1979): 121) But there are not demonstratives in the two sentences considered a moment ago—'The President gives a party for Independence Day' and 'President Barack Obama gives a party for Independence Day 2011'. The two sentences aren't equivalent respectively with, say, 'This President gives a party for our Independence Day' and 'This President Barack Obama gives a party for our Independence Day 2011'. In a historical text the author might not substitute the one for the other, if he were writing of a President different from the actual one, and if the author were not a U.S. citizen, as I am not. If there are no expressions whose meaning depends on their utterance role, a sentence isn't demonstrative.

⁷ A misplacement is indicated by the offer or the request of a repair—especially of a reformulation, adding or taking something back.

red pulp, and perhaps there is a covert constituent *with a red Y*. «Tom's car is very expensive» may speak of the car Tom owns, of the one he sells, of the car he judges the best one, of the car he spoke with us yesterday at dinner, etc.⁸ Whatever be the unpronounced constituent, it may stand for many very different relations between Tom and the car, the car Tom Xies.⁹ For each expression—*ready*, *red*, the genitive forms, etc—the range of possible covert constituents is indefinitely large,¹⁰ of any level of complexity, and selecting among them seems to depend on side information—about human being's height and age, about fruit and colors, about Tom and cars, etc.

- b. Rather than a covert constituent, the case at hand can be accounted via a free enrichment, an optional expansion adding one or more constituents that were articulated at no linguistic level.¹¹ The enrichment aims at making explicit the proposition actually communicated, which is different (and possibly richer) than the one literally expressed.¹² (Here 'communicated' doesn't imply success. The proposition communicated, as the one meant, may not be recognized.) What's the proposition actually communicated is intuitive, and unexplained. Yet, in this perspective there being no rule for the expansion is no problem, because we can conjecture that the enrichment depends on no language structure but on knowledge databases, and largely on that part of knowledge databases people believe shared with their co-conversationalists. Whatever element in the database is activated by what is literally expressed may originate an unarticulated constituent.

⁸ A champion of the covert unpronounced constituent idea is Stanley (2000 and 2007). Writes Stanley:

for each alleged example of an unarticulated constituent [mentioned in the literature], there is an unpronounced pronominal element in the logical form of the sentence uttered, whose value is the alleged unarticulated constituent. (2000: 410).

⁹ Stanley speaks of indexicality in the broad sense (2007: 38) and before him Strawson claimed there to be an indexical element. The idea looks more fitting for incomplete descriptions, which are the case Strawson originally had in mind, though I think there are better ideas even for that case.

¹⁰ The covert constituent is a variable, or better a function variable which applies to an object variable «The value of 'i' is an object provided by the context, and the value of 'f' is a function provided by the context that maps objects onto quantifier domains.» (Stanley 2007: 115).

¹¹ See Recanati (2004: 18): «Insofar as it is pragmatically rather than linguistically controlled, free enrichment is taken to be irrelevant to 'what is said', on the non-pragmatic construal of what is said.»

¹² Writes Recanati:

On this view the enrichment process through which, in context, we reach the proposition actually communicated [...] is not linguistically but pragmatically required; hence it is not an instance of saturation, but an optional process of 'free enrichment'. (2004: 10).

- c. The unarticulated constituent view is more sophisticated than the unpronounced constituent one. The «process operates in the (sub-personal) determination of what is said»¹³ and its explicit expansion is optional rather than mandatory. Although these are the standard options, notice that:
- c1. Without linguisticizing the context a priori (Stanley) or a posteriori (Recanati) one could suggest that Adelaide's utterance "compose" with one element of its context, as if Marco's evident childhood filled an argument position in 'tall'. The suggestion is not that of a thing matching a linguistic constituent with no phonetic realization, but a non-linguistic element playing directly a linguistic role, as if one had literally *said* it.¹⁴
 - c2. There is an important precedent of propositions and sentences including the semantic values of expressions, namely singular propositions and valuated sentences.¹⁵ Yet, this option is rarely chosen.

Some comments and reactions to *a* and *b*, above.

The issue is whether what is said individuates what object, quality, event, state of affairs etc it talks about. Contextualists are ambiguous on the matter. On the one side, giving a role to context seems to deny that what is said by itself individuates what's about. On the other side, assuming covert constituents, sometimes with a kind of indexicality, or conjecturing unarticulated optional constituents, contextualists too try to have what is said determine what it talks about, the unarticulated constituents alternative expanding what is literally expressed to an abstract linguistic representation summoned by contextual elements—any relevant piece of contextual information. The abstract representation is the proposition communicated.¹⁶ Yet, I would resist both the notion of a proposition made fully explicit and that of 'the proposition actually communicated', and the idea that the cognition communication affords have a full propositional format.¹⁷ What is interesting in communication is that there is a focus of explicit and a surrounding area of implicit understanding which glides into areas of non-understanding, even mutually

¹³ Recanati (2004: 17).

¹⁴ The idea of a statement composed of strictly linguistic elements and non-linguistic ones is analogous to the text that mix words and images, these playing a linguistic role without being given a linguistic profile.

¹⁵ Kaplan (1975) introduced singular propositions looking at Russell (1903). Writes Russell:

But a proposition, unless it happens to be linguistic, does not itself contain words: it contains the entities indicated by words. (47)

Kaplan introduced valuated sentences in (1986): 245f.

¹⁶ See Recanati (2004: 56).

¹⁷ The unarticulated constituent view was originally championed by Perry (1986 and 1998) and later, as we are seeing, by Recanati (2002, 2004, 2010). The proclivity to linguisticize the move was already there in Wilfrid Sellars' talk of proposition (1954).

contrasting understandings, of different information and beliefs hardly compatible with what the communication let us grasp.

The propositional claim is favored by two related aspects. Speakers can verbalize the side information they use, making it explicit, and making this explicit is exactly what theoreticians do. But people may grasp a point without explicitly representing it to themselves. Ernesto saw Ann's crossing the lane both if he saw her crossing the lane and if he saw that she crossed the lane. Only in the second case, however, Ernesto is attributed explicit note of the fact.

Moreover, not only there are indefinitely many different elements that covert or unarticulated constituents may make explicit, but there are indefinitely many aspects of the context that may be relevant, and which the contextualist view would have to deal with. This is a second reason to doubt that both unpronounced constituents and optional expansions of the proposition expressed yield a full representation of the situation at hand. «Ann is ready for the test» may be wanting—is she ready for the proficiency test, or for the medical test? Ready for the foreign language proficiency test or for the English literature one? For the foreign language proficiency test first grade or the second grade one, etc. If «Ann is ready» expresses an incomplete proposition, «Ann is ready for the test» expresses another incomplete proposition, «Ann is ready for the proficiency test» a third incomplete proposition, and so on. What is expressed, it's not fixed by some longer sentence that the speaker, the hearer, the theorist have in mind. At the same time, people are acknowledged cognition of the context, and hence in some way there is no need of any expansion, there is no need to decide what proposition exactly is communicated, or whether exactly a proposition is communicated. Besides, if indefinitely many elements of the context are cognized, there is no indeterminacy problem either, because those elements are already individuated.¹⁸

Indeed, when we refer, predicate or state something (or perform any other speech act) we fix our attention on what we are talking about, and our wording calls attention to (rather than determining) the same objects, qualities, events or states of affairs.¹⁹ What we say is added to the situation that, for us and for our audience, is already structured, even if not identically so. Taking into account different perceptual perspectives, if we are in the same place, or different perspectives, if we are not, and different (perceptual and linguistic) memories, different fantasies, we always assume something in common, and what we say is to

¹⁸ Words determining what they speak of is the most classical problem in philosophy of language since before Socrates, and is connected with that of how word and thing connect. The problem is different from another, intriguing one, which is that words telling what they state or stand for, stating their own truth, look paradoxical. Austin 1950 (1979) hints at either issues, to the first when he speaks of demonstrative conventions. (See, respectively, p. 122 and p. 126, fn. 2.)

¹⁹ Donnellan makes the point for reference of definite descriptions [see his (1966) and (1977)]. The remark can be extended to any form of reference—that of predicative names as well as that of sentences describing an event or a state of affairs.

tune our and others' attention on the objects and the aspects explicitly mentioned on that slowly fading background.

Had Adelaide said that Marco is tall *for a nine-years old child*, the addition would have been misleading—why stressing what's evident? The child is in front of us! If it were not evident, if Marco had not been in front of us, if I had not known who Marco is, or that he is a child, Adelaide's *Marco is tall* would instead have been somehow incomplete, as a fragment we can overhear listening to two persons chatting in the subway, or as a remnant of an ancient inscription some parts of which have been lost. In the envisaged circumstances the envisaged sentences would have been inappropriate but switching the two sentences in the two situations would have yielded appropriate contributions.

It is not a question, I think, of articulating the *propositional* representation of the situation, not even resorting to singular propositions or the like. What the speaker says is *placed* and *changes* the context—what is explicitly mentioned becomes evident or more so, and what isn't mentioned becomes less or not evident. For instance, uttering «Ann is ready» changes the context, from waiting for Ann's activity or action or for her facing a situation, to Ann performing the ones or facing the other.²⁰ The change is a change in the state of affairs—words come into the scene—(aiming at) inducing an attentive tuning on some specific elements²¹

I am denying that people involved in talk exchanges have a propositional knowledge but of a part of what is going on. Of course, I am not denying that a theory of all this, and the theorist would articulate a propositional theory for semantics and pragmatics. (Although, then too, at a different moment, non-propositional cognition has to be involved to ground and integrate, so to speak, the propositional representation of how people succeed in being appropriate.)

3 Language and Context-Shift

Apart from my understanding of the change, the idea isn't original. In the last forty years, context has been taken into account the more and more, acknowledging mutual interaction between text and context, and in some views developing a theory of meaning as a *context change potential*.²² In my understanding, the addition is a propositional component, and it integrates with the cognitive

²⁰ If the utterance of 'Ann is ready' were redundant, it would still change the context, though in a different way. For instance, it might suggest that someone else is not ready.

²¹ Notice that words too are elements of the state of affairs.

²² After Heim (1983). The change can be smooth or abrupt, see Lewis (1979)—by abrupt I mean what Lewis describes as follows: «If what I say requires that, then straightway it is so. By saying what I did, I have made Bruce more salient than Albert.» (1979 [1983]: 242).

representation that is propositional only is so far as it involves other linguistic or symbolic items.²³ The background stays non-propositional.

The integration of linguistic and not linguistic cognition is much wider than the case at stake. Besides the fact that in speaking and in any event of making explicit an element of the context we have an independent grasp of what we refer to or predicate of, there are general aspects such as that (a) we understand what we are told or what we read *by means of perception and memory*. (b) An utterance is a chain of expressions and «a word is an expression along with its meaning»,²⁴ hence the utterance doesn't *say* what *words* are uttered (as it does not say what language it speaks).²⁵ (c) our understanding the meaning of most words depends on our knowledge of things—in this way we catch the difference, for instance, between «Tom opened the door» and «The carpenters opened the wall».²⁶

Whether this or that datum is entertained in a linguistic form is a tricky question, because it presents the data in that very form. That language can describe almost any element of its context of use doesn't make anything linguistic the context, and a propositional representation, which is abstract and modeled on language, brings no improvement.

Graphs and pictures contain finitely many data, extrapolating them in linguistic form requires much space and time and it is convenient only when we want to have the datum explicit. Here is one example (from Tufte 2001: 42). It shows with immediate evidence Los Angeles area pollution through the different counties and the different time of the day. The data are offered visually and not linguistically, coding about 28.000 pollutant readings. If each datum put linguistically required half a line, we would have a text that we wouldn't get at a glance. My suggestion is that we get part of the contextual data the same way, at least in as much as perception or perceptual memory is involved.

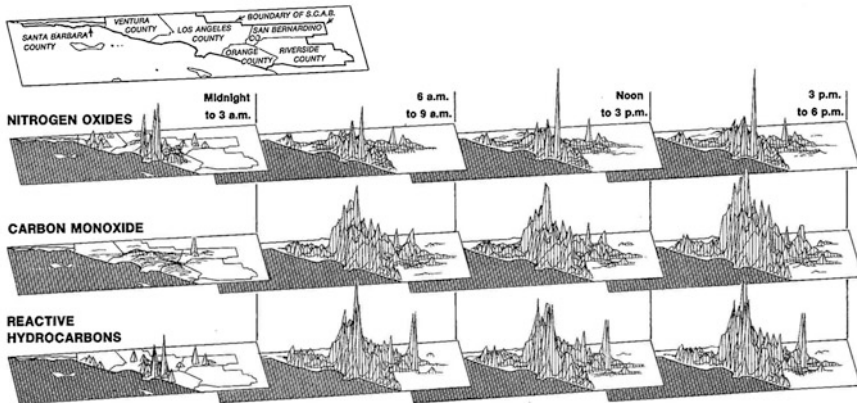
²³ Burge (2010) suggests that our knowledge is integrated at the propositional level, and that this is at the origin of objectivity. Although I think that the matching problem between our representation and how things are, whatever that means, is at the origin of objectivity, I wouldn't have it that that requires a propositional representation, i.e. a representation run on abstract elements.

To limit my citations to another few, already Stalnaker (1974), (1978) and (1998), Ballmer (1978) and (1981), Lewis (1979), Kamp (1981), Kamp and Reyle (1993), Heim (1982), (1983) and (1992), Sbisà (2002) have suggested something of what I have in mind, which is not that linguistic meaning is context dependent but the other way around. All of them, anyway, share a propositional understanding of the phenomenon perhaps apart from Sbisà.

²⁴ See Kaplan (1989: 559).

²⁵ *Ibid.* This, whether or not an utterance says or seems to say what words in what language it utters.

²⁶ Searle (1983: 145).



Los Angeles area pollution July 22, 1979

4 No Modification Without Aberration

If there were an unpronounced or an unarticulated constituent, wouldn't it be possible to articulate and pronounce it? Would what Alan says—«Rain is over»—possibly be expanded to state in full what he means?

Alan and me are having a coffee in Marciana—does he mean that *here* it was raining some moment ago and doesn't rain anymore? If we were talking about Dalia, who is in Barcelona, and about the bad weather (over there), and if Alan was just having a phone call with Dalia or with someone else in Barcelona, he may mean that rain is over *in Barcelona*. Of course, anyway I understand him, I may believe to have taken him right and be happy, or just not care and be happy the same. If not, I may inquire and come to know more about where rain is over—I didn't notice it was raining here in Marciana, is it over *here*, in Marciana or in Barcelona? Etc. (Even the no rain interval could pose problems: *for a while*, *for some days*, *for months*? Etc. Yet, I'll deal only with the location problem.)

Now, the difficulty is that no one of these forms is equivalent to the first one, as indicates the fact that the first one can be *repaired* by one of those non-equivalent ones (and precisely because they aren't equivalent), and that using any of these does *not* say what the first one says. If Alan had said «Rain is over in Marciana», he *would* have possibly been contrasting Marciana with Poggio, or with the place the person at the other end of the phone call is at, or with a third place they or we were talking about, or with a third place Alan would have introduced straightaway. The same would be true had Alan said «Rain is over in Barcelona» or the like. But even «Here rain is over» would have contrasted *here* with some other place. Hence, saying «Rain is over» isn't equivalent to any other saying. If there were a

covert or unarticulated constituent, articulating them wouldn't have to make the difference it does.²⁷

Is this a peculiar quality of 'Rain is over'? I mean, do other classical examples discussed in the literature behave differently? Of course, they do—one thing is the location aspect, which is pervasive, and another, say, the ongoing activity aspect.²⁸ «Ian is ready» Michelle says. Ready for what—for going to school, for dinner, for getting married, etc? Again, the phrase isn't equivalent to any form that can repair it. Had Michelle said instead «Now, Ian is ready» or «In half an hour Ian is ready», Michelle would have contrasted 'now' with some moments in the (near) past or in the (near) future, or a measure of waiting, half an hour, with a shorter or a longer one. Had she said that *Ian is ready for the exam*, his exam would have to be relevant, or made such immediately after, or contrasted with Ian unreadiness for, say, going to Milan for the job interview.

The case looks even more complicated. If Michelle says that Ian is ready, has she to mean, at least implicitly, what he is ready for? I don't believe she has to. Ian's readiness may be all that's relevant.²⁹ As soon as Ian is ready we, Michelle and myself, say, may have dinner, go to the movies, etc. The point may be Ian's readiness—it doesn't matter, it is irrelevant, what he is ready for. This suggests that 'Ian is ready' may not need any completion. And, therefore, that 'Ian is ready' may not have any covert or unarticulated constituent, were there³⁰ or not an absolute value for readiness.

A third argument can be devised for 'Marco is tall'. I'll not detail it, not even minimally as I did in the two cases above. I'll content myself with a different maneuver—showing that there is an absolute understanding of 'tall'. Everyone is tall if his height is above average than that of the persons of the same sex, same age cohort, perhaps same ethnic group, and perhaps also same historical time (because nutrition and health varies from one century to another in the same country). Hence, Marco may be tall at 9 years of age, and average at 18. Why is it relevant to tell the age? Because you may not know the age, and were added an absolute value of height, you may wonder how to reconcile the data. If there is an absolute value to height, of course, telling someone tall doesn't require adding anything. Of course, out of context, may be relevant to tell that Marco is a tall

²⁷ Adding anything to «Rain is over» may reverberate on the speaker. If it is misplaced, the conditions being different from any case in which an addition is proper, the misplacement would reverberate on the speaker, who would be deemed mouthful or boring, or incompetent.

²⁸ A good symptom of the difference is that in most, if not all, languages there is no indexical for the on going activity.

²⁹ My claim may remind you of analogous ones by Cappelen and Lepore (2005). My point is not specific to supposedly contextual expressions. As Cappelen and Lepore, I don't see any difference between 'Ian is ready' and 'Ian went to the movie'. If Ian went to the movie, he went to movie X or he went to movie Y, or ... But I believe that a report of what has been said in different circumstances uttering 'Ian is ready' may require saying rather different things.

³⁰ As I suspect.

nine-years old child. We always take context into account, in remarking on a situation as in reporting one such remark.

You might agree: an expansion isn't required, and it is even inappropriate. But don't have all the *rain is over* examples the same truth conditions, once indexicals and demonstratives are discharged? Aren't «Rain is over», «Rain is over in Marciana», «Here, rain is over», «Here in Marciana, rain is over»? No, they aren't, appropriateness conditions aside. The first example doesn't hint at any location, the second utterance mentions one, the third speaks of the place where the speaker is (then, instances of the second and instances of the third examples occasionally speak of the same location), the fourth utterance names a location and says the speaker is there. Now, if, where the speaker is, rain is over but he is not in Marciana (perhaps, wrongly assuming he is), and in Marciana is still raining, the second and the third examples have different truth-values, and the second and the fourth examples too have different truth-values. If the speaker lies about the weather where he is, and in Marciana rain is over, (where we, perhaps wrongly, assuming he is), the second and the fourth examples, as well as the third and the fourth ones, have different truth-values. Besides, if rain is over everywhere, it is over also here and in Marciana, or here in Marciana. But rain might be over in indefinitely many a place without being over everywhere. Hence, the truth-values of the first example and those of the other three ones may well diverge too. Besides, if we turn the appropriateness conditions into truth-conditions—as conditions speakers imply to be true—, they diverge even more because the assumptions are different and their truth values may change from context to context.

5 Closings

Summing up, an unpronounced or unarticulated indexical constituent isn't what links what is said to the context of utterance for at least five reasons. (a) As the expansions considered by (Recanati 2002) show, there are many alternatives among which we haven't any linguistic cue for choosing the one over the others. (b) If we tried to make explicit the covert or the unarticulated constituent, no explicit form would be fine-grainedly equivalent to the supposedly covert or unarticulated one, because each says something different from the supposedly covert or unarticulated form. (c) Having individuated an element of the context, expanding the text adding a phrase to make explicit that the text refers to that context is redundant, showing already determined what is spoken about. (d) The utterance isn't there to mirror its context, but to change it, and (e) anyway what is said can never fully mirror its context of utterance. Now, I will dwell only on aspects of (d).

«This is red!» say I. Red? A red apple is a ripe one with red skin and a red watermelon is a ripe one with red pulp, a red tomato has red skin and pulp, the red inkpen has either a red “skin” or writes in red ink, the communists have a red flag, human beings feeling shame have a red face, red blood cells are red looked at sideways on. There are indefinitely many ways to be red. An unpronounced

constituent could specify how and when a red thing is red only very generically (very unspecifically). An unarticulated constituent would have to be expanded as much generically and unspecifically.³¹

If I say «This is red!», I have in mind (I have my attention fixed on) what I claim to be red, and choose the utterance appropriate for calling my audience to what I classify as red, and making them understand under what aspect it is red.³² This can be put in general terms. What is said is placed, and it is produced and understood from where it's placed.³³ If it isn't understood, it is somehow misplaced—the speaker assumed a background cognitive state that her audience didn't have. The claim needs further work to be properly placed, and a repair is produced or requested and produced.³⁴ If there is no repair, the audience has a problem solving strategy—what the speaker can, here and now, call my attention to saying *this*? If the speaker misuses some words, that's often accountable for. If she is wrong, that is, her mistake has some motives, as it has her forcing the language, speaking ironically, metaphorically, or using a form which could be disputed. (People beliefs, even when wrong, are not unmotivated.) When, in introducing this paper I claimed my solution to be context sensitive, I meant that *placing* an utterance is a context sensitive action, though not context sensitive in the sense in which an indexical or a demonstrative, or tense are.

The context includes perceptions, perceptual and linguistic memories, fancies, etc. What's linguistically stored has a propositional structure, what not most likely doesn't. *What is said aims at tuning speaker's and audience's exactly on the elements mentioned.*³⁵ The standard strategy is identical to the repair one, but under two aspects: How can I call their attention to this? Here and now, what can she call our attention to saying this?³⁶

³¹ One way to conceive of such an expansion is Recanati (2004: 27 and 107-9). Writes Recanati:

The effect of an 'expansive' variadic function of the sort contributed by adverbial modifiers is to add an argument-role. The output relation therefore contains *the same argument-roles* as the input relation, plus the extra argument-role provided by the variadic function. (2004: 107).

³² This extends to predication, assertion etc., what Donnellan (1966) and (1968) argues for reference of definite descriptions.

³³ See Bach (2001: 29): «It is one thing for content to be determined *in* context and quite another for it to be determined *by* context.»

³⁴ On the notion of repair see Schegloff, Jefferson and Sacks (1977).

³⁵ The most careful of my readers asked me at this point: «Isn't it more than just focusing attention? Isn't it often aimed at getting people to believe specific things?» Yes, it is, because taking notice recalls memories, allows to reason about, etc. As my reader would have had it in other cases, here I go minimalist.

³⁶ The placement of what is said allows to determine what language and what words (*à la* Kaplan, what things the expressions used are associated with) are used. Meanings of expressions are generated and evolve *historically*, and the history of an expression is in the many heres and nows of its uses.

References

- Austin, John L. 1950 (1979). Truth. *Proceedings of the Aristotelian Society. Supplementary* 24: 111–128.
- Austin, John L. 1979. *Philosophical Papers*. 3rd edn. Oxford: Oxford UP.
- Bach, Kent. 2001. You don't say? *Synthese* 128: 15–44.
- Ballmer, Thomas. 1978. *Logical grammar*. Amsterdam: North Holland.
- Ballmer, Thomas. 1981. Context change and its consequences for a theory of natural language. In *Possibilities and limitations of pragmatics*, eds. H. Parret, M. Sbisà, and J. Verschuereen, 17–55. Amsterdam: John Benjamins.
- Bezuidenhout, Anne. 2002. Truth-conditional pragmatics. *Philosophical Perspectives* 16: 105–134.
- Burge, Tyler. 2010. *The origins of objectivity*. Oxford: Oxford UP.
- Cappelen, Herman, and Ernie Lepore. 2005. *Insensitive semantics*. Oxford: Blackwell.
- Donnellan, Keith. 1966. Reference and definite descriptions. *The Philosophical Review* 75: 281–304.
- Donnellan, Keith. 1968. Putting humpty dumpty together again. *The Philosophical Review* 77: 203–215.
- Donnellan, Keith. 1977. The Contingent *apriori* and rigid designators. *Midwest Studies in Philosophy* 2: 12–27.
- Fara Delia, G. 2000. Shifting sands: an interest-relative theory of vagueness. *Philosophical Topics* 28: 45–81.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. PhD. Dissertation, Amherst: University of Massachusetts.
- Heim, Irene. 1983. File change semantics and the familiarity theory of definiteness. In *Meaning, Use and the Interpretation of Language*, eds. R. Bauerle, C. Schwarze and A. von Stechow, 164–190. Berlin: Walter de Gruyter.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9: 183–222.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In *Truth, representation and information*, eds. J.A.G. Groenendijk, T.M.V. Janssen and M.B.J. Stokhof, 277–322. Mathematische Centrum University of Amsterdam.
- Kamp, Hans, and UweReyle. 1993. *From discourse to logic*. Berlin: Springer.
- Kaplan, David. 1975. How to Russell a Frege-Church. *The Journal of Philosophy* 72: 716–729.
- Kaplan, David. 1986. Opacity. In *The philosophy of W.V. Quine*, eds. E. Hahn and P.A. Schilpp, 229–289. La Salle, IL: Open Court.
- Kaplan, David. 1989. Demonstratives. In *Themes from David Kaplan*, eds. J. Almog, J. Perry, and H. Wettstein, 481–563. Oxford: Oxford UP.
- Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8: 339–359. Reprinted in Lewis 1983.
- Lewis, David. 1983. *Philosophical papers*, vol. I. Oxford: Oxford UP.
- Perry, John. 1986. Thought without representation. *Supplementary Proceedings of the Aristotelian Society* 60: 137–152.
- Perry, John. 1998. Indexicals, contexts and unarticulated constituents. In *Computing natural language*, eds. A. Aliseda-Llera, R. van Gabeek, and D. Westerståhl: 1–11. Stanford CSLI.
- Recanati, François. 2002. Unarticulated constituents. *Linguistics and Philosophy* 25: 299–345.
- Recanati, François. 2004. *Literal meaning*. Cambridge: Cambridge UP.
- Russell, Bertrand. 1903. *The principles of mathematics*. Cambridge: Cambridge, UP.

(Footnote 36 continued)

I want to thank you for criticism and suggestions Delia Belleri, Paul Horwich, Sebastian Moruzzi and Marco Santambrogio.

- Recanati, François. 2010. *Truth-Conditional Pragmatics*. New York: Oxford UP.
- Sbisà, Marina. 2002. Speech acts in context. *Language and communication* 22: 421–436.
- Schegloff, Emmanuel, Harvey Sacks, and Gail Jefferson. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53: 361–382.
- Searle, John. 1983. *Intentionality*. Cambridge: Cambridge, UP.
- Sellars, Wilfrid. 1954. Presupposing. *The Philosophical Review* 63: 197–215.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and Philosophy*, eds. M. Munitz and P. Unger, 197–214. New York: New York UP. Reprinted in Stalnaker, 1999, 47–62.
- Stalnaker, Robert. 1978. Assertion. In *Syntax and semantics 9: pragmatics*, ed. P. Cole, 315–322. New York: Academic Press. Reprinted in Stalnaker 1999, 78–95.
- Stalnaker, Robert. 1998. On the representation of context. *Journal of Logic, Language and Information* 7: 3–19. Reprinted in Stalnaker 1999, 96–113.
- Stanley, Jason. 2000. Context and logical form. *Linguistics and Philosophy* 23: 391–434.
- Stanley, Jason. 2007. *Language in context*. Oxford: Clarendon Press.
- Travis, Charles. 1985 (2008). On What Is Strictly Speaking True. *Canadian Journal of Philosophy* 15: 187–229. Reprinted in his *Occasion-sensitivity: selected essays*, 19–64. Oxford: Oxford UP 2008.
- Tufte, Edward. 2001. *Visual Display of Quantitative Information*. Cheshire CT: Graphics Press 2nd ed.

Squaring the Circle

Kepa Korta and John Perry

Abstract Making the distinction between semantics and pragmatics has proven to be a tricky task, leading to several problems that look like Gordian knots, or worse; perhaps semantics and pragmatics are so tangled that separating them is impossible, like squaring the circle. A widespread, plausible, Grice-inspired view of the distinction is threatened by what (Levinson *Presumptive meanings*. MIT Press/Bradford Books, Cambridge, Mass, 2000) called ‘Grice’s circle.’ Gricean inferences to derive the pragmatic content of the utterance (such as conversational implicatures) require the determination of what is said (also known as the ‘semantic content’ or the ‘literal truth-conditions’); but determining what is said (by processes of disambiguation, precisification, reference fixing, etc.) requires pragmatic inference. In a nutshell, pragmatic inference both requires and is required by the determination of what is said. Thus, there is no way to unravel semantics and pragmatics. In this paper, we will show how to square Grice’s circle. We untie the semantics/pragmatics knot, without using any of Alexander’s methods: slicing it with a sword or removing the (semantic) pin around which it was bound. The approach consists in assuming a minimal but truth-conditionally complete notion of semantic content (Perry *Reference and reflexivity*. CSLI Publications, Stanford, 2001), which doesn’t constitute what is said by the utterance, but does provide the required input for pragmatic reasoning.

K. Korta (✉)

Institute for Logic, Cognition, Language and Information (ILCLI), University of the Basque Country (UPV/EHU), Donostia, Spain
e-mail: kepa.korta@ehu.es

J. Perry

Department of Philosophy, University of California, Riverside, USA

1 Introduction

Grice (1967) famously distinguished between what a speaker said by an utterance, on the one hand, and what she implicated by saying what she said, on the other. This was widely taken, at the time, as providing a clear-cut distinction between semantics and pragmatics. Semantics would deal with ‘what is said’, also known as ‘the proposition expressed’ by the utterance (or sentence in context), its ‘propositional content’.¹ (See Kaplan (1989) for a classic statement). Pragmatics would be concerned with implicatures and any other aspect of meaning not pertaining to the truth-conditional content of the utterance. As Gazdar (1979) famously put it

PRAGMATICS = MEANING – TRUTH – CONDITIONS.

But things were soon shown to be not so simple. Figuring out the reference of many natural language expressions seems to require reasoning the same sorts of pragmatic reasoning involved in the generation and understanding of Gricean implicatures. We have pragmatics intruding on the realm of semantics.²

Grice’s initial impact on pragmatic theory was due to the wide range of applications for his theory, but as with all philosophical theories, considerable critical attention was paid to fundamental ideas and distinctions. Interest first focused on his distinctions between conventional and conversational implicatures, and particularized and generalized ones. With time, attention turned to the concept of ‘what is said’ and the extent to which this is affected by “pragmatic intrusion,” that is, the extent to which understanding what is said by the speaker requires the intervention of pragmatic processes.

The issue of the amount of pragmatic intrusion into the semantic content of the utterance is at the heart of the debate held by literalists (or minimalists) and contextualists and a number of positions in between (indexicalists, situationalists, and radical and moderate of all sorts). Despite their differences, they seem subject to what Levinson (2000) dubbed ‘Grice’s circle’:

Grice’s account makes implicature dependent on a prior determination of ‘the said.’ The said in turn depends on disambiguation, indexical resolution, reference fixing, not to mention ellipsis unpacking and generality narrowing. But each of these processes, which are prerequisites to determining the proposition expressed, may themselves depend crucially on processes that look indistinguishable from implicatures. Thus what is said seems both to determine and to be determined by implicature (Levinson 2000: 186).

Thus, within the Gricean picture, the processes of intention-recognition that invoke the Cooperative Principle and the conversational maxims seem to be

¹ The category of conventional implicatures doesn’t fit this picture, however, since being the result of the semantics (of certain words) of the sentence uttered, they do not contribute, according to Grice, to what is said.

² For a recent defense of Gazdar’s view on the semantics/pragmatics divide that attempts to avoid Grice’s circle, see Capone (2006).

needed to determine the proposition expressed by the utterance; but, at the same time, the proposition expressed seems to be required for these processes to get started. Many working practitioners accept that semantics and pragmatics are irretrievably entangled in the determination of utterance content. There is no clear-cut delimitation between semantics and pragmatics. They ignore the Gordian knot, rather than trying to untie it.³

We have a bit of terminology we find useful in thinking about this. What is said, or the proposition expressed, plays a central role in the classic picture. We call pragmatics in the service of figuring what is said, “near-side pragmatics” (see Korta and Perry (2006b)). This is the pragmatic reasoning that seems to be intruding into something that is none of its business on the classical conception. Pragmatic reasoning that starts with what is said, and seeks to discover what one is doing in saying or by saying it—what illocutionary and perlocutionary speech acts were performed, we call “far-side pragmatics”. So the question is, how does near-side pragmatics make sense?

Some tried a somewhat different approach: to assume that, all in all, semantics does *not* yield fully truth-conditional content. The “output” of semantics is the “input” to pragmatics, but this output is not a fully determined proposition, and not *what is said*. Instead of talking about ‘pragmatic intrusion’, we should accept that *what is said* is systematically a pragmatically determined content and not the input to pragmatics; the input to pragmatics is just what semantics gives us. This approach has been seen as undermining truth-conditional semantics, as cutting through the knot rather than untying it, to pursue one of our metaphors. But of course some are happy to cry “Truth conditional semantics is dead! Long live to truth-conditional pragmatics!”⁴

In this paper, we’ll argue that there are two false assumptions that generate the circle. One is that we need to identify what is said, in a canonical way, before Gricean considerations can be applied. The other is that simply because semantics underdetermines what is said, it does not provide a propositional content that can be the basis of reasoning. We’ll show that these assumptions are wrong and that abandoning them gives way to a natural account of the semantic content of an utterance, what is said, and the semantics/pragmatics distinction. So we claim, in terms of our metaphors, to square Grice’s circle, or untie the knot without slicing

³ According to Bach (2011), Levinson makes a mistake to the extent that he sees Grice himself as having the views that lead to the circle. According to Bach, Levinson conflates two senses of ‘determine’: one related to what the grammar delivers in combination with context, the other with the psychological process of ascertaining the content by a hearer:

“He [Grice] never claimed that the hearer’s inference proceeds from first identifying what the speaker says to *then* considering whether there is any ostensible breach of the maxims and, if so and assuming the speaker is being cooperative and is aiming to communicate something, to seek a plausible candidate for what that could be.”

This is doubtless a correct point about Grice, and shows that the circle need not be temporal. But it does not explain how Gricean considerations are brought to bear prior to identifying what is said; this is what we try to do.

⁴ See Recanati (2010).

it, that is, without undermining truth-conditional semantics. We'll start by considering Grice's concept of what is said.

2 Grice on What is Said

Grice (1967/1989) famously distinguished between *what a speaker says* and *what she implicates* by uttering a sentence. Think about Anne and Bob talking about their common friend Carol, who both know that she recently started working in a bank. Anne asks: "How is Carol getting on in her job?" Bob replies: "Oh quite well... She hasn't been to prison yet." Bob is clearly suggesting something here; something related to Carol's tendency to yield to the temptation provided by her occupation, as Grice would put it. However, that's not something he said, but something he implicated in saying what he said.⁵ But what did he say?

Grice's remarks suggest that his concept of 'what is said' can be taken as equivalent to 'the proposition expressed' or 'the content' of the utterance.⁶ He claims that to know what someone said by uttering a sentence one has to know

- (1) the conventional meaning of the sentence uttered;
- (2) the disambiguated meaning of the sentence in that particular occasion of use; and
- (3) the referents of referential expressions (Grice 1967/1989: 25).

This view of what is said fits well with (Kaplan's 1989) distinction between the character of a sentence, and the content of an utterance (or sentence-in-context). This also results in a seemingly perfect match between semantics and pragmatics: semantics deals with what is said; pragmatics deals with implicatures. The input to pragmatics is the output of semantics, in the form of disambiguated meaning plus context and fixing the reference of names.

But Grice's concept of what is said turned out to be less simple and clear as initially thought. Perry (1986) argued that someone who utters "It is raining" normally expresses a proposition that includes the place of the raining event, even if the sentence does not include any expression *articulating* that element of the proposition expressed. A variety of phenomena arguably showed a similar point: quantifier domain restrictions, comparative adjectives, assertions about taste, and a long list of other phenomena seem to involve constituents of the proposition

⁵ Grice included in his overall picture of meaning and communication non-linguistic 'utterances' like gestures and movements, but we will limit the discussion to linguistic utterances.

⁶ These more technical terms used by philosophers are not without problems, since they can suggest that implicatures are not contents of the utterance, or they are not propositional. Gricean implicatures (at least conversational particularized ones) are also full-blown truth-conditional (though more or less indeterminate) contents of the utterance, but we ignore this issue here, and follow common practice using 'content' only to talk about the contents that are on the 'what-is-said' part of the Gricean divide.

expressed that are not articulated by any element in the sentence uttered. A significant part of the debate between minimalism and contextualism, and all the *isms* around them, concerns the analysis of these elements: whether they are actually part of what is said or should be relegated to some other category, like the category of *implicatures* (Bach 1994) or generalized conversational implicatures; whether they are actually unarticulated or they are values for some ‘hidden’ indexicals in the logical form of the sentence uttered.

One might think that this kind of pragmatic intrusion into determining what is said is all there is to the circle; and that just a minimalist approach as found in, for instance, (Cappelen and Lepore 2005), is enough to square it. The product of sentence meaning plus disambiguation and reference-fixing of names and indexicals might not yield what intuitively is said by the speaker, but it would have the merit of giving us a clear-cut notion of *semantically expressed proposition*. If you say, “It is raining,” the semantically expressed proposition is simply *that it is raining*; something true if it is raining anywhere (roughly). That proposition is trivial, so pragmatics takes the hearer to a more promising one as what the speaker conveys: *it is raining in X*, where the identity of X is determined pragmatically.

Another view is that semantics alone doesn’t (always) get us to a proposition at all, even if we assume disambiguation and add context and reference fixing. We can only rely on obtaining a ‘proposition template’ (Carston) or a propositional radical (Bach). Moreover, whatever exactly the result is, it does not seem to constitute the right ‘input’ for implicatures; their derivation requires an ‘enriched’ proposition to serve as ‘what is said’: the content of the utterance with its *implicatures*.

On this issue, we side with Cappelen and Lepore, although with important differences. Like them, our view is that semantics provides us with propositions, and these propositions are not what is said. But our account is even more minimalist a conception than the one Cappelen and Lepore provide, and, we think, much more intuitive in a wide variety of cases.

To see the merits of our position, it is helpful to note that a Cappelen and Lepore style minimalism does not actually get us out of the Gricean circle. This is because the factors they fold into semantics: disambiguation and reference-fixing in particular, are often only resolved using pragmatic methods. As Levinson notes, reference fixing and indexical resolution and disambiguation, ‘which are prerequisites to determining the proposition expressed, may themselves depend crucially on processes that look indistinguishable from implicatures.’

Here is an example involving reference-fixing. Suppose that the Stanford Philosophy Department is meeting in the 1980s, with John Perry, John Etchemendy, John Dupre and Jon Barwise all in attendance. John Perry has been talking at length, while Jon Barwise has been waiting impatiently to say something. The chair, Nancy Cartwright says, “/djon/ needs some time to develop his views”. Is she referring to John Perry, using the name “John”, or one of the other Johns, or Jon Barwise, using his name? It seems that the listener will most likely try to figure out whether she is speaking literally, referring to John Perry and implicating that Barwise should calm down (unlikely), being non-literal, referring

to John Perry, and implicating that he should shut up, (more likely, postulating a sarcasm within Cartwright's repertoire), or speaking literally, referring to Jon Barwise (using the name 'Jon'), and implicating that time is running out so he needs to be given a turn (most straightforward), or referring to Etchemendy or Dupre, and implicating that they should quit dozing and get involved (a distinct possibility).

What seems beyond question is the principle of underdetermination,

The linguistic meaning of a sentence underdetermines what is said by a speaker uttering that sentence.

If all there is to semantics is to give the linguistic meaning of sentences (types), then there is an obvious sense in which that claim is true. Sentences say nothing; utterances do or, better, speakers do by uttering sentences. The semantic meaning of a sentence type uttered on a particular occasion is not (often, according to 'moderate' contextualists; always, according to 'radical' ones) enough to determine what the speaker said by the utterance. She might not have *said* anything, but have been asking a question, giving an example, or rehearsing a line for a play. More relevantly for our purposes, the linguistic meaning of the uttered sentence does not seem to yield a fully truth-conditional or propositional content.

3 Minimal but Complete Semantic Contents

It's helpful to start with indexicals. Suppose, out of the blue, in a crowded room, you hear the utterance

(1) I am French.

You don't even see who is doing the talking. Having no clue about who is talking, you would not really be able to say what the speaker said. Your semantic competence gets you only so far; it seems you wouldn't grasp a proposition, but just a propositional template or propositional 'radical'; a predicate like

(2) x is French.

Since you don't know who is speaking, you don't know who this predicate has to be true of, for the utterance to be true.

On the other hand, semantics does provide you with enough to get started. You know that (1) is true iff,

(3) $\exists x$ (x is the utterer of (1) and x is French)

(3) Gives you a proposition, which may serve to get some reasoning started. Perhaps you reason that most likely (1) is true, and whoever said it is French; you can't think off hand of any reason why someone would be claiming to be French if they weren't; it didn't sound sarcastic, but had more of an informative tone. This may be rather weak pragmatic reasoning, but it is pragmatic. You might look

around for some reliable signs of a French person—someone who is smoking Gauloises cigarettes for example. If you spot such a person, and are right, you will have used pragmatic reasoning and semantics (what is required for the utterance to be true, plus what seems to be the intention behind the utterance) to help you figure out who said it, which allows you to figure out *what is said*.

This illustrates our basic strategy. Semantics provides slots, which provides truth conditions for utterances; by existentially quantifying we get a proposition; this proposition usually won't be *what is said*, but it provides us with what is needed to start what we call “near side reasoning”; that is reasoning that gets us *from* perception of an utterance of a sentence and a grasp of the semantics, *to* what is said.

This picture is pretty much in tune with Borg's (2004) minimalist view of semantic content, and with Grice's comments on what one understands from an utterance in virtue of knowing the language. About an utterance of ‘He was in the grip of a vice’, he makes the following remark:

Given a knowledge of the English language, but no knowledge of the circumstances of the utterance, one would know something about what the speaker had said, on the assumption that he was speaking standard English, and speaking literally. One would know that he had said, about some particular male person or animal x , that at the time of utterance (whatever that was), either (1) x was unable to rid himself of a certain kind of bad character trait or (2) some part of x 's person was caught in a certain kind of tool or instrument (approximate account, of course) (Grice 1967/1989: 25).⁷

As Grice's remark suggests, we don't even need disambiguation to get these sorts of truth-conditions. Given the meanings in (British) English of “He was in the grip of a vice”, there is a fully propositional if utterance-bound content of **u**, namely,

- (4) $\exists x \exists t$ (x is the person or animal the speaker of **u** is talking about and t is the time of utterance & at t either x was unable to rid himself of a certain kind of bad character trait or some part of x 's person was caught in a certain kind of tool or instrument)

Suppose you receive an unsigned postcard that reads

- (5) I am having a good time here.

Without identifying the writer, the time and the place of the writing, you would not know what she or he was saying to you. But this does not mean that you would fail to grasp a complete proposition from the postcard on the basis of your semantic competence. You would understand that the utterance would be true if

⁷ In American English this is an example of two words, ‘vice’ and ‘vise,’ both pronounced/vais/. Parallel considerations would apply. But we follow British English, and Grice, in taking it to be ambiguity.

and only if its author was having a good time at the time of the utterance at the place of the utterance. This is a perfectly truth-conditionally complete content.

These propositions are not what is said, and they aren't even *about* the same things the speaker is talking about. (4) is about the utterance **u**, not about the person the speaker refers to with 'he', or even about the speaker himself or herself. We call such propositions *utterance-bound* contents. We want to emphasize three things about such contents before going further:

- (a) Utterance-bound contents are utterance-bound with respect to the utterance they are about. (4) is the utterance-bound content of **u**, but it is not the utterance-bound content of (4).
- (b) We do not claim, to repeat ourselves, that the utterance-bound content of an utterance is what the utterance expresses, or what the utterance says, or what the speaker of the utterance says.
- (c) We do claim that the utterance-bound content gives the truth-conditions of the utterance. That is, it gives the conditions that the utterance must meet, in order to be true; there must be various things, speakers, things the speaker refers to and the like, that stand in various relations to the utterance and fulfill further conditions. But saying this can be misleading. In philosophy, at least, one typically uses the term "truth-conditions" for the *counterfactual* truth-conditions, the conditions that a situation or world must satisfy, to be one in which the proposition expressed, or what is said, is true. The counterfactual truth-conditions usually line up with what we call the *referential* truth-conditions, which is what you get when you identify the witnesses of the various existential quantifiers and plug them in for the variables. For example, with respect to our last example, this might be the proposition

That Hiram is/was having a good time on March 4, in Hawaii.

This proposition isn't about an utterance, and could be true in worlds in which the utterance **u** did not occur.

Although potentially misleading, we think our use is a correct and literal use of the term 'truth-conditions'. You get different truth-conditions for an utterance, depending on what you hold fixed and what you allow to vary. The truth-conditions are *what else* has to be the case, *given* what is held fixed, for the utterance to be true. Both utterance-bound and referential truth-conditions are truth-conditions; we aren't replacing the ordinary philosophical concept, but noticing that it is part of a system of truth-conditions, or contents, that an utterance can have.

4 Utterance Contents and Implicatures

Even admitting that an utterance has a truth-conditionally complete content before disambiguation, reference fixing and any other near-side pragmatic considerations, it can be argued that, since this content does not amount to what the speaker said

by her utterance, we need to perform disambiguation and reference assignment and even further pragmatic processes (like so-called ‘free enrichment’ processes) to get the content that can appropriately be called what is said (or the ‘explicature’). This would be required by Grice’s picture of the inference of implicatures. Minimalists like Cappelen and Lepore share this view with contextualists like Carston (2002):

We agree with her that you need a contextually shaped content to generate implicatures in all of the cases she discusses. (...) What’s needed in order to derive the implicature in these cases is a contextually shaped content, i.e., a contextually shaped what-is-said.

(...)

More generally: We are happy to agree with Carston that an appropriate notion of what the speaker said must allow for contextual influences that go far beyond what the speaker said (Cappelen and Lepore 2005: 180–181).

So, if this morning someone invited you to a coffee and you uttered

(6) I’ve had breakfast,

trying to implicate a negative answer, the implicature wouldn’t have gone through, had he only got its utterance-bound content (7), or what Cappelen and Lepore take to be the proposition semantically expressed (8) (the result of semantic meaning plus disambiguation and reference assignment to referential expressions). Something like (9) which, arguably, includes elements that are not articulated in the uttered sentence, needs to be determined to infer the implicature:

(7) There is a time previous to the utterance of (6), when the speaker of (6) has had breakfast.

(8) X has had breakfast at some point or other.

(9) X has had breakfast this morning.

In this and other examples, the concept of what is said required by Grice’s theory of implicatures seems to go beyond our fully truth-conditional but minimal level of content. Hence, we seem condemned to Grice’s circle, after all.

But we are not. First, and most importantly, we have a truth-conditionally complete content whose determination is independent of pragmatic reasoning, and thus keep us out of the circle. And, second, because, contrary to what Cappelen and Lepore suggest, it is not generally the case that the hearer needs arriving at a ‘contextually shaped’ what is said to understand the implicatures of an utterance. It is not even necessary that he gets what is said or what they call the proposition semantically expressed. It is sometimes enough to get the utterance-bound content of the utterance, or a content that with some undetermined (but existentially bound) element.⁸ Suppose Kepa asked John for suggestions about whom to invite for an upcoming pragmatics conference. The conversation runs like this:

⁸ See (Korta and Perry 2006a, 2008 and 2011). In the latter we distinguish utterance-bound, speaker-bound, network-bound, referential and designational contents. We contend that any of those can be and is the right ‘input’ for the inference of implicatures.

J: He is rather unreliable, doesn't have much to say, and always takes a long time to say it.

K: Next.

In this case, Kefa need not resolve the referent for John's use of the demonstrative 'he', and maybe John doesn't intend him to resolve it. He trusts that grasping the utterance-bound truth-conditions of his utterance,⁹ Kefa will infer that that guy would not be an appropriate candidate for lecturing at the conference. Or take our postcard example. You can guess, say, who the author is, but not the place he is talking about. You can, however, understand that he is implying he is postponing his trip back home. The moral is that even without determining what is said in the sense of disambiguating the expressions used and fixing the referents involved, Gricean inference of implicatures is often possible.

To sum up, Grice's circle is avoided once a fully truth-conditional minimal semantic content is provided. This is our utterance-bound content, which does not require any pragmatic process of disambiguation, anaphora resolution or any other process that looks practically indistinguishable from implicature inference. The utterance-bound content itself can constitute a sufficient input for figuring out an implicature in some cases, while in other cases more facts about the utterance may need to be fixed. In some cases, the referential content of the utterance may suffice, in others one may need "the contextually-shaped what is said". But even in these cases the pragmatic reasoning can begin with the utterance-bound content, and can be used to arrive at the more specific contents that are required. Suppose you take your watch to be repaired and the watchmaker tells you:

(10) It will take some time to fix this watch.

The referential content of (10) is trivially true; any human act takes some time to perform. The enriched content of (10) (with the *implicature* within brackets) would be something like

(11) It will take some time [more than you might expect] to fix this watch,

which would allow you to infer the intended implicature, say, that you should take it easy. Now, some other shopper who overhears the conversation without being in position to fix the reference of the speaker's use of this watch would easily understand the implicature and implicature at issue. The inference of pragmatic contents starts from the utterance-bound content with or without reference fixing, disambiguation and other near-side pragmatic processes.

⁹ Or, more precisely, the speaker-bound truth-conditions. See Korta and Perry (2011, 2013).

5 Truth-Conditional Semantics and Pragmatics

In our view semantics has to do with the conventional meanings of words and modes of combination, and its most central part is truth-conditional semantics. Truth-conditional semantics gives us the truth-conditions of utterances in terms of the constraints imposed by these meanings. They give us the utterance-bound or reflexive truth-conditions in terms of the utterance itself. They do not give us what is said or the proposition expressed by the speaker—speakers most often attempt to talk about things in the world, not about their own utterances. So our view of semantics is minimalist.

But being minimalist without allowing pragmatic ‘intrusion’ into semantic content does not necessarily involve sacrificing truth-conditional semantics, and being pushed into Grice’s circle. Our approach allows a clear-cut distinction between semantics and pragmatics that avoids it and offers the ideal toolkit to account for the relation between our knowledge of language and its use in communication.

Acknowledgments The first author is grateful to his colleagues of the group ‘Language, Action and Thought’ at the Institute for Logic, Cognition, Language and Information (ILCLI) for their comments and criticisms, especially to Eros Corazza, Joana Garmendia, María Ponte and Larraitz Zubeldia. This work was partially supported by a grant of the Basque Government (IT323-10 and IT780-13) and the Spanish Ministry of Science and Innovation (FFI2009-08574 and FFI2012-37726).

The second author appreciates the support of the philosophy department at the University of California, Riverside, the Center for the Study of Language and Information at Stanford University, and conversations with many colleagues and students.

Both authors want to thank Alessandro Capone for his infinite enthusiasm and patience.

References

- Bach, Kent. 1994. Conversational implicature. *Mind and Language* 9: 124–162.
- Bach, Kent. 2011. Review of truth-conditional pragmatics (François Recanati). Notre Dame philosophical reviews. <http://ndpr.nd.edu/news/25657-truth-conditional-pragmatics-2/>. Accessed 26 Sept 2011.
- Borg, Emma. 2004. *Minimalist semantics*. Oxford: Oxford University Press.
- Capone, Alessandro. 2006. On Grice’s circle (a theory-internal problem in linguistic theories of the Gricean type). *Journal of Pragmatics* 38: 645–669.
- Cappelen, Herman, and Ernie Lepore. 2005. *Insensitive semantics*. Oxford: Blackwell.
- Carston, Robyn. 2002. *Thoughts and utterances. The pragmatics of explicit communication*. Oxford: Blackwell.
- Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.
- Grice, H. Paul. 1967. Logic and conversation. In *The logic of grammar*, eds. D. Davidson and G. Harman, 1975, 64–75. Encino: Dickenson. Also published in *Syntax and Semantics 3: Speech Acts*, eds. P. Cole and J.L. Morgan, 1975, 41–58. New York: Academic Press. Reprinted in Grice 1989: 22–40.

- Grice, H. Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.
- Kaplan, David. 1989. Demonstratives. In *Themes from Kaplan*, eds. J. Almog, J. Perry and H. Wettstein. New York: Oxford University Press, 561–614.
- Korta, Kapa, and John Perry. 2006a. Three demonstrations and a funeral. *Mind and Language* 21: 166–186.
- Korta, Kapa and John Perry. 2006b. Pragmatics. *The Stanford Encyclopedia of Philosophy (Revised March 2011)*, ed. Zalta, Edward, N. <http://plato.stanford.edu/entries/pragmatics/>.
- Korta, Kapa, and John Perry. 2008. The pragmatic circle. *Synthese* 165: 347–357.
- Korta, Kapa, and John Perry. 2011. *Critical pragmatics. An inquiry into reference and communication*. Cambridge: Cambridge University Press.
- Korta, Kapa, and John Perry. 2013. Highlights of critical pragmatics: reference and the contents of the utterance. *Intercultural Pragmatics* 10(1): 161–182.
- Levinson, Stephen. 2000. *Presumptive meanings*. Cambridge, Mass: MIT Press/Bradford Books.
- Perry, John. 1986. Thought without representation. *Proceedings of the Aristotelian Society* (Supplementary volume), 60: 137–51. Reprinted in Perry 2000: 171–188.
- Perry, John. 2000. *The problem of the essential indexical and other essays expanded*. Stanford: CSLI Publications.
- Perry, John. 2001. *Reference and reflexivity*. Stanford: CSLI Publications.
- Recanati, François. 2010. *Truth-conditional pragmatics*. Oxford: Oxford University Press.

Irregular Negations: Pragmatic Explicature Theories

Wayne A. Davis

Abstract I will examine negations that are “irregular” in that they are not used in accordance with standard logical rules. These include scalar-, metalinguistic-, specifying-, and evaluative-implicature denials; presupposition-canceling denials; and contrary affirmations. The principal questions are how their irregular interpretations are related to their regular interpretation, and whether their ambiguity is semantic or pragmatic. I argue here that pragmatic “explicature” (Carston) or “implicature” (Bach) theories have few advantages over implicature theories (Grice, Horn, Burton-Roberts), and that clear examples of pragmatic explicatures involve indexicality or syntactic ellipsis, which are not involved in irregular negations. I argue against claims that any interpretation can be “pragmatically derived” using either Gricean or Relevance theory. With one class of exceptions, I argue for a semantic ambiguity thesis maintaining that irregular interpretations are *idioms* that plausibly evolved from generalized conversational implicatures. The exceptions are evaluative-implicature denials, which are still live implicatures.

1 Irregular Negations

We will be focusing on negations as a species of *sentence*. The term ‘negation’ can also be used in different senses to denote a propositional operator, a proposition with that operator applied, or the speech act of asserting such a proposition. When regular, the form ‘Not-p’ stands for a sentence expressing the result of applying the negation operator to the proposition expressed by ‘p.’

W. A. Davis (✉)

Philosophy Department, Georgetown University, 20057 Washington, DC, USA
e-mail: davisw@georgetown.edu

Horn (1989: 362–4; 370–5) observed that many negations have an “exceptional” interpretation. Six types have been identified.

(1) (a) The sun is not larger than some planets: it is larger than all planets	<i>Scalar-Implicature Denial</i>
(b) That’s not a <i>tomāto</i> : it’s a <i>tomāto</i> ^a	<i>Metalinguistic-Implicature Denial</i>
(c) Vulcan is not hot: it does not exist	<i>Presupposition-Canceling Denial</i>
(d) The sky isn’t partly cloudy, it’s partly sunny	<i>Evaluative-Implicature Denial</i>
(e) Mary did not meet a <i>man</i> at the bar, she met her husband	<i>Specifying-Implicature Denial</i>
(f) John does not believe there is a god, so he is an atheist	<i>Contrary Affirmation</i>

^a ‘ä’ represents the vowel sound in ‘ah’ and ‘ma,’ ‘ā’ the vowel sound in ‘pay’ and ‘ate’

If the negation clause in (1)(a)—(1)(a)_n—had its regular (“ordinary”) interpretation, the conjunction as a whole would be contradictory. It would mean *the sun is larger than no planets: it is larger than all planets*. And if (1)(f)_n were regular, the conclusion would not follow. These negations would more charitably be interpreted in these contexts as having the meanings of the negations in (2)¹:

- (2) (a) The sun is not larger than just some planets: it is larger than all planets.
 (b) That’s not properly called a *tomāto*: it’s called a *tomāto*.
 (c) Vulcan is hot is not true: Vulcan does not exist.
 (d) It is not bad that there are some clouds, it is good that there is some sun.
 (e) Mary did not meet an unrelated man at the bar, she met her husband.
 (f) John believes there isn’t a god, so he is an atheist.

One of the most important properties of the irregular interpretation of these negations is that familiar logical rules governing negations do not hold. I use this property to define *irregular* (Davis 2010). For example, on the irregular interpretation, (1)(a)_n is true even though its root ‘The sun is larger than some planets’ is true, not false. Similarly, if John is agnostic, ‘John does not believe there is a god’ is just as false on its irregular interpretation as ‘John believes there is a god.’ There are no violations of laws of logic, of course. When a sentence ‘Not-p’ is an irregular negation, the proposition it expresses is not the negation of the proposition expressed by its root ‘p.’ Thus ‘The sun is not larger than some planets’ in

¹ See Horn (1989: 385); Geurts (1998); Davis (2010, 2011). The formulation of the negation in (c) is colloquial. In formal writing, we would enclose ‘Vulcan is hot’ in quotation marks. There are no quotation marks in speech, but the subordinated sentence would receive a distinctive intonation contour. The most explicit formulation is ‘The proposition that Vulcan is hot is not true’.

(1)(a) expresses not the negation of the proposition that the sun is larger than some planets (which negation is true only if the sun is larger than no planets), but the negation of the proposition that the sun is larger than just some planets (which negation is not true if the sun is larger than no planets). And ‘John does not believe there is a god’ in (1)(f) expresses the proposition that John believes there isn’t a god rather than the negation of the proposition that John believes there is a god.

Because they have both regular and irregular interpretations, the six types of negation illustrated by (1) have some sort of *ambiguity*. We will be concerned with the nature of this ambiguity. Is it semantic or pragmatic? If semantic, is it due to an ambiguity in the word ‘not’ or something else? If pragmatic, how does it arise?

Horn also observed that scalar-implicature, metalinguistic-implicature, and presupposition-canceling denials are normally “marked” by unusual non-semantic properties.² They typically have a distinctive intonation (a “fall-rise contour”), and are typically echoic (responses to affirmations of their roots). Specifying- and evaluative-implicature denials typically have the same marks. These marks may be absent, however, and may be present in regular negations like (3)_n:

(3) That’s not a violin, it’s a viola.

Contrary affirmations, moreover, rarely if ever have the fall-rise intonation, and are less echoic than the other irregular negations.

Irregular negations also tend to have reversed polarity licensing, allowing positive polarity items while blocking negative polarity items. Whereas ‘any’ can replace the first ‘a’ in (3) without changing the sense, it cannot replace ‘some’ in (1)(a) without making the negation regular. Irregular negation also resists morphological incorporation. ‘Not’ typically incorporates morphologically as the ‘un-’ or ‘in-’ prefixes, as in ‘unnoticed’ and ‘immaterial.’ Yet a sentence like ‘Vulcan is immaterial’ lacks the ambiguity of ‘Vulcan is not material,’ having only a regular interpretation. Contrary affirmation is again exceptional. As a contrary affirmation, ‘John does not believe there is a god’ is equivalent to both ‘John *dis*believes there is a god’ and ‘John does not believe there is *any* god.’

Some generalizations apply to all irregular negations, not just those typically marked. One universal is *limited substitutivity of synonyms*.³ ‘The sun *failed to be* larger than some planets, it is larger than all planets’ can only be heard with the regular, contradictory meaning of (1)(a). Replacing ‘not hot’ with ‘non-hot’ in (1)(c) has the same effect. Even in (1)(f), ‘not believe’ can be replaced by ‘fails to believe’ only if the negation is regular. For a more complex example, note that regular negation-correction conjunctions of the form ‘s is not P: it is Q’ like (3) can be paraphrased either ‘S is not P but Q’ or equivalently ‘s is not P, but it is Q.’ In example (1), however, the first sort of paraphrase is possible, but not the second.

² Horn (1989: 368; 374–5; 381, 392–413). See also Kempson (1986: 88); Burton-Roberts (1989: 118); Seuren (1990: 449–52); Chapman (1996: 390–1); Israel (1996: 621n1); Carston (1998: 332ff); Geurts (1998: 275, 278–80, 303); van der Sandt (2003: Sect. 7); Ramat (2006: 560).

³ Cf. Seuren (1988: 183; 1990: 443); Geurts (1998: 279). Contrast Grice (1981: 271).

‘The sun is not larger than some planets, but it is larger than all planets’ is incoherent; to obtain a paraphrase, ‘for’ should replace ‘but.’ Note too that when ‘The sun and Sirius are not larger than some planets’ is irregular, it is not equivalent to ‘Neither the sun nor Sirius is larger than some planets.’ Failures of morphological incorporation are other examples of substitutivity failure. The fact that substitutivity of synonyms is limited in irregular negations has as a corollary that they are only *partially compositional*. Replacing ‘larger’ with ‘smaller’ in (1)(a)_n changes its meaning in the predictable way. But replacing ‘some’ with the non-synonymous phrase ‘just some’ produces a sentence with the same meaning, although it is no longer irregular. Replacing ‘some’ with ‘no’ changes the meaning radically, but not in the predictable way: the results are contraries but not contradictories.

Horn (1989: 363, 374) called marked negations “metalinguistic.” Only some are metalinguistic in the logician’s sense. Whereas denials like (1)(b)_n actually refer to words, denials like (1)(a)_n and (1)(c)_n do not.⁴ They are not using English to talk about English. Horn used ‘metalinguistic’ because he took marked negations to formulate *objections to previous linguistic utterances*.⁵ Metalinguistic negation is

a device for objecting to a previous utterance on any grounds whatever, including the conventional or conversational implicata it potentially induces, its morphology, its style or register, or its phonetic realization. (Horn 1989: 363)

Horn (1989: 381) took a marked negation to “echo” a prior assertion of its root, in order to criticize it, as in (4).

- (4) Dick: The sun is larger than some planets.
Ed: The sun is not larger than *some* planets: it is larger than *all* planets.

Carston concurs that “The correct generalization about the metalinguistic cases is that the material in the scope of the metalinguistic operator, or some of it at least, is echoically used” (1986: 320). This is “the single essential and unifying property” (1998: 318).

As I show more fully in Davis (2011), being used to object to a previous utterance is not *essential* even to marked negations. There is no doubt that the negations in (1) *can* be used echoically to object to a previous utterance. Indeed, the fact that (1)(a)_n–(e)_n cannot *naturally* or *feliculously* be used out of the blue to initiate a conversation is the strong echoicity that marks them in a distinctive way. But initial use is nonetheless *possible*, without violating any rules of grammar or

⁴ Cf. Kempson (1986: 85ff); Geurts (1998: 275).

⁵ See also Gazdar (1979: 67); Horn (1985: 135); (1990: 495, 500); (1992a: 265); (2004: 10); Carston (1988: 320); (1996: 312, 320–322, 325n7, 327); (1998: 317–8); Burton-Roberts (1989: 111, 118ff); Van der Sandt (1991: 334–5, 337); (2003: §3, §4); Chapman (1996: 389, 392); Levinson (2000: 212); Huang (2007: 44). Horn (1989: 363, 374–7, 420–434) discusses many others who have held similar views, and credits the term ‘metalinguistic’ to Ducrot (1972). Contrast Geurts (1998: 278, 294).

semantics. Horn himself provided an attested example of a discourse-initial metalinguistic-implicature denial.

When Ronald Reagan carried 49 states and won 525 electoral votes, it was not an historic victory. Walter F. Mondale's poor showing wasn't an historic defeat. Mr. Mondale's choice of Geraldine A. Ferraro as his running mate wasn't an historic decision, either. None of these was an historic event. Each was a historic event. (John Chancellor, *New York Times* op-ed column). (Horn 1989: 391)⁶

Horn observes that Chancellor was clearly expecting the negations to be interpreted as regular until the last sentence forced a reinterpretation. Contrary affirmations are the least echoic. 'John does not believe there is life on Pluto' may lead us to imagine that someone said there is, but it is not at all infelicitous or unnatural if no one did.

More importantly, being used to object to a previous utterance is not *distinctive* of the marked negations, and does not explain their semantic irregularity at all. Thus, even though it is regular, (3)_n can be used to object to a previous utterance of 'That is a violin' on the grounds that the referent is a viola. Burton-Roberts (1989: 121) and Horn (1989: 397) thought Horn's thesis explained why negative polarity items cannot occur in the scope of a "metalinguistic" negation. But negative polarity items can appear in regular negations that are used to object to previous utterances, as in 'That is not a violin at all, it is a viola.' Furthermore, 'The president did not lose *any* hair, he lost *some* hair' is perfectly grammatical even though it contains a metalinguistic-implicature denial with a negative polarity item. Finally, even though we can object to an utterance of 'The sun is larger than any planets' on the grounds that it is a positive sentence with a negative polarity item, we cannot use 'The sun is not larger than any planets' as an irregular negation to lodge that objection.

What is essential to and distinctive of irregular negations is *being used to deny an implicature of their root* (Davis 2010: Sect. 7; 2011: Sect. 6). 'The sun is larger than some planets' has what Grice called a *generalized conversational implicature*. The sentence can be said to implicate "The sun is larger than just some planets" because it is conventional for English speakers to use sentences containing 'some' with such an implicature. As indicated in (2)(a), 'The sun is not larger than some planets' denies this implicature of its root when it is a scalar-implicature denial. Similarly, 'That is not a *tomäto*' implicates "That is not properly called a *tomäto*," which is what (1)(b)_n denies. Presupposition-canceling denials like (1)(c)_n are unusual in that the implicature it denies is something entailed by its root, namely "Vulcan is hot is true." Contrary affirmations are unusual in the same way. 'John believes there is a god' both implicates and entails "John does not believe there isn't a god," which is what (1)(f)_n denies. Since every irregular negation has a root, which invariably has some generalized implicatures, there is always something for it to deny—even if no one has uttered the root in the speaker's context, or

⁶ See also Horn's (1992b) birthday card example, discussed by Carston (1996: 312) and Chapman (1996: 395, 401–2).

used it with that implicature. When the root has more than one implicature, as is generally the case, the speaker determines which type of irregular negation is used by selecting which implicature to deny. Irregular negations are only partially compositional because they deny implicatures of their roots, which are not determined by the meanings of their roots.

The fact that irregular negations are used to deny an implicature of their root is the element of truth in Horn's thesis that they are devices for objecting to utterances on any grounds whatsoever. Horn's thesis is overly restrictive in one respect—the root need not have been uttered by anyone, and overly general in another—irregular negations can only be used to deny implicatures of their roots. A negation denying what its root asserts is regular (cf. Van der Sandt 2003, 2006). More significantly, an irregular negation cannot be used to object to a prior utterance of its root on the grounds that the speaker used it with a non-generalized implicature—something its root does not itself implicate. Thus, if A utters 'There is a gas station around the corner' implicating that B can get gasoline there, B can object to A's utterance if the station has no gasoline. But 'There is not a gas station around the corner' cannot be interpreted as an irregular negation denying that B can get gasoline there. For even though A implicated "B can get gas there," this is not an implicature of the sentence A used. Finally, what an irregular negation can be used to deny depends on what is within the scope of the negation. Even though 'It was the king of France who kissed the Canadian empress' presupposes both that France has a king and that Canada has an empress, 'It was not the king of France who kissed the Canadian empress' cannot be used to assert that 'It was the king of France who kissed the Canadian empress' is not true on the grounds that Canada has no empress. For the phrase generating the empress presupposition does not come within the scope of 'not.'

2 Implicature Theories

Given that the negations in (1) each have at least two conventional interpretations, it is natural to suspect that they are *semantically* ambiguous. This theory maintains that an irregular negation has at least one linguistic meaning in addition to its regular meaning. An immediate problem is to account for the source of the ambiguity. It seems clear that the various interpretations are not due to a *lexical* ambiguity.⁷ No ambiguity in 'not' could account for all the specific differences in meaning among (1)(a-f). Moreover, if 'not' were ambiguous, we should find an ambiguity in 'That is not the case' or 'Not everything flies'; but we do not. Finally, we would expect to find languages in which the difference is lexicalized (Gazdar

⁷ See Gazdar (1979: 65–6); Levinson (1983: 201); Burton-Roberts (1989, 1999); Horn (1989: 366, 385); Atlas (1989: 69); Seuren (1990: 449ff); Van der Sandt (1991: 333); Carston (1996: 327); Geurts (1998: 287). Contrast Seuren (1988: 196ff, 222).

1979: 65–6; Horn 1989: 367). It is equally hard to see how *syntactic* ambiguity could account for the different interpretations. The Russellian scope distinction might account for the ambiguity of ‘The present king of France is not bald,’ but not that of any of the negations in (1), nor the lack of a similar ambiguity in ‘The king of France hasn’t lost any hair yet.’ The ambiguity of ‘John does not believe there is a god’ looks like it is due to different clauses being within the scope of ‘not,’ but it can only be interpreted as applying to the main clause. Besides, parallel sentences like ‘John is not afraid that it will rain’ have no parallel ambiguity. The Aristotelian distinction between contradictory predicate-denial and contrary predicate-negation (Horn 1989: 107) might account for the ambiguity of ‘Vulcan is not hot,’ but not that of ‘The sun is not larger than some planets,’ ‘That’s not a *tomäto*,’ or ‘Mary did not meet a man at the bar.’ The irregular interpretations of these negations seem to introduce concepts not expressed by their roots, and thus result from more than structural reorganization. If the different interpretations of the negations in (1) are not due to an ambiguity in ‘not’ nor to an ambiguity in their grammatical structures, how can those interpretations be meanings of the sentences?

A natural alternative to the semantic ambiguity theory is the *implicature theory*. This theory maintains that one interpretation of each negation in (1) is its linguistic meaning, which results compositionally from its grammatical structure and the meanings of its component words. The other meanings are implicatures, things speakers mean but do not say when they use the negations. What speakers say is what the sentence means. What speakers implicate is a different proposition. Different types of implicature theory can be distinguished depending on whether the regular or irregular interpretation is taken to be an implicature (Table 1).

Burton-Roberts (1989: 117–9) holds the I-implicature theory, maintaining that when speakers use (1)(a)_n as a scalar-implicature denial, what they *say* is that the sun is larger than no planet. What they *implicate*—that is, what they mean by saying that—is that the sun is not larger than just some planet. Grice (1981) and Horn (1989: 486–90, 514) propose the R-implicature theory for presupposition-canceling denials. They maintain that the irregular interpretation of (1)(c)_n is what it literally means, while the regular interpretation is ordinarily an implicature. A speaker using ‘Pluto is not hot’ as a regular negation *says* that it is not true that Pluto is hot, while *implicating* that Pluto is non-hot. Bach (1994: 153–4) may hold the radical implicature theory. On all implicature theories, negations are *pragmatically* ambiguous (Horn 1985: 132, 1989: 370–7). One interpretation is their linguistic meaning, others are “pragmatically generated.”

Table 1 Implicature theories

	I-implicature theory	R-implicature theory	Radical implicature theory
Sentence meaning	N means R but not I	N means I but not R	N means neither I nor R
Irregular use	S says R and means I	S says and means I	S says O and means I
Regular use	S says and means R	S says I and means R	S says O and means R

The implicature theory seems right for evaluative-implicature denials like (1)(d), ‘The sky isn’t partly cloudy, it’s partly sunny.’ Here what the speaker *says* is contradictory, but what she *means* is not. What she means is that things are good because there is some sun, not bad just because there are some clouds. The fact that it cannot be partly sunny without being partly cloudy shows that what she said is false, but does not undermine the point of what she said. The obvious contradiction serves as a signal for implicature. The other irregular negations, however, lack the indirection characteristic of implicature. Speakers using (1)(a), for example, are not making any contradictory claims because they do not *say* that the sun is larger than no planet. Nor are speakers *implicating* that the sun is not larger than just some planets; that is what they are saying. It seems especially clear that when ‘Pluto is not hot’ is used as a regular negation, the speaker is saying that Pluto is non-hot, not implying it. In the case of (1)(c), we cannot interpret the speaker as implicating that Vulcan is non-hot. For that implies that Vulcan exists, which the speaker explicitly denies. The correction clause would cancel the alleged implicature of the negation. But then it is hard to see why we can interpret (1)(c) as a contradiction. Horn’s garden-pathing example from Chancellor is unusual in that hearers initially interpret the speaker as intending the regular interpretation. With the exception of evaluative-implicature denials, the regular interpretation of an irregular negation is usually not heard at all, so no “double processing” is required.

Grice (1975: 39) observed that implicatures are typically “nondetachable” from what is said. “The implicature is nondetachable insofar as it is not possible to find another way of saying the same thing (or approximately the same thing) which simply lacks the implicature in question” (Grice 1978: 43). Manner implicatures represent a large class of exceptions, but irregular negations are clearly not manner implicatures. For one thing, manner implicatures are always particularized, whereas irregular negations would be generalized if they were implicatures. The fact that synonyms cannot be freely substituted in irregular negations means, however, that they are detachable. If the regular interpretation of (1)(a)_n is its meaning, then what it says is that the sun is not larger than any planets. But saying that would not implicate that the sun is not larger than just some planets. Similarly, if the regular interpretation of (1)(f) is its meaning, then what it says is that John fails to believe there is a god. But that can be said without implicating that John believes there is no god.

The detachability of irregular negations is not a decisive objection to the implicature theory, in my view, because many generalized implicatures are detachable, as tautology implicatures illustrate (see Davis 1998: Sect. 5.3). But it does undermine one of the main advantages claimed for pragmatic ambiguity theories. Horn (1989: 365, 383) and Burton-Roberts (1989: 107ff) assert that implicature theories are more economical than semantic ambiguity theories because they need to postulate fewer senses. The other interpretations can allegedly be derived from the designated senses using conversational rules such as Grice’s Cooperative Principle and its maxims. Grice (1978: 47) originated this popular form of argument, calling it “Modified Ockham’s Razor.” To the extent

that irregular negations are detachable, this economy argument cannot be used. The other interpretations cannot be derived from the specified senses if other sentences have the same senses but lack those interpretations.

The main consideration showing that implicatures in general cannot be derived from the senses of sentences together with conversational principles is that there are always other possible implicatures compatible with those principles (Davis 1998). Grice's Cooperative Principle enjoins the speaker to *contribute what is required by the accepted purpose of the conversation*. The key premise of a Gricean derivation is that *the speaker could not have been observing the Cooperative Principle unless he believed (or implicated) the implicatum*.⁸ To pragmatically derive the irregular interpretation of (1)(a)_n from the regular interpretation, the Gricean must show that a competent speaker who utters (5) and says that the sun is not larger than any planets could not have been observing the Cooperative Principle unless he believed or implicated (5)(b), that the sun is not larger than just some planets. This must be true for all the varied contexts in which (5) could be used to mean or implicate (5)(b).

- (5) The sun is not larger than some planets.
- (a) Regular meaning: *The sun is not larger than any planets.*
Possible meanings or implicata:
 - (b) Irregular negation: *The sun is not larger than just some planets.*
 - (c) Unprecedented: *The sun is not larger than just one planet.*
 - (d) Unprecedented: *The sun is not larger than just half the planets.*
 - (e) Ironic regular negation: *The sun is larger than some/all planets.*
 - (f) Hyperbolic regular negation: *The sun is not larger than all planets.*
 - (g) Understated regular negation: *The sun is much smaller than any planet.*
 - (h) Contrary affirmation: *The sun is smaller than any planet.*
 - (i) Specifying implicature: *The sun is not even half as large as any planet.*
 - (j) Limiting implicature: *Other stars are larger than some planets.*
 - (k) Limiting implicature: *The sun is larger than some asteroids.*
 - (l) Ironic irregular negation: *The sun is larger than just some planets.*
 - (m) Metaphor: *The king is not more powerful than any subjects.*
 - (n) Metonymy: *The spacecraft headed for the sun is not larger than any planets.*

The “possible meanings or implicata” listed as (b)–(n) under (5) are all things a competent speaker could mean or implicate by using sentence (5) on a particular occasion. Most are instances of widely used types of implicature: irony, hyperbole, understatement, and so on. To establish that the irregular negation interpretation (b) can be derived from the Cooperative Principle together with the assumption that the semantics of English assigns the regular meaning to (5), the Gricean needs to show that the speaker using (5) would violate the Cooperative Principle if he

⁸ See Grice (1975: 31); Levinson (1983: 134–5); Horn (1992b: 260–2). Griceans sometimes weaken the premise to “It is unlikely that the speaker was observing the Cooperative Principle unless he believed (or implicated) I”; see Bach and Harnish (1979: 92–3); Levinson (1983: 115–6); Leech (1983: 30–44), 153. The same problems arise whether alternative implicatures are supposed to be impossible or improbable.

believed or implicated any of the possible interpretations other than (b). The metaphorical and metonymic interpretations can surely be ruled out given that our goal is to derive a *generalized* conversational implicature. A speaker using (5) metaphorically to mean that the king is not more powerful than any subjects would be obeying the Cooperative Principle only if the goals of the conversation happened to involve the king and his subjects. The utterance of (5) is not restricted to such contexts. The rest of the possible implicata, however, would be relevant in any context in which the regular meaning is. Without knowing the particular goals of the conversation in which (5) is uttered, and the astronomical beliefs of the speaker, there is no basis for the claim that the cooperative speaker would have to mean (b) rather than (c)–(l). The correction clause in (1)(a) would be an important contextual feature, providing a good reason to rule out (f)–(i). If the speaker believes what he says in the correction clause—the sun is larger than all planets, he cannot believe (f)–(i), and would violate the maxim of Quality. If the speaker were using (5) as an ironic regular negation, moreover, he would be unlikely to follow it immediately with a literal statement of what he just conveyed ironically; such a sequence would arguably violate the maxim of Manner. But the correction clause in (1)(a) would provide no reason to rule out (d), (e), (j), (k), or countless other possible implicatures (*The sun is not larger than just a few planets; Sirius is larger than some planets; Other yellow dwarf stars are larger than some planets; The sun is larger than some comets; etc.*). Nor would the mere presence of the correction clause rule out (l). For the correction clause itself could be ironic. Moreover, (5) has the irregular scalar-implicature denial interpretation, and can be used with that interpretation, even when there is no correction clause. The Cooperative Principle provides no more basis for predicting that (5)(b) is a generalized implicature of (5) than that (5)(c)–(k) are. It provides no explanation of why (5) has (b) as one of its readings but not (c)–(k). A speaker could use (5) to mean or implicate any of these things in particular contexts. But the sentence itself does not mean or implicate (c)–(k). The fact that (b) is special in this respect remains to be explained.

3 Neo-Gricean Pragmatics

Horn (1989, 2004) and Levinson (1983, 2000) use modified versions of Grice's maxims of Quantity and Relation as their principal explanatory and predictive tools.⁹ Consider the curious fact that sentences with very similar meanings like (6) and (7) are commonly used and interpreted differently.

⁹ See also Huang (2006).

(6)	He broke a finger (a) ✓ He broke his own finger (b) × He did not break his own finger.	“R-based Implicature”
(7)	He entered a house (a) × He entered his own house (b) ✓ He did not enter his own house	“Q-based Implicature”

A speaker uttering ‘He broke a finger’ would typically mean “He broke his own finger” rather than its negation, whereas a speaker uttering ‘He entered a house’ would typically mean “He did not enter his own house.” Horn takes (6) to have the meaning “He broke at least one finger,” and takes (6)(a) to be an implicature based on his R Principle: *Say no more than you must*. The idea is that the speaker does not need to say that he broke his own finger. It suffices for him to say that he broke at least one finger, because the context allows him to convey thereby the more specific information that he broke his own finger. Horn takes (7) to have the meaning “He entered at least one house,” and takes (7)(b) to be an implicature based on his Q Principle: *Say as much as you can*. The idea is that since the utterer of (7) does not say any more than that he entered at least one house, he must believe that he cannot make the more informative claim that he entered his own house.

An immediate problem for Horn is that if his Q principle correctly predicts (7)(b) it would also seem to incorrectly predict (6)(b). And if R correctly predicts (6)(a) it would also seem to incorrectly predict (7)(a). Horn claims that R “prevails” in one case while Q prevails in the other, but provides no explanation as to why or how. Applying Horn’s reasoning to the negation (5), R would seem to predict the specifying implicature (i) while Q would seem to predict the limiting implicatures (j) and (k). Neither principle predicts the specifying-implicature denial interpretation we are trying to explain.

In fact, the clash between R and Q prevents any prediction. Levinson (2000: 157–61) holds that Q takes priority over R, which obviously gives the wrong result for (6). In more complete formulations of their principles, Horn and Levinson avoid clashes by relativizing each principle to the other. R becomes *Say no more than you must given Q*, and Q becomes *Say as much as you can given R*. The fact that these relativized principles form a circle means that nothing can be predicted from them either. You cannot know what R predicts until you know what Q predicts, and vice versa.

4 Relevance Theory

Sperber and Wilson (1987: 699) observed that the derivations offered by Grice and his followers have an *ex post facto* character. While they may seem plausible when applied to implicatures we know in advance to exist, they would work just as well for all sorts of implicatures that are not found, and thus have no real explanatory or predictive power. Sperber and Wilson thought the defect could be remedied by replacing Grice's Cooperative Principle and maxims with their "Principle of Relevance."

We have proposed a definition of relevance and suggested what factors might be involved in assessments of degrees of relevance. We have also argued that all Grice's maxims can be replaced by a single principle of relevance—that the speaker tries to be as relevant as possible in the circumstances—which, when suitably elaborated, can handle the full range of data that Grice's maxims were designed to explain. (Wilson and Sperber 1986: 381)¹⁰

Wilson and Sperber stipulate that "the relevance of a proposition increases with the number of contextual implications it yields and decreases as the amount of processing needed to obtain them increases."¹¹ "Contextual implications" are defined as propositions that can be deduced from it together with a set of "contextual assumptions," and that cannot be deduced from it or the contextual assumptions alone. Contextual assumptions are items of background knowledge relevant to the conversation, including the propositions that have most recently been processed. Relevance is sometimes said to be directly related to the truth value and degree of confidence of contextual implications as well as their number.¹²

Very little can be predicted using this Principle. "Relevance" appears to be an efficiency ratio. But no specific function of output to effort is specified. Furthermore, the output variable normally cannot be determined: contextual assumptions are vaguely defined, and no unique set is determined by the context¹³; moreover, given the rich body of background knowledge relevant to everyday conversations, it is hard to see how the set of contextual implications could be finite.¹⁴ Degrees of

¹⁰ See also Wilson and Sperber (1981: 168–71), (2004: 612); Sperber and Wilson (1986: 381–2), (1987: 702–04), (1995: 258, 270); Kempson (1986: 89ff); Carston (1987), (1988: 42–4); Blakemore (1987: 54–71); (1992: 24–37); Yus (2006: 513–4). For critiques by a number of prominent authors, see *Behavioral and Brain Sciences*, **10**, 1987, 697–754, especially the articles by Bach and Harnish, Clark, Hinkelman, Levinson, Macnamara, Russell, and Seuren. See also Walker (1989), Levinson (1989), Davis (1998: Sect. 3.12), Bach (2010: 136).

¹¹ Wilson and Sperber (1986: 382). See also Wilson and Sperber (1981: 168–71); (2004: 609); Sperber and Wilson (1986: 125); (1987: 703), (1995: 265); Carston (1998: 336); Yus (2006: 514).

¹² Wilson and Sperber (1986: 385–6); Carston (1988: 42); Sperber and Wilson (1995: 263–6); Wilson and Sperber (2004: 608); Yus (2006: 514).

¹³ Sperber and Wilson (1986: 132–142); Kempson (1986: 90).

¹⁴ In some places, Sperber and Wilson (1986: 95–6, 1987: 702) try to prevent an infinite set of contextual implications by restricting the rules of inference to "elimination" rules. It is not clear that this effort succeeds (see e.g., Davis 1998: 102). But if it does, it calls into question the

confidence are non-denumerable, and no rule is specified for trading off differences in the various “positive cognitive effects.” The effort variable cannot be determined either: the cost depends on which context set the hearer accesses; and there is no accepted way of measuring processing cost. So there is no hope of calculating the most Relevant proposition, or even of determining that one is more Relevant than others. A separate problem is that a generalized implicature is found in too wide a variety of contexts to make it plausible that any such ratio was maximal or greater in every one.

Wilson and Sperber acknowledge that “relevance theory does not provide an absolute measure of mental effort or cognitive effect, and it does not assume that such a measure is available to the spontaneous workings of the mind” (2004: 626; see also 610). It suffices, they believe, to be able to compare the Relevance of two inputs. Even that is rarely possible, however, for two reasons. First, the comparative rules for both effect and effort have clauses requiring “other things” to be equal, which can rarely be assumed. Second, we are unable to make even comparative judgments of cognitive effect and processing effort.

The Principle of Relevance is presumably intended to imply that if two possible interpretations of an utterance are equally easy to process, then the stronger interpretation is correct. For the stronger proposition implies everything the weaker one does and more (see Wilson and Sperber 2004: 609). But it is not clear that the stronger proposition will have any additional “contextual implications” as defined, for background assumptions may entail that the weaker proposition is true only if the stronger one is. Moreover, both numbers may be infinite: if A and B both have an infinite number of implications, the fact that A has an implication B lacks does not entail that A has a greater number of implications.¹⁵ If this “corollary” is nonetheless applied to (5), it would seem to yield the prediction that speakers use (5) to mean (d) rather than (b). For (d) is stronger than (b), yet seems no harder to process; so (d) would seem to be more Relevant. When applied to the other possible meanings or implicata in (5), the special rule yields no prediction because (b) is neither stronger nor weaker. Note too that it is not evident that a stronger interpretation can have the same processing cost.

When justifying predictions, Relevance theorists generally use a “satisficing” rather than a “maximizing” principle. But a satisficing formulation cannot be used to derive specific implicatures. There could always be more than one implicata with “enough contextual effects to be worth his attention” that “puts him to no gratuitous processing effort” (Sperber and Wilson 1987: 743).¹⁶ Moreover,

(Footnote 14 continued)

importance of “Relevance” so defined. Why would people care about this oddly defined measure?

¹⁵ This is one of the paradoxes of infinity. To see the point, note that there are as many integers greater than five as there are greater than four, which is shown by the fact that the two sets can be put in a one-to-one correspondence.

¹⁶ See also Sperber and Wilson (1986: 158, 1987: 704, 1995: 270); Carston (1987: 714, 1998: 336, 2010: 247); Wilson and Sperber (2004: 611). Wilson and Sperber (2004: 614) claim that

determining that an implicata provides satisfactory Relevance also requires an absolute measure of Relevance.

Sperber and Wilson believe it follows from the Principle of Relevance that “the rational way to go about interpreting an utterance... is to follow a path of least effort and stop at the first interpretation that satisfies one’s expectation of relevance” (1995: 272).¹⁷ Again, this cannot be done without an absolute measure of Relevance. Furthermore, how is a person (or module) to know what the path of least effort is? Given an utterance of sentence (5), which of the over 87 billion orderings of (a) through (n) is the path of least effort? Sperber and Wilson seem to assume that the first interpretation that comes to mind is the easiest to process (e.g., 1995: 277). There is no reason to assume this; we often do things the hard way at first, and different things come to different hearers’ minds first.¹⁸ Sometimes the strangest things pop into our heads—like (m). Moreover, the Principle of Relevance provides no justification for choosing the path of least effort over another path if the extra effort is compensated by enough extra effects (Morgan and Green 1987: 727). Finally, there is no reason to think that an interpretation satisfying a particular hearer is one intended by the speaker or possessed by the sentence, even if there happens to be one and only one hearer.

We are evaluating the claim that implicature theories of irregular negations are more economical than semantic ambiguity theories because the irregular interpretations can be derived from the designated senses using conversational rules. I have argued that the Principle of Relevance provides no reason to select the irregular interpretation rather than other possible implicatures. An equally critical problem is that the Principle provides no reason to think that a speaker ever uses a negation with anything other than its regular interpretation. The Principle of Relevance implies that (5) is used to mean something other than its literal meaning (a) (*The sun is not larger than any planets*), namely (b) (*The sun is not larger than just some planets*), only if (b) has a larger number contextual implications per unit

(Footnote 16 continued)

“there should be no more than one” interpretation with a satisfactory degree of Relevance because “an utterance with two apparently satisfactory competing interpretations would cause the hearer the unnecessary extra effort of choosing between them.” But such extra effort is not unnecessary if, as is usually the case, it is impossible for the speaker to prevent the existence of such competing interpretations.

¹⁷ See also Sperber and Wilson (1986: 168–9, 1987: 705); Carston (1998: 341, 2010: 218); Wilson and Sperber (2004: 613, 626); Yus (2006: 514). Sometimes they say “order of accessibility” rather than “path of least resistance.” Bach’s (2010: 130) gloss is significantly different: “consider hypotheses about what the speaker means in the order in which they occur to you—how else?—and... stop as soon as a sufficiently plausible one comes to mind.” ‘Relevant’ does not mean “plausible.” Note that we often reject what we initially take to be plausible for an interpretation that is more plausible.

¹⁸ Relevance theorists sometimes assume that inferring an implicature takes less effort than “decoding” what the speaker said (e.g., Carston 1998: 337–8). Sometimes they assume the opposite (e.g., Wilson and Sperber 1986: 383). There is no basis for either assumption given our inability to measure processing effort.

processing cost. But for the reasons indicated above, there is no basis for predicting that (b) has more contextual implications than (a) in any let alone all of the wide variety of contexts in which (5) is used to mean (b), nor that if (b) does have more contextual implications, then the increase is more than proportional to any increase in processing cost.

Given these problems, it is not surprising that the derivations offered by Relevance theorists have the same *ex post facto* character Relevance theorists decried in Gricean derivations.¹⁹ In response to such objections, Sperber and Wilson (1995: 266) remark that the problems are not unique to Relevance theory, but common in cognitive psychology. That may be, but it does not excuse unfounded claims that a theory can account for pragmatic phenomena when it cannot.

5 Presupposition-Canceling Denials

Despite the fact the presupposition-canceling interpretation of a negation like (1)(c)_n ‘Vulcan is not hot’ is “marked” and “exceptional,” Horn (1989: 486–90, 514) follows Grice (1981) in taking it to be the semantic meaning of the negation. The “regular” interpretation, with its existential presupposition, is a pragmatically derived implicature. However, when we use ‘Pluto is not hot’ as a regular negation, Horn’s theory wrongly predicts that we do not assert that Pluto is non-hot, that what we say does not entail that Pluto exists, and that ‘Pluto is not hot: it does not exist’ should not seem literally contradictory. Furthermore, it is hard to see how ‘Pluto is not chilly’ could be interpreted as a scalar-implicature denial if it had the same meaning as ‘The proposition that Pluto is not chilly is not true’ or ‘Nothing is both Pluto and chilly,’ which have no such interpretation.

Carston (1998: 325; 1999: 368) faults Horn for not saying more about how other interpretations of a negation are “pragmatically derived” from the one that is supposed to be semantically encoded. Burton-Roberts (1999: 357) pointed out a large lacuna in a representative derivation of the existential presupposition of a regular negation from the presupposition-canceling interpretation.²⁰

The Gricean derivation goes something like this. Assume, as Griceans do, a semantics for (8) that amounts to the disjunction of (9) and (10)

- (8) The king of France is not bald.
- (9) There is no king of France.
- (10) There is a king of France and he’s not bald.

¹⁹ See for example Sperber and Wilson (1995: 273, 278); Carston (1996: 312–5, 1998: 340); and Sect. 5 below.

²⁰ See e.g., Wilson (1975: 99–100, 106); Atlas (1975, 1979: 273); Boër and Lycan (1976: 27–8) (but contrast 60–1); Levinson (1983: 218, 222); Lycan (1984: 84). See also Kempson (1975: 178–9). Grice’s (1981: 273–6) own derivation was even less successful.

The idea is that (8) would be an obscure, vague, non-optimal, way of conveying (9). Since (9) entails (8) but not vice versa, it is more specific. So to communicate (9) the speaker would have to SAY (9), not the much vaguer (8). So if (8) is uttered but not in order to convey (9), it must be being uttered in order to convey (10). That there is a king of France, then, is conveyed as a quantity implicature.

The difficulty is that the calculation can be shown, with equal or greater plausibility, to go exactly the opposite way. (10) entails (8) as much as (9) [if the Gricean semantics is correct]. Speakers are expected to give as much information as is compatible with their beliefs. Hence the speaker who believes that (10) is true should SAY (10) rather than the much vaguer (8). (Burton-Roberts 1999: 357; my numbering)

If the Gricean derivation worked, moreover, it could just as well be used to show that the disjunction of (9) and (10) itself implicates that there is a king of France; but it does not. Indeed, Burton-Roberts (1999: 357) went on to observe that ‘Either there is no king of France, or there is and he’s not bald’ implicates “I do not know which alternative is true”; (8) has no such implicature.

Boër and Lycan (1976: 60–61) would reply that the speaker could not have uttered (8) in order to convey (9), since her mention of baldness would then violate the maxim of Relation. But Grice’s central idea was that implicatures enable us to conform to the Cooperative Principle and its maxims even when what we say by itself does not. Moreover, their explanation fails completely in the most natural case, in which the speaker is responding to someone who previously affirmed the root of (8). Atlas (1979: 273) would suggest that (10) is the most informative (and relevant) additional claim one could make. But since (9) and (10) are logically independent, their relative informativeness and relevance cannot be compared except in very special contexts. And if Atlas’s claim were true, it would seem to preclude a presupposition-canceling interpretation of (8).

Carston (1999: 368) claims that her Relevance-theoretic account provides sound derivations, using (11)_n as her example.

(11) The king of France is not bald: there is no king of France

Schematically, Carston represents the interpretation of negations like (11)_n as follows (1998: 340):

(12)	semantically:	<i>not</i> [<i>The F is G</i>]
	via pragmatic processing (a):	[<i>The F is not-G</i>]
	via pragmatic processing (b):	<i>not</i> [“ <i>The F is G</i> ”]

This means that when a speaker says (11), the hearer first linguistically decodes ‘The king of France is not bald,’ arriving at the wide-scope, presupposition-canceling interpretation, which she represents *not*[*The king of France is bald*]. Next, the hearer applies pragmatic processing, arriving at the narrow-scope, presupposition-preserving interpretation [*The king of France is not-bald*]. Finally, since the second step generates a contradiction when the correction clause of (11) is processed, the hearer applies a second round of pragmatic processing to arrive at the

“meta-representational” or “echoic” interpretation which she represents *not*[“*The king of France is bald*”]. Carson (1998: 310) says that this has the approximate meaning “‘The king of France is bald’ is inappropriate/unassertable”. Optionally, the third step may take the hearer back to *not*[*The king of France is bald*].²¹

Let’s take example (11)_n again, thinking of it now in communicative rather than semantic terms, as an utterance produced and processed over time... The semantics is the wide-scope negation which has no entailments (that is, it cancels so-called presuppositions). Since this is too weak/uninformative, in most contexts, to meet the criterion of optimal relevance, the scope of the negation is narrowed so as to achieve sufficient cognitive effects. The negation could, in principle, be taken to target the existential entailment, but processing effort considerations mediate strongly against this, since it would leave the predicate ‘is bald’ no role to play in the interpretation; the concept will have been activated pointlessly, since it will not enter into the derivation of effects. So the preferred interpretation, preferred because of its relevance-potential, is the narrow scope, ‘presupposition’-preserving, interpretation. This is just one of many instances of the process of pragmatic strengthening at the level of the proposition expressed. We would be bound to leave it at that, were it not that some milliseconds later, when we have processed the next (juxtaposed) utterance, we find ourselves with a contradiction: there is a king of France and there isn’t a king of France. The overall interpretation of the two clauses is not consistent with the expectation of optimal relevance and a reanalysis is sought. This may be a move to an echoic (metalinguistic) analysis, as in B-R’s account... [or] it might be a ‘return’, as it were, to the descriptive wide-scope, ‘presupposition’-cancelling interpretation. (Carston 1998: 340, see also 1996: 318–9, 1999: 368ff, 376)

Carston thus denies that (11) is semantically contradictory, while maintaining that Relevance theory accounts for both “the initial P-preserving interpretation,” and for the unacceptability of the resulting contradiction. “Contradictions give rise to no cognitive effects” (Carston 1998: 340).²²

A contradiction, not derived solely through linguistic decoding, does play its part in my account as in B-R’s; the difference is that in mine it is underpinned by a pragmatic theory, built around the central definition of optimal relevance, which accounts for why a contradiction is unacceptable. The very same pragmatic theory accounts for the initial P-preserving interpretation, without having to invoke any other principles to do the work. (Carston 1999: 375ff)

Carston’s account will not survive critical scrutiny. If her derivation of multiple interpretations for (11)_n from the presupposition-canceling interpretation were sound, it would apply just as well to (13)_n.

²¹ Carston seems to be assuming a Russellian interpretation of definite descriptions, although she makes no such claim. She does imply that the wide-scope presupposition-canceling interpretation involves “Predicate denial (or sentence negation)” rather than “term (constituent) negation” (cf. Horn 1989: 107). But it is hard to see how one can deny the predicate ‘is bald’ of the king of France without presupposing that France has a king. Horn generally does not distinguish the presupposition-canceling interpretation from the metalinguistic. See for example Horn (1989: 374–5, 1992a: 265); but contrast Horn (1989: 488–9, 1990: 496, 500).

²² This is clearly true in standard logic, in which a contradiction entails every proposition all by itself; hence nothing is entailed by the conjunction of a contradiction and the set of contextual assumptions that is not entailed by either alone. The claim may not be true in a relevance logic.

- (13) It was not the king of France who kissed the Canadian empress: Canada does not have an empress.

But (13) has no consistent reading. So there must be something wrong with Carston's derivations. Indeed, defects abound.

First, the claim that the wide-scope, presupposition-canceling interpretation has no entailments at all is obviously false if 'entailment' has its usual meaning. If this were true for any reason, it would seem to follow absurdly from the Principle of Relevance that speakers never try to communicate wide-scope negations; for they would have zero Relevance. Provided the first-pass interpretation has any entailments at all, the claim that it is too uninformative in most contexts to meet the criterion of optimal Relevance is indefensible. In addition to the general reasons noted in Sect. 4, why should anyone think that wide-scope negations as a class are not very informative? 'It is not the case that the unique current British monarch is a dictator' is highly informative even on the wide-scope Russellian interpretation. If Carston's uninformativeness claim is nevertheless accepted, then her assertion that the final interpretation could optionally be a return to this insufficiently informative interpretation is undermined.²³

Second, the concept "is bald" was not activated pointlessly. The fact that [*There is no king of France*] entails *not*[*The king of France is bald*] independently of the fact that the predicate concept in the latter is "bald" does not mean that activating that concept was pointless. Among other things, the hearer could not recognize that *not*[*The king of France is bald*] has any logical properties without conceiving that proposition and all its component concepts. Furthermore, the hearer is trying to determine what the speaker (or his utterance) meant. In order to determine that the speaker intended any of the three interpretations indicated above, the hearer must activate the concept "bald." Most importantly, "Relevance" in Relevance theory is defined as the ratio of contextual effects to processing cost. The utterance of a sentence containing 'bald' will typically have countless contextual effects containing the concept "bald," whether or not those effects contribute to the purposes of the conversation. 'Relevant' for Relevance theorists does not have the conventional sense "related to the purpose at hand," which is what Grice's maxim of Relation concerns.

Third, if as Carston has plausibly claimed, a contradiction has zero contextual effects, then the ratio of contextual effects to processing costs will consequently be zero, the lowest possible. If the Principle of Relevance is the only principle of interpretation, this would mean that the correct interpretation of a sentence or passage can never assign it a contradictory meaning, which is absurd. People can

²³ Carston's claim that the process of reinterpretation might return to the descriptive wide-scope interpretation also appears to contradict her claim in the next paragraph that "The semantic level differs from the two pragmatic levels in that no final interpretation will ever involve it alone..." (1998: 341) This complements the contradiction in her claim that marked negations are "essentially" but not "absolutely" metalinguistic (echoic, metarepresentational) (1996: 319, 320; 1998: 335ff, 1999: 380).

and do contradict themselves. A sentence like (13) only has a contradictory interpretation. In the case at hand, it would mean that pragmatic processing using the Principle of Relevance cannot yield the contradictory interpretation of a negation-correction conjunction *on any pass*. If the negation clause is processed first with a presupposition-preserving interpretation, then the literal interpretation of the correction clause that would result in a contradiction will be blocked. The hearer would be forced to give the correction clause a figurative interpretation (attributing irony perhaps), or give it a different semantic analysis (perhaps taking ‘king’ there to mean “preeminent person” instead of “male monarch”). If the correction clause is processed first with its literal interpretation, then the presupposition-preserving interpretation of the negation that would contradict it will be blocked.

Fourth, Carston’s single principle can account for two interpretations only by focusing on a sequence of cognitive states of the hearer as they change over time because the context changes. It gives us no insight into why we in our current linguistic context can hear sentences like (1)(c)_n and (8) as having two different interpretations, or why speakers can properly use such sentences to mean either of two different things in any given context.

6 Pragmatic Explicature Theories

Carston agrees with Horn that the presupposition-preserving interpretation of negations is pragmatically derived from the presupposition-canceling interpretation, but denies that it is an implicature. Carston (1996: 316) explains how her theory differs from Horn’s in connection with examples like (14).

(14) Mike didn’t eat *three* cakes, he ate *four*.

Horn theorized that a numeral ‘n’ means “at least n,” and implicates “exactly n.” On Horn’s view, what the speaker says is what the sentence means. So the speaker used the negation in (14) to *say* that Mike did not eat at least three cakes, while *meaning* that Mike did not eat exactly three cakes. That Mike did not eat exactly three cakes is thus an “*implicature*” (Grice 1975: 24–5): *something meant by saying something else*, namely that he did not eat at least three cakes.²⁴ As Horn analyzes (14), it appears to be a standard scalar (quantity, limiting) implicature denial followed by a correction clause. Horn is thus advocating an I-implicature theory for these irregular negations.

Following Sperber and Wilson (1986: 182–3), Carston (1988: 33) uses “*explicature*” as the correlative technical term for *what is said* or *directly meant*. A speaker therefore implicates something by “*explicating*” something else.

²⁴ See also Grice (1969: 87); Davis (2007). Compare and contrast Saul (2001), (2001).

Carston (1988: 35, 46–7; 1996: 315–6) maintains that the difference between the two interpretations of (14)_n is a difference in explicature. That is, she maintains that the speaker uses the negation in (14) to say (or directly mean) rather than implicate that Mike did not eat exactly three cakes. Hence what the speaker said is true if Mike ate four cakes. Carston agrees with Horn that ‘three’ does not mean “exactly three” in (14), while also denying that it means “at least three.” Carston maintains that ‘three’ has the same meaning in (15)(a) and (b). She might observe that (15)(a) would be a gross non-sequitur if ‘three’ meant “at least three” while (15)(b) would be a Moorean paradox if ‘three’ meant “exactly three.”

- (15) (a) Elaine ate three cakes, so she ate one less than four.
 (b) Elaine ate three cakes, and perhaps four.

In Carston’s view, the speaker of (15)(a) nevertheless said that Elaine ate exactly three cakes, while the speaker of (b) said that she ate at least three. Carston maintains that (14) and (15) both involve “pragmatic enrichment” or “development”: saying something more complex than what the sentence one uses means. They involve enrichment that is “free” in the sense of being linguistically optional and uncontrolled. According to Carston’s “semantic underdetermination” thesis, a great deal of what is said is pragmatically inferred rather than linguistically encoded.²⁵

Let us say that an explicature is *semantic* if it is what the sentence used means, and *pragmatic* otherwise. An explicature is semantic when what is said is “linguistically encoded.” An explicature that is not linguistically encoded is pragmatic. If a speaker uses ‘2 > 1’ to say that 2 is greater than 1, his explicature is semantic: ‘2 > 1’ means “2 is greater than 1.” If a speaker uses ‘He is a Democrat’ to say that Obama is a Democrat, then his explicature is pragmatic. ‘He is a Democrat’ does not mean “Obama is a Democrat,” even on this occasion. This pragmatic explicature arose from the use of an indexical. The speaker may pragmatically explicate the same thing by giving the elliptical answer “Obama” to the question “Who is a Democrat?” Then the speaker said and meant that Obama is a Democrat, but the word ‘Obama’ does not mean “Obama is a Democrat.” Whether pragmatic explicature results only from indexicality or ellipsis, or sometimes from “free enrichment,” is an open question.

In the case of (14), Horn maintains that “Mike didn’t eat at least three cases” is the explicature—a semantic one; Carston thinks that proposition is neither an explicature nor an implicature. Carston maintains that “Mike didn’t eat exactly three cases” is a pragmatic explicature, whereas Horn takes it to be an implicature.

Applied to the difference between regular and irregular negations, the *pragmatic explicature theory* maintains that it is a difference in what is said or directly meant without a difference in sentence meaning. Speakers who use the sentences in (1) with their irregular interpretations seem clearly to be saying something

²⁵ See Atlas (1977, 1977, 1977); Sperber and Wilson (1986: 182); Carston (1988: 41, 2004b: 66, 2010: 218); Recanati (1989, 2004); Bach (1994); Bezuidenhout (2002: 274).

consistent, even though they could use the same sentences to make contradictory statements. Similarly, an expert witness may use (16) to assert under oath that in her opinion, drugs were not a factor.

(16) I do not believe drugs were a factor in his death.

She could use the same sentence to declare that she merely lacks an opinion on the matter. So the pragmatic explicature theory shares one advantage of the ambiguity theory over the implicature theory. The pragmatic explicature and ambiguity theories both maintain that the regular and irregular interpretations of negations are things the speakers say (assert, declare) and mean directly. They differ in whether the sentences used have just one meaning or more than one. Pragmatic explicature and implicature theories are alike in maintaining that negations are unambiguous; they differ over whether the non-semantic interpretation is said rather than implicated—directly rather than indirectly expressed.

We can distinguish three types of pragmatic explicature theory depending on whether the meaning of a negation is taken to be the regular interpretation, one of the irregular interpretations, or neither. The “I-explicature” theory holds that irregular interpretations are pragmatic explicatures. The “R-explicature” theory holds that the regular interpretation is a pragmatic explicature. The “radical explicature” theory holds that both regular and irregular interpretations are pragmatic explicatures. These alternatives are set out schematically in Table 2, enabling comparison with the implicature theories in Table 1.

No one, to my knowledge, holds the I-explicature theory. We will see in Sect. 10 that Atlas (1989: 145) holds the radical explicature theory, as does Carston for scalar-implicature denials involving cardinal numbers like (14). Carston holds an R-explicature theory, however, for presupposition-canceling denials. As explained in Sect. 5, she holds that the semantic meaning of a negation is the presupposition-canceling interpretation she represents *not[s is P]*; the regular interpretation, and the other irregular interpretations given to implicature denials and contrary affirmations, are “pragmatic enrichments” of *not[s is P]*. Since Carston shares Horn’s view that the presupposition-canceling interpretation is the only meaning of a negation, her view shares one of the most serious problems reviewed in Sect. 5: some negations, such as (13), have no such interpretation. Moreover, it is hard to see how scalar-implicature denials like (1)(a)_n or ‘Pluto is not chilly, it’s frigid’ could exist if negations had the semantics Carston assigns to account for presupposition-canceling denials. Carston can account for the different interpretations of (14) only because of her theory of number-term semantics.

Table 2 Pragmatic explicature theories

	I-explicature theory	R-explicature theory	Radical explicature theory
Sentence meaning	N means R but not I	N means I but not R	N means neither I nor R
Irregular use	S says and means I	S says and means I	S says and means I
Regular use	S says and means R	S says and means R	S says and means R

A problem for both Horn and Carston is that (14)_n seems to have two distinct *regular* interpretations. On the one both recognize, (14) would be interpreted as making the contradictory claim “Mike didn’t eat at least three cakes, he ate four.” But it is also possible to interpret (14) as a regular negation saying “Mike didn’t eat exactly three cakes, he ate four.” On this interpretation, (14) is just like (3); this would be the natural interpretation if the speaker and hearer were trying to identify the number of cakes each person ate—that is, if they were counting. This interpretation would be very natural even with no emphasis on the number words and no fall-rise intonation. The problem for both Horn and Carston is to explain what the difference between the regular and irregular interpretations consists in when the speaker means “Mike didn’t eat exactly three cakes, he ate four” in both cases, if (14) has the same meaning. Since this problem arises from their view about the semantics of cardinal-number terms rather than negations, we will not pursue it.

I have glossed ‘explicature’ to mean “what is said or directly meant.” There are two problems with this technical term. First, there is unclarity in the definiens because ‘said’ is used in a variety of ways (see below) and is not completely equivalent to ‘directly meant.’ Second, markedly different definitions are often given or suggested, sometimes by the same author in the same article.²⁶

- (17) What is said (Carston 1988: 33; 2004b: 65).
- (18) The proposition expressed (Carston 1988: 33; 1996: 316; 2004b: 65, 81).
- (19) The truth-conditional content of the utterance (Carston 1996: 316, 2004b: 81, 2010: 243, 244).
- (20) The explicitly communicated proposition (Carston 1988: 41; 2004b: 67; 2010: 218).
- (21) The directly communicated proposition (Carston 2010: 227).
- (22) The primary speaker meaning (Carston 2010: 270)

These descriptions differ in meaning and often in reference. (18) is defective as a definition of explicature because implicating is also a way of expressing a proposition. In Grice’s familiar gasoline example, the speaker who utters ‘There is a station around the corner’ expresses and communicates both the proposition that there is a gasoline station around the corner and the proposition that his addressee can get gasoline there. The second proposition is something the speaker implicates. Definiens (19) is undefined when the speaker expresses more than one proposition: we cannot tell which proposition it applies to in that case. In the gasoline example, we know that what the speaker *said* is true iff there is a station around the corner, whether the addressee can get gas there or not. But ‘the truth conditions of the utterance’ is undefined when the utterer has expressed more than one proposition.

Definiens (20) depends on ‘explicit.’ On a conventional understanding (see Sect. 8), something is communicated explicitly only if it is expressed by the

²⁶ See also Wilson and Sperber (1981: 159, 2004: 615, fn. 18); Sperber and Wilson (1995: 256–8); Carston (1999: 373, 2002: 124, 2004a: 633, 635, 648, 2010: 265, 266, 269). Carston appears to officially reject (19) (see 2010: 258, 234, 269–70) as well as (2004a: 634); see also Wilson and Sperber (1981: 159) and Sperber and Wilson (1995: 257–8).

sentence used: speaker meaning and sentence meaning coincide. On this understanding of (20), there could be no pragmatic explicature. In Relevance theory, however, ‘explicit’ is a technical term. Sperber and Wilson (1986: 182) stipulate that “An assumption communicated by an utterance U is *explicit* if and only if it is a development of a logical form encoded by U.”²⁷ They add that the more of the assumption that is encoded, the more explicitly it is communicated. Sperber and Wilson define an *implicature* as “any assumption communicated, but not explicitly so.” As with the other definitia, the uniqueness presupposition of (20) will typically be false (cf. Carston 2002: 169; Chaves 2010: 117). More importantly, the difference between explicatures and implicatures given these definitions does not seem to match any distinction linguists, philosophers, or ordinary speakers have been interested in.

- (23) There is a station around the corner.
 (a) The addressee can get gasoline at the station around the corner.
 (b) There is a station where the addressee can get gasoline around the corner.

In the gasoline example, Sperber and Wilson’s definitions rule that (23)(a) is an implicature because it is not a development of the logical form of (23). But (23)(b) is an explicature, not an implicature, because it is a development of that logical form. The concept *station* has been “narrowed” to *station where the addressee can get gasoline* (Carston 2004b: 76ff). Yet (23)(a) and (b) are both things the speaker meant by saying that there is a station around the corner. Both are cases of indirect meaning and implicature as Grice defined it. In the standard quantity implicature case, *The sun is larger than just some planets* would count as an explicature when a speaker says ‘The sun is larger than some planets,’ even though *The sun is not larger than all planets* is an implicature. Carston (2002: 140–1) herself argued that *B has invited a man to the wedding* is an implicature of ‘B invited his brother to the wedding’ (Romero and Soria 2010: 9), which it definitely is in Grice’s sense. But that proposition is a “development” of the logical form expressed by the sentence uttered: the concept *a man* is a “broadening” of *his brother*. So for Carston it should count as an explicature. Carston (1988: 48; 2002: 191) and Chaves (2010: 120) claim that whether a given proposition counts as an explicature or implicature is determined by the Principle of Relevance. But no conversational principle can determine whether a proposition is or is not a development of the logical form of the sentence uttered. For that depends on a structural or conceptual relationship.²⁸

²⁷ There is one difference in Carston’s (2002: 116, 124) definition, but it will not matter here.

²⁸ Carston (2002: 118) offered a reply to Levinson’s (1987: 723) presentation of this objection: “This implicature is not derived by a process of pragmatically developing the decoded content of B’s utterance; plainly, it is derived purely inferentially, by a straightforward deductive inference, one of whose premises is the assumption which *is* derived by development of the encoded content...” First, it is not clear what “development” means when it denotes a process rather than a structural relationship, nor whose performance of that process would be relevant to whether something B expressed counts as an implicature. Second, Carston seems here to be contradicting the Relevance theorist’s claim, separating them from Griceans, that explicatures and implicatures

Both (20) and (21) differ from ‘what is said’ and (18) in requiring the proposition to be *communicated* and not just *expressed*. If (20) or (21) define ‘explicature,’ there is no explicature unless some hearer understands the speaker and recognizes what proposition the speaker expressed. If we combine elements of (19) and (21), however, we can formulate an alternative description that is close in meaning to ‘what is said’: *the proposition directly expressed*. This is equivalent to (22) *the primary speaker meaning* if ‘primary’ means “direct” rather than “most important,” and we restrict attention to propositional meaning. In the gasoline example, the speaker expresses the proposition (and means) that the addressee can get gasoline at the station around the corner *indirectly*, by expressing the proposition (and meaning) that there is a station around the corner.²⁹ What he said is the proposition he expressed directly: that there is a station around the corner. In the case we are interested in, irregular negations, the main problem for implicature theories is that there is no indirection. In example (1)(a), the speaker expresses the proposition that the sun is not larger than just some planets directly, not by expressing the proposition that the sun is not larger than any planets. There are cases in which what is directly expressed is not said, as when someone answers a question by nodding his head, or uses a nonce word. There are also cases in which what is said is not something the speaker meant, as when there is a slip of the tongue.³⁰ But these cases of divergence are not relevant here. For our purposes, the pragmatic explicature theory can be defined equally well as the thesis that *the irregular interpretation of a negation is directly expressed/meant but not linguistically encoded*. In example (1)(a)_n, the speaker directly expresses the proposition that the sun is not larger than just some planets. The pragmatic explicature theory accepts that while denying that (1)(a)_n has “The sun is not larger than just some planets” as one of its meanings.³¹

7 Pragmatic Explicature Versus Implicature

Pragmatic explicature theorists must show that negations have more than one interpretation because of pragmatic explicatures rather than either implicatures or ambiguity—that is, because what is said and expressed directly may differ from

(Footnote 28 continued)

are both derived by pragmatic inferences guided by the Principle of Relevance (see e.g., Carston 1988: 37; 2004a: 633, 636, 643; 2010: 242–3, 246–7, 270; Wilson and Sperber 2004: 615).

²⁹ I identify and define different types of speaker meaning in terms of expression in Davis (2003).

³⁰ Grice was idiosyncratic in taking ‘S says that p’ to entail ‘S means that p.’ So he was forced in the case of metaphorical usage to say that the speaker “makes as if to say,” and would deny that the speaker even makes as if to say in the case of verbal slips. Cf. Neale (1992: 523–4, 549); Bach (2001: 17, 2010: 134); Davis (2007); Carston (2010: 220).

³¹ Neither ‘what is said’ nor ‘what is directly meant’ fits Carston’s use of explicature given that she allows “higher order” explicatures. See e.g., Carston (2010: 223).

what the sentence used means on that occasion. We will focus on implicature in this section, and ambiguity in [Sects. 10](#) and [11](#).

Carston's (1988: §§2–3) general argument against implicature theories begins with the Gricean premise that calculability is an essential property of implicature. She then infers that “the explicature is distinct from the implicatures of the utterance; they do not overlap in content.” From this she infers further that an implicature cannot entail the explicature, “since otherwise the explicature... is redundant, playing no independent role in inference” or “mental life.” She therefore rejects Horn's view that a speaker who utters ‘He ate three cakes’ says that he ate at least three cakes and implicates that he ate exactly three on the grounds that the latter entails the former. Since “He ate exactly three” is not an implicature, it must be the explicature. “[W]hat is said (the explicit) and what is implicated (the implicit) exhaust the (propositional) significance of the utterance...” (1988: 33) If Carston wants to apply this reasoning to (14) too, she will have to infer from her non-overlap premise that the explicature cannot entail the implicature either. For negation reverses the order of entailment: “Mike did not eat at least three cakes” entails “Mike did not eat exactly three” rather than vice versa. Of all the irregular negations we have identified, only the two interpretations of metalinguistic-implicature and evaluative-implicature denials fail to “overlap” each other.

Carston's general argument is far from compelling. (i) Calculability, as we have seen, is not even a typical property of implicatures. This is true whether implicatures are to be derived from Grice's Cooperative Principle and Maxims (§2), the Neo-Gricean principles of Horn or Levinson (§3), or Sperber and Wilson's Principle of Relevance (§4). Implicatures are like explicatures in having to be inferred, but neither are always discoverable, and when they are we are not restricted to using a particular reasoning process. (ii) An implicature is something meant by saying something else. So an implicature must be *distinct* from what is said, but that does not mean that it must *not overlap with* or *entail* what is said in any way. The gasoline example (23)(b) is a standard example in which what is implicated overlaps with and entails what is said. All specifying implicatures and understatements are like this. Carston's non-overlap assumption may come from her defining the explicature/implicature distinction in terms of Sperber and Wilson's explicit/implicit distinction. But as we saw in [Sect. 7](#), that distinction does not track anything of interest. (iii) The assumption that implicatures are inferred by a non-deductive process from explicatures would not mean that implicatures cannot entail explicatures. Enumerative induction is a classic example of a non-deductive inference in which the conclusion entails the premise. The premise is not redundant because without it we would have no evidence for the conclusion.³² More

³² Another reason Carston (1988: 38) gives for her “functional independence” rule is that if the implicature entails the explicature, the latter need not be stored in memory too, “since all the information given by the latter is also given by the former.” This falsely assumes that a person is aware of all the information entailed by a proposition as soon as that proposition is recalled, and that adding to total knowledge is the only goal of our mental life.

importantly, what needs to be inferred is not the proposition implicated from the proposition said (explicated), but the fact that the speaker implicated the former from the fact that the speaker said the latter.³³ What needs to be inferred in the petrol example is not “There’s a station where you can get gas around the corner” from “There is a station around the corner,” but “The speaker meant that there is a station where you can get gas around the corner” from “The speaker said that there is a station around the corner.” Such inferences will always be inductive. Carston’s general argument provides no good reason, therefore, to prefer her radical explicature theory of limiting-implicature denials of numerical claims or her R-explicature theory of presupposition-canceling negations over implicature theories.

Let us look at one of Carston’s (1988: 40) paradigm examples of explicature, to see if it might nonetheless provide a better model for irregular negations than implicature theory can offer.

(24) A ran to the edge of the cliff and jumped.

In a typical context of utterance, I take the following to be uncontroversial

(25) S said and meant that A jumped. ✓

(26) S meant that A jumped off the cliff. ✓

(27) Sentence (24) does not mean or entail that A jumped off the cliff. ✓

One thing in dispute is:

(28) S said that A jumped off the cliff. ?

Carston maintains that (28) is true—that S explicated rather than implicated that the man jumped off the cliff. What S said, on Carston’s view, is a pragmatic enrichment of what the sentence she used means.

One reason Carston (1988: 40) offers for (28), criticized above, is that S cannot implicate “A jumped off” by saying “A jumped” because the former entails the latter. Another reason she cites is that what the hearer will remember from the utterance is the message that the man jumped off the cliff. This is completely compatible, however, with that message being an implicature. Hearers are often more interested in what is implicated than in what is said. Suppose Ted asks Monique, whom he just picked up in a bar, “What would you like to do now?” If she answers “Test your sheets,” I’d suspect Ted would be much more likely to remember what Monique implied than what she literally said.

Following Wilson and Sperber (1981: 159), Carston might say more specifically that what S implicates (e.g., “A died”) will be worked out from “A jumped off” rather than “A jumped.” But this begs the question in assuming that “A jumped off” is not (also) an implicature worked out from “A jumped.”

³³ For a more recent example of this act-object confusion, see (Carston 2010: 225).

A third argument might be based on characterizing what is said or explicated as “the truth-conditional content of the utterance”—the proposition “p” that is true iff the utterance is true. As we saw in Sect. 6, this characterization is problematic because in any utterance, the speaker typically expresses a number of different propositions (see e.g., Wilson and Sperber 2004: 623), and the sentence uttered may express yet another proposition. Even if we characterize what is said as the referent if “*Is that true?*” were asked after the utterance, ‘that’ could refer to *any* of the propositions expressed by the speaker or the sentence, depending on what was most salient to the person asking the question. If Ted asked “*Is that true?*” after Monique uttered ‘Test your sheets,’ he would in all likelihood have been referring to the proposition that she wants to have sex with him, not the proposition that she wants to subject his sheets to a test. So the fact that after (24) is uttered, we would take “*Is that true?*” to be asking whether A jumped off the cliff does nothing to establish that “A jumped off the cliff” was an explicature rather than an implicature.

Probably the most influential reason Carston (1988: 44–46) offers for the truth of (28) is that it allegedly explains why the proposition that A ran to the edge of the cliff and jumped off appears to be expressed when (24) embeds in conditionals and other compounds.³⁴ Consider:

- (29) (a) If A ran to the edge of the cliff and jumped, he is probably dead.
 (b) If A ran to the edge of the cliff and jumped off, he is probably dead.

Someone who used (29)(a) in all likelihood meant what (29)(b) means. If “A jumped off” is a mere implicature of (24), Carston believes, it would not be part of the antecedent of the conditional proposition expressed by uttering (29)(a). Carston might cite as evidence here that even though ‘She dresses neatly’ can be used to implicate “She is a lousy philosopher,” speakers cannot use ‘If she dresses neatly, then we should not hire her’ to express a proposition whose antecedent contains that implicature.

This hypothetical induction is problematic. First, an alternative explanation is ready to hand: generalized conversational implicatures embed in the way illustrated while particularized conversational implicatures do not. This can be verified with scalar implicatures: ‘If some students got As, then some didn’t’ can be used to express the logical truth ‘If some but not all students got As, then some didn’t,’ whose antecedent contains the upper-bound implicature ‘Not all students got As.’ We need to be clear, though, on what sort of “embedding” is illustrated by (29)(a). Quite correctly, Carston is not maintaining that the proposition expressed by the *sentence* (29)(a) is that expressed by (29)(b). It is easy to imagine cases in which sentence (29)(b) is true while (29)(a) is false. The sentences do not have the same truth conditions. There is no “pragmatic intrusion” here (Levinson 2000: 198). It

³⁴ Cf. Recanati (1989); Neale (1992: 536–7); Levinson (2000: 214); Wilson and Sperber (2004: fn. 18); Carston (2004a: 646, 2004b: 74–98); Romero and Soria (2010: 5). Contrast García-Carpintero (2001: 113).

would nonetheless be normal for *speakers* to use (29)(a) to express that proposition. It is therefore as natural to hold that (29)(b) is a generalized implicature of (29)(a) as it is to hold that “He ran to the edge of the cliff and jumped off” is a generalized implicature of ‘He ran to the edge of the cliff and jumped.’ In general, it is plausible that “If p’ then q” is a generalized implicature of ‘If p then q’ when “p’” is a generalized implicature of ‘p.’ This rule appears to generalize to other compound propositional forms.

Second, on Carston’s account of what is said, it is not clear how the truth of (28) explains why (29)(a) can be used to express what (29)(b) expresses. For Carston’s view breaks the tight connection Grice maintained between what is said and what a sentence means. If what is said is related to what a sentence means by the same sort of pragmatic factors as what is implicated, why should one embed and not the other? If what a sentence says is what it means, in contrast, then embedding is a consequence of semantic compositionality: the meaning of ‘If p then q’ is a function of the meanings of ‘p’ and ‘q,’ and the proposition expressed by ‘If p then q’ contains the proposition expressed by ‘p’ as its antecedent and that expressed by ‘q’ as its consequent. But if what is said can be as distinct from what the sentence means as Carston maintains, there is no more reason to expect embedding of what is said than there is to expect embedding of what is implicated.

Carston (2004b: 95–7) hypothesizes that explicatures embed because they are “enrichments” of the logical form of the sentence uttered. She notes that in (31), C’s negation denies B’s suggestion (to use a neutral term) that Mary hurt her knee as a result of falling over, whereas in (30), C’s negation does not deny B’s suggestion that Bill has a girlfriend.

- (30) A: Does Bill have a girlfriend these days?
 B: He visits New York every weekend.
 C: No, he doesn’t. He goes there to see his ill mother.

- (31) A: Why is Mary crying?
 B: She fell over and hurt her knee.
 C: No, she didn’t. She hurt her knee and fell over.

The reason for this difference, on my account, is that B’s suggestions are within the scope of C’s denial only if the negations are irregular. Irregular negations can only be used to deny implicatures of their roots, which are generalized implicatures. Whereas B’s suggestion is a generalized implicature of the sentence he uttered in (31), it is not in (30). Carston’s explanation is that B’s suggestion is within the scope of C’s negation in (31) but not (30) because it is a “pragmatic enrichment” of the logical form of the sentence B utters in (31) but not of that in (30). Hence Carston classifies B’s suggestion as an explicature in (31) and an implicature in (30). If this explanation held up, it would mean that the Relevance Theorist’s distinction between “explicit” and “implicit” communication does have some theoretical interest. But in fact, it is mere coincidence that what is denied is a pragmatic enrichment in (31) but not (30). This can be seen by varying the examples.

- (32) A: Does Bill get to see girlfriend these days?
 B: He goes to New York every weekend.
 C: No, he doesn't. He goes there to see his ill mother.
- (33) A: Why is Mary being taken away in a stretcher?
 B: She fell over and hurt her knee.
 C: No, she didn't. She hurt her knee but not badly.

In both (32) and (33), B's suggestion is an "enrichment" of the logical form of the sentence B uttered: "He goes to New York every weekend *to see his girlfriend*" in (32), and "She fell over and *badly hurt* her knee" in (33). Neither suggestion can be interpreted as within the scope of C's negations. The reason, on my view, is that B's suggestions are both non-generalized implicatures: a relevance implicature in (32) and an understatement in (33). An irregular negation cannot be used to deny an enrichment of its root unless that enrichment is an implicature.

8 'Say' and 'Implicate'

I have been showing that Carston's arguments for the truth of (28) are unsound. Prima facie evidence for (28) is provided by the fact that people will commonly describe the speaker as having said that A jumped off, and that includes the speaker himself. That is, (28) would be a natural thing to say in many contexts. But this can be interpreted as *loose usage*. In contexts in which the subtle distinction between saying and implying is not important, S is close enough to having said that A jumped off for the purposes of the conversation. S could always raise the standard, however, by insisting "I didn't *say* that A jumped off, all I said was that A jumped." Of course, S would open herself up to the charge that she *misled* her audience if she knew A just jumped up and came down on the cliff. But could we say that S *lied*? Or that what she said was *false*? I do not think so.

In other contexts, it should be noted, (28) is clearly false. Suppose a helicopter was hovering that A could jump on to. S might have meant something different by (24) in that case. But what he said was the same. Carston has to maintain that S said different things in these two cases, which does not seem correct, unless 'said' is being used loosely.

Grice (1975: 24–5; 1969: 87) was aware that 'say,' when followed by a that-clause, can be used more or less narrowly. He was using the term fairly strictly, tying what is said closely to the conventional meaning of the sentence the speaker uttered. The man who said "There is a station around the corner" did not say in Grice's sense that one can get gas at the station because 'There is a station around the corner' does not itself mean or imply that one can.³⁵ Carston is using the term

³⁵ It may be that Grice took all explicature to be semantic. But I believe he was simply ignoring indexicality for simplicity, as was common at the time. An additional complication is that 'S said that... NP...' has both transparent and opaque interpretations, the former being weaker than the

more loosely, allowing what is said to go beyond what the sentence means. Since both took implicatures to be things meant but not said, it is to be expected that some of what Grice calls an implicature Carston would call an explicature. Since ‘say’ and ‘implicate’ are correlative terms, the more loosely ‘say’ is used, the more strictly ‘implicate’ will be used. There does not appear to be any substantive issue here, just a terminological choice. The fact that conversational principles play as significant a role in inferring what people say as in inferring what they implicate (Wilson and Sperber 1981: 155; 2004: 615; Carston 1988: 37; 2004a: 633, 636, 643; Bach 2010: 128) can and should be granted even by ambiguity theorists.

Bach would deny (28) too, because he uses ‘say’ in a very strict sense, to mean “strictly, literally, and explicitly say” (1994: 160). The three adverbs here have different meanings. The first two express ways for a speaker to use an expression and mean something by it. ‘Strictly’ is opposed to ‘loosely’: a speaker can use ‘S said p’ strictly to mean “S said p” or loosely to mean “S was close enough to having said p.” ‘Literally’ is opposed to ‘figuratively’: a speaker can use ‘S said p’ literally to mean “S said p” or figuratively to mean any number of other things, such as “S did not say p” (irony), “S did something like saying p” (metaphor), and so on. What a speaker meant when using ‘S said p’ loosely or figuratively may be true even though ‘S said p’ is false. When ‘VP’ is headed by a verb other than ‘use’ or ‘mean,’ ‘literally VPed’ has a mixed meaning. Thus, ‘S literally died,’ means “S died, and I mean that literally.” ‘S strictly speaking violated the law’ means “S violated the law, and I mean that strictly.”

‘Literally’ also has a sense in which it describes a way for an expression to mean something, or for a speaker to say something. ‘Kick the bucket’ literally means “strike the bucket with ones foot” even though it also means “die.” Someone who used ‘Jim kicked the bucket’ said that Jim died while literally saying that Jim struck the bucket with his food. In the same way, to *explicitly say* that p, is to say that p in a particular way: by using a sentence that literally means “p.” One can say something without doing so explicitly in at least three ways: by using ellipsis, by using an indexical, or by using an idiom. Even Carston should agree that (24) cannot be used to *explicitly say* in this sense that A jumped off the cliff. For ‘jumped’ does not mean “jumped off the cliff.” This can be very confusing because, as we observed, Carston follows Sperber and Wilson (1986: 182) in using ‘explicitly communicate p’ in a way that does not entail explicitly saying p. For the Relevance Theorists, ‘explicit’ seems intended to mean “direct” rather than “linguistically encoded.”

When we describe the expert witness using (16) as having said under oath that in his opinion, the drug was not a factor in the victim’s death, our usage of ‘say’ cannot be described as loose. We take the witness to have *testified* to that. Nevertheless, we must acknowledge that he did not *explicitly say* that, and that what he

(Footnote 35 continued)

latter. If John says “I saw Obama,” then “John said that he saw the president” is true when transparent but false when opaque.

literally said was weaker. So Bach is on to something that seems to favor the implicature theory. I will have to accommodate this observation when I defend the ambiguity theory below. The regular interpretation of a negation is its literal meaning; the irregular interpretations are non-literal.

Bach (1994: 140–1) also says that implicatures must be completely separate from and conceptually independent of what is (explicitly) said.³⁶ This is false on the Gricean definition of implicature I am using. If a speaker means that *there is a station where the addressee can get gas around the corner* by saying that *there is a station around the corner*, then the speaker implicated the former even though it is an “expansion” of what is said. The same is true when ‘some’ is used to implicate “just some.” But I believe Bach is just giving the term ‘implicature’ a narrower definition, like Carston but for a slightly different reason. Bach introduces the term ‘*implicature*’ (with an ‘i’ rather than an ‘a’) to cover cases of speaker meaning that are neither explicitly said nor implicated in his sense. Bach’s view of sentences like (24) does not differ substantively from Carston’s (see Bach 2010), which in my terminology is an implicature theory.³⁷ Bach (1994: 153–4) also appears to hold an implicature theory of irregular negations, but his discussion is too brief to determine whether he holds an R-, I-, or radical implicature theory.

There is one respect in which (24) provides a poor model for most irregular negations, which is independent of whether (28) is true or false, and which counts firmly against the implicature theory. In the case of (24), S means both what the sentence means (“He ran to the edge of the cliff and jumped”) and something stronger (“He ran to the edge of the cliff and jumped off”). In the case of (1)(a), in contrast, we do not mean, and are not taken to mean, what the first clause means (“The sun is not larger than at least some planets”); all we mean is something weaker (“The sun is not larger than just some planets”). The two levels of speaker meaning found in the cliff example are not present. In this respect, scalar-implicature denials differ from evaluative-implicature denials (§2). There could be two levels of speaker meaning in a contrary affirmation, but only if the indirection were the reverse of what we find in (24). The expert witness might have implied that he lacked the belief that the drug was responsible by saying that in his opinion the drug was not responsible. Users of (24) clearly do not imply that A jumped by saying that he jumped off. We will see another respect in which (24) differs from negations when we discuss Atlas in Sect. 10.

³⁶ Bach (2001: 19) sometimes says implicatures must be “external” to what is said. This would seem true as long as the implicature is not a part of what is said.

³⁷ Bach (2010: 132) later said “In implicating something, a speaker means one thing and conveys something else in addition... To ‘implicite’ something (if I may coin a term) is to say it, but only partially, since one is leaving part of what one means implicit.” Bach’s definition of ‘implicature’ here conforms to mine. But if part of what one means is not said, then there is something one meant but did not say—hence something one implicated.

9 Indexical and Elliptical Explicature

Perhaps a clearer example of pragmatic explicature would provide a better model for irregular negations. Wilson and Sperber's (1981: 157–9) paradigm was 'A plays well,' which is used on different occasions to say different things, depending on what A plays. Bach (1994) offered a similar paradigm of "implicature":

- (34) A finished.
 (a) A finished the sonata.
 (b) A finished dinner.
 (c) A finished his homework.

On different occasions, (34) is used to express different propositions as indicated by (a), (b), and (c). Nevertheless, sentence (34) itself does not mean any of these things, even on those occasions. Speakers can nonetheless *say* these things by uttering (34). This is true even when 'say' is used strictly. Suppose the expert witness is asked, "Did you finish your examination?" If she answers "I finished," then she has testified that she finished her examination. If the witness has finished something but not her examination, she cannot defend herself from the charge of lying by saying "All I said was that I finished." So in one respect, (34) is a better model of irregular negations than (24) is. Example (34) is also better because there is no indirection. The witness did not *imply* that she finished her examination. She did not mean that she finished her examination *by* meaning (or by saying) that she finished. The witness did mean (and say) that she finished, but that is just to say that she meant (said) that she finished her examination. The relation between (34) and (34)(c) is like the relation between *A finished it* and *A finished her examination*. In the witness's context, to mean that she finished it is to mean that she finished her examination. Whereas jumping off and jumping up are different ways of jumping, finishing dinner and finishing the sonata are not different ways of finishing (or of finishing it).

Sentence (34) differs markedly from (24) in that it does not itself express a complete, truth-evaluable proposition. Like 'A finished it,' (34) is used to express different propositions on different occasions. It is not even true that 'A finished' means or expresses the proposition that *A finished something* (Bach 1994: 130). It is not like 'A ate' or 'A married.' Note that only in highly unusual contexts could the witness use 'I finished something' to mean that she finished her examination. Indeed, if she uttered that in response to "Did you finish your examination of the victim?" the implication would normally be that he finished something *else*.

I believe that sentence (34) provides a good illustration of pragmatic explicature because it is *indexical*. Whereas 'finished' is a transitive verb in (34)(a), (b), and (c), it is an intransitive verb in (34). 'Finished_{intr}' means something like "finished_{trans} the contextually indicated object(s)." The suggestion that (34) is elliptical is plausible. But if it were truly elliptical, then 'finished' would have to be a

transitive verb in (34). Yet it cannot be. ‘Finished’ differs markedly from ‘completed,’ as Bach (1994: 128) observed.

- (35) A completed.
 (a) A completed the sonata.
 (b) A completed dinner.
 (c) A completed his homework.

Sentences (35)(a), (b) and (c) are synonymous with their counterparts in (34). But unlike (34), (35) is not a complete sentence, and cannot be used as one.³⁸ ‘Completed’ cannot be used as an intransitive verb. Negations, in contrast, are complete sentences, expressing complete thoughts, on their compositional interpretation as well as their irregular interpretation. They do not have to be “completed” in any way.

Carston (2010: 235ff) argues against an indexical analysis of expressions like ‘finished_{intr}’ on the grounds that the hypothesized covert indexical does not behave like an indexical.³⁹ She claims that overt indexicals like ‘she,’ ‘it,’ ‘this,’ and ‘there’ “must be assigned a specific value rather than be merely existentially closed,” whereas the hypothetical indexical expressed by ‘finished_{intr}’ has an existentially closed interpretation in a sentence like ‘After A finished, A celebrated.’ Her premise about overt indexicals in this case is faulty. ‘She’ has an existentially closed interpretation in sentences like ‘When a woman was attacked, she fought back.’ Indexicals can similarly be bound by a universal quantifier, as can the indexical expressed by ‘finished_{intr},’ as ‘Whatever A is doing, if A finishes, A celebrates.’ ‘Finished_{intr}’ is even more similar to indexicals like ‘neighbor’ and ‘foreign.’ In ‘If someone annoys a neighbor, there will be trouble,’ the indexical implicit in ‘neighbor’ may refer back to ‘someone.’

Bach denies that ‘A finished’ can be used to say strictly, literally, and explicitly that A finished his homework. This is correct, but only because the last adverb does not apply. If the defense attorney asks the expert witness “Did you finish your examination of the victim?” she can strictly and literally give an affirmative answer by uttering ‘I finished’ if she can do so by uttering ‘I finished my examination.’ But only by uttering the latter can she *explicitly* say that she finished her examination. For ‘I finished’ does not mean “I finished my examination.” The witness was not merely implying that she finished her exam, she asserted it. Because she did literally and strictly speaking say that she finished the examination, she can be accused of lying if she finished something but not her examination. Lying does not require explicit saying. Similarly, if the witness says “I finished the examination,” and we say “Dr. Smith did too” or “Dr. Smith finished it too,” we are saying—and asserting rather than implying—that Dr. Smith finished the examination. We are being completely literal and speaking strictly. But we are not

³⁸ (35) can be used to express a complete thought in special contexts, by saying “A completed...” and then pointing to what it is A completed. The demonstrative element completes the thought.

³⁹ She cites Martí (2006).

explicitly saying that Dr. Smith finished the examination. Furthermore, if we say it in any of these ways, we are being completely literal. Nothing figurative is going on. “Did you mean it literally?” can only be answered “Yes.” Because we did literally and strictly speaking say that Dr. Smith finished the examination, we would be lying if we knew she finished something but not her examination.

To support his view of what is said, Bach would observe that if (34) were used to say (strictly and literally) that A finished his homework, then it would have to be either *indexical* or *elliptical*. He rules out indexicality on the grounds that an indexical must be “there in the sentence” (1994: 133). He observes that there is no pronoun, demonstrative phrase, or temporal or locational adverb in (34). But not all indexicality is carried by these lexical elements, as illustrated by ‘It is raining’ and ‘Bob bought a foreign car.’ The former refers to the present time via the verb tense; the latter refers to the country of utterance via the adjective ‘foreign,’ which roughly means “from a country other than this one.” ‘Finished_{instr}’ is as indexical as ‘foreign,’ although in a different way. Bach similarly rules out ellipticality in part on the grounds that omitted material necessarily “corresponds to something in the sentence” (1994: 132). But if A says ‘Mary sings,’ and B says ‘Jane too,’ nothing in the expression B uttered corresponds to the omitted verb ‘sing.’ The expression B uttered is an incomplete sentence precisely because it lacks a verb.

Bach’s position turns on his “intuition” that “*the elements of what is said must correspond to the constituents of the utterance*” (1994: 137, my emphasis; see also 2001: 15), where the correspondence is “syntactic” (160). The elements must be phonologically or syntactically realized (2010: 127).⁴⁰ If this principle were true, it would be hard to see how Bach’s Perot example could be used to say either of the two different things Bach allows.

- (36) I know a richer man than Ross Perot.
 (a) I know a man who is richer than Ross Perot.
 (b) I know a man who is richer than any man Ross Perot knows.

Bach properly counts (36) as elliptical: because of a syntactic ambiguity, it can be used to say the two different things expressed by (a) and (b).⁴¹ But the two different suppressed elements are no more in the one sentence (36) than the two meanings of (36) are. Since the two things said do not have the same structure, they cannot both have the same structure as (36), which is just a sequence of words. “Suppressed material” may be in a *meaning, deep structure, or logical form* of a sentence, but is not in the sentence itself.

⁴⁰ Bach (2001: 17) nonetheless claims that not all elements of an uttered sentence need be uttered. But if a speaker fails to utter any word or morpheme in (36), she does not utter (36). Someone who utters ‘I know a richer man than Ross’ does not utter (36), even though he might say the same thing.

⁴¹ Since (36) is syntactically ambiguous, (a) and (b) are semantic rather than pragmatic explicatures.

Even a simple sentence like ‘2 is prime’ would seem to violate Bach’s correspondence constraint, since it can be used to say that 2 is divisible only by 1 and itself, and yet does not contain a syntactical element referring to 1. If Bach allows that this example satisfies the constraint because the whole saying-element “divisible only by one and itself” corresponds to the expression ‘prime’ (on the relevant interpretation), or because an element in the deep structure or logical form of the sentence refers to 1, then (34) should satisfy the principle too because the whole saying-element “finished his homework” corresponds to the expression ‘finished,’ and the deep structure and logical form of the sentence contains an indexical element referring in the context to his homework. Finally, it is hard to decide whether an uncontroversial idiom like ‘He kicked the bucket’ does satisfy Bach’s correspondence constraint because the saying element “died” corresponds to the expression ‘kicked the bucket,’ or does not because the expressions ‘kicked,’ ‘the,’ and ‘bucket’ do not individually correspond to any elements in what is said.

For a variety of reasons, then, I do not believe Bach’s correspondence principle can be used to support the claim that the irregular interpretation of a negation is not something said. I suggest that his necessary condition for saying should be recast as a comparative principle of explicitness: the greater the syntactic correspondence between a sentence and what it says, the more explicitly it says that. We will return to this principle in [Sect. 11](#).

If irregular negations were indexical, that would support the pragmatic explicature theory. But it is hard to see how any known form of indexicality could enable a sentence with a fixed meaning to express propositions that differ in as many different ways as regular and irregular negations do. The contextual variation possible with ‘not p’ is not nearly as delimited as with (34). Given that an irregular negation is always used to deny an implicature of its root, we might hypothesize that ‘Not-p’ in every case means something like “A *contextually selected implicature of ‘p’ is not true.*” But ‘not-p’ cannot have this meaning when it is a regular negation. And any given root ‘p’ typically has generalized implicatures that cannot be denied by any irregular negation. For example, ‘S knows p’ implicates “S does not know not-p” just as ‘S believes p’ implicates “S does not believe not-p.” But ‘S does not know p’ cannot be used to deny that implicature (that is, to affirm “S knows not-p”). And in addition to its scalar implicature, ‘Some S are P’ has the ignorance implicature ‘I do not know how many S are P.’ But ‘Some S are not P’ does not have “I know how many S are P” as an irregular interpretation. (34) is thus a poor model for irregular negations too.

Some irregular negations are plausibly elliptical: scalar-, metalinguistic-, and specifying- implicature denials. For example, ‘The sun is not larger than some planets’ seems elliptical for ‘The sun is not larger than just some planets’; the ‘just’ is omitted but understood.⁴² It is not clear, though, that this is an example of pragmatic explicature because it is not obvious that the first sentence has the same

⁴² Bach (1994: 132) would deny that this is ellipsis because the omitted element is not syntactically determined.

meaning when used elliptically in this way. Furthermore, the other irregular negations are clearly not elliptical. So the fact that some clear examples of pragmatic explicature involve ellipsis provides little support for the pragmatic explicature theory of irregular negations.

10 Generality Versus Ambiguity

Kempson (1975: 99ff) and Atlas (1989: 73–7) use the standard identity test to argue that negations are *general* in sense rather than ambiguous. Certain forms of ellipsis, pro-formation, and coordination, along with words like ‘too’ and ‘either,’ allow qualitative and quantitative diversity in application but require identity of sense: ambiguity produces zeugma or syllepsis.⁴³ The word ‘plane,’ for example, is ambiguous because it means either “airplane” or “wood plane” (among other things). In either sense, the term is general. ‘Plane’ applies in the first sense to all airplanes, whether they be Boeing 747s or Sopwith Camels. In the second sense, it applies indifferently to five-inch pocket smoothing planes and two-foot jointer planes. Consequently, the sentences in (37) could be used to describe a situation in which Jack saw a Boeing 747 and Jill saw a Sopwith Camel, or a situation in which Jack saw a small smoothing plane and Jill saw a long jointer plane, but not a situation in which Jack saw a Boeing 747 and Jill saw a long jointer plane. In the third case, unlike the first two, ‘plane’ does not apply in the same sense, so these grammatical constructions cannot be used without zeugma. “Cross-readings” are prohibited. Similarly, (38) could correctly be used if neither of the two saw any airplanes, or neither saw any wood planes, but not if one saw an airplane but no wood plane, and the other saw a wood plane but no airplane.

- (37) Jack saw a plane; Jill saw one too.
 Jack saw a plane; Jill did too.
 Jack and Jill saw a plane.

- (38) Jack didn’t see a plane; Jill didn’t see a plane either.

Kempson and Atlas observe that (39) is acceptable, and conclude that negation is general rather than ambiguous.

- (39) The king of France is not wise (since France is not a monarchy), and the same thing goes for the queen of England (who is a typical Windsor).

However, the first parenthetical clause forces the first negation to have the weaker irregular interpretation, which is then naturally given to the second negation. The second parenthetical clause does not similarly force the stronger regular

⁴³ See Chomsky (1965: 35–6, 1972: 33); Lakoff (1970); Zwicky and Sadock (1975: 17–36); Lyons (1977: 405–9); Atlas (1977: 327–30, 1979: 277–8); Levinson (1983: 201).

interpretation, and so does not induce a cross-reading, as Atlas and Kempson assumed. We can get a cross-reading if we reverse the order of the conjuncts, but then the result is decidedly zeugmatic.

- (40) The queen of England is not wise (since she is a typical Windsor), and the same goes for the king of France (since France is not a monarchy).

On first pass, we naturally read the first conjunct of (40) as regular, and are then perplexed by the second clause, which cannot have the same reading but must. If we accept (40) at all, we have to go back and give both clauses the irregular reading.

Burton-Roberts argues that since the identity of sense requirement can be met by the more general presupposition-canceling interpretation, the ambiguity hypothesis cannot be verified.

[T]he two putative senses of the negation are semantically related to each other in such a way that it is demonstrably impossible (as shown by Zwicky and Sadock 1975 [pp. 23–5]) to establish empirically that there is indeed a genuine semantic ambiguity. The more specific internal understanding entails the more general external understanding; for a semantics that includes the external operator, it is impossible to demonstrate that it also includes the internal operator. (Burton-Roberts 1989: 100)⁴⁴

This reasoning is unsound. First, the identity test is not the only source of evidence for ambiguity. So its failure to demonstrate ambiguity would not mean that ambiguity cannot be established. Second, the same results are obtained with terms that beyond reasonable doubt have both a specific and a generic sense, like ‘animal,’ ‘New Yorker,’ and ‘act.’ Thus, (41) can be used when Jack saw a *homo sapiens* and Jill saw a *felis domestica* because the generic sense applies in both cases.

- (41) Jack saw an animal; Jill saw an animal too.

This does not prove that ‘animal’ is unambiguous. Indeed, if a second test is administered, this time imagining that Jack saw a cat and Jill a human, then the result is liable to be zeugmatic, confirming ambiguity. As Zwicky and Sadock (1975: 24) note, we are not *forced* in the second case to give the first instance of ‘animal’ the more specific interpretation. For the general reading is available and applicable. But we *can* do so, quite naturally. And when we do, the results are positive for ambiguity. Third, only presupposition-canceling and specifying-implicature denials are weaker than their regular interpretation. As (1)(a), (b), (d) and (f) illustrate, contrary affirmations are stronger than the corresponding regular negation, while scalar-implicature, metalinguistic-implicature, and evaluative-implicature denials are neither stronger nor weaker. The examples in (42) show

⁴⁴ Cf. Kempson (1979: 285–6; 291–2). Contrast Horn (1989: 317, 365ff, 561n1); Martin (1982: 252ff); Seuren (1990: 433–4).

that when these irregular negations are crossed with regular negations, the results are clearly zeugmatic.

- (42) (a) Carnegie did not control *some* of the steel market; Rockefeller did not control some of it either.
 (b) Malibu Beach is not a *beech*; an oak is not a beech either.
 (c) The sky isn't partly cloudy; the sky in the Atacama desert isn't partly cloudy either.
 (d) Democrats do not believe social security should be abolished; dead republicans do not believe that either.

Finally, because two different kinds of irregular negations are weaker than the corresponding regular interpretations, we can get a zeugma even with them. Consider:

- (43) Mary did not meet a *man* at the bar; the tooth fairy did not meet a man at the bar either.

So the identity test supports the ambiguity theory.

Atlas argued that negations failed a second test that generic-specific ambiguities must pass. "If the expression is truly ambiguous, it ought to be possible to assert the general case and deny the specific case without contradiction" (1989: 71).⁴⁵

Consider:

- (44) When ordering the invasion of Grenada, Ronald Reagan was acting but he was not acting.

Since the default presumption is that different occurrences of the same word in one sentence will have the same meaning, (44) will initially strike us as contradictory. But with a little effort, we can hear the two occurrences of 'acting' as having different meanings ("doing something" versus "performing as an actor"), making the sentence non-contradictory and true. Compare now:

- (45) Vulcan is not hot, but it is not not hot.

Atlas found such a sentence to be "semantically out of bounds," so he concluded that negations fail the non-contradiction test. Atlas's claim is reasonable, because on the default presumption the sentence is contradictory. Moreover, if the first clause is true because Vulcan does not exist, the second clause cannot be true if both negations are regular. Nevertheless, there is a coherent interpretation of (45): let the first two 'not's be irregular, and the third regular (so let the second 'not hot' mean "non-hot"). This interpretation can be forced by adding 'either' to (45). So the non-contradiction test is positive for ambiguity.

⁴⁵ See also Zwicky and Sadock (1975: 7–8); Atlas (1977: 324).

Since Atlas (1989) believed that negations failed the tests for ambiguity, he concluded that negations are “sense-general” rather than ambiguous. They are “sense-unspecified” for scope (1989: 69). His model is “a General term like ‘neighbour’, which is not ambiguous between the readings “nearby dweller who is male” and “nearby dweller who is female” but instead has a sense that is grammatically unspecified for gender” (Atlas 1989: 123). Since negations are sentences rather than terms, an even better model is Carston’s example (24) above, in which ‘A jumped’ does not specify where A jumped, and is thus more general than ‘A jumped off’ or ‘A jumped up.’ Like Carston, Atlas relies on pragmatic inference to “construct” the specific meanings on the basis of the sense-general meaning and contextual factors, although he relies on Gricean principles rather than the Principle of Relevance. Since he maintains that both the regular interpretation R and the irregular interpretation I is constructed from a more general meaning M, Atlas endorses the radical explicature theory, although he does not use that term.

Atlas’s theory is hard to comprehend because it is not evident what the generic meaning M could possibly be in the case of negations. What else does ‘Vulcan is not hot’ mean other than its regular and irregular interpretations? How in particular does the “sense-general” meaning differ from the presupposition-canceling interpretation? Atlas appears to be positing a Lockean substratum. One of Atlas’s theses is that the generic meaning of negations “is not a bearer of truth-value; it is not a proposition or a logical form” (1989: 69). But since I and R are propositions, how could they be *more specific* than M unless M is also a proposition, one that is entailed by both I and R? The meaning of the indexical sentence ‘It is not hot,’ for example, is not a truth-evaluable proposition. But its meaning cannot be described as more general or more specific than the meaning of ‘Vulcan is not hot.’ The meanings are not comparable in this way. There does not, however, appear to be any proposition expressed by negations other than the regular and irregular interpretations we have discussed. Even if we ignore the issue of whether M is a proposition completely, it is hard to see how a content of any sort could be more general than all the different interpretations we have observed for the six different negations in (1). These different interpretations are not determinates of any single determinable. In each case, the relationship between the irregular and regular interpretations differs markedly from that between ‘A finished the sonata’ and ‘A finished something.’ The many different interpretations of negations are not simply special cases in which a single more general sense applies.⁴⁶

Terms that clearly have a sense-general meaning are poor models for the phenomenon we are studying. Ambiguity and generality place different constraints on the speaker. Thus, anyone who uses ‘That is a plane’ has to have intended one

⁴⁶ Atlas might be claiming that M is a propositional *form* of which I and R are specific instances—like ‘A finished NP,’ whose substitution instances include ‘A finished the sonata’ and ‘A finished dinner.’ But there is no one form of the regular interpretation and all the irregular interpretations of a negation such that all the interpretations are just different substitution instances. Besides, on no interpretation do any of the negations in (1) seem to represent a form.

of the senses of the word ‘plane’ if he was using English. He had to have meant “That is an airplane” or “That is a wood plane” (or one of the other things the sentence means). That determines what the sentence means on that occasion, and what the hearer must do to understand it. If the speaker tells us that he was not trying to say the object was an airplane or a wood plane, or that these alternatives did not even occur to him, or that he does not care which it is as long as it is a plane, then we could only conclude that he was using the word ‘plane’ without understanding it. There is no similar requirement that the speaker mean either “That is a Sopwith Camel” or “That is a Boeing 747” if he means “airplane” by ‘plane,’ or even that the speaker has considered these alternatives. Indeed, if he is speaking English, then he cannot *mean* “Sopwith Camel” or “Boeing 747” *by* ‘plane.’ For ‘plane’ does not mean either of these things in English. If ‘plane’ is used as a word of English, it can never happen that ‘That is a plane’ asserts a falsehood because ‘That’ referred to a Sopwith Camel while ‘plane’ meant “Boeing 747.” In contrast, the sentence would assert a falsehood if ‘that’ refers to a jointer plane while the speaker means “airplane” by ‘plane.’ Atlas (1989: 90) says that “It is precisely because the meaning of the ‘not’-sentence is sense-general that the same string of words... can be differently interpreted as making different statements.” But when a sentence is genuinely general because it does not semantically encode a certain condition, it cannot be interpreted as stating that that condition obtains. ‘That is my neighbor’ cannot be interpreted as stating that the person is male or female.

In these respects, negations are like ambiguous terms rather than general terms. Anyone who uses (1)(a) in conventional English, for example, must intend either the regular interpretation or one of the irregular interpretations. That is, the speaker must mean either “The sun is not larger than any planets” or “The sun is not larger than just some planets” (etc.). That will determine what the sentence means and states on that occasion of use, and what hearers must recognize if they are to understand the utterance. If the speaker said “Neither of those thoughts crossed my mind; I don’t care about that issue; all I said was that the sun isn’t larger than some planets,” we could only conclude that she did not know what she was saying.

A common reason for rejecting ambiguity claims is what I call “gut incredulity.” People simply find it hard to believe—absurd even—that sentences have so many meanings.⁴⁷ The question is, why should this be hard to believe? Everyone acknowledges that all the different readings of negations we have identified are possible, and that they are conventional. Once this much is acknowledged, the additional claim that the readings are meanings should be just a question of detail.

I believe the argument that has been most influential in leading philosophers and linguists to reject ambiguity claims is the appeal to Grice’s Razor. The thought is that it is more economical to postulate conversational implicatures rather than senses, because conversational implicatures are derivable from independently

⁴⁷ See for example Katz (1972: 92); Sperber and Wilson (1986: 188); Kittay (1987); Bach (1994: 150). Contrast Davis (2003: Sect. 10.6).

motivated generalizations. As indicated in §§2–4, though, sentence implicatures are never calculable. There are always alternatives, and the conversational rules generally conflict. Because generalized conversational implicatures are conventional, postulating them is no more economical than postulating senses. The same goes for explicatures and implicatures. Even if Grice’s Razor were true, the greater economy would not outweigh the linguistic evidence for ambiguity.

A corollary is Bach’s (1994: 151ff) “missed generalization” argument. If all the different interpretations of a negation could be pragmatically derived from a single meaning, we would miss a generalization by claiming that the different interpretations represented independent semantic conventions. This argument mistakenly assumes that conversational principles can generate or predict the different readings of negations. Moreover, another generalization can be offered: all the different irregular senses arose from conversational implicatures that became conventional and then direct (Sect. 11).

11 Idiom Theory

In addition to their regular, fully compositional interpretation, negations have at least six irregular interpretations illustrated in (1) and glossed in (2). On their irregular interpretations, negations are used to deny a generalized conversational implicature of their roots. The main question we have been investigating is the nature of these interpretations. Are negations strictly speaking ambiguous, with up to seven distinct linguistic meanings? Or are all but one of these interpretations implicatures or pragmatic explicatures?

Given that negations are commonly and properly used and understood with all seven interpretations, the ambiguity or semantic explicature theory is at least *prima facie* attractive. This theory is supported by the fact that negations test positive for ambiguity using both the identity and non-contradiction tests. Moreover, there is no viable alternative to semantic ambiguity theories except for evaluative-implicature denials.

The principal objections to ambiguity theories are (1) that they are uneconomical, and (2) that there is no way to account for the ambiguity. The economy argument, we have seen, is based on the fallacious idea that implicatures or pragmatic explicatures can be derived from senses using conversational principles. The second objection is more serious. Given that the irregular interpretations of negations are not fully compositional, they cannot arise from lexical or syntactic ambiguities—from different meanings of the word ‘not’ or differences in the scope of ‘not.’ How then can they be ambiguous?

The answer can be discovered by observing that many expressions have, in addition to fully compositional meanings, other meanings that do not result compositionally from the meanings of their components. Such expressions are *idioms*. Standard examples are *kicked the bucket* and *end of one’s rope*. In addition to its compositional meaning: *performed the act of kicking with the bucket as its*

object, ‘kicked the bucket’ has another quite unrelated and totally noncompositional meaning: *died*. Idiomatic meanings are conventional and direct: it is conventional for speakers to use ‘kicked the bucket’ to mean “died,” and they do so directly, not by meaning “performed the act of kicking with the bucket as its object.” The theory that irregular negations are idiomatic immediately explains some of their most characteristic properties, such as limited substitutivity of synonyms. Irregular negations are like ‘end of one’s rope’ rather than ‘kicked the bucket’ in being partially compositional. Replacing ‘rope’ with a synonym like ‘cord’ eliminates the idiomatic meaning entirely, while changing possessive pronouns changes the meaning of the phrase predictably. Similarly, replacing ‘some’ with ‘any’ in (1)(a) eliminates the implicature-denial meaning, while replacing ‘planets’ with ‘asteroids’ changes the meaning in predictable ways.

Many idioms are known to have evolved from metaphors, euphemisms, and other figures of speech,⁴⁸ which are forms of conversational implicature.⁴⁹ For example, ‘cut and run’ originated as a nautical term in the 18th century meaning “cut the anchor cable and leave immediately.” It was subsequently used metaphorically to mean “make a quick escape,” and is now a completely dead metaphor. When the metaphor was alive, the now idiomatic meaning was conveyed indirectly, and the literal meaning directly. As metaphors gradually die through repeated use, less and less attention is paid to the literal meaning. When dead, the formerly metaphorical meaning is conveyed directly, usually without conveying the literal meaning. People who use ‘cut and run’ today are not saying that the subject is cutting the anchor and sailing away, and are rarely even thinking about ships. The lexicalization of generalized conversational implicatures is a similar and even more common phenomenon. A relatively recent example was the process that gave ‘virus’ a meaning applicable to computers. At first the computer virus meaning was an particularized conversational implicature, a biological metaphor. It caught on, becoming a generalized implicature. Eventually, the metaphor died. That is, speakers began conveying “computer virus” directly, making this another meaning of ‘virus’ in English.

Evaluative-implicature denials are still “live” implicatures, so the implicature theory works for them. I hypothesize that the other irregular interpretations of

⁴⁸ Ammer (1997), Siefring (2004), and “Euphemism” (*Wikipedia*) are filled with fascinating examples. See also Sadock (1972); Grice (1975: 58); Searle (1975: 76ff); Morgan (1978); Horn (1989: 344–5); Hopper and Traugott (1993: 75–93); Cowie (1994); Davis (1998: Ch. 6, 2003: Ch. 8).

⁴⁹ Both Bach and Carston depart from Grice in not classifying what is meant in figurative speech as implicature. This is partly a terminological difference (Sect. 8). But they also make a substantive mistake in denying that figurative uses involve indirection (Bach 1987: 71; 1994: 144; 2001: 41, n. 3; 2006: 27–8; Carston 2010: 220). When a speaker uses ‘It is a beautiful day’ ironically, for example, the speaker means and expresses the belief that it is not a beautiful day *by saying* that it is a beautiful day (though not by meaning that it is a beautiful day). Moreover, it is essential to this being irony, rather than speaking in a code, that the speaker means that it is not a beautiful day *by using a sentence by which he means* “It is a beautiful day.” I sort out these complexities in Davis (2003: Ch. 2).

negations became senses as what was originally conveyed by implicature became conventional and then direct; they are dead implicatures. Given that speakers standardly used sentences of the form ‘S Vs some O’ to mean “S Vs just some O,” it would have been natural for them to begin using ‘S does not V some O’ to mean “S does not V just some O.” It is also plausible that with repetition, this usage became first conventional and then direct. When that happened, sentences of the form ‘S does not V some O’ acquired the irregular sense in addition to their regular sense. Implicature involves using words to say one thing and thereby mean another. It involves both a direct meaning and an indirect meaning. As long as this indirection exists, the second meaning will remain an implicature rather than a sense. When it becomes conventional to express the second meaning directly, it becomes a second sense (Davis 2003: Part II; 2005: Ch. 5).

Direct irregular negations are unusual idioms because of their productivity. I explore this property elsewhere (Davis 2011: Sect. 9). In other respects, though, they have the typical properties of idioms. For example, the conventions assigning idiomatic meanings to sentences are like conversational implicature conventions, and unlike lexical conventions, in not being completely arbitrary. We always perceive or expect some antecedent connection between the idiomatic meaning or implicature and the literal meaning (Davis 1998: Sect. 6.5). The same is true of the irregular and regular interpretations of negations. Idiom theory also accounts for the “special,” “marked,” and “unnatural” character of irregular negations (Kempson 1975; Horn 1989: 362–4, 370–5); Burton-Roberts 1989: 100). Semantic, grammatical, and even phonological irregularity is part of what makes idioms idioms. Consider ‘He doesn’t know which end is up.’ This is partially but not completely compositional. Standard entailments fail (‘He either forgot or never learned’). And only some components can be emphasized (‘He *doesn’t* know which end is up’ is okay, but not ‘He doesn’t *know* which end is up’ or ‘He doesn’t know which end is *up*.’) The idiom theory also helps to explain why “marked” negation fails to incorporate morphologically (Horn 1989: 392). In marked negations, the expression composed of ‘not’ and the predicate ‘P’ is not a compositional unit, and so does not express a contrary of the concept “P.” Contrary affirmation does incorporate because ‘not V⁺’ expresses the contrary concept “V⁻”; the irregular meaning of ‘does not believe’ is thus “disbelieves.”

Following Bach (1994: 153–4), one might think that an irregular interpretation I is not a meaning of negation N because N does not *literally and explicitly* say I. There is a more explicit way of saying it. It is true, for example, that (1)(a) is literally a contradiction, and that someone who uses it to say that the sun is larger than all planets, not just some, is not being fully explicit. These same things can be said about idioms, however. ‘Kicked the bucket’ literally means “performed a kicking action whose object was the bucket” rather than “died.” But ‘kicked the bucket’ also means “died” in English. Consequently, a speaker can use ‘Steve kicked the bucket’ to say that Steve died; but she will not be literally or explicitly saying that. ‘Literal’ is here opposed to ‘figurative,’ whether the figures of speech

Table 3 Idiom theory

	Idiom Theory
Sentence meaning	N means both R and I
Irregular use	S says and means I
Regular use	S says and means R

are dead or alive. A literal meaning is a *fully compositional* meaning.⁵⁰ Since an idiomatic meaning results from a convention to use a compound with a meaning other than its compositional meaning, it is a non-literal meaning. When a negation is used with an irregular interpretation, the speaker is not being maximally explicit because there is a more explicit way of expressing the same proposition. Thus, ‘The sun is not larger than just some planets’ is a more explicit way of saying what ‘The sun is not larger than some planets’ is used to mean when it is a scalar-implicature denial.

The theory that direct irregular negations are idioms is schematized in Table 3. Because direct irregular negations are idioms, negations are semantically ambiguous. Because indirect irregular negations are living implicatures, some negations are also pragmatically ambiguous. The evaluative-implicature denial interpretation is a generalized implicature, as maintained by the I-implicature theory. The scalar-, metalinguistic-, and specifying-implicature denial interpretations, along with the presupposition-canceling and contrary affirmation interpretations, are idiomatic meanings. Given the arbitrariness of conventions, there is no predicting when or even if the evaluative-implicature denial interpretation will become an idiom too, but it could.

References

- Ammer, Christine. 1997. *The American heritage dictionary of idioms*. Boston: Houghton Mifflin Co.
- Atlas, Jay. D. 1975. Presupposition: A semantico-pragmatic account. *Pragmatics Microfiche* 1(4): D13–G9.
- Atlas, J.D. 1977. Negation, ambiguity, and presupposition. *Linguistics and Philosophy* 1: 321–336.
- Atlas, J.D. 1979. How linguistics matters to philosophy: Presupposition, truth, and meaning. In *Syntax and semantics, 11: Presupposition*, eds. Oh Choon-Kyo, and David A. Dinneen, 265–281. New York: Academic Press.
- Atlas, Jay D. 1989. *Philosophy without ambiguity*. Oxford: Oxford University Press.
- Bach, Kent. 1987. *Thought and reference*. Oxford: Oxford University Press.
- Bach, Kent. 1994. Conversational implicature. *Mind and Language* 9: 124–162.

⁵⁰ ‘Literal’ is also used to describe the etymology of a word, as when it is said that ‘hippopotamus’ literally means “river horse.” In this case, the literal meaning is not something the word means in English at all.

- Bach, Kent. 2001. You don't say? *Synthese* 128: 15–44.
- Bach, Kent. 2006. The top 10 misconceptions about implicature. In *Drawing the boundaries of meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, eds. Betty J. Birner and Gregory Ward, 21–30. Amsterdam: John Benjamins.
- Bach, Kent. 2010. Implicature vs. explicature: what's the difference? In *Explicit communication*, eds. Belén Soria and Esther Romero, 126–137. Houndmills: Palgrave Macmillan.
- Bach, Kent, and Robert Harnish. 1979. Indirect speech acts and illocutionary standardization. In *Pragmatics: Critical concepts*, ed. Asa Kasher, 682–723. London: Routledge.
- Bezuidenhout, Anne. 2002. Generalized conversational implicatures and default pragmatic inferences. In *Meaning and truth: Investigations in philosophical semantics*, ed. Joseph K. Campbell, 257–283. New York: Seven Bridges Press.
- Blakemore, Diane. 1987. *Semantic constraints on relevance*. Oxford: Basil Blackwell.
- Blakemore, Diane. 1992. *Understanding utterances*. Oxford: Basil Blackwell.
- Boër, Steven, and William Lycan. 1976. *The myth of semantic presupposition*. Bloomington, IN: Indiana University Linguistics Club.
- Burton-Roberts, Noel. 1989. On Horn's dilemma: Presupposition and negation. *Journal of Linguistics* 25: 95–125.
- Burton-Roberts, Noel. 1999. Presupposition-cancellation and metalinguistic negation: A reply to Carston. *Journal of Linguistics* 35: 347–364.
- Carston, Robyn. 1987. Being explicit. *Behavioral and Brain Sciences* 10: 713–714.
- Carston, Robyn. 1988. Implicature, explicature, and truth-theoretic semantics. In *Mental representations: The interface between language and reality*, ed. Ruth Kempson, 155–181. Cambridge: Cambridge University Press. Reprinted in *Pragmatics: A reader*, ed. Steven Davis, 33–51. Oxford: Oxford University Press (1991).
- Carston, Robyn. 1996. Metalinguistic negation and echoic use. *Journal of Pragmatics* 25: 309–330.
- Carston, Robyn. 1998. Negation, 'presupposition', and the semantics/pragmatics distinction. *Journal of Linguistics* 34: 309–350.
- Carston, Robyn. 1999. Negation, 'presupposition' and metarepresentation: A response to Noel Burton-Roberts. *Journal of Linguistics* 35: 365–389.
- Carston, Robyn. 2002. *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell Publishing.
- Carston, Robyn. 2004a. Relevance theory and the saying/implicating distinction. In *The handbook of pragmatics*, eds. Laurence R. Horn, and Gregory Ward, 633–656. Oxford: Blackwell Publishing.
- Carston, Robyn. 2004b. Truth-conditional content and conversational implicature. In *The semantics/pragmatics distinction*, ed. Claudia Bianchi, 65–100. Stanford, CA: CSLI Publications.
- Carston, Robyn. 2010. Explicit communication and 'free' pragmatic enrichment. In *Explicit communication*, ed. Belén Soria, and Esther Romero. Houndmills: Palgrave Macmillan.
- Chapman, Siobhan. 1996. Some observations on metalinguistic negation. *Journal of Linguistics* 32: 387–402.
- Chaves, Jose E. 2010. Explicature, what is said, and Gricean factorisation criteria. In *Explicit communication*, eds. Belén Soria and Esther Romero, 109–125. Houndmills: Palgrave Macmillan.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press. Ch. 1 reprinted in *Readings in the philosophy of language*, eds. Jay Rosenberg and Charles Travis, 324–364. Englewood Cliffs, NJ: Prentice-Hall (1971).
- Chomsky, Noam. 1972. *Language and mind*. New York: Harcourt Brace Jovanovich, Inc.
- Cowie, Anthony P. 1994. Phraseology. In *The encyclopedia of language and linguistics*, ed. Ron Asher, 3168–3171. Oxford: Pergamon Press.
- Davis, Wayne A. 1998. *Implicature: Intention, convention, and principle in the failure of Gricean theory*. Cambridge: Cambridge University Press.

- Davis, Wayne A. 2003. *Meaning, expression, and thought*. New York: Cambridge University Press.
- Davis, Wayne A. 2005. *Nondescriptive meaning and reference*. Oxford: Oxford University Press.
- Davis, Wayne A. 2007. How normative is implicature? *Journal of Pragmatics* 39: 1655–1672.
- Davis, Wayne A. 2010. Irregular negations: implicature and idiom theories. In *Meaning and analysis*, ed. Klaus Petrus, 103–137. Houndsmills: Palgrave Macmillan.
- Davis, Wayne A. 2011. “Metalinguistic” negation, denials, and idioms. *Journal of Pragmatics* 43: 2548–2577.
- Ducrot, Oswald. 1972. *Dire et Ne Pas Dire*. Paris: Hermann.
- García-Carpintero, Manuel. 2001. Gricean rational reconstructions and the semantics/pragmatics distinction. *Synthese* 128: 93–131.
- Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.
- Geurts, Bart. 1998. The mechanisms of denial. *Language* 74: 274–307.
- Grice, H. Paul. 1969. Utterer’s meaning and intentions. *Philosophical Review* 78: 147–77. Reprinted in *Studies in the way of words*, ed. H. Paul Grice, 86–116. Cambridge MA: Harvard University Press (1989).
- Grice, H. Paul. 1975. Logic and conversation. In *Syntax and semantics, 3: Speech Acts*, eds. Peter Cole and Jerry Morgan, 41–58. New York: Academic Press. Reprinted in *Studies in the way of words*, ed. H. Paul Grice, 22–40. Cambridge, MA: Harvard University Press (1989).
- Grice, H. Paul. 1978. Further notes on logic and conversation. In *Syntax and semantics, 9: Pragmatics*, ed. Peter Cole, 113–28. New York: Academic Press. Reprinted in *Studies in the way of words*, ed. H. Paul Grice, 41–57. Cambridge, MA: Harvard University Press (1989).
- Grice, H. Paul. 1981. Presupposition and conversational implicature. In *Radical pragmatics*, ed. Peter Cole, 183–98. New York: Academic Press. Reprinted in *Studies in the way of words*, ed. H. Paul Grice, 269–282. Cambridge, MA: Harvard University Press (1989).
- Hopper, Paul J., and Elizabeth C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Horn, Laurence R. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61: 121–174.
- Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Horn, Laurence R. 1990. Showdown at truth-value gap: Burton-Roberts on presupposition. *Journal of Linguistics* 26: 483–503.
- Horn, Laurence R. 1992a. Pragmatics, implicature, and presupposition. In *International encyclopedia of linguistics*, vol. 2, ed. William Bright, 260–266. New York: Oxford University Press.
- Horn, Laurence R. 1992b. The said and the unsaid. *Ohio State University Working Papers in Linguistics* 40: 163–192.
- Horn, Laurence R. 2004. Implicature. In *The handbook of pragmatics*, eds. Laurence R. Horn, and Gregory Ward, 3–28. Oxford: Blackwell Publishing.
- Huang, Yan. 2006. Neo-Gricean pragmatics. In *The encyclopedia of language and linguistics*, 2nd edn, ed. Keith Brown, 231–238. New York: Elsevier.
- Huang, Yan. 2007. *Pragmatics*. Oxford: Oxford University Press.
- Israel, Michael. 1996. Polarity sensitivity as lexical semantics. *Linguistics and Philosophy* 19: 619–666.
- Katz, Jerrold J. 1972. *Semantic theory*. New York: Harper and Row.
- Kempson, Ruth. 1975. *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.
- Kempson, Ruth. 1979. Presupposition, opacity, and ambiguity. In *Syntax and semantics, 11: Presupposition*, eds. Choon-Kyo Oh and David Dinneen, 283–297. New York: Academic Press.
- Kempson, Ruth. 1986. Ambiguity and the semantics-pragmatics distinction. In *Meaning and interpretation*, ed. Charles Travis, 77–103. Oxford: Basil Blackwell.

- Kittay, Eva. 1987. *Metaphor: Its cognitive force and linguistic structure*. Oxford: Oxford University Press.
- Lakoff, George. 1970. A note on vagueness and ambiguity. *Linguistic Inquiry* 1: 357–359.
- Leech, Geoffrey. 1983. *Principles of pragmatics*. London: Longmans.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, Stephen C. 1987. Implicature explicated? *Behavioral and Brain Sciences* 10: 722–723.
- Levinson, Stephen C. 1989. Review of relevance. *Journal of Linguistics* 25: 455–472.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Lycan, William. 1984. *Logical form in natural language*. Cambridge, MA: MIT Press.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Martí, Luisa. 2006. Unarticulated constituents revisited. *Linguistics and Philosophy* 29: 135–166.
- Martin, John N. 1982. Negation, ambiguity, and the identity test. *Journal of Semantics* 1: 251–274.
- Morgan, Jerry L. 1978. Two types of convention in indirect speech acts. In *Syntax and semantics, 9: Pragmatics*, ed. Peter Cole, 261–280. New York: Academic Press. Reprinted in *Pragmatics: A reader*, ed. Steven Davis, 242–253. Oxford: Oxford University Press (1991).
- Morgan, Jerry, and Georgia M. Green. 1987. On the search for relevance. *Behavioral and Brain Sciences* 10: 725–726.
- Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509–559.
- Ramat, Paolo. 2006. Negation. In *Encyclopedia of language and linguistics*, 2nd ed, ed. Keith Brown, 559–567. Amsterdam: Elsevier.
- Recanati, François. 1989. The pragmatics of what is said. *Mind and Language* 4: 293–329.
- Recanati, F. 2004. *Literal meaning*. Cambridge: Cambridge University Press.
- Romero, Esther, and Belén Soria. 2010. Introduction: Explicit communication and relevance theory pragmatics. In *Explicit communication: Robyn Carston's pragmatics*, eds. Belén Soria, and Esther Romero, 1–24. Houndmills: Palgrave Macmillan.
- Sadock, Jerrold M. 1972. Speech act idioms. In *Papers from the eighth regional meeting of the Chicago linguistic society*, eds. Paul M. Peranteau, Judith N. Levi, and Gloria D. Phares, 329–339. Chicago: Chicago Linguistic Society.
- Saul, Jennifer. 2001. Review of implicature: Intention, convention, and principle in the failure of Gricean theory by Wayne Davis. *Noûs* 35: 630–641.
- Saul, Jennifer. 2002. Speaker meaning, what is said, and what is implicated. *Noûs* 36: 228–248.
- Searle, John. 1975. Indirect speech acts. In *Syntax and semantics, 3: Speech Acts*, eds. Peter Cole and Jerry L. Morgan, 59–82. New York: Academic Press. Reprinted in *Pragmatics: A reader*, ed. Steven Davis, 265–277. Oxford: Oxford University Press. (1991).
- Seuren, Pieter A.M. 1988. Presupposition and negation. *Journal of Semantics* 6: 176–228.
- Seuren, Pieter A.M. 1990. Burton-Roberts on presupposition and negation. *Journal of Linguistics* 26: 425–453.
- Siefring, Judith, ed. 2004. *The Oxford dictionary of idioms*. Oxford: Oxford University Press.
- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.
- Sperber, Dan, and Deirdre Wilson. 1987. Précis of relevance: Communication and cognition. *Behavioral and Brain Sciences* 10: 679–754.
- Sperber, Dan, and Deirdre Wilson. 1995. *Relevance: Communication and cognition*, 2nd ed. Oxford: Blackwell.
- Van der Sandt, Rob A. 1991. Denial. *Papers from the Chicago Linguistics Society: The Parasession on Negation*, 27(2): 331–344.
- Van der Sandt, Rob A. 2003. Denial and presupposition. In *Perspectives on dialogue in the new millennium*, eds. Peter Kühnlein, Hannes Rieser and Henk Zeevat. Amsterdam: John Benjamins.
- Van der Sandt, Rob A. 2006. Negation: Semantic aspects. In *Encyclopedia of language and linguistics*, 2nd edn, ed. Keith Brown, 570–577. Amsterdam: Elsevier.

- Walker, Ralph C.S. 1989. Review of relevance: Communication and cognition, by Dan Sperber and Deirdre Wilson. *Mind and Language* 4: 151–159.
- Wilson, Deirdre. 1975. *Presupposition and non-truth-conditional semantics*. New York: Academic Press.
- Wilson, Deirdre, and Dan Sperber. 1981. On Grice's theory of conversation. In *Conversation and discourse*, ed. Paul Werth, 155–178. New York: St. Martins Press.
- Wilson, Deirdre and Sperber, Dan. 1986. Inference and implicature. In *Meaning and interpretation*, ed. Charles Travis, 45–76. Oxford: Basil Blackwell. Reprinted in *Pragmatics: A reader*, ed. Steven Davis, 377–393. Oxford: Oxford University Press. (1991).
- Wilson, Deirdre, and Dan Sperber. 2004. Relevance theory. In *The handbook of pragmatics*, ed. Laurence R. Horn, and Gregory Ward, 607–663. Oxford: Blackwell Publishing.
- Yus, Francisco. 2006. Relevance theory. In *Encyclopedia of language and linguistics*, 2nd edn, ed. Keith Brown, 512–519. Amsterdam: Elsevier.
- Zwicky, Arnold M., and Jerrold Sadock. 1975. Ambiguity tests and how to fail them. In *Syntax and semantics*, 4, ed. John P. Kimball, 1–36. New York: Academic Press.

The (in)Significance of the Referential-Attributive Distinction

Anne Bezuidenhout

Abstract This paper tackles the question as to whether or not the referential-attributive (RA) distinction has any information-structural significance. That is, does this distinction mark a contrast between different ways a speaker might package the informational content of utterances that use definite descriptions? Furthermore, are there any overt markers of this distinction? In this paper I focus on referential and attributive uses of definite descriptions, and leave the issue of indefinites for another occasion. I answer both of the above questions negatively. While definite descriptions have an information-structural role to play, the RA distinction does not. Moreover, after examining several potential candidates for markers of the RA distinction, I conclude that there are no such overt markers. This is in fact to be expected if one accepts my proposal to see referential and attributive uses of definite descriptions as having the same general function, namely to single out something as a center of interest. The difference is simply that referential uses focus on role bearers, whereas attributive uses focus on role properties. Which of these the speaker intends will depend on context and hearers will have to rely on contextual assumptions to recover the intended message.

1 Introduction

Almost a half-century ago, Donnellan (1966) introduced a distinction between referential and attributive uses of definite descriptions. His intention was to show that the two main contenders for a theory of descriptions, namely those of Russell (1905) and Strawson (1950), were at best incomplete, since neither could account for both uses. The philosophical debate has raged ever since, with no clear winner

A. Bezuidenhout (✉)
University of South Carolina, Columbia, USA
e-mail: anne1@sc.edu

emerging. Moreover, it isn't even established what kind of claim Donnellan was making. Was he making an ambiguity claim about English 'the' or simply appealing to facts about pragmatic usage? Could a univocal semantics coupled with a Gricean theory of conversational implicature account for the facts just as well? If we do accept a univocal semantics for 'the', should this be a referentialist or a quantificational one? Perhaps instead we should allow that definite descriptions are semantically underspecified?

One thing that is striking about this philosophical debate is how detached it is from empirical considerations of the sort that linguists make use of in their semantic and pragmatic theorizing. For example, definite descriptions are frequently discussed without any contrastive reference to indefinite ones, one notable exception being Ludlow and Segal (2004). And when the contrast is noted, almost no philosopher discusses the fact that definiteness and indefiniteness is not marked in the same way in all natural languages. For example, there are many languages that have no definite or indefinite articles. Moreover, even when languages do have articles to mark the definite/indefinite contrast, it is by no means clear that patterns of usage conform to those in English. There are corpus studies (for example comparing translations between English and Norwegian) that show that speakers of different languages use articles in pragmatically different ways. Not every English use of a definite expression has a Norwegian definite as its translational equivalent. For details, see Fretheim (2010).

In the linguistics literature, the definite/indefinite contrast is most often discussed in the context of issues having to do with discourse and information structure. Definites are said to be associated with information that is entailed by common ground information or somehow readily identifiable (i.e., information that is "given"), whereas indefinites introduce new information. Even though it is by no means clear that the definite/indefinite contrast lines up perfectly with this given/new contrast, it is certain that definites and indefinites play different roles in "conversational tailoring", i.e., in how speakers organize and present information to their interlocutors in typical real-time conversational exchanges.

The philosophical literature on the referential/attributive (RA) distinction makes almost no reference to this literature in linguistics on information structure of discourse. And on the other side, linguists concerned with the information structure component of natural language rarely worry about the philosophers' RA distinction.¹ I will attempt to illuminate both the RA distinction and discussions about discourse structure by connecting these two issues. The RA distinction cannot by itself settle the question as to the correct semantics for definite descriptions, since expressions of many sorts (e.g., proper names, indexicals, quantifier phrases) can be used either referentially or attributively, yet we would not want to say they all belong to the same semantic category. For a defense of this

¹ There are linguists who have addressed the issue of the RA distinction, such as Rouchota (1992). However, they tend to debate philosophers on the philosophers' turf, which I have indicated does not engage with issues about the information structure component of language. Fretheim (2010) is an exception.

claim, see Bezuidenhout (1997) and for a different slant see Nunberg (2004). However, an investigation of the differences in discourse function between referential and attributive uses of descriptions (both definite and indefinite) has the potential to throw new light on discussions about discourse structure and on the (interminable) philosophical debate about definite descriptions.

In this paper I will first address the issue as to the pragmatic or discourse functions of referential and attributive uses of descriptions (Sect. 2) and then ask what differences these make in the structuring of information in conversational exchanges (Sect. 3). Here my focus will be on definites rather than indefinites. I will ultimately conclude that both referential and attributive uses of definite descriptions fall under a more general category of usage, and that the RA distinction per se has no information structural significance.

2 The Function of Attributive Versus Referential Uses of Descriptions

As noted above, there has been a fair amount of discussion in the linguistics literature about the definite/indefinite contrast and about the information-structural (I-S) role of this contrast, but not much about the I-S role of the RA contrast. The definite/indefinite contrast has been posited to line up (more or less) with the given/new contrast. So one way to think about the issue I'm concerned with is to ask how to accommodate the RA contrast into the classification scheme illustrated in Table 1.

The contrast between definites and indefinites is frequently illustrated by a simple stereotypical discourse, such as (1), which begins with the use of an indefinite expression, such as 'a man', and then continues in the next discourse segment with the use of a definite expression, such as 'the man':

(1) A man entered the room. The man sat down.

The claim is that the indefinite signals new information and serves to introduce a new discourse entity. This discourse entity can then be picked up as a referent in a subsequent discourse segment by the use of an appropriate definite expression. Exactly how this works is not my main concern in this paper. In the literature this process is frequently characterized in metaphorical terms. For example, Heim

Table 1 How the given/new and definite/indefinite contrasts intersect

	Given/old information (either discourse- or hearer-old)	New information (both discourse- and hearer-new)
Definites	<u>The man</u> sat down	<u>The first person to walk on Mars</u> will be an accomplished musician
Indefinites	<u>A certain person</u> wants to meet you at the bar	<u>A man</u> entered the room

(1982) treats indefinites as instructions to open a new mental “file card”, which represents the discourse entity that has just been introduced. The associated descriptive information is then entered onto that card. For example, ‘a man’ creates a file card on which the information [+male], [+human], [+adult] is recorded (along with the information presented by the following predicate ‘entered the room’—which in turn contains a definite description that will require accessing a file card for an already represented discourse entity). Subsequent uses of definite expressions signal the need to search for an existing file card and to enter any additional information onto the appropriate, already existing file card. Givón (1995) in a similarly metaphorical vein talks of indefinites as signals to add nodes to hierarchical tree structures and of definites as instructions to search an appropriate tree structure (mental model) for an antecedent node in which to ground the current reference and to update the information associated with that node.

The use of the definite expression signals that the entity it refers to is already “given”—that is, that it belongs to the set of currently cognitively available discourse entities. A cognitively available entity may be one that has been previously mentioned in the discourse, one that is mutually perceptually available in the current discourse context, or one whose representation can be retrieved from memory without undue effort. In the third case, the entity might better be described as inferable rather than given. However, for the purposes of this paper these fine-grained distinctions don’t much matter. For more details, consult Clark and Haviland (1977), Heim (1982), Givón (1995), Kadmon (2001), Lambrecht (1994) and Prince (1981).

Of course, things are not quite as simple as this suggests, and this too is widely recognized. Abbott (2004) does an excellent job of summarizing the facts about the definite/indefinite contrast and the extent to which it does or does not line up with the given/new contrast. One thing is clear, however, namely that definite expressions can serve to introduce new information and indefinites can be used to refer to entities that are already given. The examples in Table 1 illustrate the point.

Firstly, a definite description such as ‘the first person to walk on Mars’ can be used discourse initially, even though the hearer has no prior cognitive fix on a relevant entity. Simply by understanding the descriptive information (*person-walking-on-Mars*) and combining this with the ordinal information (*first*), the hearer can acquire a cognitive fix on a unique discourse entity. This is not to deny that some prior information that is mutually manifest to the conversational partners will be required. Such information will be needed, for example to understand what counts as a person for present purposes (e.g., do robots whose AI surpasses ordinary human intelligence count as persons?), or to understand what counts as walking for present purposes (e.g., does it count as walking on the surface of Mars if you are inside a robot that is walking on the surface of Mars but no actual human leg movements are required to move the robot?). The point just is that no representation of a discourse entity corresponding to the definite description needs to be already in the mutual cognitive environment of the interlocutors (what I’ll call the common ground from here on out). Fraurud (1990) has a sustained discussion of the class of “first mention” definite descriptions. Frazier (2006) reports

experimental data on the processing of novel definites. Poesio and Vieira (1998) report on a corpus study, based on articles from the Wall Street Journal. They found that a large percentage of uses of definite descriptions introduce novel discourse entities.

Secondly, an indefinite expression such as ‘a certain person’ can be used to refer to an entity that has the discourse status of a given entity. For example, suppose Jack and Jill are at a cocktail party. They are gossiping about one of Jack’s friends who Jill badly wants to date. Later in the evening Jack approaches Jill, who is now talking to a group of their work colleagues, who also know the friend but are unaware of Jill’s crush. In order not to betray her secret to their work colleagues, Jack says to Jill: ‘A certain person wants to meet you at the bar’. She will know that he is referring to the friend who they were talking about earlier.

Turning now to the question of how to fit the RA distinction into the scheme of Table 1, we need to address the question as to the discourse functions of these two types of uses of definite expressions. When philosophers want to illustrate the RA distinction, they come up with rather contrived descriptions of circumstances to show how a single sentence containing a definite description (‘Smith’s murderer is insane’) can be interpreted either attributively or referentially, depending on the background context. And these examples certainly do appear to illustrate a real difference in function. But what exactly are these two functions? And how frequently and for what conversational purposes do ordinary speakers use descriptions according to one function rather than the other?

The function of referential uses seems fairly clear. I favor an account offered by Strawson (1971) according to which referential uses provide the hearer with identifying information that allows him/her to select an appropriate entity from among a set of currently available discourse entities. Strawson calls such an entity an “antecedent centre of interest” (Strawson 1971: 90–91). He argues that conversation in general is governed by several principles (Strawson 1971: 76, 92–3), including:

- (PPI) The Principle of the Presumption of Ignorance
- (PPK) The Principle of the Presumption of Knowledge
- (PR) The Principle of Relevance

The PPI articulates the idea that one goal of conversation (not the only one of course) is to give information to one’s audience. Since one can’t inform someone about what he or she already knows, this means that there is a sense in which a speaker must presume ignorance of certain matters on the part of her audience (Strawson 1971: 76). The PPK articulates the idea that the information content of a statement isn’t in general something that stands by itself. The speaker intends the information conveyed by her statement to be processed in a context in which her audience is already presumed to know certain things (Strawson 1971: 92). Strawson gives as an example the fact that when a speaker uses a referring term, she invokes identifying knowledge on the part of the audience of the thing referred

to by the referring expression (Strawson 1971: 92–3; see also 78–9). With respect to the PR he says that we don't "direct isolated and unconnected pieces of information at each other, but on the contrary intend in general to give or add information about what is a matter of standing interest or concern" (Strawson 1971: 92).

These three principles give us some idea of how we can identify a center of interest. The PR dictates that it must be something that is connected to "matters of interest" to the conversational participants. Strawson qualifies this by saying that these will be matters of "standing" interest, and presumably this qualification is intended to cover a range of different cases.² The PPI and PPK also help us to understand what a center of interest might be. Matters of interest will include things about which we desire more information (information that we currently lack—by the PPI) and these things we desire information about will be things that are already known to us in some way or other (by the PPK). This prior knowledge could be merely dispositional. For instance, perhaps we have identifying knowledge of something but we are not currently thinking about that thing. However, this identifying knowledge can be invoked by a speaker's use of an appropriate referring expression, in this case a definite description. The thing will in this way be brought to the center of attention. Matters of interest can be conceptualized as a cognitive domain and a center of interest will then be an element in the domain whose representation is currently in the spotlight of attention.

For more contemporary views about how discourse referents are accessed and on the role of an entity's information status in determining its accessibility, see Ariel (1990), Givón (1995), Grosz et al. (1995), Gundel (1996), Gundel (2012), Gundel et al. (1993). There is also work in computational linguistics and in psycholinguistics on the processing of definites. It is common in this work to assume that definite descriptions are anaphoric expressions whose reference must be determined by searching for an antecedently introduced discourse entity, an assumption that we've already seen is problematic, as it ignores discourse-initial uses of definite descriptions. For a small sampling, see Almor (1999), (2000), Gordon and Hendrik (1998), and Vieira and Poesio (2000).

Most linguists, especially those concerned with the given/new contrast and the I–S role of definites, simply assume that the function of definites is referential. So looking to the literature in linguistics is not much help in determining the function

² I suggest that we would need to make at least the following distinctions: The matter of interest might be one of long standing or one that is conversation-specific. And it might be a latent interest or one that has been explicitly addressed in the prior conversation. Since these two distinctions crosscut one another, there are four possible types of standing interest. The latter distinction (latent vs. explicit) is analogous to the distinction that has been made between hearer-old vs. discourse-old information. The notion of a standing interest and the notion of old information are not synonymous, since not all old information amounts to something of standing interest.

of attributive uses of definite descriptions.³ I propose that attributive uses are ones in which the speaker means to focus on a *role property* rather than a *role bearer*. Consider the following example:

(2) The mayor ought to be impeached.

We are to imagine that we're driving through a small town along its main street, which is full of potholes. We have no idea who the mayor of the town is. Nevertheless, there is a shared background assumption that it is one of the duties of a mayor of a small town such as this to take care of its roads, by imposing and collecting city taxes, hiring road crews, etc. (Of course, in a bigger city it is unlikely to be part of the duties of a mayor to take care of these things. There would be a city works manager to hire work crews and decide which roads are in need of repair, and there'd be a budget officer to handle tax collection and spending). The point is that the speaker is not intending to refer to the actual role occupier but rather intends to say that whoever currently possesses that role property deserves to be impeached. So (2) in its attributive use can be paraphrased as follows:

(2') The functionary possessing the role property of mayor of this town ought to be impeached.

This should make it clearer that the focus is on the role, not the role occupier. Of course, it is presupposed by speaker (and so by the hearer) that there is an occupier of the role. But it is not the occupier *qua individual entity* that is the center of interest. Rather, it is the functionary—someone or something with *that role property*—that is the center of interest. If one found out that the role occupier was a committee whose joint decisions and actions were construed as the mayoral decisions and actions, or if one discovered that a robot had been programmed to make the decisions and actions allowed by the role, this would not make one retreat from the claim that the functionary possessing the role property of mayor of the town ought to be impeached.

In other words, the speaker is interested in focusing on a certain role property and information about the characteristics of the individual who happens to currently occupy that role is irrelevant to the current conversational purposes, except inasmuch as this information tends to undermine the assumption that the role property *being current mayor of this town* is instantiated in the conversational context. For example, if it turns out that the “town” is actually a facade created for a movie set and the “mayor” is an actor pretending to be a mayor of a town, then

³ Russellians would of course say that attributive uses are to be given a quantificational analysis. However, in the text I am not talking about the correct semantics for definite descriptions but about the pragmatic function of such descriptions when used attributively. I have already expressed skepticism that the mere fact that descriptions can be used attributively settles anything regarding the semantics of definites. But debating this issue here would take me too far afield.

the role property the speaker intended to bring into focus is not instantiated and his/her conversational move is undermined.⁴

Thus to sum up the last few paragraphs, I am proposing that referential uses of definite descriptions have the function of invoking identifying knowledge of a particular individual and are intended to help the hearer pick out a (contextually) unique entity from among a set of discourse entities. Attributive uses of definite descriptions on the other hand are intended to invoke identifying knowledge of a role property and to help the hearer to focus on the contextually instantiated role property the speaker has in mind.

3 How is the Distinction Between Referential and Attributive Uses Marked?

I turn now to the question as to how speakers mark the RA distinction in ordinary conversation. Are there any overt markers of this distinction that a hearer might be cued to? In the philosophical literature that was spawned by Donnellan (1966), the examples discussed are invariably simple declarative sentences, such as ‘Smith’s murderer is insane’, ‘The man drinking a Martini is a history professor’, etc. And the whole point is that a single sentential form can be understood either referentially or attributively, depending on context. This would suggest that there is no overt marker of the RA distinction.

One might think that definite descriptions that entail uniqueness, such as (3) and (4) below, can only have attributive uses, so the use of expressions such as ‘first’, ‘greatest’, ‘highest’, etc. can be regarded as overt markers of attributive uses:

- (3) The first person to walk on Mars will be an accomplished musician.
- (4) The first boy in line will get the prize.

It may be a little hard to imagine a referential use of the definite description in (3) (mainly because it is hard—but not impossible—to construe (3) as a remark about a particular identifiable person), but (4) can certainly be understood referentially, as shown by the dialog in (4’):

(4’) Anne: Can you point out who’ll get the prize? Bob: The first boy in line will get the prize.

Assume that Anne and Bob are parents attending a school awards ceremony and the students are lined up along the side of the hall in which the ceremony will be held. In (4’), Bob uses ‘the first boy in line’ as a way of invoking perceptual identifying knowledge in Anne of the boy in question. This allows him to single

⁴ I discuss the issue of presupposition failure in more detail in Bezuidenhout (2010) and in an unpublished paper titled ‘Presupposition failure and assertoric inertia’, which draws inspiration from Abbott (2000), von Stechow (2004), and Horn (2002).

out the boy without any physical gesture, which perhaps would seem rude in the circumstances. So, expressions such as ‘first’ are not after all markers of attributive use.⁵

Another possible suggestion is that the richer the descriptive material contained in a definite description, the more likely it is that the description is being used referentially. Consider the difference between using the minimal description ‘the boy’ as opposed to the richer description ‘the boy on the left with the baseball cap turned backwards’. In the latter case there is a stronger presumption that the speaker is trying to provide identifying information that will aid the hearer in picking out the intended referent from amongst a set of potential discourse referents. This tendency—for interpreters to treat additional information as intended for identificatory purposes—can be seen in the results of some experimental work in psycholinguistics. Research on the processing of so-called “garden path” sentences with reduced relative clauses, such as ‘The horse raced past the barn fell’, has shown that people opt initially for an interpretation according to which ‘raced’ is taken as a main clause verb in the past tense, rather than as a past participle in a subordinate relative clause.⁶ However, if such a garden-path sentence is processed in a context where there is more than one horse, and hence a need to use a relative clause to distinguish one horse from another (say to distinguish the horse that was raced past the barn from the horse that was raced across the field), then people are less likely to be garden-pathed. For a discussion of the processing of garden-path sentences and the ameliorating effects of context, see Crain and Steedman (1985), Spivey-Knowlton et al. (1993) and Pickering and Traxler (1998). For a more skeptical view, see Binder et al. (2001).

However, whether or not there is a tendency to take such additional information as supplying identifying information for picking out a discourse referent, the quantity of descriptive information in a definite description cannot by itself be taken as a marker of referential use. This is because there is no reason to deny that the addition of descriptive information could also help when a definite description is used attributively to focus on a role property. For example, take (2) above to be uttered in a context in which we’ve driven through two towns, presumably both with mayors, and that one town had beautifully maintained streets whereas the other had streets riddled with potholes. In such a context it may be necessary to add descriptive information to enable the hearer to discern the contextually relevant role property—it is the role property of *being mayor of the town with crumbling infrastructure* that is at issue.

⁵ To get an attributive use of (4), imagine a scenario in which the US President has declared a competition to find the most enterprising 5th grader. After many elimination rounds, the finalists have been flown to Washington DC and are assembled on the Rose Garden lawns. The President makes an announcement as to who the winner will be: “The first boy in line will get the prize”. There is a frantic scramble as the boys fight to be first in line. The President’s utterance here seems to be focused on a role property rather than a role occupier.

⁶ For a cartoon about garden path sentences, go to <http://www.qwantz.com/> and search for comic number 204.

So we seem to be back to the position that there are no overt markers of the RA distinction. However, even if declarative sentences with definite descriptions do not overtly signal one reading over the other, there may be linguistic forms that *are* biased to one or the other reading. For example, consider the following sentences, using left dislocation, contrastive stress, it-clefting and pseudo-clefting respectively. For a discussion of the information structural role of such constructions, see Ward and Birner (2004).

- (5) As to the boy, he will get the prize.
- (6) THE BOY will get the prize; HIS PARENTS will receive a commendation.
- (7) It is the boy who will get the prize.
- (8) Who will get the prize is the boy.

In all of (5)–(8), there is a presupposition that there is a set of referents from amongst which the boy is being selected. In other words, there is a set of discourse entities, and the boy is selected from that set, as being what (Strawson 1971) called the current “center of interest”. So, it is easy to get referential readings for (5)–(8), and moreover, the fact that the speaker is using some sort of focusing device (left-dislocation, contrastive stress, clefting) makes it even more likely that the speaker’s intention is to use the definite descriptions referentially.

Still, even though the default readings may be referential, perhaps we can imagine scenarios in which the intention is to use the definite descriptions in (5)–(8) attributively? Suppose that the US President has declared a competition to look for the ideal American family who can be used for an advertising campaign. There is prize money to be awarded. The winning family must consist of a mother, a father, and two children under the age of ten, one being a boy and the other a girl. In such a context, it may seem that we can get attributive readings, especially if we add the phrase ‘whoever he/they may be’, as in (5′)–(8′) below:

- (5′) As to the boy, whoever he is, he will get the prize.
- (6′) THE BOY, whoever he is, will get the prize, but HIS PARENTS, whoever they are, will receive a commendation.
- (7′) It is the boy, whoever he is, who will get the prize.
- (8′) Who will get the prize is the boy, whoever he is.

However, it seems just as plausible to think of these as involving a straightforward referential use, where the referents differ only in their degree of vividness from the referents in the original examples, which, we can suppose, were entities that speaker and hearer were personally acquainted with. The background scenario assumed for (5′)–(8′) (namely, the one about the President’s announcement of a search for the ideal family) defines a set of discourse referents, namely a family consisting of two parents and two children. Then subsequent uses of definite descriptions select referents from amongst this set. It is true, we don’t know too much about these referents (beyond their biological relations to each other and their genders). But what we know is sufficient for us to be able to tell them apart from each other, and that is all the identifying information needed in this context to get a cognitive fix on the entities in question.

In other cases we may need richer identifying information to pick out an entity from amongst a set of discourse entities in order to make it the center of interest. In some cases, for example, the speaker may rely on the hearer's ability to survey the visual scene or expect the hearer to retrieve information from memory rather than simply relying on prior linguistic context. But although this introduces more complexity into the task of identifying a center of interest, the mechanism at some general level is the same. There is a set of potential discourse referents, and the focusing constructions point to which of these entities is currently the center of interest.⁷

So, although it may seem that we've isolated attributive uses of definite descriptions in (5')–(8'), this is not so. What we have in fact uncovered is a continuum of cases where the identificatory task varies in complexity, depending on the amount and kind of information needed to get a cognitive fix on the entity that is currently the center of interest. Sometimes purely linguistic information will do. In other cases, perceptual or memorial information may be needed in addition. It is also worth noting that the phrase “whoever he/she/it/they may be”, which is often taken as a marker of attributive use, would turn out to be no such thing. If the argument of the last couple of paragraphs stands, this qualifying phrase may simply mark the fact that the referent is not a vivid object of acquaintance, but one that is identifiable on purely linguistic grounds, or other minimal information available in the context.

If (5')–(8') are not truly attributive uses, then can we conclude that it-clefting and these other non-standard syntactic forms are overt markers of referential use? Not quite yet, since it seems that such information structuring devices can be used to bring a role property into focus, just as much as (5)–(8) bring a role occupier into focus. Take (2) again, repeated here, and a possible follow-up comment from someone who disagrees with the initial speaker's assumption that the task of repairing roads is one of the duties of a mayor:

(2) The mayor ought to be impeached.

(9) It is the city works manager who should be impeached.

In this context, (9) carries the presupposition that there is a set of competing role properties (the properties of being mayor, of being city works manager, being

⁷ Kent Bach has in several places, e.g., (Bach 2004), railed against the idea of a discourse referent. Bach appears to believe that discourse referents have a different ontological status from “real” ones, existing only in discourse representation structures rather than being worldly entities. However, this is not so. While discourse referents are indeed *represented* in discourse models, they may be real-world entities (if the discourse is about the real world). We need to distinguish the entity from its representation. This is equally true of a referent in Bach's favored sense. Even if we restrict ourselves to real world entities in our talk (although it isn't at all clear that we in fact do restrict ourselves in this way or why we should be so restricted), Bach still has to grant that these real world entities have to be cognitively represented somehow for us to think and talk about them. We don't use the entities themselves as discourse representations.

district council member, etc.) and the it-cleft construction picks out one of these properties from the set.

As I argued in Sect. 2 of this paper, attributive uses of definite descriptions can be thought of as cases where the role rather than the role-occupier is the center of interest. On this way of looking at things, there is a more general picture that unites referential and attributive uses under one umbrella. From the information structural perspective, definite descriptions always have the function of singling out a center of interest. Whether it is a role occupier or a role property that is the center of interest is typically determined on the basis of shared (common ground) assumptions that are operative in the conversational context.

We've seen that when an entity (or role-occupier) is the center of interest, more or less complex descriptive information may be needed for the task of identifying the intended referent, and the hearer may be expected to use information available from memory or perception in addition to linguistic information in order to complete the task.

In cases where it is the role that is the center of interest, there will also be a range of cases where more or less descriptive information is needed to focus on the relevant role. In some cases it may be that the descriptive phrase by itself provides all the identifying information needed to pick out the role that the speaker intends to place at the center of attention. Frequently, however, additional contextual information will be needed to single out the relevant role, since speakers typically use incomplete definite descriptions. This is the case with my example (2) above, which used the definite description 'the mayor'. In the imagined context, the hearer had to figure out that it was the role property of *being the mayor of this town that we are currently driving through* that was the center of interest. A similar remark can be made about the example of an attributive use of 'the first boy in line' mentioned in note 5. The President's use of this definite description will be understood in context to single out the role property of *being the boy from among the group of finalists then assembled on the lawns in the Rose Garden who is first in line*.⁸

Thus, whether we are talking about referential or attributive uses of definite descriptions, there is always something that the speaker is trying to bring into focus as the center of interest. However, the speaker will provide differing amounts of explicit descriptive detail and the hearer will be expected to retrieve differing amounts and kinds of implicit identificatory information. And this applies whether the center of interest is a role occupier or a role property.

⁸ And of course there is always the need for additional background information to fully understand what the speaker said. In the mayor case, it is knowledge about the tasks of a mayor of a small town. In the case of the first boy in line, context is needed to understand where the head of the line starts. Perhaps the President used a hand gesture to indicate the head of the line. Or perhaps a spot had already been determined in advance by the placement of a sign.

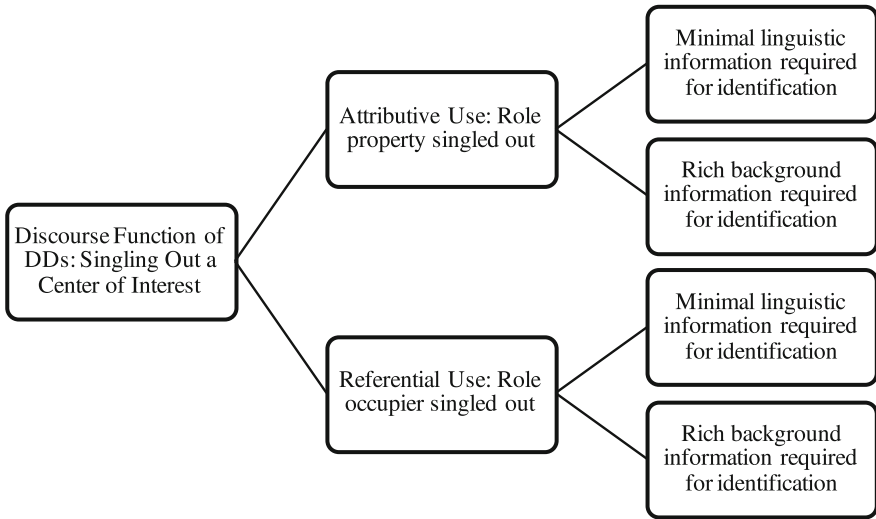


Fig. 1 The discourse functions of definite descriptions

4 Conclusions

In this short paper my aim was to see to what extent the philosophers’ distinction between referential and attributive uses of definite descriptions can be made to mesh with discussions in linguistics about the information structure component of natural languages. I can summarize my conclusions by reference to Fig. 1.

An examination of the central kinds of examples used to illustrate the RA distinction (as it applies to definite descriptions) led me to propose that the function of referential uses of definites is to single out a role occupier as the center of interest. On the other hand, the function of attributive uses of such descriptions is to focus on a role property. This led me to propose that these uses can be united under a more general category, namely the category of expression whose function it is to single out a center of interest (in the Strawsonian sense that was spelled out in Sect. 2). This would apply even for cases of definites used discourse initially.

I noted that the sorts of non-standard syntactic constructions (such as it-clefting) that have been shown to play a role in structuring information can be used both with referentially and attributively used definite descriptions. This means that the RA distinction *per se* has no information structural significance. What has significance from the I–S perspective is the general function of singling out a center of interest. Whether that center of interest is a role occupier or a role property is not marked in any special way in discourse structure.

One point about markers (or lack thereof) in particular is worth repeating, as it will be the one that philosophers of language are likely to zero in on. I argued that the qualifying phrase “whoever he/she/it/the may be”, which is often assumed to be one indicator of an attributive use of a definite description, is in fact no such

thing. Such a qualifying phrase can equally be a marker of the fact that the intended referent(s) are not entities of personal acquaintance to either the speaker or the hearer, and hence are things that are to be cognitively identified merely in some “thin” or minimal way.

I gave illustrations to support my claim that the quantity of descriptive information a speaker uses can vary, both when the speaker’s intention is to single out a role occupier and when the intention is to focus on a role property. Sometimes all that is needed to single out a center of interest is minimal linguistic information. On the other hand, usually perceptual or memorial information will be needed in addition to purely linguistic information. The important point is that even fairly minimal information can suffice to allow a hearer to identify a discourse entity from among a candidate set of entities (as in the case of the ideal family example discussed in Sect. 3). Similar remarks apply to cases where the speaker is trying to single out a role property. Minimal linguistic information may be all that is needed.

I also reiterate a point made at the outset, namely that this paper is concerned with the information-structural role of the RA distinction, and remains neutral on the question as to what the correct semantics for definite descriptions is, as well as on the issue as to whether the RA distinction marks a semantic ambiguity or merely a pragmatic duality of use. The mere fact that we recognize a RA distinction does not settle the semantic question, as the RA distinction is compatible with a semantic ambiguity account as well as with an account that treats all uses of definite descriptions in a semantically univocal way. Furthermore, if one treats definite descriptions univocally, there is still a further issue as to whether the semantics is referentialist or quantificational or one that argues for semantic underspecificity (the main alternatives that have been argued for in the literature).

Finally I should note that the discussion in this paper is incomplete, as I have not tackled the issue of indefinite descriptions or how to incorporate the RA distinction into what we know about the I–S role of indefinites. Indefinites have referential and attributive uses just as much as definites do, as argued by Bezuidenhout (1997). Thus my aim of showing how the RA distinction fits into the scheme of Table 1 is not yet complete. Moreover, as far as indefinites are concerned, people have distinguished referential from specific uses of indefinites. Consult Abbott (2004) for a survey of these usage facts. So this would add another layer of complexity in trying to figure out how the RA distinction intersects with attributive, referential and specific uses of indefinites. This is a task that I leave for another occasion.

References

- Abbott, B. 2000. Presuppositions as non-assertions. *Journal of Pragmatics* 32: 1419–1437.
Abbott, B. 2004. Definiteness and indefiniteness. In *Handbook of pragmatics*, eds. L. Horn, and G. Ward, 122–149. Oxford: Blackwell.

- Almor, A. 1999. Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review* 106: 748–765.
- Almor, A. 2000. Constraints and mechanisms in theories of anaphor processing. In *Architectures and mechanisms for language processing*, eds. M. Pickering, C. Clifton, and M. Crocker, 341–355. Cambridge: Cambridge University Press.
- Ariel, M. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Bach, K. 2004. Descriptions: Points of reference. In *Descriptions and beyond*, eds. M. Reimer, and A. Bezuidenhout, 189–229. Oxford: Oxford University Press.
- Bezuidenhout, A. 1997. Pragmatics, semantics and the referential/attributional distinction. *Mind* 106: 375–410.
- Bezuidenhout, A. 2010. Grice on presupposition. In *Meaning and analysis: Themes from H. Paul Grice*, ed. Klaus Petrus, 75–102. Basingstoke: Palgrave MacMillan.
- Bezuidenhout, A. Presupposition failure and assertoric inertia. Unpublished ms., University of South Carolina.
- Binder, K., S. Duffy, and K. Rayner. 2001. The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language* 44: 297–324.
- Clark, H., and S. Haviland. 1977. Comprehension and the given-new contrast. In *Discourse production and comprehension*, ed. R. Freedle, 1–40. New York: Ablex.
- Crain, S., and M. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In *Natural language parsing: Psychological, computational and theoretical perspectives*, eds. D. Dowty, L. Karttunen, and A. Zwicky, 443–467. Cambridge: Cambridge University Press.
- Donnellan, K. 1966. Reference and definite descriptions. *Philosophical Review* 75: 281–304.
- Von Stechow, P. 2004. Would you believe it? The king of France is back! (Presuppositions and Truth—Value Intuitions). In *Descriptions and beyond*, eds. M. Reimer, and A. Bezuidenhout, 315–341. Oxford: Oxford University Press.
- Fraurud, K. 1990. Definiteness and the processing of noun phrases in natural discourse. *Journal of Semantics* 7: 395–433.
- Frazier, L. 2006. The big fish in a small pond: Accommodation and the processing of novel definites. Paper presented at the OSU Presupposition Accommodation Workshop, Oct 13–15, in Columbus, Ohio.
- Fretheim, T. 2010. Demonstratives and reference to individuated objects vs. reference to properties of objects: A contrastive analysis. In *The Role of Data at the Semantics-Pragmatics Interface*, eds. E. Nemeth, and K. Bibok, 95–121. Berlin: De Gruyter Mouton.
- Givón, T. 1995. Coherence in text vs. coherence in mind. In *Coherence in spontaneous text*, eds. M. Gernsbacher, and T. Givón, 59–115. Amsterdam: John Benjamins Publishing.
- Gordon, P., and R. Hendrik. 1998. The representation and processing of coreference in discourse. *Cognitive Science* 22: 389–424.
- Grosz, B., A. Joshi, and S. Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21: 203–226.
- Gundel, J. 1996. Relevance theory meets the givenness hierarchy: An account of inferrables. In *Reference and referent accessibility*, eds. T. Fretheim, and J. Gundel, 141–153. Amsterdam: John Benjamins & Co.
- Gundel, J. 2012. Pragmatics and information structure. In *The Cambridge handbook of pragmatics*, eds. K. Allan and K. Jaszczolt. Cambridge: Cambridge University Press.
- Gundel, J., N. Hedberg, and R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274–307.
- Heim, I. 1982. The semantics of definite and indefinite noun phrases, PhD diss., University of Massachusetts, Amherst.
- Horn, L. 2002. Assertoric inertia and NPI licensing. *Proceedings of the Chicago Linguistic Society* 38: 55–82.
- Kadmon, N. 2001. *Formal pragmatics*. Oxford: Blackwell.
- Lambrecht, K. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

- Ludlow, P., and G. Segal. 2004. On a unitary semantical analysis for definite and indefinite descriptions. In *Descriptions and beyond*, eds. M. Reimer, and A. Bezuidenhout, 420–436. Oxford: Oxford University Press.
- Nunberg, G. 2004. Descriptive indexicals and indexical descriptions. In *Descriptions and beyond*, eds. M. Reimer, and A. Bezuidenhout, 261–279. Oxford: Oxford University Press.
- Pickering, M., and M. Traxler. 1998. Plausibility and recovery from garden-paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 940–961.
- Poesio, M., and R. Vieira. 1998. A corpus-based investigation of definite description use. *Computational Linguistics* 24: 183–216.
- Prince, E. 1981. Toward a taxonomy of given-new information. In *Radical pragmatics*, ed. P. Cole, 223–256. New York: Academic Press.
- Rouchota, V. 1992. On the referential/attributive distinction. *Lingua* 87: 137–167.
- Russell, B. 1905. On denoting. *Mind* 14: 479–493.
- Spivey-Knowlton, M., J. Trueswell, and M. Tanenhaus. 1993. Context effects in syntactic ambiguity resolution: discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology* 47: 276–309.
- Strawson, P. 1950. On referring. *Mind* 59: 320–344.
- Strawson, P. 1971. Identifying reference and truth-values. In *Logico-linguistic papers*, 75–95. London: Methuen.
- Vieira, R., and M. Poesio. 2000. Processing of definite descriptions in corpora. In *Corpus-based and computational approaches to discourse Anaphora*, eds. S. Botley, and M. McEnery, 189–212. Amsterdam: John Benjamins.
- Ward, G., and B. Birner. 2004. Information structure and non-canonical syntax. In *Handbook of pragmatics*, eds. L. Horn, and G. Ward, 153–174. Oxford: Blackwell.

Quotation and the Use-Mention Distinction

Paul Saka

Abstract Quote marks, I claim, serve to select from the multiple ostensions that are produced whenever any expression is uttered; they act to constrain pragmatic ambiguity or indeterminacy. My argument proceeds by showing that the proffered account fares better than its rivals—the Name, Description, Demonstrative, and Identity Theories. Along the way I shall need to explain and emphasize that quoting is not simply the same thing as mentioning. Quoting, but not mentioning, relies on the use of conventional devices.

Quotation is one mechanism by which we can *mention*; as such, it is used for attributing exact words and thoughts to others; for distancing oneself from a given word choice (as in scare quotes); for indicating titles; for expressing irony (Sperber and Wilson 1981; Jorgensen et al. 1984; Groening 1996); and for explaining truth (the disquotation theory), meaning (truth-theoretic semantics), external negation (Horn 1989), and indirect discourse (Cappelen and Lepore 1997; Seymour 1994, and the inscriptional theory of Carnap 1937 and others). Finding out how quotation works, therefore, is highly important.

I shall suggest that every expression token (e.g. this particular inscription: cat) ambiguously or indeterminately refers to itself and to various items associated with it (including the inscription-*type* “cat”, the pronunciation /kæt/, the concept CAT, and the extension of cats). Quote marks—by which I mean double apostrophes as used in the USA, single apostrophes as used in Britain, double angles as used in parts of Europe, italicization, or any like conventional device—help to disambiguate the intended reference, although they are usually neither necessary nor sufficient for doing so.

This approach to understanding quotation is supported by its intrinsic plausibility and by the fact that it fares better than its rivals. In defending it, therefore, I shall summarize the extant theories of quotation and argue that they are

P. Saka (✉)

University of Texas-Pan American, Edinburg, TX 78539-2999, USA

e-mail: sakap@utpa.edu

inadequate. §1 and §2 rehearse the Name and Description Theories and their familiar flaws. §3 recounts the Demonstrative Theory of Donald Davidson and raises a new objection, that it cannot handle iterated quotation. §4 untangles the Identity Theory and raises an analogous objection, plus two others: one charging circularity and the other concerning the quotation of abstract types. §5 expounds my own account of quotation as the formal disambiguation of multiple ostension, and enumerates the ways in which it improves on the preceding theories.

1 The Name Theory

You can refer to anything, including a linguistic expression, by using its name; and you can assign a name to any object present by exhibiting a label along with an act of ostension. For example:

1. green eggs and ham

The label “(1)” now refers to the phrase “green eggs and ham”. In this case the label is well motivated. First—given the scholarly practice of using parentheses in the dubbing of expressions—the parentheses suggest (but do not entail) that an act of dubbing is at hand. Second, the numeral n within parentheses suggests (but does not entail) that the exhibited expression is the n th expression to be named in the present paper. Although these elements—the parentheses and numeral—possess mnemonic value, they are arbitrary from the point of view of compositional semantics. The meaning of “(1)” is not a function of the “meanings” of “(...)” and “1”. “(1)” could have referred, if my dubbing had been more haphazard, to any expression at all. “(1)” is not a description but a name.

In like manner, according to the Name Theory, quoted expressions name their referents. Tarski writes:

Quotation-mark names may be treated like single words of a language... the single constituents of these names... fulfil the same function as the letters and complexes of successive letters in single words. Hence they can possess no independent meaning. (Tarski 1933, p. 159)

In the same vein Quine writes:

From the standpoint of logical analysis each whole quotation must be regarded as a single word or sign, whose parts count for no more than serifs or syllables... The meaning of the whole does not depend upon the meanings of the constituent words. (Quine 1940, p. 26)

For Tarski, Quine, and scores of textbooks in logic, quotations are names.

1.1 *The Forward Productivity Problem*

Of course the Name Theory is an utter failure, since the quote mark is a systematically productive device that can be applied to expressions that we have never heard quoted before. This argument is so obvious that one wonders whether

Tarski and Quine could have seriously meant that quotations in natural language function just like proper names. Richard (1986) and Bennett (1988) think that they used “name” in the sense of a denoting phrase or singular term, and I would too if it were not for the fact that Tarski and Quine are both so clear about denying that quotations contain meaningful structure. Perhaps Quine meant that quotations function like names only in so far as mathematical logic is concerned, but this position does not appear to be Tarski’s, who claims that the Name Theory “seems to be the most natural one and completely in accordance with the customary way of using quotation marks” (1933, p. 160).

1.2 The Reverse Productivity Problem

Just as we can productively go from knowing any expression to knowing its quotation, we can go from knowing the quotation of any expression to knowing the expression itself. Consider Anscombe’s puzzle, which generalizes beyond personal names to all linguistic labels:

If I am told “That man’s name is ‘Smith’”, his name is mentioned, not used, and I hear the name of his name but not his name... [Hence] It is impossible to be told anyone’s name [or the word for anything]. (Anscombe 1957, p. 49)

The patent falsehood of the conclusion establishes the falsity of the premise, that to mention a name is to name the name.

1.3 The Simultaneity Problem

Another argument against the Name Theory is that expressions can simultaneously be used and mentioned. Quine (1943) notes that the mention of “Giorgione” in (2) provides for the interpretation of “so-called” while the use of the very same word provides for the interpretation of “his size”:

2. Giorgione was so-called because of his size.
3. Giorgione was called “Giorgione” because of his size.

This would pose a problem for the Name Theory, which imposes a rigid distinction between use and mention, but Quine tries to escape by claiming that (2) is elliptical for (3).

Quine not only fails to describe transformational procedures by which statements of type (2) might derive from statements of type (3), he fails to give any reason for supposing that the necessary kinds of principles even exist. Suffice it to say that other, more problematic, cases of simultaneous use-mention exist:

4. Quine says that quotation “... is weird”.

As Davidson (1979) observes, and Cappelen and Lepore (1997) emphasize, the material inside the quote marks (minus the ellipsis, to be strict) is mentioned in so far as I am attributing exact words to Quine. At the same time, the words are being used in so far as they form a predicate rather than a noun phrase, singular term, or name.

The generally acknowledged inadequacy of the Name Theory leads to the Description Theory.

2 The Description Theory

Linguistic expressions can be described in various ways:

EMPIRICALLY: The first sentence of *Tale of Two Cities*.

LEXICALLY (with or without an associated phrase structure): The sentence beginning with “It”, then “was”, then “the” ...

ORTHOGRAPHICALLY: The sentence beginning with *I*, then *t*, then space, then *w*, then *a*, then *s*,...

PHONOLOGICALLY: The sentence beginning with /I/, then /t/, then /w/, then /ə/, then /z/,...

The so-called structural descriptions—those based on the formal (orthographic/phonological) elements or the substantive (lexical) elements—are used in different versions of the Description Theory of Quotation.

According to the orthographic version of the Description Theory, the quotation of an expression describes, letter for letter, what is being quoted. Thus, “cat” = the expression formed by cee followed by ay followed by tee. (Tarski (1933) and Quine (1960) advocate this approach in addition to the Name Theory, and Richard (1986) in effect defends it as well.) This account relies upon naming in so far as cee is the name of “c”, ay is the name of “a”, and tee is the name for “t”, and so on for all characters used in the language. However, the Description Theory is not limited as the Name Theory is when it comes to constructing and interpreting quotations of novel expressions.

Geach (1957, §18) holds a lexical version of the Description Theory, according to which the quotation of an expression describes, word for word, what is being quoted. This improves on the orthographic version in that it explicitly treats “the cat in the hat” not just as a sequence of meaningless marks but as a string of language. At the same time, it is less plausible than the orthographic version in so far as it postulates a primitive name for every lexeme in the language.

2.1 Problems Revisited

Clearly, both versions of the Description Theory fall to the same arguments that defeated the Name Theory. (1) The Description theory fails to explain how

simultaneous use and mention is possible. (2) Although the Description Theory, especially the orthographic version, can productively deal with some novel expressions, it still cannot deal with foreign and non-linguistic marks. I can say “ α ” is a Greek letter, and you can understand me even if you had never encountered Greek before in your life. Further, as Searle (1969, p. 76) notices, I can say “The sound of a California jay is...” and you can understand me even though “...” cannot be described by any names for phonic, orthographic, or lexical items in any human language. Moreover, apropos the lexical theory, I can quote nonsense, both comic (5) and tragic (6):

5. “Twas brillig and the slithy toves/Did gyre and gimble in the wabe...”
6. Upon expiring, he said “ggrmph”.

Again, you would understand what I have said. Disputing the relevance of cited nonsense and animal sounds, Geach and Bennett respectively write:

Properly speaking, these are not quotations at all; we cannot *quote* sheer nonsense, we can only parrot it. (Geach 1957, §18)

One might occasionally put a scrawl or doodle between quotation marks, and be understood to have referred to its shape, but only as a joking extension of conventional quotation. (Bennett 1988, p. 405)

The former response nakedly asserts that there is a difference between quotation and parroting, which is precisely the point in contention between the Description Theory and the picture theories (to be discussed in the remaining sections of this paper). The latter response relies on the intuition that citation of scrawls is fundamentally humorous or non-serious. While I do not share this judgment myself, I can see why you would want to hold it if, like Bennett, you subscribe to a hybrid theory that combines elements of description and demonstration. According to Bennett, every quotation “X” means something like.

The type whose every token resembles, in respects $R_1 \dots R_n$, this: X.

Respects $R_1 \dots R_n$, are language-relative. In English, the relevant respects for identifying tokens as types involve (for instance) top–bottom orientation (“u” versus “n”), but not ink-color. Reference to $R_1 \dots R_n$ provides the Bennett account with descriptive content; reference to “this: X” provides it with demonstrative content. The theory avoids some of the problems of the pure Description Theory, but is susceptible to the same criticisms that apply to the “paratactic” Demonstrative Theory.

3 The Demonstrative Theory

We can refer to material objects via deictic demonstration or pointing. When we do this to the ink or soundwaves that make up linguistic tokens, we quote. This is the theory of Davidson (1979); his followers Partee (1973), Goldstein (1984), Garcia-Carpintero (1994), and Cappelen and Lepore (1997); and, uncited by them,

Christensen (1967). According to Davidson's version of the theory, quote marks refer to the shape of the quoted material.¹ Thus (7) is equivalent to (8).

7. "Cats" is a noun.
8. Cats. That complex of shapes is a noun.

In (8), *cats* is semantically inert while the demonstrative pronoun does the referring.² Likewise, in (7), *cats* is semantically inert while the quote marks do the referring. This approach explains how quotation can be productive, and it holds promise for explaining how simultaneous use and mention is possible (a promise on which Cappelen and Lepore (1997) begin to make good).

3.1 The Punctuation Problem

Washington (1992) claims that spoken quotation does not require quote marks, and that (7) can be read without saying "quote-unquote", without quotative intonation, and without finger-quote gestures. In effect, (9) is a grammatical and true sentence:

9. Cats is a noun.

The point can be made even stronger, I might add. Quote marks are often omitted in writing as well (contra Reimer 1996): it is downright *normal*, outside of scholarly writing, to exclude quote marks, especially in constructions like "The word cats is a noun"; and even in logic publications, where one might expect the greatest exactitude, it is common for quote marks to be omitted.

If the quoted word is semantically inert and thus can be taken out of the sentence, and if the quote marks are merely optional punctuation with no real semantic role, then (7) ought to be equivalent to (10).

10. Is a noun.

Yet it is not. Therefore, according to Washington, (7) cannot be analyzed as the Demonstrative Theory claims.

In response, Garcia-Carpintero (1994) grants the legitimacy of (9) and he concurs with Washington and current linguistic theory that (10) contains no implicit subject, demonstrative or otherwise. Nonetheless, he says, some demonstrative element is understood by context and/or conversational implicature.

¹ Washington (1992) criticizes Davidson's theory on the grounds that quotation may refer to size and other features aside from shape, and Garcia-Carpintero (1994) concurs. Granted, I can truthfully say "a" is smaller than "a". However, "shape" can refer to all sorts of formal properties, including size. After all, size is a geometrical feature, speaking in the strictest, mathematical sense.

² For the sake of clarity, I sometimes use double apostrophes and italicization as interchangeable marks of quotation. When I put a word in italics and a set of apostrophes, it is equivalent to being inside of two sets of apostrophes.

Garcia-Carpintero's adversion to conversational implicature is left undeveloped, and I have no clue how it might go. The reference to context, however, is clearer. We often mix language with manifest context in our attempts to communicate. As Christensen (1967) says, for instance:

Standing in front of a Studebaker, there is thus no need of saying, "This car is a Studebaker." We can simply point at it and say "A Studebaker!" without use of any name or description [or verbal deictic]. (Christensen 1967, p. 360)

However, comparing quotation to contextual deixis in this way goes against the evidence. Only a foreign speaker would ever point at a car and say "is a Studebaker". If you were to omit the subject "this" you would omit the copula "is" as well. The same principle should apply to metalinguistic reference, absent any account that predicts the contrary.

As a matter of terminology, I would suggest that quotation tautologically requires quote marks. This is not to deny the existence of statements like (9), where mentioning occurs without quote marks, but only to insist that mentioning is not the same thing as quoting. An expression without quote marks, even if it is being mentioned, is linguistically distinct from an expression with quote marks, and it will prove convenient below in §5 to reserve a special term for the latter—"quotation" being the most appropriate. Of course, none of this undermines the thrust of Washington's punctuation argument against the Demonstrative Theory.

3.2 *The Recursion Problem*

The Demonstrative Theory also fails to account for iterated quotation. Just as we can refer to the word in (11), a verb, by means of forming the metaword in (12), a noun phrase, we can refer to the word in (12) by means of forming the meta-metaword in (13), ad infinitum. (It is clear that quotations are noun phrases, as they function as grammatical subjects.)

- 11. Sit
- 12. "Sit"
- 13. "'Sit'"

But the Demonstrative Theory translates the perfectly good (14) into (15), which in turn translates either into the uninterpretable (16) or else into (17).

- 14. "'Sit'" is a noun phrase.
- 15. "Sit." That is a noun phrase.
- 16. Sit. That that is a noun phrase.
- 17. Sit. That. That is a noun phrase.

The problem is that (17) fails under every possible construal. (i) The second "that" cannot refer to the first "that". Although the first "that"—being a pronoun—does in fact constitute a noun phrase, this move gives *all* iteratively quoted expressions

the same reference: “‘sit’” and “‘eat’” for instance would each denote the interior quote marks; (18) would wrongly come out false while (19, 20) would wrongly come out true.

18. “‘Sit’” contains 5 characters (including two quote marks).
19. “‘Sit’” contains 2 characters (just interior quote marks).
20. “‘Sit’” is a pair of quote marks.

(ii) Nor can the second *that* in (17) refer to the entire complex “‘Sit. That.’” For this sequence, being other than a noun phrase, would contradict (14). (iii) Nor can the second “that” in (17) refer to the referent of the first “that”, to “‘Sit’”. For this verb, being other than a noun phrase, would also contradict (14). Besides, there is no linguistic reason for thinking that such a demonstrative in such a position could refer in this transitive way; and even if there were, we would again end up having to deny (18), this time having to say:

21. “‘Sit’” contains just 3 characters.

The only way that I see for the Demonstrative Theory to avoid outright falsehood is for the recursive semantics of the language to freeze when it reaches quotations within quotations. But this move is empirically unsatisfactory, for English does not distinguish between exterior quote marks and interior quote marks except as a stylistic device to remind the reader when there is a quote within a quote; the interior quote marks do not, intuitively, possess a sense distinct from the exterior marks. Furthermore, this move is theoretically ad hoc, as there is no independent motivation for treating quotation as the sole exception to the rule that syntactically recursive constructions possess recursive semantics.³

4 The Identity Theory

Frege (1892), Quine (1940, pp. 26, 40), Wittgenstein (1953, §16), Tajtelbaum (1957), Whately (1957), Searle (1969, p. 75), Washington (1992), and Reimer (1996) can all be taken as advocating the Identity Theory, according to which quotation is “autonomous” (not, as Davidson (1979) repeatedly writes, “autonomous”). Unfortunately the brevity and imprecision of these works pose problems of interpretation. Washington, who gives a comparatively clear account, says: “in quotation, expressions are used to mention themselves” (1992, p. 583). This is open to at least three construals. (1) It might mean that when an expression is a quotation then it refers to itself: (a) refers to (a).

³ Cappelen and Lepore (1997, p. 439) note the non-iterativity of quotation within their Demonstrative Theory, though they do not recognize it as a problem. Grandy (1973, p. 106) notes an analogous recursion problem for Davidson’s paratactic theory of indirect discourse.



The problem here is that it makes the Identity Theory circular: it explains quotation (a) in terms of quotation (a), which we cannot understand until we already understand what is being explained. (ii) Alternatively Washington might mean that, inside the context of quotation marks, an expression refers to itself: (b) refers to (b). The problem here is that although it explains the semantic value of (b), it does not explain the semantic value of either the quote marks or the quotation (a) as a whole. The claim that (b) refers to (b) might be true, but unless it is conjoined to some account of the double apostrophe it hardly contributes to the theory of quotation. (iii) Finally, Washington might mean that (a) refers to (b), and that this counts as self-reference because (a) is the same as (b), semantically speaking. This, I believe, is what the Identity Theorist means to say and must say: while quote marks add a pragmatic flourish, they possess no semantic value of their own.

Whereas the Demonstrative Theorist regards quote marks (or context) as referential and the quoted material as an inert adjunct, the Identity Theorist conversely regards the quoted material as (self-)referential and the quote marks as semantically empty. Despite these differences, the Identity and Demonstrative accounts can both be called picture theories, for both claim that quotation resembles its referent, the quoted material. Picture theories are supported by the prevalence of belief in word magic. Word magicians of what I will call Type I obscure the distinction between the use of a word and its referent. In both pre-literate and modern scientific societies, for instance, people commonly believe in the potency of “jinxes”: referring to something bad will make it more likely to happen. Word magicians of Type II obscure the distinction between the mere mention of a word and its use. Thus even the highly educated—teachers and editors—generally refuse to publicly quote obscenities. They omit them altogether or else print ciphers like “f***”. Now what is the difference between “f***” and “fuck”? Whereas the Description Theory is forced to mistakenly hold that there is little difference between the two, picture theories can explain why one is more offensive than the other: the one pictures the actual obscenity whereas the other does not.⁴

⁴ While it is clear that word magic of Type I is irrational, Type II is a doubtful case. In any event, I apologize if I offend any sensibilities, and hope you see that my point might have been lost had I stuck to ciphers and circumlocutions.

Perhaps too I ought to clarify that “word magic” is a standard anthropological term (cf. Ogden and Richards 1923), and I am not using it as an implied ad hominem against anybody. Obviously I am not trying to taint the Name and Description Theorists by this terminology, for they deny the existence, or at least the significance, of word magic. Just as obviously I am not trying to taint picture theorists, for the picture theory, or something like it, is what I am advocating. What I do want to do is to emphasize some robust properties of quotation that get ignored in the philosophical literature.

Like the Demonstrative Theory, the Identity Theory avoids the productivity problem and (perhaps) the simultaneity problem. But again like the Demonstrative Theory, it faces its own version of the recursion problem—plus problems about speech-acts and multiple ambiguity.

4.1 The Recursion Problem

Since the Identity Theory treats quote marks as semantically empty, “‘sunset’” must have the same meaning and reference as “sunset”. Duplicating quote marks is like ending an interrogative with two question marks. It may add emphasis, but that is all. Yet “‘sunset’” and “sunset” have distinct referents; thus, the Identity Theory must be wrong.

The Identity Theorist may respond that “‘sunset’” and “sunset” are not equivalent simpliciter; rather, the *use* of “‘sunset’” is equivalent to the *mention* of “sunset”. This claim may be right so far as it goes, but it is rather anemic in two different ways. First, it hardly seems characteristic of the Identity Theory; I doubt that any other theory would deny it. Second, it invites an explication of the use-mention distinction, which the Identity Theory never even attempts to give. But this point anticipates the next one.

4.2 The Speech-Act (Circularity/Vacuity) Problem

According to the Identity Theory, “Kim” refers to itself, that is to “Kim”. But we also know that “Kim” refers to Kim. Since “Kim” and Kim are not the same, we are forced into concluding that “Kim” possesses distinct senses or uses. This leaves the Identity Theorist with the task of giving a theory of use. It would be nice to know what “uses” are (even a pretheoretic characterization would be helpful); the criteria for individuating uses; the relation between a word’s having multiple uses and its being polysemous; and so forth.

Considering the link between the Identity Theory and the speech-act tradition, we might wonder whether we can find an answer in terms of illocutionary force. Unfortunately the prospects look dim. To begin with, we cannot simply distinguish use and mention by introducing a “quotative” illocutionary force, as this move is either vacuous or circular. It is vacuous if the quotative force is left undefined, uncharacterized, and unexplained; it is circular if mentioning is explicated by the quotative force, and the quotative force is explicated by mentioning. At the same time, we do not want to contrast quotations and non-quotations in terms of the presence or absence of the familiar illocutionary forces either. We do not want to say, for instance, that whereas (22) makes a genuine assertion, (23) is a mere locution, lacking illocutionary force.

22. The earth is flat.

23. “The earth is flat.”

Although it is true that I am not asserting (23), I am quoting someone—perhaps hypothetical, perhaps not—as having made an assertion rather than query or command. Only by acknowledging that the quotation contains assertive force within it can we understand that (23) calls for assent or dissent, that (23) is true or false. Besides which, if quotations lacked illocutionary force, there would be no way to treat iterated quotation. If (23) is a mere locution, distinguished from (22) by lack of assertion, what would distinguish (23) from (24)?

24. “‘The earth is flat.’”

In summary, neither illocutionary forces nor anything in the literature of Identity Theory provides means for distinguishing between quotation and non-quotation.

4.3 *The Multiple Ambiguity Problem*

Tokens refer (or are used to refer); types do not (are not). This is the case for all sorts of words, but is most evident for indexicals: the token of “I” in the abstract of this paper refers to Paul Saka, whereas other tokens in other places and times refer to other individuals. Abstracted from particular speakers, “I” does not refer to anyone. If we accept this view that tokens refer while types do not, and if we agree with the Identity Theorist that quoted expressions refer to themselves, then we must conclude that while expression tokens may be quoted, expression types may not. To put it another way: tokens have the capacity to refer, hence the capacity to refer to themselves, hence the capacity to be used in quotations. Types do not refer, hence do not refer to themselves, hence cannot be used in quotations.

Yet this is not how quotation works (cf. Garver 1965; Christensen 1967; Goldstein 1984; Seymour 1994). Quotation is multiply ambiguous or indeterminate: you can stick quote marks around a token to refer to that very token (25); to refer to some other token of the same type (26); or to refer to its type (27).

25. “I” refers to me.

26. “I”, as said by you, refers to you.

27. “I” does not refer to anyone in particular; only tokens of it do.

Moreover, quotations can refer to words understood as form-content pairings (28); to lexemes understood as words abstracted from their inflectional paradigms (29); to forms, that is spellings or pronunciations (30); and to content both immediate (31) and translated (32).

28. “Run” is used in the third-person plural but not singular.

29. “Run” refers to *run*, *runs*, *ran*, *running*.

30. “Run” consists of three letters.

31. The concept “premise” is the same as the concept “premiss”.

32. Galileo (who spoke no English) said, “The earth moves!”.

It is even possible for quotations to “refer” to non-existents given that types, lexemes, and content are all somewhat controversial theoretical entities. In the event that lexemes, say, do not really exist, “the lexeme ‘air’” would refer to a certain lexeme, that is to nothing; it would not refer to air. Thus, examples (28–32) show that there really are different kinds of quotation, regardless of whether types, lexemes, and content are real.

For most purposes it is not necessary to distinguish between token-quotation and type-quotation, or between word-quotation and lexeme-quotation, or between form-quotation and content-quotation; the intended interpretation is either immaterial or else obvious. When it matters for scholarly purposes, however, writers will sometimes explicitly stipulate a system of distinct quote marks. Lyons (1977), Atlas (1989), and Horn (1989), for instance, all preface their works by glossing double apostrophes, single apostrophes, italics, and small-capital letters, each with one or more different functions. What this again proves is that quoted tokens can, *contra* Identity Theorists, refer not only to themselves but to related sorts of linguistic items as well.

Now the Identity Theorist could argue that using an expression is like using a hammer in that it implicates several parts, plus the whole, all at once. Just as in hammering you use a hammer, its handle, its head, your arm, and perhaps a nail, so in speaking you use a word token, a word type, a lexeme token, a lexeme type, etc. By using a word token in mention mode, you quote a word token; by using a lexeme type in mention mode, you quote a lexeme type; and so forth. Any sort of linguistic item at all can thus be referred to by quotation.

But in hammering, you simultaneously use many implicated components. By analogy, the use of an expression in mention mode ought to collectively refer to types, tokens, forms, and content. This is not what happens, of course; statements (25–32) must be read in highly specific ways to be true, or even to be grammatical (each is third-person singular). Therefore the Identity Theory fails regardless of whether or not it invokes the hammer metaphor.⁵

⁵ Boolos (1995) discusses another kind of ambiguity, one due to scope. For instance, (i) denotes either (ii) or (iii):

- (1) “a” followed by “b”
- (2) ab
- (3) a” followed by “b

Incidentally, Boolos proposes a way of resolving such ambiguity: in an ideal language you could use subscripts so that each left quote mark gets explicitly and uniquely paired up with some right quote mark.

5 Formal Disambiguation of Multiple Ostension

My own understanding of quotation begins with the notion of deferred ostension (Quine 1968, Nunberg 1978, Fauconnier 1985; cp. Goldstein 1984, Reimer 1996). *Direct* ostension associates a term with its referent via immediate experience, for instance by deictic demonstration or by the simple exhibition of an item. *Deferred* ostension to an absent object X may be secured by pointing at or describing something present that is saliently related to X. For example, by pointing at a newspaper you can refer to a newspaper office or a newspaper company:

33. [Pointing at paper] I've got to get over there today and place my ad.
 34. [Pointing at paper] They were bought out by Murdoch.

Pointing is not necessary; you can also verbally describe one thing in order to refer to another:

35. The school [meaning the school building] burned down today.
 36. The ham sandwich [the customer] is waiting for her check.

The “pragmatic functions” that map from description to intended referent are, as you would expect, squishy. None the less, research by Norrick (1981) identifies a number of recurring associations, for example between merchandise and manufacturer (34), institution and physical instantiation (35), process and product, cause and effect, part and whole, container and content; while research by Nunberg gets at such general principles as “recoverability”.

I claim that every use of language is an act of multiple ostension, partly direct and partly deferred, of at least the following kinds of items.

- (a) orthographic form: *cat*
- (b) phonic form: /kæt/
- (c) lexical entry: <*cat*, /kæt/, count noun, CAT>
- (d) intension: CAT
- (e) extension: {x: x a cat}

First, exposure to the written label *cat* (a) or the spoken label /kæt/ (b) evokes the corresponding lexeme (c) in every competent speaker of English, where a lexeme is an arbitrary ordered *n*-tuple including orthographic form, phonic form, syntactic category, meaning, register, etc.⁶ This evocation happens automatically and spontaneously as a result of human cognitive architecture trained in English (see Fodor 1983 on the reflex-like nature of language-processing). Second, the lexeme <*cat*, /kæt/, noun, CAT> specifies the intension CAT (d) according to the pragmatic function WHOLE-PART. Third, CAT determines the extension {x: x a cat} (e) according to some mysterious but widely assumed function. Thus, the utterance of “cat”—which directly ostends or exhibits the phonic token /kæt/—deferentially ostends the corresponding form type, the lexeme <*cat*, /kæt/, noun, CAT>, the

⁶ For simplicity I subsume the phonetic and the phonological under the “phonic”.

concept CAT, the customary referent $\{x: x \text{ a cat}\}$, etc. These items form a package deal in which you cannot get the label without getting the rest. With this understanding of deferred ostension in mind, I propose that use and mention can be understood as follows.⁷

(u) Speaker S uses an expression X iff:

1. S exhibits a token of X;
2. S thereby ostends the multiple items associated with X (including X's extension);
3. S intends to direct the thoughts of the audience to the extension of X.

(m) Speaker S mentions an expression X iff:

1. S exhibits a token of X;
2. S thereby ostends the multiple items associated with X;
3. S intends to direct the thoughts of the audience to some item associated with X other than its extension.

For example, let X = the phrase "every cat". Then (1) S exhibits a token of X by either writing *every cat* or uttering /Evri kæt/ or even by pointing at someone else's tokening, as happens in charades. (2_a) Such an exhibition directly ostends the token form *every cat* or /Evri kæt/; and it deferringly ostends the types *every cat* and /Evri kæt/ and the noun phrase sequence made up of the two lexemes <every, /Evri/, determiner, EVERY> + <cat, /kæt/, count noun, CAT>—that is, the noun phrase structure [every_{DET} cat_N]_{NP} and its intension EVERY(CAT). (2_b) In addition it deferringly ostends, as reference, every cat. (3) In the case of use, S intends "every cat" to refer to every cat; in the case of mention, S intends "every cat" to refer to some item that is saliently associated with "every cat" other than its extension, that is to one of the items listed under (2_a).

Although I have characterized the use-mention distinction in terms of extension, referentialist semantics is by no means essential to my account. I myself prefer the anti-realist "internal semantics" of Jackendoff (1983), which construes "reference" as a solipsistic mental state; others are invited to adjust (u) and (m) according to their own favored theories of meaning.

The account offered possesses a number of virtues. To begin with, it allows for the existence of mentioning without quote marks. In addition, it reveals how much use and mention have in common. Although the reference clauses (u-3) and (m-3) differ from each other, this difference is made possible by the same background conditions: the existence of a conventionalized language (2) and its exercise (1). Finally, it follows from formulations (u) and (m) that use and mention, though distinct, are perfectly compatible: you can intend to direct the thoughts of your

⁷ I use "use" and "mention" in their technical senses (to utilize an expression with customary reference versus to utilize it in reference to itself). It seems that I need to make this clear because some writers insist that, in order to mention an expression, you need to use it (Geach 1950; Ziff 1960, p. 27; Garver 1965; Davidson 1965 and 1979). The sense in which this is obviously true is irrelevant.

audience to the customary reference of X while at the same time intending to direct the thoughts of your audience to other things as well, for example to X itself. Thus, my account opens a path for treating simultaneous use and mention.

With the foregoing understanding of use and mention in place, we are now prepared to consider the definition of quotation. Syntactically, a pair of quote marks is a discontinuous determiner (a complex symbol which, applied to an argument expression, produces a noun phrase). Semantically, a pair of quote marks is a concept or intension, QUOT, which ambiguously or indeterminately maps its argument expression X into some linguistic item saliently associated with X other than the extension of X. Although quote marks generally do not specify among token, type, form, and concept, they still serve to partially disambiguate, for they rule out customary reference as the intended interpretation. Thus, the speaker who uses quote marks announces “I am not (merely) using expression X but am mentioning it”.

So far we have seen, in the explanation of (u) and (m), how use and mention *without* quotation work. To see how use and mention *with* quotation work, suppose X = the singly quoted expression “*cat*”.⁸ Then (1) S exhibits a token of X by writing “*cat*” or by uttering /kwot kæt ənkwot/ or by uttering /kæt/ with distinctive intonation or accompanying finger gesture or by just uttering /kæt/. (2_a) Such an exhibition directly ostends the orthographic token “*cat*” or some spoken counterpart; and it deferringly ostends the corresponding type, and the noun phrase structure [“_{DET} [cat]_N _{DET}”]_{NP} and its intension QUOT(*cat*). (2_b) The latter ambiguously specifies the extension: the token *cat*, the type *cat*, the token /kæt/, the type /kæt/, the structure [cat]_N, or the concept CAT. (3) In the case of use, S intends “*cat*” to refer to the extension of “*cat*”, namely to one of the items listed under (2_b). In the case of mention, S intends “*cat*” to refer to some item saliently associated with “*cat*” other than its extension, namely to one of the items listed under (2_a).

As a consequence of my account, the mention of X and the use of “X” quite properly secure the same reference. Further, whereas the use-mention distinction is characterized in terms of the speaker’s intentions, the distinction between a quotation and a non-quotation is a formal, grammatical affair. Distinguishing between use and mention in a language without quote marks is a purely pragmatic affair, but in a language with quote marks mentioning can be explicitly marked (although such marking is not obligatory).

Now suppose X = the doubly quoted “‘*cat*’”. Then (1) S exhibits a token of X by writing “‘*cat*’”. (3_a) Such an exhibition directly ostends the token form “‘*cat*’”, and it deferringly ostends the type “‘*cat*’” and the noun phrase structure [“_{DET} [“_{DET} cat_N ”_{DET}]_{NP} ”_{DET}”]_{NP} and its intension QUOT(“‘*cat*’”). (2_b) The latter ambiguously specifies as extension: the token “‘*cat*’”, the type “‘*cat*’”, the structure [“_{DET} cat_N ”_{DET}]_{NP}, or the concept QUOT(*cat*). (2) In the case of use, S intends “‘*cat*’” to refer to one of the items listed under (2_b). In the case of mention, S intends “‘*cat*’” to refer

⁸ I myself am using two layers of quotation—a pair of double apostrophes plus italics—to describe S’s speech. In S’s speech, however, only a single layer is being used.

to some item listed under (2_a). In short, the Disambiguation Theory correctly handles iterated quotation.

My account differs from the Demonstrative Theory in that I do not take quote marks as pointing. Rather, I compare quote marks to the subscripts you see on ambiguous words in dictionaries and sometimes in philosophy, except that quote marks help us to steer around a case of systematic ambiguity. For example, “warm” has a number of senses including PRODUCER OF HEAT (“this jacket is warm; it keeps me comfortable”) and PRODUCT OF HEAT (“this jacket is warm; it has been sitting on the stove”). If we agree to mark the instrumental sense with the subscript “inst” then a number of consequences follow. First, “warm_{inst}” means PRODUCER OF HEAT; second, “warm” remains ambiguous (does it lack a subscript because the speaker deliberately intended “not in the instrumental sense”, does it lack a subscript because the speaker was sure that “in the instrumental sense” was so clear as not to require subscripting, or does it lack a subscript because the speaker just did not stop to think about the ambiguity?); third, novel expressions such as “hot_{inst}” PRODUCER OF TOO MUCH HEAT become unambiguously interpretable. Quote marks work analogously, functioning to single out (or rather narrow down) the intended reading. Their presence indicates a metalinguistic use; their absence does not necessarily indicate anything at all; and they generalize to new expressions.

But, you might object, in specifying that the metalinguistic use is operative, don’t the quote marks serve to point at the quoted expression? It is easy to think so, because the quoted word is always present during the act of quotation. Nonetheless, I do not think that the quote marks function deictically (and hence referentially). In the first place, it is possible for a quotation to refer to something that is not physically present, as in the case of word meanings and word types. Second, it is possible to forego the use of quote marks altogether.

My account differs also from the available versions of the Description Theory in that it does not use names for elements. The quote marks, when they are used, do describe which aspect of an ostension is operative; but the panoply of ostension is provided by picture (in so far as the exhibited linguistic label resembles itself) and by other pragmatic relations (in so far as the form token evokes a type or associated content, etc.).

My account bears some resemblance to the Identity Theory: in both use and mention, the speaker exhibits the same form. However, there are critical differences. To begin with, quotation may refer not only to a given word-form token but also to the corresponding word-form type, to other tokens of that type, to lexemes, to concepts, and so forth. Thus, my account qualifies as a picture theory in its analysis of the quotation of forms (since the inscription “*cat*” really does picture the inscription *cat*). But it is not entirely a picture theory, as my analysis of the quotation of content involves no iconic relation at all (since the inscription *cat* does not resemble the meaning CAT in any visual, acoustic, or other physical manner). Another difference is that my account treats an expression as a quotation or not depending on its linguistic structure, whereas the Identity Theory treats quotation as a matter of function (as the use-mention distinction is for me). Further, on my account the quotation as a whole (quoted material plus quote marks) is

referential. For the Identity Theorist, the quoted material alone refers, and for the Demonstrative Theory the quote marks alone refer.

The Disambiguation Theory differs from the Tajtelbaum version of the Identity Theory in particular. For Tajtelbaum (1957), a word refers to itself because “we have the (tacit) convention that a name and its name are denoted by the same word.” According to the Disambiguation Theory, in contrast, mention is not a matter of convention. Whether we use single apostrophes or double—indeed, whether we use any formal device at all—is a matter of convention. But the fact that we can use a word to refer to itself or to associated items is a natural result of the fact that the direct ostension of a form token deferringly ostends type, content, etc. That is, the exhibition of a word form makes the corresponding word type, lexeme, sense, and referent manifest to any human mind that speaks the language and is versed in the notions of word type, lexeme, sense, and referent.

Some may object that the Disambiguation Theory counts as a version of the Identity Theory, or that the Identity Theory is at least compatible with my proposals, in spirit if not in letter. I do not know how to judge this claim for I do not see any “spirit” in the Identity Theory beyond what it actually says. At any rate, it makes no difference to me whether my account is regarded as a version of the Identity Theory, or distinct from it, so long as my substantive points are acknowledged, for example that mentioning is not the same as quoting.

On this point, in fact, my account distinguishes itself from all others. For the Identity Theory, quote marks are but window dressing, and deploying them or removing them changes nothing. For the Demonstrative Theory, quote marks are so crucial that they are postulated as logically present even when physically absent. My theory, on the other hand, treats an expression with quote marks as distinct from the same expression without quote marks.

Is this really a virtue, though? Some critics have denied that it is, claiming that quoted mention and unquoted mention obviously mean the same. But this is wrong, for it is just false that (37, 38) mean the same.

37. Chicago has seven characters.

38. “Chicago” has seven characters.

The fact is that (37) and (38) are ambiguous or indeterminate in different ways. The full explanation of (37) is that it has at least the following two readings, the first one true and the second one false.

39. The word “Chicago” has seven characters.

40. The city Chicago has seven characters.

In contrast, (38) has the following two readings, one true and one false.

41. The word “Chicago” has seven characters.

42. The expression “‘Chicago’” has seven characters.

While (39) = (41), the total range of possible readings for (37) differs from the range of possible readings for (38). Thus, (37) and (38) demand different analyses, and on this score my account is indeed attractive.

Another distinctive feature of my account is that it explicitly treats customary reference as a species of ostension on a par with antonymous reference. Use and mention are two sides of one coin; although distinct, they implicate one another. You simply cannot have a language without the potential for both use and mention (although you can have a language without quotation).

Yet problems remain. If you may refer to whatever you ostend, and if the utterance of a word-form deferringly ostends all of the associated linguistic material, then why can't you quote a word in order to refer to its associated language or dialect, as in (43)?

43. *“Warshboard” is a stupid dialect.

44. “Warshboard.” Now there’s a stupid dialect.

The acceptability of (44) proves that the quotation of “warshboard” suffices to ostend or make manifest a certain dialect. So if ostension were equivalent to reference, (43) should be acceptable.

The solution lies in rejecting any simple-minded equivalence of ostension and reference. Sometimes one reference is blocked by a different prior reference.

45. *The ham sandwich, which was inedible, left without tipping.

Here “ham sandwich” ostends both a comestible and a customer. Once the expression is established in the relative clause as referring to the food, it cannot later be taken as referring to the person. Yet hard-and-fast rules elude us.

46. ?“Ghosttown” has more letters in it than people.

47. Lee’s dissertation, which weighs five pounds, has already been refuted.

48. A: What dialect does he speak? B: “Tomahto”.

Although (46) strikes me as somehow anomalous, as necessarily jocular, it does not seem as bad as (45). In (47) “dissertation” ostends both a material token and an abstract type; reference to the former does not preclude reference to the latter (Fauconnier 1985). And in (48) the citation of a word does pragmatically refer to a dialect. The general lesson is that developments in the psychology of association and in the pragmatics of reference will better enable us to judge the thesis that quotation formally disambiguates multiple ostension.

6 Conclusion

The Disambiguated Ostension Theory of Quotation consists of two components. First, it assumes that the capacities for both use and mention stem from the same source, namely from the fact that the human mind associates a multiplicity of deferred ostensions with any exhibited token, thus giving rise to pragmatic ambiguity. This thesis is a plausible consequence of the findings of Nunberg (1978) and Norrick (1981), and it is further supported by the fact that surely all languages contain a use-mention distinction. (Even formal languages that officially

lack a use-mention distinction still, in practice, possess it. This suggests that the distinction is made not by particular languages per se but rather by what is common across English, Eskimo, the predicate calculus, etc.—namely, human cognitive agency.) While this first component allows for a pragmatic use-mention distinction, the other component of the Disambiguation Theory claims that quote marks formally announce that mentioning is taking place. This view too is plausible if you think that linguistic elements normally signify something, and it is further suggested by the fact that (37) and (38) are ambiguous in distinct ways.

Another consideration in favor of the Disambiguation Theory is that it is more successful than rival theories in dealing with the facts which we have seen any theory of quotation ought to respect. These facts are summarized and extended below.

CONVENTIONALITY: Although the capacity for mention inheres in the natural connection between a linguistic token and its associated items, the existence of a device for explicitly marking mention—for quoting—is conventional. Such a device can be found within most modern writing systems, and in the grammatical systems of a few languages (e.g. the indirect discourse mood of German), but it appears to be absent throughout most of history in most languages. Furthermore, since convention rests upon a marriage of form and function, an expression with conventional quote marks possesses a range of possible significance distinct from that of an expression without quote marks. *This latter fact disproves the Identity Theory.*

PRODUCTIVITY: If you are party to the convention, then knowing any expression will enable you to know its quotation; and knowing any quotation will enable you to know what is quoted. *The fact of productivity rules out the Name and Description Theories.*

ITERATIVITY: Quotations are themselves expressions that may be quoted in turn, giving yet a different reference. *Iterativity rules out the Demonstrative and Identity Theories.*

PICTORIALITY: Belief in word magic is powerful throughout the world, in both stone-age and industrial cultures—the mere mention of an obscenity is often taboo or at least somewhat shocking. Furthermore, constructions like “Pigs are called ‘pigs’” possess a sort of a priori truth. Both of these observations suggest that quotation resembles what is quoted. *These facts rule out the Name and Description Theories.*

SIMULTANEITY: In some contexts you may both use and mention an expression at the same time. In some contexts you may anaphorically ground an expression for both use and mention. *The fact of simultaneity poses a mortal problem for the Name and Description Theories. It remains to be seen how other theories fare.*

MULTIPLICITY: With quotation you can refer to both linguistic expressions and non-linguistic vocalizations and imprints; to both form and content; to both types and tokens. On the rare occasions when we need to make these distinctions explicit, we take advantage of the diverse devices for marking quotations: double apostrophes, single apostrophes, italicization, underlining... *The fact of multiplicity disproves the Identity Theory and, at least in Davidson’s version, the Demonstrative Theory as well.*

OMISSIONS: Philosophers, linguists, logicians, and lay writers commonly omit quotational markings on mentioned expressions (sometimes with a statement to the effect that “I won’t be fussy”)—for example “Diabetes comes from Greek.” Instead of treating this as an error, an outright falsehood, it would be more charitable to take it as indeterminate/ambiguous between the false “The disease diabetes itself comes from the Greek language” and the true “The word ‘diabetes’ comes from Greek”. *The fact that there are omissions rules out the Demonstrative Theory.*

On all of these scores, I have argued, the theory of quotation as formal disambiguation of multiple ostension fares better than the alternatives. A fuller understanding of it now awaits further research in psychopragmatics.⁹

References

- Anscombe, G.E.M. 1957. Report on analysis problem #10. *Analysis* 17: 49–52.
- Atlas, Jay. 1989. *Philosophy without ambiguity*. Oxford: Oxford University Press.
- Bennett, Jonathan. 1988. Quotation. *Nous* 22: 399–418.
- Boolos, George. 1995. Quotational ambiguity. In *Paolo Leonardi and Marco Santambrogio*, ed. On Quine, 283–296. Cambridge: Cambridge University Press.
- Cappelen, Herman, and Ernie Lepore. 1997. Varieties of quotation. *Mind* 106: 429–450.
- Carnap, Rudolf. 1937. *The logical syntax of language*. New York: Harcourt and Brace.
- Christensen, Niels. 1967. The alleged distinction between use and mention. *Philosophical Review* 76: 358–367.
- Davidson, Donald. 1979. Quotation. *Theory and decision*. Reprinted in *Inquiries into truth and interpretation*, 79–92. Oxford: Oxford University Press, 1984.
- Fauconnier, Gilles. 1985. *Mental spaces*, 2nd ed, 1994. Cambridge MA: MIT Press. Cambridge University Press.
- Fodor, Jerry. 1983. *The modularity of mind*. Cambridge: MIT Press.
- Frege, Gottlob. 1892. On sense and reference. In *Translations from the philosophical writings of Gottlob Frege*, 3d edn., eds. Peter Geach and Max Black, 56–78. Oxford: Basil Blackwell, 1980.
- Garcia-Carpintero, Manuel. 1994. Ostensive signs. *Journal of Philosophy* 91: 253–264.
- Garver, Newton. 1965. Varieties of use and mention. *Philosophy and Phenomenological Research* 26: 230–238.
- Geach, P.T. 1950. On names of expressions. *Mind* 59: 388.
- Geach, P.T. 1957. *Mental acts*. London: Routledge & Kegan Paul.
- Goldstein, Laurence. 1984. Quotation of types and types of quotation. *Analysis* 44: 1–6.
- Grandy, Richard. 1973. Review of *words and objections: Essays on quine*. *Philosophical Review* 82: 99–110.
- Groening, Matt. 1996. Life in hell. *Village Voice*, April 9, p. 3.
- Horn, Laurence. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge: MIT Press.
- Jorgensen, Julia, George Miller, and Dan Sperber. 1984. Test of the mention theory of irony. *Journal of Experimental Psychology: General* 113: 112–120.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Norrick, Neal. 1981. *Semiotic principles in semantic theory*. Amsterdam: John Benjamins.
- Nunberg, Geoffrey. 1978. *The pragmatics of reference*. City University of New York linguistics dissertation. Relevant portion is available in “The non-uniqueness of semantic solutions”. *Linguistics and Philosophy* 3(143–84): 1979.
- Ogden, C.K., and I.A. Richards. 1923. *The meaning of meaning*. New York: Harcourt, Brace and World.

⁹ Most of this paper was written under a University of Illinois University Fellowship in 1995–96, and revisions were made while I held a Beckman Institute Cognitive Science Fellowship in 1997. For reading and commenting on various drafts, I gratefully acknowledge Steve Wagner, Tim McCarthy, Mark Sainsbury, and an anonymous reviewer.

- Partee, Barbara. 1973. The syntax and semantics of quotation. In *Festschrift for Morris Halle*, eds. Stephen Anderson and Paul Kiparsky, 410–418. New York: Holt, Rinehart and Winston.
- Quine, Willard. 1940. *Mathematical logic*, 3rd ed. Cambridge: Harvard University Press.
- Quine, Willard. 1943. Notes on existence and necessity. *Journal of Philosophy* 40: 113–27. Most of this material is available in “Reference and modality”, *From a logical point of view*, Cambridge: Harvard University Press, 1961.
- Quine, Willard. 1960. *Word and object*. Cambridge: MIT Press.
- Quine, Willard. 1968. Ontological relativity. *Journal of Philosophy* 65: 185–212. Reprinted in *Ontological relativity*, 26–68. Columbia University Press, 1969.
- Reimer, Marga. 1996. Quotation marks: Demonstratives or demonstrations? *Analysis* 56: 131–142.
- Richard, Mark. 1986. Quotation, grammar, and opacity. *Linguistics and Philosophy* 9: 383–403.
- Searle, John. 1969. *Speech acts*. Cambridge: Cambridge University Press.
- Seymour, Michel. 1994. Indirect discourse and quotation. *Philosophical Studies* 74: 1–39.
- Sperber, Dan, and Deirdre Wilson. 1981. Irony and the use-mention distinction. In *Radical pragmatics*, ed. Peter Cole, 295–318. New York: Academic.
- Tajtelbaum, Al. 1957. It is impossible to be told anyone’s name. *Analysis* 17: 52f.
- Tarski, Alfred. 1933. The concept of truth in formalized languages. In *Logic, semantics, metamathematics*, 152–178. Oxford: Oxford University Press, 1956.
- Washington, Corey. 1992. Quotation. *Journal of Philosophy* 89: 582–605.
- Whitely, C.H. 1957. Names and words. *Analysis* 17: 119–120.
- Wittgenstein, Ludwig. 1953. *Philosophical investigations*. New York: Macmillan.
- Ziff, Paul. 1960. *Semantic analysis*. Ithaca: Cornell University Press.

Indirect Reports and Pragmatics

Nellie Wieland

Abstract An indirect report typically takes the form of a speaker using the locution “said that” to report an earlier utterance. In what follows, I introduce the principal philosophical and pragmatic points of interest in the study of indirect reports, including the extent to which context sensitivity affects the content of an indirect report, the constraints on the substitution of co-referential terms in reports, the extent of felicitous paraphrase and translation, the way in which indirect reports are opaque, and the use of indirect reports as pragmatic vehicles for other speech acts such as humor, insult, or irony. Throughout I develop several positions: (1) that a semantic analysis of indirect reports is insufficient, (2) that the distinction between direct and indirect reports is not clear and that indirect reports are the predominate way of reporting while direct reports may be a para-linguistic variation on them, (3) that most questions about the semantics and pragmatics of indirect reports will rely on a full understanding of the nature of what is reported and how it gets reported, (4) that an analysis of reporting requires the pragmatic tools of metarepresentation and a social, inter-personal understanding of relevance and shared knowledge.

1 Introduction: Direct and Indirect Reports

The philosophical literature is rich with discussion of quotation and indirect reports. It has played a starring role in the philosophy of language at least since Davidson’s (1968) “On Saying That.” Much of this literature has concluded that the analysis of indirect reports should be largely pragmatic. Unfortunately, there has been a scarcity of accounts of what a pragmatic theory of indirect reports

N. Wieland (✉)
California State University, Long Beach, CA, USA
e-mail: nellie.wieland@gmail.com

should look like, how, if at all, it should complement a semantic account of indirect reports, and how such a pragmatic account can respond to some of the most interesting philosophical questions about reporting. Whereas quotations are used to directly report the utterances of another, the indirect report has far more latitude. The result is that indirect reports have most of the same features that make direct reports so philosophically interesting and many unique features that make them especially pragmatically interesting.

An indirect report typically takes the form of a speaker using the locution “said that” to report on an earlier utterance (e.g., “Galileo said that the earth moves”). The original utterance could come from the speaker herself or a third party; the contexts of the original and reported utterances could be identical or distinct. Indirect reports differ from direct reports in that they typically do not use quotation (in written form with quotation marks or in spoken form with implied quotation marks) (e.g., “Galileo said, ‘the earth moves’”). Or, if they do, they take the form of mixed-quotation where part of the report is direct and part indirect (e.g., “Galileo said that the earth ‘moves’”).

In one notable philosophical discussion of indirect reports, Cappelen and Lepore (1997b) argue that semanticists will have little to say about the most linguistically interesting features of indirect reports. The reason for this is that semanticists are interested in the most general and systematic features of linguistic expressions, whereas speakers, when they report, are interested in reporting something narrow and particular about a context and an utterance in that context. It is no surprise to Cappelen and Lepore that the aims of a semanticist in analyzing ‘to say’ and a reporter in using ‘said that’ differ. It’s worth considering their claim at length:

[Semanticists] tend to agree that semantics is a discipline that aims to characterize systematically certain features of linguistic expressions and to do so in a way that captures general truths about our linguistic practice, not just truths about particular speakers in particular contexts. Indirect speech, on the other hand, is a device reporters use for characterizing acts (utterances) performed by other speakers. In so doing, reporters are interested neither in systematicity nor in generality; they aim to convey something about a particular act in a particular context C to a particular audience situated in a different context C*. Reporters draw on information about the specific intentions of, knowledge about, and the history of a reported speaker in C and (maybe) similar features of an audience in C*. These are features one does *not* want to solicit when the aim is systematic and general.¹

The present discussion both accepts and rejects Cappelen and Lepore’s central claims in this passage. I will accept that the aims of the semanticist and the speaker

¹ Cappelen and Lepore (1997b), p. 289. Note that Cappelen and Lepore’s aim is totally different from the aim of this paper. They say that it is *because* reporters are doing these things that we should not expect a semantic analysis to capture reports. Their goal is to protect semantics from what look like threatening examples from indirect reports. They acknowledge the semanticist’s inability to explain indirect reports, but maintain that this is because reporting is something that speakers do for non-semantic reasons. They do not offer or entertain a full-fledged pragmatic theory to explain the relevant linguistic features of indirect reports.

differ, and I will accept that the semanticist does not have a comprehensive analysis of ‘to say’ available to her. But I will reject that there are no general and systematic claims that can be made about the pragmatics of indirect reports. Although speakers are attempting to “convey something about a particular act in a particular context C to a particular audience situated in a different context C*” when they indirectly report, they still do so in systematic ways, and under identifiable constraints.

In what follows, I introduce the principal philosophical and pragmatic points of interest in the study of indirect reports. I will argue that indirect reports are more pragmatically complicated than direct reports. They lack explicit, conventional markings in written text (such as italics or quotation marks), and they permit a great deal of communicative deviation. Presumably, a felicitous indirect report needs to convey the content communicated in the original utterance—but there is a fair amount of reasonable disagreement as to what this might constitute. Felicitous indirect reports allow for the substitution of co-referential terms, context shifting, translation across languages, paraphrase, reporting by inference, and the use of demonstratives to stand in for utterances. Indirect reports also work as standard pragmatic vehicles for humor, irony, implicature, insult, praise, etc.

I will discuss several possible pragmatic approaches to the analysis of reports. In particular, I will discuss the ways in which reporting practices are similar to other forms of metarepresentation. I also assess the analyses available depending on whether language is conceived of as a verbal or written phenomenon. Philosophers have tended to analyze language in its written form, largely ignoring verbal communication. Linguists tend to do the opposite. Given that direct reports are explicitly marked only in written texts, we can see that a verbal orientation would suggest that analyses of direct and indirect reports should largely be the same. If this is the case, then much of the philosophical literature on quotation should be re-assessed.

2 Puzzles About Indirect Reports

2.1 *Indirect Reports and the Use/Mention Distinction*

It is standard in philosophical discussions of quotation to distinguish between the use and mention of a term or expression. The idea is straightforward on its face: some language is used meta-linguistically to talk about other language. A speaker can say (1) and (2) coherently and felicitously:

- (1) Boston is a large American city.
- (2) ‘Boston’ has six letters.

In (1) the term ‘Boston’ is being used; in (2) it is being mentioned. One is not saying in (2) that the city itself has six letters, but rather that the name of the city is comprised of six letters. (2) is a claim about language; (1) is a claim about the world.

Given this distinction it is natural to think of reports as special, meta-linguistic uses of language since they involve mentioning the speech of others.² Consider direct reports:

(3) The President said, “Osama bin Laden resisted.”

In this case, the speaker of (3) is directly reporting the President’s speech; the use of quotation marks indicates that the speaker of (3) is merely mentioning and not using the words “Osama bin Laden resisted.” Now consider a case of an indirect report:

(4) The President said that Osama bin Laden resisted.

Direct written reports employ an arguably semantic tool³ to mark them as such, i.e., quotation marks, italics, or angle brackets. Indirect reports have a less explicit tool: the phrase ‘said that’ or its equivalent. One possibility to consider is that the phrase ‘said that’ has a similar function to written quotation marks insofar as it indicates that the words to follow are being mentioned and not used. As we will see in what follows, this is a tough case to make given the varieties of indirect reports. The use/mention distinction is typically directed at strings of words and not at linguistic content.

Could there be an alternative to the use/mention distinction for indirect reports? In this case, there would be a distinction between a speaker using content (e.g., the communication of linguistic content that takes place in the original utterance) and a speaker mentioning content (e.g., the reporting of linguistic content that takes place in the indirect report). Consider (5).

(5) Osama bin Laden resisted.

When the President utters (5), he is using this expression in order to mean something or other. When a reporter utters (4), she is merely mentioning whatever it was that the President meant in uttering (5), but is not attempting to mean

² It is not widely agreed upon that quoting is equivalent to mentioning. It is clear that there are forms of mentioning that do not involve quoting. This leaves open the possibility that quoting is a subspecies of mentioning. Some quotation theorists [see Saka (1998) and Washington (1992)] argue that the analyses of quotation and mentioning will be closely related, if not the same. Others [see Cappelen and Lepore (2009)] disagree that quotation and mentioning correspond in this way.

³ In the quotation literature it is not universally accepted that quotation marks (and their corollaries) are semantic devices. I am glossing over these disagreements here.

whatever it was that the President himself meant with his utterance of (5). The reporter in (4) distances herself from the meaning conveyed by the President given that her speech act is one of *reporting* and not one of *describing* what unfolded in Osama bin Laden's compound.⁴ This becomes clearer when we consider another case of indirectly reporting (5).

(6) The President said that the leader of al-Qaeda resisted.

In this case, the speaker is using a co-referential substitution for the name 'Osama bin Laden' (see Sect. 2.4). In doing so the speaker is mentioning—in a manner of speaking—what the President communicated, although not his exact locution. Admittedly, there is a bit of clumsiness in applying the standard use/mention distinction to indirect reports in this way. But this should highlight that there are interesting questions to be asked about what exactly is being reported in an indirect report as well as what is doing the reporting (see Sect. 2.9).

2.2 *Indirect Reports and Context Sensitivity*

Indirect reports are linguistic objects that can convey content from one context to another. They need not always do this (e.g., consider the case of a speaker reporting on her immediately prior utterance, "I said that we're going to be late!"). When reports do convey content from context to context, it is essential that relevant features of the reported context are assumed in the reporting context. Presumptively, these features of the reported context would be constituted in shared knowledge.

Given this, we can say that indirect reports are themselves context sensitive; a complete understanding of the content of the report requires knowing *something* about the context in which it was uttered. It has also been argued that indirect reports themselves provide a test for context-sensitivity.⁵ An aspect of the debate on the appropriate divide between pragmatics and semantics has been the extent to which the meanings of terms are sensitive to context. Another way of thinking about this is that if an expression is identified as 'context-sensitive', then its meaning is underdetermined without knowledge of context. As an example, consider a clearly context-sensitive expression such as 'today' embedded in an indirect report:

⁴ Indirect reports have thus been described as containing "two voices" and as being micro-narrations. In reporting, the speaker is using her own voice in order to convey the voice of another (or from another context); in doing so she is establishing a kind of narrative with the locution "said that" that presupposes both common knowledge between the reporter and audience of the reporter's goal, and common knowledge of relevant features of the context of the original utterance. See Capone (2010).

⁵ See Cappelen and Lepore (2004).

(7) Marylee said that she will arrive today.

As reported, the content of Marylee's original utterance is underdetermined. It depends, at the very least, on the day of her original utterance. (7) is a felicitous report if Marylee's utterance of "I will arrive today" occurred on the same day as the utterance of (7). Likewise, (7) is a felicitous report if Marylee's utterance of "I will arrive tomorrow" occurred the day before the utterance of (7), etc. The way in which indirect reports can test for the context-sensitivity of an expression is by seeing if it survives being reported (without quotation marks but with the same actual expression) in a new context. We can see that obvious indexicals such as 'I', 'you', 'he', 'she', 'today', and 'tomorrow', and demonstratives such as 'this' and 'that' do not survive this test and so must be context-sensitive. On the other hand, other candidates for context-sensitivity such as adjectives do gradably pass this test—at least at first glance. For example, consider an expression such as 'is tall'. Consider whether, when embedded in an indirect report, this expression compromises the felicity of said report:

(8) She said that her niece is tall.

It's arguable that an utterance of "My niece is tall" by the original speaker contains the context-sensitive expression 'is tall'. After all, there are many ways of being tall, and many comparison classes against which height can be measured (e.g., tall for a seven year old, tall for a WNBA player). Given this context sensitivity, we might expect (8) to be underdetermined in the same way that (7) is. Yet, this is not obviously the case.⁶ This tells us that indirect reports themselves might provide a tool for determining the extent to which language is context-sensitive, and thereby provide insight to the boundary between semantics and pragmatics.

2.3 Indirect Reports and Inter-substitutability of Co-referential Terms

Co-referential terms and expressions can be substituted for one another in indirect reports felicitously. For example:

(9) A: This morning I had pancakes, toast, and coffee.
 B: A said that she had breakfast this morning.

⁶ See Wieland (2010a) for an opposing view.

As we can see, inter-substitutable co-referential terms in felicitous indirect reports fail strict semantic analysis and yet seem to reflect our practices of indirect reporting. Consider this *co-referential substitution*:

- (10) A: My favorite *tapa* is *patatas bravas*.
 B: A said that her favorite *tapa* is the third item on your menu.

In this case, the term ‘*patatas bravas*’ is substituted with a definite description with a value that can only be determined in the reporting context. It would be implausible to suggest that the original speaker meant anything like *the third item on your menu* in the original context of utterance. Nevertheless, ordinary reporting practices take advantage of this sort of inter-substitution. In other cases, *co-referential substitution* will fail. Consider this case:

- (11) A: The word ‘mendacious’ means *lying*.
 B: A said that the word ‘lying’ means *lying*.

While ‘mendacious’ and ‘lying’ are semantically equivalent and thereby intersubstitutable, it would be infelicitous to report A’s utterance in this way. (The reason for this is partly because (11) involves the use of quotation marks, marking this as a mixed report. See Sect. 2.6.) Co-referential substitution depends largely on maintaining the propositional content of the original report. But indirect reports do not rely on strict propositional preservation; in everyday felicitous reporting contexts, propositional content varies so much that it casts into doubt the idea that reports of any kind *require* such preservation.

2.4 Indirect Reports and Propositional Content

Co-referential substitution is not the only way for an indirect report to felicitously modify an original utterance. An examination of the varieties of indirect reports suggests that indirect reports need not represent the propositional content of the original utterance in its entirety or without alteration. This, perhaps, is the primary reason why a semantic analysis of indirect reports is of limited utility in comparison to a pragmatic account. For example, consider what we can call *conjunction elimination*⁷:

- (12) A: I went to the taco stand and bought a soda.
 B: A said that she went to the taco stand.

⁷ See Cappelen and Lepore (1997b), p. 282. They refer to these and some of the following examples as those with “partial semantic overlap” and suggest that such examples will require a pragmatic analysis.

The propositional content of *A*'s utterance minimally includes two conjuncts. *B*'s report of *A*'s utterance leaves out one of the conjuncts, so *A*'s utterance and *B*'s report do not contain the same propositional content. Consider another case. Call this *modifier elimination*:

- (13) *A*: I had some delicious nachos for dinner.
B: *A* said that she had nachos for dinner.

This is also a case where the propositional content does not survive from *A*'s utterance to *B*'s report. In cases such as these, the felicity of a report that alters a modifier cannot be determined along semantic lines. To see this, consider a related possibility, call it *modifier introduction*:

- (14) *A*: I met a woman at the party.
 **B*: *A* said that she met a beautiful woman at the party.

Most (but not all) of the time, *B*'s report will not be felicitous. Similarly, not all cases of *modifier elimination* will be pragmatically felicitous. Consider:

- (15) *A*: I had some low-fat nachos for lunch.
B: Did *A* have anything healthy to eat today?
 **C*: *A* said that she had nachos for lunch.

Some *modifier eliminations* and *modifier introductions* alter the original utterance in a pragmatically infelicitous way and some do not. These are governed by pragmatic constraints on relevance and not semantic rules (or so I argue in the final section of this paper).

Another example of how an indirect report can modify original propositional content is when the report involves an *inference* based on the content of the reported utterance and knowledge of context. Consider this example of an *inferential indirect report*⁸:

- (16) *A*: I didn't fail any students.
B: Professor *A* said Maryanne passed her exam.

Just as long as *B* knows that Maryanne is one of *A*'s students, then *B* can felicitously report *A*'s utterance in this way. The fact that the intersubstitutability of co-referential terms and paraphrase on the basis of inference are not only possible but commonplace suggests that an indirect report does not function to replicate the original utterance, and it does not even function to convey content that is identical to the original utterance, but rather its pragmatic function is to convey whatever is relevant about the original utterance to the reporter and audience given new facts about the reporting context. Relevant features include the

⁸ This example is from Cappelen and Lepore (1997b), p. 285.

word choice of the original speaker, propositional content of the original utterance, intonation, stress, or accent of the original utterance, the broad structure or word order of the original utterance, etc. Similarly, facts about the reporting context are just as relevant to the nature of the report. For example, if the original utterance is in French, and the audience only speaks Italian, the report will be a translation.

Finally, one further category of report to consider is the implicature. Consider the following report:

- (17) A: It's awfully stuffy in here.
 ?B [to C]: A said that she would like for you to open a window.

What is interesting about this report is that it affords a natural felicitous reading and an equally natural infelicitous reading. The implicature *that A would like for you to open a window* is natural and readily available to B. It's less clear that it's felicitous to say that A said such a thing. This case differs from the *inferential indirect report* because while speaker A did not utter the words attributed to her by B in either case, in an inferential indirect report, B's report of what A said logically follows from what A said (and other facts known in context). It would be reasonable for B to object *that she did not say this* in the case of a reported implicature. For either of these cases, reading them as felicitous requires a robustly pragmatic interpretation of 'said that'—one that many would not be willing to accept.

2.5 Are Indirect Reports Opaque?

Given the above discussion, we can address another corollary between indirect and direct reports. Direct reports are typically described in the philosophical literature as 'opaque'. This means that when synonymous expressions are substituted into the original utterance in quotation, the quotation can fail to preserve truth-value. Given the discussion in Sects. 2.3 and 2.4, we can see that indirect reports are opaque—*some* of the time. This suggests that the locution 'said that' and quotation marks function similarly. However, opacity itself is tricky to identify. An expression is said to be opaque if exchanging synonymous terms alters the truth-value. A standard example looks like this:

- (18) 'Bachelor' has eight letters.
 (19) 'Unmarried man' has eight letters.

The substitution of 'unmarried man' for 'bachelor' results in a failure to preserve truth-value in the move from (18) to (19). But as we know from examples (9) and (10), not all substitutions result in a failure to preserve truth-value. (The preservation of truth-value in examples (9) and (10) is not merely coincidental but

is, in part, a consequence of the flexibility inherent in indirect reports.) However, consider the following indirect report:

- (20) A: Superman can fly.
 ?B: A said that Clark Kent can fly.

Is this report felicitous? The answer here is not obvious. On one hand, it is plausible to say that *B*'s report in (20) is not felicitous because, at the very least, *A* would not accept *B*'s paraphrase of what she said.⁹ Moreover, it would be false to ascribe the belief that *Clark Kent can fly* to *A* because, despite the co-reference of 'Superman' and 'Clark Kent', *A* does not believe this.¹⁰ Although *B*'s report has the same propositional content as *A*'s original utterance, the report is nonetheless false because it does not reflect what *A* believes. Given this, it looks as if indirect reports are in fact opaque.

On the other hand, *B* is not reporting what *A* believes; *B* is reporting what *A* said. It seems plausible that in the context of *B*'s utterance, *B* and her audience know that Superman = Clark Kent. They might know about Clark Kent's secret identity *and* know that *A* does not possess such knowledge. Imagine for a moment that, in the reporting context, it is common knowledge that Superman = Clark Kent, but the extent of Superman's powers is not common knowledge. Under such a description, *B* might be making a reasonably informative report. This suggests that indirect reports are *not* opaque.

The discrepancies in indirect reports prompt us to conclude that the opacity of indirect reports cannot be semantically analyzed in the same way as direct reports. Whether or not indirect reports are opaque depends on pragmatic principles and contextual features. In particular, it depends on the common knowledge that exists between reporter and audience and the communicative goal that the reporter has.

2.6 Mixed Reports

It is difficult to get far in a survey of indirect reports without recognizing that a great deal of reporting takes the form of a mixed report, or a mixed quotation. The analysis of mixed reports borrows from analyses of quotation and analyses of indirect reports. Consider this standard example:

- (21) Quine said that quotation "has a certain anomalous feature."

⁹ Cf. Capone (2010).

¹⁰ There is a substantial literature on belief ascription given co-referential substitution. For a good discussion please see Saul (1998).

One might claim that a mixed report such as (21) does not have the same pragmatic latitude as a pure indirect report. Further, one might claim that (21) requires, at the very least, that Quine used the particular locution, “has a certain anomalous feature,” and not just that Quine uttered an identical speech act type with translatable propositional content. Is a pragmatic account committed to denying this?¹¹ This is an interesting question, and the answer to it turns on *what* is being reported in any given report—direct or indirect (see [Sect. 2.9](#)). It seems arbitrary to presuppose that quotation marks are supposed to capture locutional accuracy rather than accuracy of content, motivation, tone, or purpose. What is being aimed for in any given report might depend on the conversational goals of the reporter and audience. This is not to deny that locutional accuracy is not a normative standard in many settings (e.g., journalism or academic papers). This is compatible with it being far less normative in everyday discourse, particularly when that discourse is verbal (see [Sect. 3.1](#)) and the quotation marks are implied, gestured, or marked in some other way.

While mixed reports might be particularly semantically troublesome, they are less so under a pragmatic analysis. The quoted phrase indicates that the reporter is establishing distance between herself and the original utterance (i.e., she is mentioning and not using this utterance), and that she is doing so in such a way as to establish that her report is accurate along some dimension relevant to the reporting context. If the reporting context is journalistic or academic, it is very likely that her report will be true only if what falls between the quotation marks is an identical locution to that which was uttered by the original speaker. Given that the starting point of a pragmatic analysis concerns the conversational goals of the speakers, and given that the principal pragmatic measure of reports is felicity (and not truth *per se*), the mixed report is analyzed as any report would be. The reading of quotation marks under this analysis of indirect reports is just that they draw special attention to (or distance from) some portion of the report. The reporter has already established that she is not using this language (by way of the phrase ‘said that’) and, further, that she is using quotation marks to re-affirm that point with respect to a particular passage.

This way of interpreting the mixed report makes less of the differences between the direct report, the mixed report, and the indirect report. Quotation in a mixed report is a means of establishing both accuracy and distance with respect to the original report. But quotation itself can operate on various levels, from locution to illocution and even to perlocution. A pragmatic analysis should be ready to recognize this. (See [Sect. 2.9](#).)

¹¹ As suggested by Cappelen and Lepore (2009).

2.7 Indirect Reports as Social Objects

The analyses offered in Sects. 2.4, 2.5, and 2.6 of propositional content in reports, opacity, and quotation marks are undoubtedly unconventional. Many of the claims are defiant of a straightforward semantic analysis of the phrase ‘said that’. And yet they may be most appropriate for a pragmatic perspective on these phenomena. One of the reasons is that pragmatics, as a discipline, is disinclined (relative to semantics as a discipline) to concern itself with the utterance in isolation. It is far more concerned with what has variously been called the *context* or the *situation* of utterance. And contexts and situations are usually social events. They involve a reporter, an audience, a place, goals, and past conversation. They also involve cognitive mechanisms, loads, limits, and predispositions. The pragmatist is more likely to situate the report in the reporting context with a simultaneous representation of the original situated speech act.

On the face of it, this seems like an appropriate approach for the analysis of indirect reports. After all, they explicitly involve the recreation and representation of what someone else said, and often what they meant, why they said it, and how they said it, all while maintaining the report’s relevance to the current speech situation. As a speech act, it’s explicitly intercontextual and metarepresentational. The challenge is to maintain accuracy of description of the speech act itself while maintaining cognitive, empirical plausibility. Talk of *context* and *situation* and multi-level representations is widespread.¹² Yet I do not think there is a decisive account of these terms that maintains empirical plausibility—for example, the account of metarepresentation needs to characterize the semantic and pragmatic function of a report without proposing a structure of representation that is implausibly complicated. (See Sect. 3.2 for further discussion.)

2.8 As Pragmatic Vehicles: Humor, Irony, Insult, Praise

Indirect reports are used as a number of different kinds of pragmatic vehicles. They are also affected by the non-literal content of the original speech act. Irony presents a clear case for the need for a pragmatic account of indirect reports. Consider this case¹³:

- (22) After a really bad philosophy talk, *A* says:
A: That was, like, *really* good.
 **B*₁: *A* said that the talk was really good.
*B*₂: *A* said that he didn’t like the talk much.

¹² For an elaborate account of the semantic and pragmatic details of metarepresentation and indirect discourse, see Recanati (2000). For an incisive criticism of this attempt, see Ludwig (2003).

¹³ Originally in Larson and Segal (1995). Quoted in Cappelen and Lepore (1997b), p. 284.

Even if we might say that *B*₁ literally reports what *A* said, we recognize that it is an infelicitous report of *A*'s utterance since *A* was using irony. A reasonable constraint on the practice of reporting is that consistency of speech act type be maintained.¹⁴ If an utterance is meant to be ironic, the report is infelicitous if it reports that utterance literally. For example, Romeo says that Juliet is the sun, and Montague reports: "Romeo said that Juliet is the sun. But she's not a giant ball of fire. So Romeo said something false." Montague's report is clearly infelicitous because it fails to treat Romeo's original utterance as figurative. Likewise, Montague's conclusion that Romeo has said something false is also clearly wrong. If a speaker uses an interrogative and is reported as having made an assertion, the report is, again, infelicitous. (Such reports could be made felicitous by substituting 'said that' with the appropriate term, 'asked whether', 'joked that', etc.)

However, some indirect reports use the act of reporting in order to insult, praise, or perform other speech acts. In doing so, the reporter does not maintain speech act consistency with the original utterance. It's arguable that if the reporter did so, the report would not serve its purpose—namely, to insult, praise, etc. Consider this use of humor, insult, and scorn in the report but not in the original speech act.

- (23) *A*: I voted for Sarah Palin.
B: *A* said that she voted for Caribou Barbie.

Now there's a clear sense in which *A* did *not* say that she voted for Caribou Barbie—after all, there is no such person. But there are also easily recognizable contexts in which *B* and her audience recognize that 'Sarah Palin' and 'Caribou Barbie' are co-referential even if 'Caribou Barbie' is not the conventional name for Sarah Palin (rather, it's a mildly insulting moniker). Is this just like any other case of co-referential substitution? *A* could readily object that she did not say that she voted for 'Caribou Barbie', especially since she likely finds that name insulting. This example suggests that there is a general constraint that consistency of speech act content be maintained; but it is the violation of this constraint that, in part, makes *B*'s utterance humorous or insulting. It is the very recognition of constraints on such reporting practices that makes these kinds of non-literal reports possible through pragmatic manipulation.

There is a related phenomenon in the use and analysis of slurs. (Note that (23) is not an example of a slur, but rather an insult or expression of scorn and contempt). There is reasonable debate¹⁵ as to whether a report of a slurring expression is itself offensive—or *as* offensive, or offensive in the same way as the original use of the slur. If a speaker were to report the slurring speech of another, it's reasonable to think that the speech is being used mentioned and not used (see Sect. 2.1). As such the reporter is not slurring in reporting the use of a slurring term. Speakers often resist this, however. In reporting on the surprising use of a slurring term, speakers might say,

¹⁴ See Capone (2010).

¹⁵ Anderson and Lepore (2013), 'Slurring words'.

“A said that she is a—I can’t bring myself to use the word,” or “A said that he is an ‘n-word’”—in each case the speaker is unwilling to even repeat the slur in question for fear of using the slur inadvertently. This may be explained by the possibility that some speakers regard slurring terms as non-referring. If they report the original utterance, using the slurring term, the speakers may regard this as an endorsement of the slurring term—at the very least as an endorsement of its referential status. This maintains the insight that reports are acts of mentioning and not using; but simultaneously explains why reporters are hesitant to even report on others’ slurs.

2.9 What Is An Indirect Report Reporting?

As discussed in Sect. 2.1, indirect reports differ from direct reports at least insofar as they lack conventional markings in written texts. (Some have suggested that implied quotation marks are essential features of verbal direct reports as well.) It is also clear that the expression ‘said that’ and its corollaries mark a report and establish pragmatic distance between the current speaker and the reported speech. This presents two related questions: what is doing the reporting and what is being reported?

The literature on quotation might be informative here, although it is not going to answer our questions directly. In the analysis of quotation, some have claimed that whatever is inside of quotation marks is a name (or is being named).¹⁶ As such, the quoted expression is unanalyzable. This improbable view of quotation has been widely rejected. Another view of quotation treats quotation as a description of an earlier utterance.¹⁷ Both the name view and the description view (as we might call them) tell us more about the report (or quote) than they do about what the report (or quote) is operating on. Neither give us clear account of what is being reported, whether locution, propositional content, speech act, etc. Neither the name view nor the descriptive view are defended as plausible accounts of quotation anymore.

No clear winner has dominated the contemporary quotation literature, but there are three major contenders: the demonstrative theory, the disquotational theory, and the use/identity theory.¹⁸ Although each of these views primarily provides a semantic analysis of quotation marks themselves, we can look to them to get an answer to the question of this section as well. The demonstrative theory tells us that quotation marks are demonstratives referring to whatever token expression is

¹⁶ See Tarski (1933) and Quine (1940, 1961).

¹⁷ See Geach (1957).

¹⁸ For the classic defense of the demonstrative theory (also called the paratactic theory), see Davidson (1979) and Seymour (1994). Accounts of the disquotational theory can be found in Richard (1986) and Ludwig and Ray (1998). Accounts of the use/identity theory can be found in Washington (1992), Saka (1998), Reimer (1996), and Recanati (2000, 2001). An overview and analysis of all of these views can be found in Cappelen and Lepore (2009).

between them. Without getting into the details of this view let me just point out one pertinent feature. Under this view, the quotation marks demonstrate *something* about the expression token, but don't indicate what that *something* is. This is a semantic weakness of the view in that it underspecifies a general rule for the use of quotation marks. But it is a pragmatic strength of the view in that it ties the *something* to the goals of the speaker, the common knowledge between speaker and audience, the relevant features salient in the reporting context, and so on. The use/identity theory of quotation also has an answer to the question of this section. It claims that by using quotation marks (implicitly or explicitly), a speaker mentions the same token being uttered between the quotation marks. (As this view has developed, its defenders have claimed that it might also mention a type or a shape of an expression, or even something more loosely related.) The central claim of this theory is that quotation marks (and perhaps all reports) just serve to indicate that mentioning and not using is taking place.

In the case of indirect reports specifically, the problem is slightly different. Given that there is no claim to direct reporting, it is unnecessary to suppose that the report looks like or names the original utterance in some exact way, or that what follows the locution 'said that' is disquotationally equivalent to the original utterance. In fact, indirect reports don't function in that way at all. Indirect reports shift with context (see Sect. 2.2), allow for paraphrase and inter-substitution of co-referential terms (see Sect. 2.4), and are not bothered by translation. Given this, the direct report is only a "picture" of the original utterance if one squints.

So what is the locution 'said that' reporting? The literature on quotation has been focused on the locution itself—quotes somehow refer to a series of words in a particular order—because this is a way of thinking about quotation that strongly resonates. Work on indirect reports has been less concerned with the locution and more concerned with the content and the speech act that delivers that content. If we think of what is being reported as a contentful speech act, there are still a number of possibilities for how to think of the report. For example, the report might be of the type instantiated by the original token expression, a type otherwise related to the original token (through paraphrase or inference, perhaps), some other token related to the reported token, or the original token itself. It could be that an indirect report is less directed toward conveying a contentful speech act than in the concepts contained in the reported expression, or some other possibility. It's not altogether clear that we can choose only one of these options, since what is reported may be a representation of the original utterance either in form (broadly speaking) or content (broadly speaking), depending on the reporter's goals. The great diversity among reports leaves the question of this section far more open than it has been treated in the literature on quotation. For example, indirect reports, in their pragmatic guise, can be used to report on the tone of voice, accent, or attitude of the original speaker. The goal in making the report may not be accurate representation, but humor, scorn, irony, etc. (see Sect. 2.8). In these cases, propositional content and replication of the original locution may be insignificant.

Some have argued that indirect reports must always report the minimal propositional content of the original utterance.¹⁹ This is to say that no matter what the report is *doing* in terms of the conversational goals of the reporter or audience expectations, the report is only a felicitous report if minimal propositional content is preserved in the reporting. Consider the following exchange:

- (24) A: The cat is on the mat.
 B: A said that the cat is on the mat.

Whatever *B* is doing by reporting *A*'s utterance, she reports that *A* said *that the cat is on the mat* and *B*'s report is true iff *A* said *that the cat is on the mat*. This seems plausible, dully so, on its face. But in making this claim the defenders of this view also make the claim that any given utterance may communicate any number of other propositions in addition to the minimal one that can be linguistically decoded from the original utterance. Here's a further elaboration of this position, from Cappelen and Lepore:

What an utterance of (1) says depends in part on the contextually salient comparison class, standards of measurement, and other such things.

(1) Serena is really smart.

...Suppose all we tell you is that Venus uttered (1). We predict the following:

- There is a sense in which you can understand what Venus said, viz. that Serena is really smart.
- You can repeat what Venus said, i.e. do what we're about to do right now, viz. say the same as Venus did: Serena is really smart.
- You can indirectly report Venus by uttering (2):

(2) Venus said that Serena is really smart.²⁰

And another example of how they think content is shared between contexts:

In some unspecified context, Herman utters: 'Napoleon was an interesting guy'. Herman thinks it is obviously true that other people in other contexts have said, could have said, and will say exactly what Herman said with this speech act.²¹

We can imagine for these claims the extension from saying the same thing in different contexts to another form of content sharing—reporting. The reason they think the content-sharing that occurs when we report is so easy is because all instances of the same locution share minimal semantic content once basic ambiguities have been resolved and indexicals have been specified. And why not? The question of what a report is reporting can't be unanswerable—this must not be the case given the ease with which we navigate indirect reports in conversation.²²

¹⁹ See Cappelen and Lepore (2004) and Borg (2004).

²⁰ Cappelen and Lepore (2006), pp. 1020–1021.

²¹ With modification from Cappelen and Lepore (2004), p. 127.

²² See Wieland (2010b) for an opposing view.

If the semantic minimalists' answer to the question of this section is just that an indirect report necessarily reports minimal propositional content and optionally reports a plurality of additional content, the pragmatic theorist on these issues should probably resist. It's easy to see why philosophers and semanticists have tended to think about reporting in terms of semantic identity. But the work of interpreting real world indirect reports has less to do with options semantically triggered by locutions such as 'said that' and more to do with speakers embedded in a context. In a given context, reporters need to convey some understanding or representation of the original context; simultaneously, they need to convey their own reasons for reporting the original utterance, and the relevance of the report to the current context. Sometimes the latter will simply be to convey information, but, oftentimes, indirect reports will be vehicles for a multitude of other speech acts. Wilson (2000) uses the expression "the exploitation of resemblances" to explain the cognitive processes in play when interpreting an indirect report. I think this is the correct way of thinking about it given the variety of reports introduced in this section. The interpreter takes into consideration some collection of clues that are linguistically encoded and takes into consideration contextual information that would lead to the most salient, least taxing interpretation (more on this in Sect. 3). I have tried to emphasize thus far that this available and salient interpretation may resemble the original utterance in a host of ways. Many of these ways may be paralinguistic (e.g., I might report on the utterance of *A* in order to practice my accent in *A*'s dialect), but they are nonetheless common to real world reporting practices.

3 Pragmatic Accounts of Indirect Reports

3.1 Corollaries with Pragmatic Accounts of Direct Reports

How we distinguish between direct and indirect reports differs depending on what we think a language is. Saka (2011) distinguishes between a 'phonocentric' and a 'graphocentric' approach to the study of language. The former he attributes to linguists, arguing that linguists by and large identify language as a natural, verbal (and signed) phenomenon. Written language is, at best, a secondary form of language, meant only to represent verbal speech, and perhaps not a properly linguistic phenomenon at all. The 'graphocentric' approach to the study of language, on the other hand, is the approach most often adopted by philosophers of language, who conceive of language as an abstract (non-natural?) phenomenon. Philosophers of language are most likely to analyze linguistic reality as found in written text and to treat verbal speech as a secondary form of language riddled with 'performance' errors. Saka convincingly argues that settling debates about direct and indirect reports requires first settling questions about the proper approach to language itself. For example, consider these two examples from Recanati (2001, p. 661):

- (25) Alice said that life is difficult to understand.
 (26) Alice said that life is ‘difficult to understand’.

Recanati’s somewhat surprising claim is that, when spoken, these two sentences are, in fact, the same. They do not differ in linguistic meaning, but rather differ only paralinguistically. From a phonocentric perspective this shouldn’t be surprising. If we were to imagine a speaker reporting on Alice’s assertion we would hear the same thing in both cases. There would be nothing linguistically marked as mentioned rather than used, or quoted directly rather than indirectly. Of course, for philosophers of language, the use/mention distinction is so salient, and the intuition that quotation marks themselves have a semantic role is so strong, that Recanati’s claim seems preposterous. If Recanati is correct, or if, more generally speaking, the phonocentric approach to language is the correct approach, the distinction between direct and indirect reports does not dissolve, but it does not rely on what it has been thought to rely on—namely, the presence or absence of quotation marks.

On the other hand, when Recanati analyzes (25) and (26) graphocentrically, he concludes that, while there is no difference between the linguistic meanings of (25) and (26), the quotation marks in (26) indicate that the phrase they contain is being demonstrated, rather than represented in some way that differs from the meaning of (25).²³ This debate is not settled. The semantic role of quotation marks and phrases such as ‘said that’ are still in question, and the pragmatics of how speakers mark off, demonstrate, or report other speech or text is underexplored.

Finally, it’s worth noting that speakers (as opposed to writers) mark the distinction between direct and indirect reports in a large variety of ways. Direct reports are marked in English with interjections such as, “and this is a direct quote,” or modifiers such as “literally,” or “verbatim,” or “word-for-word.” For example, the speaker in (26) may, in a natural speech setting say, “Alice said that life is—literally—difficult to understand.” This would indicate to the audience that the reporter is telling us, directly, what it is that Alice said.²⁴ Reporters also use air-quotes or say “quote-unquote” in order to indicate that the report is direct rather than indirect, although it’s possible that this borrowing from written text only occurs in post-literate settings. (Saka (2011) notes that the use of air-quotes and the phrase “quote-unquote” are often used in clumsy and incoherent ways. This seems right, and seems to suggest that insofar as speakers are borrowing quote-talk to set up a direct report they are doing it without relying heavily on the actual conventions of written text.) These variations in how speakers mark direct reports in English are evidence that they are committed to locutional accuracy, and don’t draw a strong distinction between direct and indirect reports. Speakers certainly distinguish between loose paraphrase and strong accuracy, but it’s not

²³ Note that he is arguing against Cappelen and Lepore (1997a), who reject this analysis.

²⁴ Here we can see how underdeveloped this analysis is. The interjection of ‘literally’ clearly sounds like it is marking off a direct quote. On the other hand, is it functioning as a demonstration? This is less clear.

obvious that they draw this distinction on the basis of locutional accuracy rather than on the basis of some other kind of representational accuracy. Getting at the heart of what is taking place in a report requires a more complete theory of what reporters and interpreters are each doing in the reporting context.

3.2 Resources for a Complete Pragmatics of Indirect Reports

One possible pragmatic framework for understanding indirect reports is to treat them as a kind of metarepresentation. The task of a reporter is to represent relevant features of the reported context to the audience and thereby convey something about that earlier context. The felicity of the report is constrained, in part, by the cognitive demands of interpreting the report given the current context and shared goals. It is plausible, to a certain extent, that the act of reporting involves having a theory of mind in order to understand what was said and meant in the original context, what the audience knows about that earlier context, and a representation of the two contexts and the respective agents' epistemic attitudes toward those contexts.

There are a few apparent problems with analyzing indirect reports in this way. The first is that understanding the locution 'said that' appears to be a relatively simple cognitive process—one not requiring robust metarepresentational abilities. After all, very young children have the ability to deploy and interpret 'said that' before they have a fully-formed theory of mind. The second, and related, problem is that analyzing indirect reports as metarepresentational suggests that speakers and hearers over-represent in a way that is empirically implausible. Grice originally proposed such a model as described here by Wilson:

For a Gricean speaker's meaning to be conveyed the speaker's intentions must be not merely recognized but transparent, in a sense that seems to be definable only in terms of an infinite series of metarepresentations. The speaker must not only (1) intend to inform the hearer of something, and (2) intend the hearer to recognize this informative intention, but (3) intend the hearer to recognize the higher-order intention in (2), and so *ad infinitum*.²⁵

This was an assumption of Gricean semantics and early intention-based semantics, albeit recognized as flawed from early on. Stephen Schiffer correctly describes this approach as "superintellectualist":

In *Meaning* I suggested that in an adequate account of declarative speaker-meaning the speaker must intend the audience to have what I called a "truth-supporting" (as opposed to a pragmatic or prudential) reason for his belief that *p*, but this now strikes me as an especially absurd instance of the superintellectualist conception of mind implicit in the whole IBS approach. I doubt that ordinary speakers can be relied on to have sophisticated beliefs about the sort of grounds they intend others to have for their beliefs. Also it now

²⁵ Wilson (2000), p. 415.

seems to me possible that there should be acceptable cases of meaning where the audience is not intended to have any reasons for the belief that the speaker intends to activate by his utterance.²⁶

In this passage Schiffer is criticizing both the explicitly representational burden as well as the discursive and truth-oriented assumption in this philosophical account. Assuming that a speaker believes *such-and-such* and that his audience recognizes that he believes *such-and-such*, and so on, where each *such-and-such* is identical in linguistic meaning and logical form is implausible. In the case of reporting, this superintellectualist conception of mind is only intensified: the reporter must convey the original speaker's intentions alongside her own, with the relevant information from both contexts, all the while expecting her audience to recognize both the original intention, the intention in reporting, and the higher-order intentions *ad infinitum*. Even if we were to find a way of preventing an infinite iteration of representations, this superintellectualist metarepresentational account of reporting seems too cognitively demanding to be plausible.

While the standard view within pragmatics is that reports are best explained as a kind of metarepresentation, the strongest (but still underdeveloped) pragmatic account avoids the problems of *over*-representation and explicit discursivity. In the case of indirect reports these acts of metarepresentation are even more salient. For example, consider the following exchange:

(27) A: I'm hungry.

B_1 [to C]: A said that she's hungry.

B_2 [to C]: A said that she would like dinner now.

The report in B_1 is more direct than the one in B_2 and presumably requires minimal cognitive resources in order to execute it. The report in B_2 is indirect and inferential. In this case, B understands A 's utterance, the meaning behind A 's utterance, and the relevance of it to C . In doing so, she infers that A would like to satisfy her hunger with dinner. But both B_1 and B_2 are candidates for felicitous reports. One possible explanation for this is that metarepresentation in general, and in indirect reports in particular, is largely non-discursively inferential and not committed to the recovery of literal meanings (the 'what is said' traditionally construed). Since it appears that young children use various forms of inferential reasoning as a means of language acquisition, the inferences themselves cannot antecedently require strong discursive abilities or the recovery of fully-formed literal meanings.²⁷ It should be clear that Grice was aware of these problems and thought it would be sufficient that speakers and hearers regard themselves and their audiences as representing in the requisite ways even if they don't actually do so. Sperber and Wilson (in various places) have defended a more plausible (if still cursory) account. Consider the following:

²⁶ Schiffer (1987), pp. 246–7.

²⁷ See Wilson (2000) and references therein.

Inferential comprehension starts from the recovery of a linguistically encoded sentence meaning, which is typically quite fragmentary and incomplete. The goal of pragmatic theory is to explain how the hearer, using available contextual information, develops this into a full-fledged speaker's meaning. The Communicative Principle of Relevance motivates the following comprehension procedure, which, according to relevance theory, is automatically applied to the on-line processing of attended verbal inputs. The hearer takes the linguistically decoded sentence meaning; following a path of least effort in the accessing of contextual information, he enriches it at the explicit level and complements it at the implicit level until the resulting interpretation meets his expectations of relevance; at which point, he stops.²⁸

The virtue of this account is that it is general but unified enough to make sense of the puzzles described in Sect. 2. It also—to its credit—makes it easy to see how communication can go wrong even when linguistically encoded sentence meaning is readily available to both speaker and interpreter and how the same process can have degrees of sophistication. When an audience misunderstands a speaker—as happens—it is possible that enrichment went wrong at either the explicit or implicit level or that the interpreter's expectations of relevance were inappropriate. After all, these basic pragmatic principles are supposed to explain why it is that complicated communication (in the form of reports, for instance) go well much of the time but also go partially or entirely wrong much of the time as well.

This account also explains the observation that I made above, namely that it seems as if young children are able to deploy and interpret 'said that' even prior to developing a robust theory of mind. In describing this comprehension procedure, Wilson notes that it is compatible with speakers of different abilities. She describes three such speakers who have interpretive strategies she labels *naïve optimism*, *cautious optimism*, and *sophisticated understanding*.²⁹ Each interpretive strategy is associated with a different level of linguistic comprehension and/or theory of mind. For example, a typical language learner progresses through these strategies from infancy to adulthood when speakers and interpreters are typically able to represent complicated theories of mind with respect to one's audience and other speakers that one may be reporting. And as Wilson points out, these categories correspond, to a certain extent, with some of the interpretive obstacles encountered by autistic people. Of course, a complete pragmatic theory will likely propose a more finely-grained non-linear account of the possible interpretive strategies an interpreter might have. But it seems to me that this account, as I've only give the bare outlines of it here, is the best pragmatic (or even bolder: linguistic) framework for the explanation of reporting practices. However, there is still work to be done in philosophy and pragmatics on this meta-linguistic and meta-representational phenomenon.

²⁸ Wilson (2000), p. 420.

²⁹ These are described in Wilson (2000), pp. 421–422. She borrows these three categories from Sperber (1994).

Looking ahead, what will be the components of a complete theory of indirect reports?

- (a) An account of the reporting context and the aims of that context. This would minimally require an understanding of the boundaries of context itself as well as an account of the common knowledge between reporter and audience. Of course, much work has been done by pragmatic theorists to account for context itself. It has been beyond the scope of this paper to describe this work. My suggestion, moving forward, is that the reporting context differs from ordinary speech contexts and, as such, will be described in different terms.
- (b) A full account of the metarepresentational processes needed in order to represent the relevant features of the original context within the reporting context. This metarepresentational cognitive task requires a triple level of embedding of linguistic understanding, common knowledge, and conversational goals given the framework for comprehension described above. As indicated above, it should be scalar both respect to age and ability level but also context.
- (c) An account of the socio-cultural/linguistic interpretive defaults that might be employed interpreting or representing the original utterance as well as the reported utterance. This will require a tandem account of the understanding of socio-cultural/linguistic defaults for the act of reporting itself.
- (d) An account, given (a), of what counts as expressing the same content. What has not been accounted for in this paper, but what is certainly of interest to philosophers, is the way in which reporting is an act that has standards of evaluation. In this paper I have used the broad term ‘felicity’ to describe a successful report. Contexts vary with respect to the required strength of fidelity between the original utterance and the report of this utterance. In some, but not all, cases it is appropriate to assess whether the report is *true*. In some, but not all, cases it is appropriate for there to be *identity* between the original utterance and the report. In other cases, the two utterances, whether in content, form, or something else entirely paralinguistic, need to resemble each other in some other way altogether. In this paper, I have argued that what counts as *saying the same thing* in the original utterance and the report will depend on the norms of discourse for the relevant domain and the conversational goals of the reporter and her audience. This implies that there is no unified analysis of ‘said that’ or its corollaries that is either available or desirable. However, this does not excuse pragmatic theory from establishing the general principles that constrain speakers in particular cases.

The principles of cooperation proposed by Grice, the general principles of relevance theory, alongside the few additional pragmatic principles aimed at indirect reports that have been proposed here and in recent work on indirect reports go a long way toward explaining many of the linguistic puzzles raised by indirect reports. Indirect reports are proving to be one of the most illuminating linguistic phenomena for understanding the pragmatic/semantic divide, the proper analysis of ‘what is said’, metarepresentation, and quotation.

References

- Anderson, Luvell and Ernest Lepore. 2013. Slurring words. *Noûs* 47: 25–48
- Borg, Emma. 2004. *Minimal semantics*. Oxford: Oxford University Press.
- Capone, Alessandro. 2010. On the social practice of indirect reports (further advances in the theory of pragmemes). *Journal of Pragmatics* 42: 377–391.
- Cappelen, Herman, and Ernest Lepore. 1997a. Varieties of quotation. *Mind* 106: 429–450.
- Cappelen, Herman, and Ernest Lepore. 1997b. On an alleged connection between indirect speech and the theory of meaning. *Mind & Language* 12: 278–296.
- Cappelen, Herman, and Ernest Lepore. 2004. *Insensitive semantics*. Oxford: Basil Blackwell Publishers.
- Cappelen, Herman, and Ernest Lepore. 2006. Shared content. In *Handbook of Philosophy of Language*, The Oxford. New York: Oxford University Press.
- Cappelen, Herman, and Ernest Lepore. 2009. Quotation. In *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, ed. N. Zalta. <http://plato.stanford.edu/archives/fall2009/entries/quotation/>.
- Davidson, Donald. 1968. On saying that. *Synthese* 19: 130–146.
- Davidson, Donald. 1979. Quotation. In *Inquiries into truth and interpretation*, 79–92. Oxford: Oxford University Press.
- Geach, Peter. 1957. *Mental acts*. London: Routledge Kegan Paul.
- Larson, R., and G. Segal. 1995. *Knowledge and meaning*. Cambridge, MA: MIT Press.
- Ludwig, Kurt. 2003. François Recanati's *Oratio obliqua, oratio recta: an essay on metarepresentation*. *Philosophy and Phenomenological Research* 66: 481–488.
- Ludwig, Kurt, and Georges Ray. 1998. Semantics for opaque contexts. *Philosophical Perspectives* 12: 141–166.
- Quine, W.v.O. 1940. *Mathematical logic*. Boston, MA: Harvard University Press.
- Quine, W.v.O. 1961. *From a logical point of view*. Cambridge, MA: Harvard University Press.
- Recanati, François. 2000. *Oratio obliqua, oratio recta: an essay on metarepresentation*. Cambridge, Mass: MIT Press.
- Recanati, François. 2001. Open quotation. *Mind* 110: 637–687.
- Reimer, Marga. 1996. Quotation marks: demonstratives or demonstrations? *Analysis* 56: 131–142.
- Richard, Mark. 1986. Quotation, grammar, and opacity. *Linguistics and Philosophy* 9: 383–403.
- Saka, Paul. 1998. Quotation and the use-mention distinction. *Mind* 107: 114–136.
- Saka, Paul. 2011. Quotation and conceptions of language. *Dialectica* 65: 205–220.
- Saul, Jennifer. 1998. The pragmatics of attitude ascriptions. *Philosophical Studies* 92: 363–389.
- Schiffer, Stephen. 1987. *Remnants of Meaning*. Cambridge: MIT Press.
- Seymour, Michel. 1994. Indirect discourse and quotation. *Philosophical Studies* 74: 1–38.
- Sperber, Dan. 1994. Understanding verbal understanding. In *What is intelligence?* 179–198. Cambridge: Cambridge University Press.
- Tarski, Alfred. 1933. The concept of truth in formalized languages. In *Logic, semantics, metamathematics*, 2nd edn., 1983, 152–278. Indianapolis: Hackett.
- Washington, Corey. 1992. The identity theory of quotation. *Journal of Philosophy* 89: 582–605.
- Wieland, Nellie. 2010a. Context sensitivity and indirect reports. *Philosophy and Phenomenological Research* 81(1): 40–48.
- Wieland, Nellie. 2010b. Minimal propositions and real world utterances. *Philosophical Studies* 148(3): 401–412.
- Wilson, Dierdre. 2000. Metarepresentation in linguistic communication. In *Metarepresentations: a multidisciplinary perspective*, 411–448. New York: Oxford University Press.

Immunity to Error Through Misidentification (IEM), ‘de se’ and Pragmatic Intrusion: A Linguistic Treatment

Alessandro Capone

Abstract In this paper, I defend the idea that pragmatic intrusion is involved in ‘de se’ constructions: the ego-concept being a component of the ‘de se’ thought. I defend this idea from a number of objections. I explore the related notion of immunity to error through misidentification and I claim that this too depends on pragmatic intrusion. I defend this view from obvious objections. I take immunity to error through misidentification in ‘de se’ thoughts to depend on the fact that the thinking subject makes an implicit use of the first-person pronominal and there is no question of attributing a referent to the pronominal, since the referent is given in the subject of the thought. In third-person ‘de se’ attributions, some form of simulation can be used to reconstruct the thinking subject as using a form of the first-person pronominal. Immunity to error through misidentification is attributed to the thinking subject through simulation.

1 Introduction

In this paper I try to reduce a philosophical topic to a linguistic one. The philosophical topic is ‘Immunity to error through misidentification’ (one who thinks ‘I remember walking in Oxford’ can be mistaken as to whether it was Oxford or Cambridge where he walked, but cannot be mistaken as to whether it was him who walked (in Oxford or Cambridge)) and the linguistic topic is the first-person pronoun (and all its semantico-pragmatic mysteries) in the context of a discussion

I would like to very warmly thank Wayne A. Davis who kindly commented in detail on this paper. Of course, all remaining errors are my own.

A. Capone (✉)

University of Messina/Palermo, San Francesco P 107, Barcellona PG, ME 98051, Italy
e-mail: alessandro.caponis@gmail.com

of pragmatic intrusion (the question whether ‘de se’ thoughts incorporate semantically or pragmatically the ‘Ego’ component is an important one). This article explores the consequences of accepting certain ideas by Castañeda and after a brief rejection of the attempt by Campbell (1999) to expunge linguistic semantics from the issue of I.E.M. (thus confining it to epistemology), I show in detail that the issue benefits enormously from the findings on pragmatic intrusion. Detailed arguments are provided on why pragmatic intrusion and not merely semantics is involved in understanding the topic of IEM.

2 Part 1

2.1 Immunity to Error Through Misidentification

Instead of summarizing the literature on IEM in detail, I will sketch with broad brush strokes the state of the discussion. The positions on IEM are mainly two. There are philosophers who claim that the phenomenon of IEM derives from semantic considerations, mainly considerations of direct reference (Shoemaker (1968); Evans (1982); Strawson (2012); Peacocke (2012); Christofidou (2000)) and authors like Recanati (2009) who relate the phenomenon of IEM to the ascription of mental or physical properties either through introspection or proprioception or through perception (in the ordinary mode). Shoemaker seems to occupy a unique position in the philosophical panorama since his approach to IEM combines both epistemology and semantic considerations. I will mainly sum up the position taken by Shoemaker and then voice some objections by Wayne Davis (personal communication) which might well be seen as deriving from the alternative camp of philosophers represented eminently by Recanati.

2.2 Shoemaker’s Position

The article by Shoemaker (1968) is clearly indebted both to Wittgenstein and to Castañeda. From Wittgenstein he takes the notion of IEM and from Castañeda he takes the idea that IEM is linked to use of first-personal pronominals in thought or assertion.

Shoemaker deals with immunity to error through misidentification and seems to imply that this immunity is linked to the first person. Basically, IEM consists in the fact that one cannot ascribe oneself a mental property (e.g. I am in pain) and ask oneself ‘Is it me who is in pain?’. In other words, the possibility of misidentifying the subject does not occur. Shoemaker thinks that the first person (the use of the pronominal ‘I’) has referential uses, a position which, as we will subsequently see, is important for showing where IEM comes from, since in using the pronominal ‘I’ the subject cannot but refer to himself and also cannot misidentify himself.

Shoemaker connects IEM with the pronominal 'I' but he is careful to stress that IEM is not connected with demonstratives in general, as I could say 'That is green' while I am in an hallucinatory state.

What is clear in the discussion by Shoemaker is that the person who attributes himself some property by using the word 'I' cannot fail to identify himself, and this applies both in cases of corrigible or incorrigible statements (Shoemaker distinguishes between incorrigibility and IEM). The subject is available to himself through **introspection** and Shoemaker distinguishes between the ordinary mode of perception and perceptions which we would ordinarily express using the verb 'feeling' by saying that in ordinary perception acts we can perceive the perception act but we cannot do so with perception acts in which feeling is involved. This concept probably needs deepening, but what Shoemaker is after is, presumably, a way to distinguish perception by the senses from introspection, in which the subject is immediately aware of himself (the object is presented to the self from the inside).

I would now like to address a problem noted by Wayne Davis (personal communication) in the formulation of IEM by Shoemaker. According to Shoemaker, at least one aspect of IEM consists in the fact that one cannot ascribe oneself a mental property (e.g. I am in pain) and ask oneself 'Is it me who is in pain?'. The utterance "I am in pain. Is it me who is in pain?" would be unquestionably odd, but an utterance of "John is in pain. Is it John who is in pain?" is also odd. The oddity of "I am in pain. Is it me who is in pain?", according to Davis, is not due to the fact that 'I' is first-personal and pain mental. It is due to the fact that the first sentence obviously entails the answer to the second. So this cannot be what one means by IEM.

Of course, Davis is right in noting the entailment relationship, which obviously makes such pairs of utterances (in which the preceding utterance entails the following utterance) odd. However, an appropriate question would be to ask how the entailment works when such utterances are embedded in a verb of propositional attitude like 'believe'. The pair (assertion and question): "John believes he is in pain. Is it John who is in pain?" is not odd, after all. On the contrary, the pair 'I believe I am in pain. But is it me who is in pain?' is odd. Suppose Davis is right and the oddity can be imputed to the entailment (or to a pragmatic implication of uttering the preceding utterance, given that now the relationship between the preceding and the following utterance no longer seems to be one of entailment). Then one could replace the pair of utterances related by entailment or a similar relation with a pair of utterances which are no longer related in this way. Suppose I say "I believe I am in pain. But if you were me, could you doubt that it is me who is in pain?". Now, it is no longer clear that the relationship between the preceding utterance and the answer to the following utterance is one of entailment. The question now seems to be an epistemological/semantic one. One who has such a belief, regardless of the way she herself expresses that belief, cannot reasonably doubt that it is her which is in pain.

Wayne Davis, after rejecting the possibility that IEM consists of the impossibility or implausibility of asking a certain kind of question, after one has expressed a certain belief, says that IEM consists in the fact that one cannot be mistaken in ascribing to oneself (*de se*) a mental property. This is the view that beliefs resulting from introspection are infallible. Davis considers cases of people who deceive themselves or people whose desires are unconscious as providing evidence against the view that introspection guarantees infallibility.

However, it should be clear that many of the authors who defend the idea of IEM do not equate it with infallibility; and they do well to do so, because, otherwise, their views could be attacked on grounds similar to Davis'. Immunity to Error through Misidentification seems to be a logical property of certain kinds of beliefs or statements, rather than only being a claim on epistemology. It seems to me that IEM derives from considerations on what one must accept if one accepts a certain belief which is held in a certain way. I propose that IEM is one of those issues where epistemology depends on the semantico/pragmatic apparatus of language.

Before proceeding with this paper, I want to deal with a problem in order to restrict the scope of this paper. It may appear that I am assuming that a subject's introspective beliefs about his or her own '*de se*' mental states cannot be mistaken. This assumption, however, might be questioned on the grounds that it denies that self-deception and unconscious mental states ever occur. Since I do not want to deny that self-deception and unconscious mental states ever occur, it is obvious that my claim concerning IEM must be restricted and qualified in such a way that it does not have this undesired consequence. In other words, we shall retain only the part of the claim that is compatible with the idea that sometimes self-deception occurs. The cases I consider in this paper are clearly not those where self-deception can occur, but those that exclude self-deception. Introspection as a safe method of knowledge of internal states is here seen as confined to those cases where one cannot be mistaken about one's own basic identity ($I = I$). Introspection of the self as self seems to be one case in which self-deception has to be excluded. In this paper I want to tie closely IEM to a semantic claim and I want to make epistemology subordinated to semantics.

Another point I want to make clear from the beginning is that although I assume that '*I*' typically refers to the speaker or to the thinker, this need not always be true (a referee mentions arguments by Castañeda), nor is this truth indispensable to my claim about pragmatic intrusion. If my views were expressed through semantic terminology or only through semantic concepts, then this defect could not be easily remedied. But since my paper is about pragmatic intrusion, all I need for my claims to get through is that the pronominal '*I*' be interpreted in those cases when IEM is discussed as first-personal, that is to say as referring to the thinking subject. Surely pragmatic intrusion suffices to guarantee this.

2.3 *Semantics or Epistemology in IEM? (Campbell 1999)*

Since my paper constitutes a slide from an epistemological to a semantico/pragmatic treatment, before proceeding with my pragmatic treatment I want to present and then refute the position by Campbell (1999), according to which there are no reasons for believing that IEM should rest on a semantic perspective. This section also makes clear how linguistic theses about the first person pronoun could explain why introspective ‘de se’ beliefs cannot be mistaken.

The paper by Campbell links the notion of IEM to the special processes involved in having experiences of mental states, which he takes to be **dedicated** ways of finding out about the world. It appears to me that Campbell is trying to find out whether the notion of IEM can be disentangled from the idea that it depends on the meaning of the first person. He rather seems to opt for the idea that it depends on the nature of the experience of the mental state, on the dedicated process of attributing oneself a mental state. The paper starts with an example, in which IEM seems to be linked with the meaning of an expression. Then he considers a different example invariably showing that IEM can depend on the contingencies of a situation, rather than on the meaning of a linguistic expression.

Campbell considers the case of a businesswoman who has a secretary keeping track of all her movements. The secretary also travels a great deal, but he never has any idea where she is (suppose she herself has a secretary to keep track of her movements for her). We may suppose that the secretary at any moment knows where her boss is, but she knows nothing about where anyone else is, including herself. If you call the secretary, she will always be able to tell you where her boss is, but she cannot give you any information about where anyone else is. You might sometimes get misinformation, but NOT misinformation that will involve a mistake of identification. It cannot happen that the secretary will inform you that her boss is in London, when, in fact, she ought to say that Henry (who is not her boss) is in London (having mistaken Henry for her boss, somehow). As Campbell says, this case of IEM depends on the contingencies of the situation and NOT on the meaning of a linguistic expression.

This “tortured” example was offered as a contrast to the view that IEM can depend on the meaning of the first person pronominal. Unlike other scholars, Campbell opts for the idea that it is the contingency of the situation of experiencing a mental state that explains IEM. Experiencing or possessing a mental state involves, according to Campbell, a self-attribution. And a self-attribution is a **dedicated** process:

In the case of the first person, what is happening is rather that the subject is using ways of finding out about the world that are, as we might say, “dedicated” to the properties of one particular object, namely that very person. They are not ways of finding out that could equally apply to any of a range of objects. It is for this very reason that although the subject using such a way of finding out can make a mistake, it could not be a mistake about who is in question (Campbell 1999: 95).

Campbell points to the asymmetry between the task of finding out the referent of 'I' when 'I' is uttered by someone different from the hearer and understanding the use of the first person by the person who uses 'I'. When you use 'I' you do not need any identification act at all.

I should note that I agree with Campbell on the idea that IEM rests on a dedicated process, however I do not agree that this dedicated process is independent of the meaning and interpretation of the pronoun 'I'. Campbell writes about self-ascription as if it could ever occur without using the word 'I', something which I tend to exclude (And Campbell himself avoids the solution that a thinker can make use of the concept 'whoever is having these experiences', because this would allow in the case of 'I am hearing trumpets' 'whoever is having this headache' as subject and one could not ensure identity between "whoever is having this headache and say the person who is experiencing trumpets (see Campbell 1999, p. 94)).

In fact, when Campbell expatiates on the difference between interpreting 'I' relative to a context (when 'I' is not uttered by the interpreter) and interpreting 'I' by the person who uses it, it is clear that matters of interpretation bear on IEM, because Campbell says that the thinker who produces 'I' cannot be mistaken about the identity of 'I' and I think (and Campbell seems to imply) that IEM is connected with the impossibility of mistaken identity. Of course, one needs to show that IEM derives from semantic considerations. But this is not impossible. Consider the process of producing 'I' in mental speech? Is it not dedicated too? This process seems to be severely constrained. You do not start with the rule: 'I' refers to the person who produces it and then look for the referent. If you did so, then you could end up retrieving the wrong referent. When you use 'I' you do not go from the context of utterance to see who produced 'I' to the referent of 'I', but you go the other way round, you start with the referent and produce a way of presentation of the referent. So, the process is dedicated in the sense that you cannot be mistaken about the identity of the person producing 'I' simply because the deictic 'I' in silent speech involves a dedicated process: Start with the referent and make sure that the same referent is what your speech is about when you use several occurrences of 'I'. But this looks very much like the description of a semantic phenomenon.

2.4 'de se' in Philosophy and Linguistics

Immunity to error through misidentification, we have claimed, stems from the use of 'I' and from certain basic semantic characteristics of 'I'. It follows logically that any treatment of 'de se' thoughts incorporating the first person inherits immunity to error through misidentification for the thought, memory or other attitude described. Whether immunity is direct or derived is not really important for us, since our task is to describe what thoughts a person has in attributing herself a propositional attitude; thus, it goes without saying that if our task is to describe what thoughts are in the mind of a person to whom we ascribe a propositional

attitude, then we also have the task of specifying whether those thoughts have IEM or not. In the following section, I discuss both ‘de se’ thoughts and immunity to error through misidentification. The two issues, in fact, seem considerably entangled. Of particular importance for us is the case of description of a ‘de se’ thought, in which we attribute a first personal mode of presentation to a thinker/believer and thus we simulate possession of IEM.

In this section I shall present what I take to be the most influential theories on ‘de se’. Higginbotham’s view is philosophical/linguistic, but I have decided to include it in this philosophical section because it is the only one that has the merit of unifying the first-personal character of ‘de se’, with phenomena such as the internal dimension of PRO and immunity to error through misidentification. I will mainly use the perspective outlined in Higginbotham (2003), because it is linguistically explicit, in making recourse to anaphoric concepts and to concepts taken from Fillmore’s theory, and I will supplement it with considerations by John Perry (the idea that the essential indexical needs to make use of the concept ‘I’ at some level of (pragmatic) interpretation).

2.5 Higginbotham (2003)

In his seminal paper, Higginbotham deepens ideas by Castañeda (1966) and Perry (1979). The basic idea is that a ‘de se’ mode of presentation differs significantly from a ‘de re’ mode of presentation, having important implications for action and decision. Higginbotham (2003) recognizes that there is something special about first-personal uses of pronominals such as those discussed by Castañeda. The merits of his discussion lie in his pointing out that constructions with PRO may even be more first-personal than uses of ‘he himself’ and in linking the issue of immunity to error through misidentification to the issue of the internal perspective in connection with PRO (in cases of verbs like ‘remember’, ‘imagine’, etc.). He claims that the propositional analysis articulated through the notion of anaphora and thematic roles is superior to the property-based view of Lewis and Chierchia. In fact, according to him, the property-based analysis of beliefs and attitudes ‘de se’ does not allow the theorist to explain (1) immunity to error through misidentification; (2) the internal dimension of PRO in complements of verbs such as ‘remember’, or ‘imagine’.

Higginbotham accepts Perry’s idea that ‘de se’ attitudes involve a first-personal mode of presentation (involving sometimes the word ‘I’ or some related notion) and reformulates such a view through considerations based on anaphora and thematic relations.

Higginbotham also accepts Peacocke’s (1981) consideration that a ‘de se’ thought involves the use of a mode of presentation ‘self’ which only the thinker can use and which nobody else can use in reporting such a thought.

Higginbotham considers cases with PRO such as:

1. John remembers PRO going to Paris
2. John remembers that he went to Paris;
3. John remembers that he himself went to Paris.

The first-personal nature of (1) is expressed through a notation which involves self-reflexive thought:

4. For $x = \text{John}$, $\exists e$, remember $(x, e, \wedge \exists e'$: go to Paris $(\sigma(e), e')$)

(1) is different from (2) because it involves an internal dimension. It is the internal dimension which, according to Higginbotham, causes immunity to error through misidentification. We can capture this 'internal dimension' through the expression in logical form of a thematic role: the person who undergoes the action in question. So we can reformulate (1) through (5)

5. For $X = \text{John}$, $\exists e$, remember $(x, e, \wedge \exists e'$: go to Paris $(\sigma(e) \ \& \ \theta(e'), e')$).

With this elucidation in mind, we can explain the following facts:

Only Churchill gave the speech
 Churchill remembers giving the speech
 :
 Only Churchill remembers giving the speech.

Surely someone who listened to the speech remembers that Churchill gave the speech or remembers his giving the speech. But are the speeches which Churchill remembers giving and which another person remembers hearing the same kind of thought? At some level of abstraction they are. At some deeper level, however, they are not. What validates the inference in the deduction above is the fact that Churchill remembers giving his speech from the inside. So in case he has forgotten giving the speech and someone else informs him that, in fact, he gave the speech, Churchill cannot (truthfully) say that he remembers giving the speech. Memory involves an internal perspective in case PRO is used in the complement clause. Thus, if one remembers falling downstairs, one must certainly have memories of sensations of pain; something which one need not have in case memory is reconstructed through an external narration.

Higginbotham discusses an interesting question. He asks whether mad Heimson who believes that he is Hume has numerically the same belief as Hume. The question, put crudely, is if the belief Heimson has in believing that he himself is Hume is the same which Hume has in believing that he himself is Hume. The answer by Higginbotham is ambivalent. On the one hand, their beliefs are different, so much so that we must say that, in believing he is Hume, Heimson has a false belief while in case Hume believes he is Hume, we shall say that he has a true belief. This is nicely expressed through Higginbotham's anaphoric treatment:

6. For $x = \text{Heimson}$, $\exists e$, believe $(x, e, \wedge \exists e'$ (be Hume $(\sigma(e) \ \& \ \theta(e'), e')$).

Since $\sigma(e)$ is anaphorically related to Heimson, there is clearly an **external** component to that thought. However, Higginbotham says that at some level of generality, we can say that Heimson and Hume have the same thought

Be Hume ($\sigma(e)$ & $\theta(e')$, e')

Higginbotham illustrates this through an analogy to two collapses of bridges. Of course, in one sense two collapses of bridges cannot be the same event, unless the bridges are the same. In another sense, we could say that the collapses of two distinct bridges are the same type of event provided that the bridges have similar characteristics.

3 Part 2: Pragmatic Intrusion

3.1 Pragmatic Treatments

In this section, I will report the pragmatic treatments by Capone (2010b), based on Relevance Theory considerations. Here I spell out certain implications of the ideas in Capone (2010b).

3.2 Capone (2010b) and the Pragmatics of ‘de se’

Capone (2010b) is an eclectic treatment mixing linguistic, cognitive and philosophical considerations in order to predict pragmatic results. His approach is rather eclectic and is a rethinking of pragmatic scales à la Levinson/Horn/Huang in terms of considerations based on Relevance Theory. His ideas, in essentials, are very simple. If one accepts Higginbotham’s considerations on the logical forms of ‘de se’ and ‘de re’ beliefs (to pick up just the most representative of the attitudes), it goes without saying that the logical forms of ‘de se’ beliefs entail the logical forms of ‘de re’ beliefs (See also Tasheck 1985). Hence the possibility of pragmatic scales. On a Relevance Theory line of thinking, the ranking of ‘de se’, ‘de re’ in terms of entailment entails a ranking in terms of informativeness. Then it goes without saying that a ‘de se’ interpretation of a pronominal (in those cases where both interpretations are possible) is informationally richer and, following the Principle of Relevance, greater Cognitive Effects, with a parity of cognitive efforts, are predicted. We may also concoct stories in which a ‘de se’ interpretation leads to some kind of action which the ‘de re’ interpretation would never cause (See Perry; see also Capone 2010b, the pill story). If this line of thought is accepted, then we can explain easily why

7. John believes he is clever

tends to be associated with a ‘de se’ interpretation. As Jaszczolt would say, this interpretation tends to be ‘default’. Of course, its default status derives from the way the mind is predisposed to calculate inferences and also from the human tendency to standardize or short-circuit familiar inferences that are probabilistically high.

A possible objection to this line of thought may be the following:

You have not given any reason for thinking that the stronger interpretation is as easy to process as the weaker one and no method of measuring cognitive effort is identified. Furthermore, the fact that one interpretation entails another does not ensure that it has greater cognitive effects.

Answer:

While the objection is correct in principle, it does not work in practice for the following reason. A ‘de se’ interpretation is stronger not only because of entailments, but because of the effects of semantic entailments. A ‘de se’ interpretation, compared to a ‘de re’ interpretation’ limits considerably the amount of objects that need to be taken into consideration for interpretation. If John believes ‘de re’ that he is happy, then in principle there is no limit to the objects which we can consider to establish which person John is thinking about (what the belief is about); we surely need a theory of context in order to limit the number of objects that can be taken into consideration. Even if the context is severely restricted, there may at least be a couple of alternatives, such that they would make the interpretation process more costly than in the case of the ‘de se’ interpretation, which is only about a single object. So it seems to me I have amply shown that processing costs are higher in the case of ‘de re’ interpretations. I agree that entailments need not be equated with positive cognitive effects, but in the case of a ‘de se’ interpretation a positive cognitive effect which is a consequence of entailments is that the thought has consequences on the person thought about. So, If I think I have to go, I will go, and I will not refrain from moving paralyzed by the thought that the person I am thinking of could be someone else. So, the fact that a ‘de se’ interpretation entails a ‘de re’ interpretation immediately translates into immediate and important positive effects (more or less the kind of effects on action Perry was thinking about in his famous paper).

According to Capone, one may also investigate scales such as the following:

8. John wants to go away;
9. John wants [him to go away];
10. John remembers going away;
11. John remembers his going away.

The use of the marked pronominal, instead of less marked PRO, tends to invite an interpretation which is complementary to that associated with PRO. This can be explained in terms of M-scales in the framework of Levinson/Horn/Huang or in terms of cognitive efforts, which tend to pick up an interpretation disjoint from the one associated with the expression involving least amount of cognitive efforts.

Capone also explains certain interesting examples by Perry, which seem to illuminate further the boundary between semantics and pragmatics. Readers are referred for these to Capone (2010b).

Perhaps the most interesting discussion found in Capone (2010b) concerns the internal dimension of PRO, which is connected by Higginbotham to immunity to error through misidentification. Capone argues that, in connection with certain verbs, such as ‘remember’ the internal dimension of PRO is guaranteed by semantic effects up to a certain point, and that at least part of the internal dimension associated with PRO is due to pragmatic effects driven by typical scenarios. With some other verbs, such as ‘expect’, Capone argues that it is less likely that the internal dimension of PRO is a semantic inference and opts for the view that it is a pragmatic increment. Other verbs such as, e.g. ‘knows how’ are examined.

The most radical part of Capone’s ideas is that Higginbotham’s semantic elucidations for verbs such as ‘remember’, ‘imagine’ etc., refined and important though it is, suffers from a certain weakness, which cannot be remedied semantically, but only pragmatically. Leaving aside formal notation, Higginbotham’s treatment of (12)

12. John remembers walking in Oxford

comes out as ‘The agent of the remembering/John remembers that the agent of the remembering was walking in Oxford’. But then John should know that he is the agent of the remembering, a grammatical expertise which may not be acquired by anyone at all (See also Davis 2013 on this problem). My discussion suggests that Higginbotham’s logical form of this sentence could be paraphrased along the lines of: There is a REMEMBERING-event *e* such that (1) the agent of *e* (intuitively, the person who remembers) is John and (2) the content of *e* (intuitively, what is remembered) is that the agent of *e* participates in an event of walking in Oxford. The main possible argument against Higginbotham’s view is that Higginbotham’s logical form does not entail that John is aware that he is the agent of the REMEMBERING-event. Higginbotham’s logical form thus seems to leave open the possibility, for example, that John, believing that he is Mary, envisions in a kind of third-person way John walking in Oxford, and in actual fact ‘John remembers walking in Oxford’ seems to exclude this. This argument might have some force (despite the fact that Higginbotham in his paper dismisses this kind of possibility as “pathological”) but depends on how we spell out REMEMBER. The possibility I mentioned appears to be pathological because it involves a quasi-memory implanted in somebody else’s brain. However, a less pathological possibility is that someone (else) remembers the walking in Oxford (that specific event) in which he was also an agent of the walking (suppose it is also true that he happened to walk in Oxford at the same time when Higginbotham was). Then the only way to prevent this from happening is to set up REMEMBER so that it applies only to events that are remembered from the inside, not from the outside. Now, if this is done through semantics (say a stipulation), full generality is lost,

because I claimed in Capone (2010b) that for other verbs also allowing ‘de se’ interpretations the internal dimension of PRO is not to be specified semantically (due to exceptions). These are the cases of the verbs ‘expect’ or ‘imagine’. So, instead of opting for a semantic stipulation, I would go for full pragmatic intrusion and also claim that there is a pragmatic presupposition that identifies the agent of the thinking/remembering/etc. event and the agent of the embedded event.

Furthermore, this analysis presupposes that there is a unique thinker of this thought and is incompatible with the possibility that someone else, say God, is having the same thought. (This difficulty was raised by Neil Feit (p.c), and is to be taken seriously). The third kind of problem is that, despite the fact that Higginbotham says that these constructions are first-personal, there is nothing in their logical form that makes them first personal, unless one allows as normative the inference ‘I = the believer of this thought’ (and here endless discussions could arise on how obvious, normative or natural this inference is or should be).

My own view is that the first personal element EGO must somehow be incorporated into the propositional form, NOT at the level of semantics, but at the level of pragmatics. We need pragmatic intrusion—and here the theories due to Levinson (2000), Carston (2002), Sperber and Wilson (1986) come to our aid. Accepting that semantics can be underdetermined, we may incorporate certain elements through pragmatic intrusion. Considerations of parsimony may even lead us to think that pragmatic intrusion, in this case, is to be preferred to incorporation of the component EGO at the level of Higginbotham’s logical form.

Now, I want to dwell on the possible replies to Neil Feit’s objection to Higginbotham. First of all, I sum up Neil Feit’s opinion:

Another reason why I do not think Higginbotham’s account can handle ‘de se’ cases adequately is this. It seems possible that somebody could believe (correctly or mistakenly, it does not matter) that he is not the only thinker of a certain thought, for example he might believe that God is thinking it too. More generally, he might think that he is not the only thinker of any of these thoughts. But, even with this, it seems he could have a ‘de se’ belief. But on Higginbotham’s view—and other similar views—such a belief amounts to ‘the believer of this thought is F’. This cannot be what the belief amounts to however, since he does not think there is a unique believer, the believer of his thought. Moreover, if someone else (God perhaps) really is having the same thought, then all Higginbotham-style beliefs are false, but he could surely have some true ‘de se’ beliefs (personal communication in Capone (2010b)).

Now, of course, when I say ‘John believes he is not crazy’ I do not have in mind other believers of this thought other than John. And, if it is somehow in the background that God and I are the only believers of this thought, I do not think I thereby express or intend to express that John believes that he and God are not crazy. Nor does Higginbotham think so (I assume). The examples by Higginbotham, such as ‘John remembers walking in Oxford’, are less vulnerable to Feit’s objection. Higginbotham’s possible reply would be that, given the anaphoric

properties of PRO, it goes without saying that the unique believer of this thought (the agent of the remembering) is John and not God (Is not anaphoric coindexation enough to make this clear?). It is not even necessary to resort to the more complicated story that makes the subject of the walking plural: the believers of this thought, assuming a kind of metaphysics in which wherever one is, God is there too. (And I suppose that if talk of God is infused into Higginbotham's story, then certain metaphysical consequences would not be completely absurd).

Of course, the problem raised by Neil Feit becomes more cogent not in the cases of constructions dear to Higginbotham, but in cases of interpretative ambiguities such as (13):

13. John thinks he is happy.

Here pragmatics is abundantly involved, as even Higginbotham has to admit, and it goes without saying that if Feit's objection has some cogency, this goes up to some point, because if, by pragmatic intrusion, we create an anaphoric identity link between the thinker of this thought and John, the uniqueness condition is valid and thoughts about God's having the same thought are out of the question. (So either we assume that some pragmatic linking between 'John' and 'the believer of this thought' is presupposed, making Feit's considerations otiose, or one needs to insert the anaphoric link explicitly into the semantics).

Another line of response to Feit could be the following. Suppose that the corresponding analysis of 'John thinks he is not crazy' would wrongly imply that John believes that he is the sole individual to have that thought. Even so one could fail to see why this is the case—it would commit one to the existence of a certain thinking event with a unique agent, but certainly there could be other thinking events around with other agents that involve a similar content, and so the agent of that first thinking event could believe that there are.

It does not seem to me that this line of response can dissolve the problem. The problem is that for the sentence 'John thinks he is happy', we do not want it to be true if John thinks that God (the thinker of 'He is happy') is happy. We want to exclude this case, and to do so we need to add something on top of Higginbotham's elucidation, presumably the Ego component, which would transform a 'de re' thought into a 'de se' thought [also along the lines of Maier (2009), who admits that a 'de se' mode of presentation in the presupposition of acquaintance of an object is what distinguishes a 'de re' from a 'de se' thought].

A 'precis' is in order. Despite seeing some problems in Higginbotham's semantic elucidation, I accept it as the best possible semantic solution to work in tandem with pragmatic intrusion. If pragmatic intrusion is granted, Higginbotham's semantic elucidation works very well to provide minimal truth-conditions on top of which full pragmatic intrusion accrues.

3.3 *EGO or not EGO?*

While Castañeda (1966) in his seminal papers disseminated some really original ideas about ‘de se’ attitudes, and provided the basic examples alighting the theoretical discussion, he was clearly at a fork in having to decide whether ‘he*’ was completely irreducible (a clearly radical and original claim) or whether it could be partially reduced, say by making use of the concept EGO, to appear somehow in the semantic/pragmatic analysis of uses of the essential indexical. The other horn of the dilemma is certainly constituted by Perry’s ideas that beliefs ‘de se’ amount to specifications of mental states in which the concept EGO appears somehow (if it could not be shown to be semantically present, it could be shown to be indispensable for a pragmatic type of analysis). While the considerations by Perry are quite straightforward and presumably presuppose the at least partial reducibility of ‘de se’ to the EGO concept, Castañeda’s considerations about the irreducibility of ‘de se’ are fully articulated and explicitly deny that recourse to the concept EGO, even if invoked through pragmatic machinery, could be useful. Now, this is unfortunate because there is a clear link between IEM and the EGO component of a ‘de se’ thought.

3.4 *Immunity to Error Through Misidentification is the Result of Pragmatic Intrusion*

Immunity to error through misidentification is a property which ‘de se’ constructions inherit from the property of the ‘first-person’. However, if my claim that ‘de se’ constructions involve use of an implicit EGO component through pragmatic intrusion is correct, it cannot be true that immunity to error through misidentification is a semantic property of ‘de se’ constructions, although we can legitimately say that it is a pragmatic property of ‘de se’ constructions, being derivative from the EGO component incorporated into ‘de se’ constructions through pragmatic intrusion.

For the time being, we must decide whether we should derive immunity to error through identification from the internal dimension of PRO (or of a ‘de se’ construction) or whether we should derive the internal dimension of PRO (or ‘de se’) from immunity to error through misidentification. This is not a trivial question. We can make this question even more complicated by asking whether the internal dimension is derivable from the implicit use of EGO in ‘de se’ constructions. After all, we could have the following logical chain:

EGO > Internal dimension > immunity to error through misidentification.

If the logical chain above has some validity, and we can establish without doubt that EGO is a pragmatically enriched component of the ‘de se’ construction, then we ‘ipso facto’ show that the internal dimension of ‘de se’ and immunity to error

through misidentification are consequences of pragmatic intrusion and, in particular, the incorporation of EGO in 'de se' constructions.

Have we got independent support for such a line of thought? Recanati (2009) has insisted that not all 'de se' constructions involve immunity to error through misidentification and that in some cases proprioception is involved in guaranteeing immunity to error through misidentification. What is proprioception? While the discussion is undoubtedly complicated, Recanati distinguishes between feeling that something is the case and seeing that something is the case. For example, I can feel that my arm is broken or I can see that my arm is broken. In case I feel that my arm is broken, proprioception is involved and there can be no case for error due to misidentification (it is proprioception that guarantees immunity to error through misidentification). If I see that my arm is broken (say I look at my arm and see an arm that is broken, but I mistake your arm for my arm and I make an identification mistake, then immunity to error through misidentification is not guaranteed). While there is some truth in this discussion, I think it deserves deepening. However, unlike Recanati, instead of placing the burden on the distinction between perception and proprioception, I want to make immunity to error through misidentification depend (at least in basic cases like 'John thinks he is clever') on the awareness of the subject of the thinking experience. Of course, awareness of the subject of experience involves some kind of self-awareness and not proprioception proper or only perception, but the kind of immunity to error through misidentification in cases like 'John thinks he is clever' is different from the cases discussed by Recanati and does not concern objects of experience but subjects of experience. (See also the discussion by Galen Strawson). Thus proprioception may not be the right concept in this case, because it is not the case that the thinking subject is engaged in proprioception in thinking (with some appropriate exceptions, of course: This thought makes me nervous; this thought makes me sad; this thought made me tremble; this thought made me faint). Thinking is the essential relation necessary for establishing a thinking subject. It is the act of thinking that establishes the subject and the identity between the subject of thinking and the subject of the thought. While the person who thinks (14)

(14) I think I am clever

is not particularly engaged in an interpretation process but provides the appropriate EGO concept by the act of thinking and this is enough to ensure immunity to error through misidentification; something different occurs in (15)

(15) John thinks he is clever.

Here the hearer/reader must simulate (as noted by Igor Douven 2013) an act of thinking and in simulating this act of thinking she supplies an EGO concept through inference. Of course, pragmatic inference, utilizing the principle of relevance, independently supports the simulation process and establishes the anaphoric link between John and 'he' and also supplies the EGO concept which is incorporated into the thought by pragmatic enrichment. Having established that John

thinks of himself as Ego and that this is guaranteed by the act of thinking in itself, the hearer can simulate John's mental state and, in particular, the internal dimension of the thought (he thinks he is clever or happy because he experiences cleverness or happiness).

In the above, I have proposed that the first-personal dimension of 'de se', as pragmatically implicated, is logically responsible for immunity to error through misidentification (we could also see this case as a case of immunity of error through misidentification being supervenient (in the sense of Chalmers 1996) on the ego-component of 'de se').

If, as I proposed, the EGO component of a 'de se' thought, is due to pragmatic intrusion, immunity to error through misidentification is a consequence of a pragmatic attribution in **reports** of 'de se' thoughts. In naturally occurring 'de se' thoughts which are not reported, it is the act of thinking and the identity between consequential acts of thinking that guarantees the EGO component, and, consequently immunity to error through misidentification.

According to Wayne Davis (p.c.) my thesis is that the semantics together with the pragmatics of first-person pronoun explains why certain beliefs about ourselves are immune to error through misidentification. This seems incorrect for Davis, as being first-personal is not what makes a belief special; it is being the product of introspection that makes it special. "Moreover, it is hard to understand how any fact about the *words we use to express* beliefs can explain why those beliefs cannot be mistaken. How can facts about the English word 'I'—a mere letter or speech sound with a meaning and use—explain why a certain propositional attitude has a special epistemological status? Propositional sentences are not sentences or strings of letters or speech sounds." (Notice, however, that I am not claiming that IEM is ONLY a semantic phenomenon but that it is an epistemological/semantic/pragmatic phenomenon).

Ok. Suppose that introspection is, as Wayne Davis proposes, the essence of IEM. Is there a way to argue that introspection is not really possible (everyone should accept at least that it is facilitated by the use of the first-person pronominal) unless a predicate is assigned to a subject and that introspection specifically necessitates assigning a predicate to a first-person subject? The role played by semantics is evident. But this does not amount to saying that the first-person pronominal is essential in self-attribution of an introspective thought; surely the fact that the thinker and the attributor are the same makes it possible that no mistake is made in the attribution even if no first-person pronominal is used; however, in expressing that thought to others it is necessary to use the first-person pronominal.

A related question is whether all beliefs and thoughts have to be expressed in words. As Davis (p.c.) says, a belief or thought about ourselves is distinct from any words we use to express it. "The belief that I myself am in pain is very different from the sentence 'I am in pain'; the belief does not contain four words, is not a sentence of English, cannot be seen, heard, written or uttered, and so on. The sentence may occur in inner speech, but inner speech is not belief or thought. An I-thought does not contain the word 'I'."

Ok. Let us grant that at least sometimes one need not use language to form a thought, in particular a thought about oneself. Let us suppose that this thought is not very different from a sensation (pain). The fact that pain occurs in myself is sufficient to make the case that a mistaken identification cannot occur. Suppose then that it is the object (my body) which prevents me from making a mistaken identification. Yet I suppose that the body is at least part of myself and, thus, the thought, if it has to be expressed, requires the logical subject ‘I’ and only this, because a definite expression could not preclude a mistaken identification. So, the argument is not that a first-personal thought requires the word ‘I’, but that it requires it if it has to be expressed. The question of misidentification, in a thought which is sufficiently similar to the attribution of a sensation, cannot occur; but it can occur when the thought is expressed. This is all we need. So I may concede that Wayne Davis is partially right about belief—even if I should say that the kind of belief he has in mind is more similar to sensation than to the propositional attitudes we normally discuss.

3.5 Why Immunity to Error Through Misidentification is Logically Independent of the Internal Dimension of PRO/‘de se’

Admittedly, the reasons which I furnish in this section against making a logical connection between the internal dimension of PRO/‘de se’ and immunity to error through misidentification depend on some previous considerations of the inferential behavior of de se/PRO, presented in Capone (2010b). There I wanted to make the provision/expression of the internal dimension of PRO/‘de se’ a pragmatic constituent of the thought. However, after some discussion I moved towards the more balanced view that, in general, especially with verbs such as ‘remember’, the internal dimension of PRO is semantically associated with the specific construction (PRO, in our case) (See Ninan 2009 for an important paper on imagining from the inside, which amounts to imagining a scenario from the point of view of one of its participants, in a centred world). Then I have speculated that the internal dimension (constituent) supplied through the semantics is only partial or gappy (in line with views by Carston (2002) on semantic underdetermination) and that pragmatics is responsible in part for supplying a partial internal dimension. For certain other verbs, such as ‘expect’, ‘know how’, etc. I have speculated that the internal dimension constituent is fully provided through pragmatics.

Now, what are the consequences of the acceptance of the views above for the plausibility of the view that immunity to error through misidentification depends on the internal dimension of PRO/de se? The most immediate consequence would be that, in the most straightforward cases, like ‘**expect**’, or ‘**imagine**’ ‘de se’ constructions (he* or she*) should not be associated with immunity to error through misidentification, if IEM depends on the properties of the internal

dimension through semantics. Thus, someone who expects to leave for Rome tomorrow may legitimately have some doubts as to whether he himself is involved in the thought that he will leave for Rome tomorrow. But this is absurd. Immunity to error through misidentification must be granted for cases such as ‘expect’ and ‘imagine’ as well and this shows that immunity is not logically dependent on the internal dimension (which is implicated in these cases, if my view in Capone (2010b) is correct).

What other consequences follow from the fact that the internal dimension of PRO/‘de se’ is only partially semantically expressed and partially pragmatically articulated in cases such as ‘remember’? If we grant the logical dependence between immunity to error through misidentification and the internal dimension of PRO, we paradoxically arrive at the result that the greater the pragmatic enrichment in connection with the internal dimension of PRO/‘de se’, the greater the immunity to error through misidentification. However, nobody says or is willing to accept that immunity to error through misidentification is a gradable notion.

My initial idea that the internal dimension of PRO/‘de se’ may serve to support immunity to error through misidentification and not to establish it is not implausible; however, the immediate consequence of having such a view is that the internal dimension of PRO is useful in establishing immunity to error through misidentification only in those cases where there can be some doubt because a sentence is ambiguous. Consider, again an ambiguous sentence similar to one example by Recanati:

(16) He thought his legs were crossed.

Depending on whether he was only seeing his legs crossed or was also feeling them (proprioception being involved), (16) presents (or does not) a case of immunity to error through misidentification. The internal dimension of the pronominal ‘his’ is clearly projected through a pragmatic enrichment and, thus, proprioception is responsible for promoting immunity to error through misidentification through the intermediation of a first-personal pronoun. The pragmatically enriched internal dimension and proprioception go hand in hand and serve to reinforce immunity to error to misidentification in the sense of disambiguating a sentence which is interpretatively ambiguous.

3.6 Wayne Davis (2013), ‘de se’ and Indexicals

In this short section, I make connections between IEM and Davis’ (2013) view that ‘de se’ is linked with deixis. Davis thinks (in essentials) that ‘de se’ attitudes are to be explained by reference to deictic concepts. The thinking subject thinks of himself through a deictic. This is similar to what I have claimed myself, although Davis is more detailed. I was content with an ‘I’ concept, while Davis distinguishes between a deictic, a demonstrative and an anaphoric use of ‘I’. The deictic

use of 'I' is probably what is involved in 'de se' thoughts, deictic uses being licensed by what Davis calls 'presentations'. The thinker thinks of himself and has a presentation of himself that gives interpretation to his use (whether mental or verbal) of 'I'.

I depart from Davis in recognizing a dichotomy between the **thinker's** use of 'I' in thought, and the **hearer's** interpretation of 'I' or 'he*' in an ascription of thought. The thinker's use of 'I' in thought needs no special act of interpretation and involves immunity to error through misidentification in that no identity is needed or established, as there is no interpretation problem from the point of view of the thinker, who surely has a presentation of himself which is perhaps tacit and who keeps track of himself and his identity through the act of thinking, rather than through the act of interpretation. Instead, the ascription of 'de se' attitudes (to someone else) involves an interpretation problem and keeping track of the referent and mode of presentation used by the thinker either through a simulation or through a pragmatic act of interpretation guided through the Principle of Relevance or both. The two perspectives are different and surely the use of 'I' in ascription of 'de se' attitudes involves both an internal and an external dimension. The deictic use discussed by Davis may be suitable to both dimensions, provided that we are clear that a 'presentation' or 'self-presentation' is involved in the thinker's awareness of EGO, while a simulation or pragmatic interpretation is involved in understanding the presentation which the thinking subject experiences. Perhaps it would not even be incorrect to say that we can speak of a deictic use when referring to the hearer's interpretation problem, while from the point of view of the thinker there is no interpretation problem and thus it is not a matter of establishing the content of the deictic thanks to contextual coordinates. All that is required is the thinking act and the thinking act is its own context and also its own content.

Before closing this section, should we be content with Davis's exposition? While surely Davis' story resolves the problems he himself raised in connection with Higginbotham's theory (along the lines of the problems I myself discussed), he does so in an ambiguous way. Is the use of the deictic a semantic or a pragmatic component? I was clear that pragmatics was involved in establishing the EGO concept in 'de se' attitudes—even in cases of PRO, which are particularly problematic for Davis since PRO does not receive content from a context and thus is not easily assimilable to a first-personal concept). If we accept the considerations by Davis, we should have a double interpretation process. The provision of an EGO concept and, then, the interpretation relevant to a context of use (but this I admitted through lavish use of anaphora). From the point of view of the thinking subject, however, there is no pragmatics, since he has direct introspective access to his/her own thoughts. Pragmatics is involved only from a third person perspective, that of a hearer who tries to reconstruct the speaker's thoughts and self-awareness.

3.7 ‘De se’ and Modularity of Mind: Cancellability?

Finally it is time to examine the issue of the cancellability or non-cancellability of the ‘de se’ inferences I have discussed at length. Non-cancellability *per se*, as Grice was well aware, does not militate against the pragmatic nature of an inference. I have claimed elsewhere that explicatures are non-cancellable and the motivation I gave for this is that explicatures tend to be motivated by problems in the logical form, when a sentence is perceived to be blatantly false or a logical absurdity and pragmatics is there to help and remedy the problem. Since the explicature is the *Deus ex machina* of the semantics, I have claimed in a number of publications that it is and should be non-cancellable. This seems to fit in with a modular view of pragmatic processes, as argued in a number of publications. (See Capone 2010b, Capone 2011 for detailed arguments).

We saw that the incorporation of the Ego concept was the Deus-ex-Machina of the semantic treatment à la Higginbotham, protecting this treatment from all the objections raised by e.g. Davis (2013).

One further reason for opting for a pragmatic treatment of ‘de se’ is, of course, the parsimony of levels that it affords us, as we can eliminate at least an important meaning component from the semantics, obtaining it for free from pragmatics.

One last reason for opting for a pragmatic level of meaning in ‘de se’ attributions is the differential mechanisms of ‘de se’ thoughts in view of what happens in the mental processes of the thinker and of what happens in the mental processes of the hearer. The hearer is in a different position, both with respect to calculation of the Ego component and of the anaphoric links within the ‘de se’ ascription and with respect to the attribution of Immunity to Error through Misidentification. The disparity between the position of the thinker and the position of the speaker/hearer in connection with pragmatic inferences was noted in an article by Jeff Speaks (2006). The disparity between the thinker and the speaker/hearers’ stance to the inference is due to the fact that **luminosity** is available in thought, introspection being a guide to one’s intended meanings, while the meanings projected by the speaker and understood by the hearer in conversation do not rely on luminosity but on an explicit effort to get intentions across through contextual clues and cues.

While immunity to error through misidentification is presupposed for the thinker in virtue of the continuity afforded by the act of thinking (thus immunity seems to be an ‘**a priori**’ category of first-personal thought, something similar to a pragmatic presupposition along the lines of Maier 2009) and by the fact that in thinking the question of misidentification cannot arise; for the hearer, immunity is a logical consequence of the pragmatic inference involved in assigning an ego component to the ‘de se’ thought. Simulation and, also pragmatic interpretation flowing from the Principle of relevance are clearly involved.

The disparity between the speaker’s perception of himself as himself and the hearer’s ascription of ‘ego’ to the thinker has interesting consequences concerning cancellability. The speaker’s perception of himself as himself is clearly non-cancellable; the hearer’s ascription of EGO to the thinker of the ‘de se’ thought is

driven by contextual clues which lead the interpretation process in a certain direction, from which it is impossible to go back. So both from the speaker's and the hearer's perspectives it is impossible to cancel the EGO component of the thought.

Implicitly, I have replied to some qualms by Coliva (2003) about the idea that Immunity to Error through Misidentification depends on the ego concept incorporated in 'de se' attitudes. Her main objection to this idea, is that the use of 'I' in de se thoughts (whether explicit or implicit) is not enough to guarantee a first-personal thought. Coliva speaks of the split between speaker's reference and semantic reference (see also Jaszczolt 2013 for a similar argument). Given this split (which has emerged especially in the discussions of Donnellan's attributive/referential distinction), it may not be correct to say that immunity to error through misidentification depends on the presence of a pronominal like 'I' in logical form. The case discussed by Bezuidenhout (1997) (Bill Clinton: The Founding Fathers invested me with the power to appoint Supreme Court justices) does justice to the ideas and doubts exposed by Coliva. In the example by Bezuidenhout 'me' is used attributively, and not referentially. Of course Coliva does well to address the issue of the pragmatic nature of the incorporation of the EGO-component into 'de se' attitudes. However, we get the impression that her skepticism on the idea of deriving immunity to error through misidentification is not completely justified, given the heavy presence of pragmatic intrusion in propositional forms. Given the non-cancellable character of the pragmatic inference which I posited in 'de se' thoughts, the fact that 'I' can be interpreted attributively, rather than referentially is not a problem. Of course, my claim that immunity to error through misidentification follows from the Ego-like nature of 'de se' should be confined to cases where Ego is interpreted referentially. But this is, of course, presupposed by the 'de se' semantic/pragmatic analysis. Again, we should distinguish between the interpretation of the construction (e.g. I believe I am happy) on the part of the speaker, which heavily relies on Mentalese (the speaker has direct access to his own thoughts, and, thus, ego as used in 'de se' constructions is clearly and directly referential) and the interpretation on the part of the hearer. When we examine the dimension of the hearer, we see that the interpretative problem of 'de se' constructions consists in assigning, through pragmatics, an inferential increment that makes the logical form more plausible than it would otherwise be. The pragmatic enrichment, thus, could not make use of an uninterpreted EGO component, but has to make use of an interpreted EGO component, a component that is referential and not attributive (as in the example by Bezuidenhout). Of course, if we accepted a view in which the EGO component were assigned at the level of the semantics (say by identifying PRO with 'I' or an EGO-concept), then Coliva's objections could be certainly and dramatically applicable. But this is one more reason for opting for a semantico/pragmatic treatment, rather than for opting for a semantic treatment only. In a sense, we owe to Coliva the intuition that pragmatic intrusion resolves problems that would otherwise be insuperable.

Before bringing this article to conclusion, I want to point out that my considerations have something in common with Maier (2009). Maier essentially

distinguishes between ‘de re’ and ‘de dicto’ attitudes deriving ‘de re’ cases by accommodating presuppositions at top level in the DR. ‘De se’ cases are clearly cases in which the object one is acquainted with is a centred object. Only the presuppositions of discourse can establish whether the object one is acquainted with is a centred object. Now, in my discussion I have made an effort to disentangle cases where the presuppositions of discourse are responsible for IEM and the cases where the hearer needs to undertake an inferential process to simulate the mental processes of the person whose thoughts are being described in an assertion.

4 Conclusion

This paper has been loaded with theoretical considerations and their consequences. Presumably we have reached a stage in which, pragmatics, which originated in philosophy and was propagated outside philosophy giving impetus to communication-oriented linguistic views, can serve to throw light on philosophical topics. I cannot exaggerate the importance of seeing the phenomenon of immunity to error through misidentification as a consequence of pragmatic intrusion. It is true; I have reached a stage in which the theory has become loaded with various consequences of previously accepted views. However, it is the nature of interconnected considerations and interlocking ideas one finds in this paper, that makes it rich, by provoking novel and perhaps radical discussions of a phenomenon of which we knew little, before putting some thought to pragmatic intrusion.

References

- Bezuidenhout, Anne. 1997. Pragmatics, semantic underdetermination and the referential/attributive distinction. *Mind* 106(423): 375–407.
- Campbell, John. 1999. Immunity to error through misidentification and the meaning of a referring term. *Philosophical Topics* 66(1): 89–104.
- Capone, Alessandro. 2010a. What can modularity of mind tell us about the semantics/pragmatics debate? *Australian Journal of linguistics* 30/4: 497–522.
- Capone, Alessandro. 2010b. “Between Scylla and Charibdis”: The semantics and pragmatics of attitudes ‘de se’. *Journal of Intercultural Pragmatics* 7/3: 471–503.
- Capone, Alessandro. 2011. The attributive/referential distinction, pragmatics, modularity of mind and modularization. *Australian Journal of Linguistics* 31(2): 153–186.
- Carston, Robyn. 2002. *Thoughts and Utterances*. The pragmatics of explicit communication. Oxford: Blackwell.
- Castañeda, Hector-Neri. 1966. ‘He’: A study in the logic of self-consciousness. *Ratio* 8: 130–157.
- Chalmers, David J. 1996. *The conscious mind*. Oxford: OUP.
- Chen, Cheryl K. 2009. Bodily awareness and Immunity to error through misidentification. *European Journal of Philosophy* 19(1): 21–38.
- Coliva, Annalisa. 2003. The First Person: Error through Misidentification, the Split between Speaker’s and Semantic Reference and the Real Guarantee. *Journal of Philosophy* 100(8): 416–431.

- Christofidou, Andrea. 2000. Self-consciousness and the double immunity. *Philosophy* 75: 539–569.
- Davis, Wayne. 2013. Indexicals and ‘de se’ attitudes. In *Attitudes de se: linguistics, epistemology and metaphysics*, eds. N. Feit, and A. Capone. Stanford: CSLI.
- Douven, Igor. 2010. The pragmatics of belief. *Journal of Pragmatics* 42(2010): 35–47.
- Douven, Igor. 2013. The epistemology of ‘de se’ beliefs. In *Attitudes de se: linguistics, epistemology and metaphysics*, eds. N. Feit, and A. Capone. Stanford: CSLI.
- Evans, Gareth. 1982. *The varieties of reference*. ed. J. McDowell. Oxford: OUP.
- Feit, Neil. 2008. *Belief about the self*. Oxford: Oxford University Press.
- Garrett, Brian. 1995. Wittgenstein and the first person. *Australasian JP* 73(3): 347–355.
- Grice, H. Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.
- Higginbotham, J. 2003. Remembering, Imagining, and the First Person. In *Epistemology of Language*, ed. Alex Barber, 496–535. Oxford: Clarendon Press.
- Horn, Laurence. 1972. *On the semantic properties of the logical operators in English*. Mimeo: Indiana University linguistics club.
- Horn, Laurence. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Horn, Laurence. 2006. The border wars: A neo-Gricean perspective. In *Where semantics meets pragmatics*, eds. Ken Turner, and Klaus von Heusinger, 21–48. Oxford: Elsevier.
- Horn, Laurence, and Gregory Ward. 2004. *The handbook of pragmatics*. Oxford: Blackwell.
- Howell, R.J. 2007. Immunity to error and subjectivity. *Canadian Journal of Philosophy* 37(4): 581–604.
- Huang, Yan. 2013. De Se Attitude/Belief Ascription and Neo-Gricean Pragmatics: Logophoric Expressions in West African Languages and Long-Distance Reflexives in East and South Asian Languages. In *Attitudes de se: linguistics, epistemology and metaphysics*, eds. N. Feit, and A. Capone, Stanford: CSLI.
- Jaszczolt, K. 1999. *Discourse, beliefs, and intentions*. Oxford: Elsevier.
- Jaszczolt, Katarzyna. 2005. *Default semantics: Foundations of a compositional theory of acts of communication*. Oxford: Oxford University Press.
- Jaszczolt, Katarzyna. 2013. Contextualism and Minimalism on ‘De Se’ Belief Ascription. In *Attitudes de se: linguistics, epistemology and metaphysics*, eds. N. Feit, and A. Capone, Stanford: CSLI.
- Kaplan, David. 1977. Demonstratives. An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In *Themes from Kaplan* (1989), eds. Josef Almog, and John Perry, Howard Wettstein, 481–564. Oxford: Oxford University Press.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized implicature*. Cambridge: MIT Press.
- Lewis, D. 1979. Attitudes ‘De Dicto’ and ‘de se’. *The Philosophical Review* 88(4): 513–543.
- Maier, Emar. 2009. Presupposing acquaintance: A unified semantics for *de dicto*, *de re*, and *de se* belief reports. *Linguistics and Philosophy* 32: 429–474.
- Ninan, Dilip. 2009. Persistence and the first-person perspective. *Philosophical Review* 118(4): 425–464.
- O’ Brien, Lucy. 1995. The problem of self-identification. *Proceedings of the Aristotelian Society* 95: 235–251.
- Peacocke, Christopher. 1981. Demonstrative thought and psychological explanation. *Synthese* 49(2): 187–217.
- Peacocke, Christopher. 2012. Explaining ‘de se’ phenomena. In *Error through misidentification. New essays*, eds. S. Prosser, and F. Recanati. Cambridge: CUP.
- Perry, John. 1979. The problem of the essential indexical. *Noûs* 13(1): 3–21.
- Prior, James. 1999. Immunity to error through misidentification. *Philosophical Topics* 26(1–2): 271–304.
- Recanati, F. 2009. De re’ and ‘de se. *Dialectica* 63(3): 249–269.
- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance*, 2nd edn. Blackwell: Oxford.
- Shoemaker, Sidney. 1968. Self-reference and self-awareness. *The Journal of Philosophy* LXV(19): 555–567.

- Shoemaker, Sidney. 1996. *The first-person perspective and other essays*. Cambridge: CUP.
- Smith, J. 2006. Which immunity to error? *Philosophical Studies* 130: 273–283.
- Speaks, Jeff. 2006. Conversational implicature, thought and communication. *Mind & Language* 23(1): 107–122.
- Stalnaker, Robert. 1981. Indexical belief. *Synthese* 49(1): 129–151.
- Strawson, Galen. 2012. I and I: immunity to error through misidentification of the subject. In *Error through misidentification. New essays*, ed. S. Prosser & F. Recanati. Cambridge: CUP.
- Taschek, William. 1985. Referring to oneself. *The Canadian Journal of Philosophy* 15(4): 629–652.
- Der Sandt Robert, Van. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9: 333–377.
- Vignemont, de, Frédérique. 2012. Bodily immunity to error. In *Error through misidentification. New essays*, eds. S. Prosser, and F. Recanati. Cambridge: CUP.
- Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Oxford: Blackwell.

Further Reflections on Semantic Minimalism: Reply to Wedgwood

Alessandro Capone

Abstract In this paper I discuss a paper by Wedgwood in which he considers the possibility that Relevance Theory and Semantic Minimalism share at least some common resources. I maintain that the two theories have different aims and different orientations and that it might be fruitful to understand why Cappelen and Lepore stick to Semantic Minimalism despite the various objections levelled to their theory. I explore certain minimalist solutions along the lines of considerations by Michel Seymour, adopting Jaszczolt's considerations on parsimony of levels of interpretation. I assume that logical forms contain certain variables which can be filled (or saturated) in context. As a final proposal, I broach the idea that pragmatic enrichment at the level of the predication can be avoided by resorting to a more complex enrichment at the level of the subject. I resort to ideas by Jaszczolt (specifically POL), to argue that parsimony considerations require that enrichments be operated at the level of subjects, if possible, thus avoiding a less parsimonious view according to which both subject and predicates should be enriched.

1 Introduction¹

In this paper, I reply to an important and also intellectually stimulating paper by Wedgwood (2007). Wedgwood's paper is of importance because it throws light on certain structural similarities between theories which belong to different conceptual domains, philosophical semantics and cognitive pragmatics. It creates

¹ Abbreviations are explained in Appendix 1. I would like to give thanks to Wayne Davis, M. Wedgwood, Yan Huang, and Noel Burton-Roberts who provided comments for this paper.

A. Capone (✉)
University of Messina/Palermo, Palermo, Italy
e-mail: alessandro.caponis@gmail.com

undoubtedly important bridges between the two theoretical constructions and provides a very intelligent critique of Cappelen and Lepore's (2005) book (from now onwards Cappelen and Lepore is abbreviated to C & L). My difference with Wedgwood is one of point of view and of emphasis. Also, in some cases I point out what I take to be the weakest parts of his paper. I am sure he will find a way to refine those points and to answer them from his point of view, which is different from mine.

The three main threads of the paper are as follows:

I disagree with Wedgwood's conclusion that there is more common ground between C & L and RT than there initially appears to be;

I disagree with Wedgwood's contention that RT's semantics is more minimal than C & L.

I hold that C & L's tests for context-sensitivity are robust, and thus that 'ready' and certain other apparently context-sensitive expressions have an abstract, context-insensitive semantics, which needs to be coupled with null prepositional phrases, which, in turn, require saturation.

Semantic minimalism, as proposed by C & L (2005), has met vigorous opposition on the part of Radical Contextualists. Semantic minimalism is the view that the semantic content of a sentence *S* is the content that all utterances of *S* share in virtue of what is contributed by their words and the syntactic relations among constituents.² 'It is the content that **all** utterances of **S** express no matter how different their contexts of utterance are. It is also the content that can be grasped and reported by someone who is ignorant about the relevant characteristics of the context in which an utterance of *S* took place' (C & L 2005, 143). Semantic Minimalism claims that semantic content is pragmatically affected only with respect to standard indexicals, and that semantic content is propositional.³ In his review of Cappelen and Lepore (2005), Recanati (2006) writes:

Like most philosophers of language, C & L are minimalists, but they defend a fairly radical version of minimalism, close to the literalist ideal. Instead of multiplying indexical variables, as other minimalists do, they maintain that there is only a short list of context-sensitive expressions and that sentences not including any of them are such that their meaning is their truth-conditional content (Recanati 2006: 21).

Recanati finds it strange that the authors should say that sentences such as 'John is ready' express complete propositions, while refusing to say what these

² The addition of 'in virtue of what is contributed...' serves to overcome the objection that—unless such a modification is added—one could hold that all utterances of a sentence *S* shared content *p* but that *p* was pragmatically enriched in ways not traceable to standard indexicals.

Wayne Davis (personal communication) considers that this definition must be conjoined with another definition which C & L provide: Semantic Minimalism holds that "the semantic content of a sentence *S* is the proposition that all utterances of *S* express (when we adjust or keep stable the semantic values of the obvious context sensitive expressions)" (C & L 2005: 2).

³ One should note that the shared content assumption is not shared by all versions of minimalism [for instance, it is rejected in Borg (2004)].

propositions are (in detail), apart from saying that e.g. ‘John is ready’ expresses the proposition that John is ready.

I agree that C & L should provide more than their mysterious story. At least one should try and provide some elucidations of the meaning of words such as ‘ready’.

The main argument against contextualism is that, once we fall into the trap of contextualism, the kind of argument used to show that the literal meaning of an expression is not sufficient to provide a full proposition will also show that the truth-conditional meaning arrived at through contextual augmentations will not be enough, as more stringent truth-conditions are required. Presumably this regress should be extended *ad libitum*. This is also another mysterious point. Even if we grant that the contextualist’s claims generalize too much (as the authors say), this does not amount to saying that an infinite regress is generated.⁴

Despite various shortcomings, my view is that we should very carefully consider what C & L say and we should at least accept the various tests they propose which will enable us to detect context-sensitivity or, alternatively, context-invariance. This is not to say that the discussion of such tests should be uncritical and it is possible that such tests will be refined considerably in the years to come. In this paper, I try to assess the validity of some criticism levelled to C & L on the part of Wedgwood (2007).

2 Some Background

In this section I shall mainly outline the main arguments of the paper by Wedgwood. The main aim is to create the context for the discussion of very specific and technical points later on in subsequent sections. In this way, it will not be possible to say that I have extrapolated excerpts from the general discussion by focusing on parts of the discussion which are not particularly crucial to the general line of the discussion.

⁴ One could cast doubt on my idea that there is a limit to what contextualist claims can say about meaning. Consider, for example, one of C & L’s examples: ‘John went to the gym’. Did he go to the vicinity of the gym? Did he enter the gym? Or suppose we grant that the example is so heavily context dependent and that we resolve for this type of context-dependence, by saying that on a plausible interpretation the speaker meant that John entered the gym. Is not there a limit to what context-dependence could say about this example? Should we continue the context-dependent claims and say that under-determination is there because we are still able to articulate the purpose of getting into the gym? Ok. Let us suppose that we also provide a purpose constituent: ‘John went to the gym to keep fit’. Let us suppose that we can go on providing implicit constituents until our imagination allows us to do so; is not there a limit to the number of constituents we can provide by free enrichment? My intuition is that it is not possible to add implicit constituents *ad libitum*, for one thing; the intentions of the speaker are finite and, thus, going on to add implicit constituents will not realistically model finite speaker’s intentions.

Wedgwood sees RT and C & L's positions as being compatible. If this is right, Wedgwood takes this to be an important result as it represents the convergence of two major distinct approaches: a philosophical perspective on semantics; a cognitive perspective on pragmatics.

For C & L, context sharing is the very foundation of communication. RT has a similar concern with content-sharing, even if this is fundamentally the result of the combination of encoded meaning and inferential processes.

In general, Wedgwood argues that C & L's outlook is essentially very similar to that of RT, despite C & L's persuasion that they differ substantially from RT.

According to Wedgwood, there is a difference between the two approaches. This is constituted by C & L's notion of minimal content. Nevertheless, Wedgwood argues that this is largely equivalent in function to RT's encoded meaning. For C & L, minimal content is propositional, while encoded meaning is not necessarily propositional. The key property of minimal semantic content is that it should be conveyed consistently across all contexts. According to Wedgwood, RT's encoded meaning does the same job.

Wedgwood agrees that RT is an example of what C & L calls radical contextualism in so far as it accepts that the influence of context on propositional meaning is pervasive and ubiquitous. According to RT, the intrinsic content of a linguistic form is systematically sub-propositional (it must be complemented by pragmatic inferences to achieve full propositionality).

Wedgwood expatiates on the basic assumptions of RT, its character as a theory that studies 'mind-reading' processes in which a hearer must attempt to infer the speaker's communicative intentions on the basis of knowing that the speaker relies on knowledge of the language and on propositions that are mutually manifest to both the speaker and the hearer.

Wedgwood stresses that RT considers that what is understood to be communicated by any utterance is heavily context-dependent and that, despite this, people succeed in communicating with each other without any special problem. This is ensured because communication is governed by some set of principles enabling people to recognize the intended meanings.

Wedgwood considers that despite the emphasis on context, RT admits that some meaning is encoded in linguistic forms. The meaning encoded in a given linguistic form is sometimes called its 'logical form' or, as Wedgwood prefers, 'encoded meaning'.

According to Wedgwood, RT assumes that a speaker relies on assumptions which he takes to be mutually manifest in his expectation that the hearer will recover the intended meaning.

According to RT, interpretation proceeds according to a presumption of optimal relevance, where relevance is conceived of as a ratio of communicative rewards to processing efforts. These rewards are what RTs call 'cognitive effects' (such as strengthening or weakening/eliminating existing assumptions or providing new assumptions that interact with existing assumptions to allow new deductions).

Words and structures of a given language can be taken to convey a variety of things in different contexts, depending on the assumptions which are brought to

bear on the interpretation process. This means that linguistically encoded forms constrain but do not (fully) determine what is conveyed on a given occasion.

According to Wedgwood, RT accepts that there are two distinct notions of content: the inputs and outputs of the communication process, that is to say encoded and inferred meanings.

These two levels are sufficiently similar to C & L's position, which they call semantic minimalism which combines with speech act pluralism, the idea that the same sentence can be used to make a variety of (different) speech acts in different contexts. C & L too recognize two important levels of meaning: minimal semantic content and speech act content.

A point of divergence between C & L and RT is that the former take content sharing to be guaranteed by the conventional means of language. RT, instead, does not agree that communication involves the sharing of (identical) thoughts, but only similarity of thoughts. Wedgwood takes RT to be a more realistic and psychologically plausible theory as it reflects our daily experience quite accurately. In RT the goal of communication is to enlarge the interlocutor's mutual cognitive environment; that is, communication is successful if two people can tell that they have some more assumptions in common than they had before. This does not require the exact reproduction of thoughts.

3 Discussion of Wedgwood's Response to C & L: General Considerations on Semantic Minimalism

The paper gives the impression that Wedgwood is more favourable to RT than to C & L while he attempts to establish the basic tenets on which the semantics/pragmatics debate ought to rest. In the attempt to reconcile the two theories, Wedgwood claims that moderate contextualism is opposed to both Semantic Minimalism and RT. It is doubtful that by creating a common enemy one can reconcile two radically different theories. For one thing, moderate contextualism is opposed to Semantic Minimalism only in so far (as proven by C & L) one slides from there into Radical Contextualism. It might be legitimate to hold the view that C & L are against all forms of contextualism and, in particular, they go to some length to distance themselves from moderate contextualists. Even if MC did not always slide into RC, MC already goes beyond their belief that one should limit contextual influence to basic indexicals (or anyway, to elements of the Basic Set of context-sensitive expressions). However, given that C & L accept pragmatic intrusion at the level of pronominals and deictic elements, in principle they should accept moderate forms of contextualism provided that they are reducible to usage of pronominals.⁵ It is the sliding from MC to RC that crucially scares them. This

⁵ Someone might say that this is something that C & L explicitly reject. Yet, as my quotation of Seymour makes clear, it is reasonable to suppose that C & L are compelled to accept that when

assertion should be properly evaluated in the background of important considerations by Seymour (2010):

In more controversial cases like “Mary has had enough” and “John is ready,” it could be argued that their incomplete character may be explained by the implicit presence of a demonstrative expression like ‘this’, so that the sentences should read “Mary has had enough of *this*” and “John is ready for *this*.” Here I follow Capone’s minimalist explanation of the nature of such incomplete propositions (Capone 2008). If they could be interpreted as implicitly containing empty slots that can be interpreted as demonstratives or discourse-deictic anaphoric expressions, sentences expressing incomplete propositions would indeed be harmless for minimalism. The reason is that they would be analysed as implicitly involving expressions belonging to the basic set. “John is tall” and “Mary is rich” would also contain implicit semantic empty slots calling for completion by a particular reference class. These sentences should perhaps be analyzed as “John is tall (relatively to *this* reference class),” and “Mary is rich (relatively to *this* reference class).” Quantified statements like “All came for breakfast” would be an elliptical form for “All of *them* came for breakfast” (or “*They* all came for breakfast”), and would thus also be implicitly containing expressions belonging to the basic set of indexical expressions. (Seymour 2010: 2872).

Now, it appears to me that the cases above are reducible to semantic items having invariant semantics combinable with implicit (or null) elements that appear to belong to the Basic Set as formulated by C & L. We do not need to extend the basic set and the assumption that sometimes expressions of the basic set are null from a phonetic point of view is needed independently of the specific issue we are confronted with here. Nobody disputes that we have null elements such as PRO or pro in English or in Italian. But even in abstraction from grammatically stipulated null categories, we have free recourse to null categories as in ‘Look’ (look at that) or as in implicit arguments (The ship was sunk to collect the insurance). We may use a technical term for such null categories. One characteristic of these silent PPs is that they are elements present in semantic interpretation that are recoverable from context and without which semantic interpretation cannot arrive at a full proposition.

One might argue that my claim that we do not need to extend the Basic Set as formulated by C & L is confusing, as the implicit demonstratives that I am advocating to account for the context-sensitivity of ‘ready’, ‘enough’, comparative adjectives, etc. must be encoded in the meanings of these expressions (thus, I must be suggesting that the Basic Set be extended to include these expressions).

Now we are at a fork. Either we say that, e.g., ‘ready’ is a context-sensitive expression, and extend the Basic Set, or we say, as I propose, that ‘John is ready’

(Footnote 5 continued)

linguistic structure mandates a gap, such a gap must be filled by saturation. Of course, it is possible to have different views of the various examples discussed here or by Stanley. However, in principle C & L cannot deny that, when there is a gap, it must be saturated; in the same way they cannot deny that a deictic requires saturation; a pronominal requires saturation: pro and PRO require semantic values, either assigned by the grammar or by the linguistic co-text (Rizzi 1982).

has a meaning equivalent to 'John is ready for that'.⁶ The semantics needs to say that 'ready' subcategorises for a prepositional phrase (PP), which can be explicit as in 'John is ready for the exam' or left inexplicit as in 'John is ready'. Is this kind of uses rare in language? I would say not. I can reply 'I am coming' to my mother who says I should go to lunch. When a constituent is left out, it is reconstructed on the basis of context and of the grammatical patterns of the expressions which are actually present in the utterance (Thus I say 'I am coming (to the kitchen)' but not 'I am coming (under the kitchen)'). Implicit arguments are not that rare in the language, as there are examples such as 'The ship was sunk to collect the insurance' (See Roeper 1987).

In other words, we need no special resources other than implicit arguments, that is null PPs, which are recognised as a grammatical category anyway. So we do not need to list, as the commentator implies, 'ready', 'happy', 'enough' etc. as part of the Basic Set. All we need to ensure is that null PPs are part of the Basic Set.

It could be argued that my proposal, in fact, amounts to accepting indexical elements at LF, in which case it would amount to accepting a position analogous to that put forward by Stanley. But SM is not a thesis according to which some contextual effects on propositional content are due to indexical elements at LF, but rather a thesis according to which propositional content is fully determined by the semantics of the over material. Such contents are admittedly impoverished and skeletal, but they are the basics of a theory of meaning.

My reply to this objection is the following. Hidden indexicals as well as PRO and pro are present in the structure of the sentence. So there may be a minimalist story according to which some minimal propositions are calculable on their basis. There may be another story according to which the full proposition is fleshed out by assigning values to variables. This story which involves saturating variables is not in conflict with the minimalist account. Even assuming that minimalists can account for problematic adjectives like 'ready' in some other way, they would have to cope with the interpretation of pro and PRO. Such a story would involve variables at LF. What I think would characterize semantic minimalists as opposed to people like Stanley is the desire to keep truth-conditions minimal and to accept

⁶ Wayne Davis (personal communication) considers that we have two uses of 'John is ready'. In one use, 'John is ready' is elliptical for 'John is ready to go to dinner'. In this case, it is not plausible to think that 'John is ready' is elliptical for 'John is ready for that'. Davis thinks that the sentence is elliptical for 'John is ready for dinner'. He thinks we need to distinguish two views: (1) 'ready' is elliptical for phrases of the form 'ready for NP'; (2) 'ready' means 'ready for that'. On (2) 'ready' is indexical; on (1), it is not indexical and does not have a meaning although it is elliptical for phrases that do have a meaning. According to Davis, there would not be any independent meaning that 'ready' contributes to the meaning of 'ready for dinner' or 'ready for that'.

My reply is that considerations of parsimony militate against Davis's story. His story amounts to positing an ambiguity. At most we can grant an interpretative ambiguity. A common denominator could be 'X is ready for x' and allow that in some contexts x has the force of a demonstrative.

that things like ‘John is ready for that’ have minimal truth conditions even if no saturation has yet occurred.

4 RT and C & L on Shared Content

Wedgwood goes on to write:

I identify two key points of confusion in C & L’s arguments. The first, dealt with in section 3.1, is the failure to distinguish two kinds of assessment of the content of utterance: one from the addressee’s subjective point of view and the other from the point of view of an omniscient third party (such as a philosopher or a linguist). The former plays a key role in RT, while crucial parts of C & L’s arguments erroneously appeal to the latter. The second point of confusion concerns different kinds of content. As I discuss in section 4, the phenomena that C & L invoke as evidence against RT involve a kind of content that is not guaranteed to be shared under any theoretical approach. Indeed, their own position on this kind of content—in so far as it is made explicit anywhere—is shown to be equivalent to that of RT in all crucial respects. As such, it is just as well for C & L that their ‘shared content’ arguments against RT fail, as any arguments of this kind that succeeded in undermining RT would also undermine their own position (Wedgwood 2007: 649).

The passage above is fairly problematic. It presents two points. Let us start with the first point.

Let us assume that this position can be summed up as follows:

The difference between RT and SM is that the former claims that propositional content ensues from inferential processes that take place in the hearer’s mind, whereas the latter maintains that such content ensues from meanings conventionally encoded in words and, hence, propositional content persists independently of particular instances of interpretation.

Now, it should be born in mind that an explicature is what is taken to express the propositional content and that, according to Carston (2002), at least, an explicature has two components: a semantically encoded component and a pragmatically derived component. Now, if things stand as Carston says, it is simply not true that propositional content ensues from inferential processes that take place in the hearer’s mind. At least part of this content is provided by the semantics. The difference between the two positions, as I see it, is that Minimal semantics stresses the conventional layer of meaning, while Relevance theorists stress the non-conventional layer of meaning. It should be conceded that at least in some cases propositional meaning is exhausted by the conventional meaning of a sentence; in some other cases, propositional meaning is not exhausted by conventional meaning.

There is one further complication in this picture. Carston is ready to accept that explicatures are developments of conventional meaning, even if she insists that usually conventional meaning under-determines semantic interpretation (and pragmatics must intervene to provide full propositional forms). However, there are cases in which one must decide whether the contribution of the semantics must be

discarded or modulated—ironic utterances are cases in which literal meanings do not reach the level of propositionality. Cases of modulation à la Recanati (2010) also show that pragmatic interpretation may reach the level of compositionality.

So, it is not necessary that the conventional meanings of words will make up the linguistic component of the explicature. Both the conventional and the pragmatic component of the explicature are subject to pre-semantic interpretation. However, it could be said that the conventional meaning of a sentence is where the pragmatic interpretation originated and that the human mind must be able to have access to the meanings from which the pragmatic interpretation originated. It is also plausible that the human mind, in constructing the propositional form, does not throw away those bits of language which served to arrive at the propositional form. To provide an example, a hearer who encounters the utterance ‘The lion stood in the middle of the square’ will proceed from its literal meaning to its modulated meaning ‘The statue of the lion stood in the middle of the square’. However, the sentence ‘The lion stood in the middle of the square’ does not completely disappear from the hearer’s mind, since the effect of surprise crucially hinges on that linguistic construct. Poetic interpretation of metaphors, likewise, does not throw away bits of linguistic structure, as these contribute to poetic effects. Furthermore, a person who uses a sentence in its modulated sense and a hearer who understands it should be able to keep in mind not only the final product of interpretation, but the process as well and the linguistic sentence from which it originated. If this is the case, then one can argue that it is methodologically sound to keep the semantic and the pragmatic component of the explicature separate, while surely one must be able to merge them (at some stage).

Wedgwood, in his paper, characterizes RT as a theory of the hearer’s interpretation. It is not clear to me whether this is a point that is crucial to Wedgwood’s paper or not. It is possible that not much in the main argumentative line of that paper changed if this point was dropped. It is worthwhile discussing it nevertheless. Wedgwood’s insistence on the hearer’s perspective does more harm than good to RT. If I were challenged to provide arguments for this persuasion, I should quote from Capone (2010a) on indirect reports. In this paper I argue that the process of interpreting indirect reports usually (albeit not always, as interpretation process must adapt to special contexts) prevents the speaker (the reporter) from using an NP (coextensive with the NP actually used in the original utterance) that would transform (and distort) the original speech act, as such a move would place excessive processing efforts on the hearer. Now, my understanding of indirect reports is that pragmatic principles constrain both the speaker’s behaviour and the hearer’s behaviour. Emphasis on the hearer would surely provide only half of the explanation required, perhaps the half which we need least, as after all it is legitimate to try and explain the speaker’s behaviour as well, and such a behaviour clearly incorporates expectations about the hearer’s behaviour.

Wilson and Sperber (2004) write:

The central claim of RT is that the expectations of relevance raised by an utterance are precise and predictable enough to guide the hearer toward the speaker's meaning (Wilson and Sperber 2004: 607).

Since an utterance creates expectations of relevance, it is licit to deduce that the speaker's linguistic behaviour is what creates such expectations. Thus, both the speaker and the hearer must be taken into account in an explanation of how relevance works. After all, when the authors say that human cognition tends to be geared to the maximization of relevance, 'maximization' is best understood as maximization by both the speaker and the hearer. Someone might take issue with this, by saying that after all, when I say that human cognition is geared to the maximization of relevance, 'maximization' is best understood as maximization by both the speaker and the hearer, this is confused. Cognition does not involve a speaker and a hearer (only communication does). Despite my difference of point of view, I find this suggestion useful. Since I am obviously talking of cognition in the context of communication (I admit it is not my intention to write a paper about the hearer's cognitive powers in isolation from the speaker's cognitive powers), it is probably useful to distinguish the speaker's cognitive processes from the hearer's cognitive processes. However, I claim that cognitive processes follow a social path of intentionality, which means that the speaker in uttering a sentence is busy predicting what is going on to happen in the hearer's mind, while the hearer is busy understanding what the speaker meant in virtue of predicting the direction of the hearer's cognitive processes. It is correct to say that communication can be guaranteed by the fact that both the speaker and the hearer share the same psychological make-up (Recanati 2010). If Recanati's intuition must be made the most of, it is useful to see that cognition involves a mirroring relation between the speaker and the hearer.

If further evidence is needed, consider the following case, from Blommaert (2005). Blommaert considers a misunderstanding between himself and a research associate. They were at a conference together and they had rooms in the same hotel. They were speaking Dutch. Blommaert wanted to say that his associate had a nice balcony by a Dutch utterance equivalent in meaning to 'You have a balcony too'. The word for 'balcony' in Dutch is ambiguous between 'balcony' and 'breasts'. It happened that when he used this utterance, there was a woman walking in the opposite direction who was wearing a deeply cut summer top exposing parts of her breasts. Blommaert writes: Unfortunately, my young female research associate had noticed the woman—she had picked up a contextualization clue—and the term 'balcony' suddenly acquired a very suggestive, sexually offensive, meaning, which called for extensive explanation and damage repair afterwards. My words had been placed in (or made to point to) a context which had altered their meaning, triggering a shift from a descriptive, neutral meaning for 'balcony' to an implicit, male sexist slang meaning (Blommaert 2005, 42). This is clearly a situation where the speaker's meanings (or interpretations) and the hearer's meanings (or interpretations) diverge. This is a situation that perfectly models Wedgwood's considerations on inferential work (on interpretation of utterances).

Yet, I would like to claim that RT, like any other pragmatic theory, must not typically model situations in which there is a discrepancy between the hearer's and the speaker's interpretation. There are ways of expanding the discussion of Blommaert's example that point to the fact that the addressee somehow goes wrong in interpreting intentions. In fact, she follows an individual path of interpretation preferring it to the social path of interpretation (see Jaszczolt 1999).⁷ Since Blommaert never pointed or looked at the woman with the deeply cut top, the hearer made reference to a contextual clue which was only (unilaterally) presumed. The context which she brought to bear on the utterance was her unilateral version of the context, as she had no evidence that the context was shared. She also followed an individual path of interpretation by choosing the pejorative meaning, when she had a choice between a meaning which was in line with the intentions compatible with the public role assumed by the speaker (a professor speaking in the context of a conference) and a meaning which could implausibly relate to sexual matters. Presumably most Dutch people who use the word for 'balcony' are in Blommaert's position; however, ordinary usage is never as ambiguous as Blommaert really wants to let us accept. In context, ambiguities are easily settled. One should always go for the social path of interpretation (as Jaszczolt 1999 says), as this is the only path of interpretation that can ensure that the speaker's and the hearer's meanings coincide. Since following the social path of interpretation involves the speaker's duty to project meaning on the basis of conventions of language and of use and the hearer's duty to interpret intentions on the basis of the same conventions of language and of use, convergence is guaranteed if both the speaker and the hearer abide by their duties. Presumably, there are also norms of interpretation besides the Gricean maxims, like the ones I tentatively put forward in Capone (2004). The most basic ones of which are the following:

Show Goodwill

Contextualise an action in such a way that it can be interpreted positively; if you do not find a context in which it can be interpreted positively, then at least allow for the possibility of finding a context in which the action can be interpreted positively.

Be Constructive

Repair your coparticipant's mistakes by attributing positive interpretations to her actions; in particular, adjust any interpretations of her actions by taking into consideration the intentions that can be plausibly attributed to her.

⁷ Intentionality is a notion that makes sense of the content of speaker's mind during the process of uttering a certain sentence as correlated to the utterance. Intentionality that follows the social path of interpretation is undispersed intentionality, according to Jaszczolt (2005). From Jaszczolt's picture, we understand that the hearer is allowed to reconstruct the speaker's intentions on the basis of conventions correlating sentences and meanings and conventions of use correlating utterances in context with interpretations. Speakers' intentions and hearers' reconstructed intentions should converge if both speaker and hearers follow the social path of intentionality.

Now, there is nothing ‘ad hoc’ about these norms for interpretations. The other day I happened to make a blatant mistake and I said ‘Hello’ to a woman in the street that looked like an acquaintance of mine. When I realized I had greeted the wrong person, I apologised. What followed was a mini interaction in which the woman intended to play down the event—categorizing it as a thing that can happen to everyone. If one wonders why such lengthy elaborations follow these incidents, the reason is clear. People want to be seen as showing goodwill—they will let errors pass and they will discard pejorative interpretations when it is possible. In the light of this, what is striking about Blommaert’s example is that it may serve to model a theory of language which is utterly different from RT—in fact, RT too needs a mechanism to ensure that in the standard case the speaker’s and the hearer’s interpretations coincide (or at least do not differ vastly). Wedgwood’s picture of RT makes it appear suitable to modeling situations like Blommaert’s, while, in fact, it is suitable to giving us a model of Blommaert’s situation that tells us why the hearer went wrong, what was wrong in the interpretation process and why the speaker was perfectly entitled to receive an interpretation in line with his original intentions.

A possible objection to my approach based on good will is that a principle such as ‘Show good will’ is surely liable to change with context. In RT terms, the attitude of one’s interlocutors is a contingent factor that may be mutually manifest or may be made so as a result of their behavior. Such attitude cannot therefore be constitutive of any general conversational principle.

My reply would be the following. Although I have not said anything about whether the principle ‘Show good will’ could be reduced to the principle of relevance (arguably it could be so reduced, given that such a principle would tend to maximize positive effects in the face of the same cognitive efforts; something to the same effect could be argued in connection with the maxim of Quality which could be subsumed, arguably, under the maxim of Good will), there are reasons to believe that ‘Good will’ is rooted in our culture and has roots in social practices. Just to give you an example, consider the example I discussed in Capone (2005), describing the evil teacher who asks a question expecting/desiring that the pupil will not answer it. In Capone (2005), I argued that this a bad teacher, one who does not conform to the standards of benevolent teachers, who are the norm in society. He does not conform to the standard because he does not conform to the maxim of Good Will.

The second point by Wedgwood is more difficult. Wedgwood says that RT, despite what C & L claim, has an approach similar to C & L on content-sharing. Wedgwood accepts that the level of linguistic meaning due to the conventional meanings of words and grammar is what is common to both semantic minimalism and RT. Yet, the assumption that linguistic meaning is largely underdetermined makes the two theories sufficiently differentiated.

5 RT and C & L on the Equation of Semantic Content with Intuitions About Speech Act Content

Wedgwood (2007) discusses the (so called) Mistaken Assumption (by C & L):

[The mistaken Assumption] A theory of semantic content is adequate just in case it accounts for all or most of the intuitions speakers have about speech act content, i.e. intuitions about what speakers say, assert, claim and state by uttering sentences (C & L 2005, 53).

What C & L apparently fail to recognize is that RT quite explicitly rejects this too (see further Sect. 5.1). It may be that relevance theorists tend to express the idea in rather different terms—for example, ‘linguistically encoded semantics falls short of truth-conditionality’—but since truth conditions are conventionally identified with what the speaker is taken to be committed to (i.e. to have asserted, claimed or stated) by uttering some sentence, this is in fact just another way of expressing the very same point (Wedgwood 2007: 654).

There are interpretative complications. What does ‘this’ refer to? Presumably it refers to The mistaken Assumption. ‘The same point’ also requires disambiguation: does the expression refer to ‘the same point as The mistaken Assumption’? Or does it refer to ‘RT quite explicitly rejects this too’? Even after such a disambiguation, there is very little sense in what is said. In any case, let us see if it is true that RT too rejects the equation of semantic content with speech act content. Carston (1991) says that a sentence such as:

- (1) If the king of France dies and France becomes a republic, I will be happy, but if France becomes a republic and the king of France dies I will not be happy

has contradictory truth-conditional import and this defect is remedied by resorting to inferential pragmatics. In other words, truth-conditional import is calculated by trying to find out what kind of assertion the speaker made, what his intentions were in making this assertion. It is implicit in her view that the truth-conditional content of (the whole of) sentence (1) is not merely based on the truth-conditional content of each of the conjuncts in each conditional, but there is a truth-conditional content at the level of the speech act that is (globally) pragmatically conveyed. Now, while surely both Carston and C & L agree that the semantics of ‘and’ is basically truth-conditional, it is not clear that for Carston the truth-conditional content of (the whole) of utterance (1) and the truth-conditional import based on sentential semantics do not come apart. Consider what Carston (1991) says about (1).

As Cohen (1971) has pointed out, if *and* is simply truth-functional and the temporal and causal connotations are captured by implicatures, then (22a) and (22b) should be contradictory at the level of explicit content (instantiating ‘If P then Q but if P then not Q’, and ‘Not P; P’ respectively) (...). However, these examples are not understood as contradictory or redundant. Those who wish to maintain an alleged implicature analysis have to say that the alleged temporal and causal implicatures contribute to the truth conditions of the utterance in which they occur (...).

So a clear difference emerges between, e.g. C & L and Carston. C & L would have to hold the view that there is a contradiction in (1) (which is then resolved at

the level of the speech act). Carston, instead, holds the radical view that the pragmatics of conjunction is integrated into truth-conditional semantics. (I discuss this problem in Capone 2006).

I understand that Wedgwood would say, at this point, that Carston, like C & L, differentiates between sentential meaning and the level of the speech act. For Carston, the sentence would have to be contradictory at the level of the sentence, but not at the level of the speech act. But the question now arises, how is this possible? Do intentions have the magical powers to make a sentence that is contradictory non-contradictory? How can this be? It can be argued that the sentence is not contradictory in the first place. At this point, it is clear that Carston's account and NOT C & L's account is confused. C & L can argue that the sentence is not contradictory because it has temporal variables that can be filled and they are filled only at the level of the speech act. Carston, instead, projects the categories of the speech act to the level of the sentence. The problem does indeed arise from the habit of confusing speech act categories with sentential categories.

Now there is the further question of how we should use 'semantic content'. If we use semantic content in the way Wedgwood presumably does, this comes to (boils down to) semantic import. If we use it the way C & L do, this comes to propositional import. What Wedgwood commits himself to is a rhetorical fallacy (called 'equivocation'; see Stati 2002, 122): he attempts to defeat C & L by shifting the semantic import of their words. For RT 'semantic content' is linguistically encoded content, which may fall short of a proposition (a proposition being recovered only once rich pragmatic processing takes place). This distinguishes it from semantic minimalism which claims that linguistically encoded content will be propositional. Thus it is clear that 'semantic import' is understood differently by RT and by C & L, the former being committed by its use only to the underdetermined logical forms of semantics, while C & L allow a modest dose of pragmatic intrusion.

Of course, on both views, it should be granted that semantics should not be equated with the level of speech act content, since this invariably involves recourse to pragmatics. But the question whether a speech act can be associated, at least in some cases, with a minimal propositional content is undoubtedly important. The fact that we can communicate some messages in virtue of language in a non-miraculous way is of some concern for the theorist of language. The very notion of conventionality relies on the idea that we can express messages by using language. There should be some tough correlations between sentences and utterances—otherwise it would be absurd to create a sophisticated language system. Despite contextualists' attempts to place the onus of signification on contextual assumptions and pragmatic strategies, there is a reason why we do not have a magic word which in different contexts can express all the possible speech acts of the language. The fact that a language consists of (usually at least) fifty thousand words serves to illustrate that we pay a price for the possibility of expressing a potentially high quantity of meanings and that these meanings are structured on the basis of words—conventional tools at the basis of the semiotic process.

Words and syntax offered us an obvious way to escape the limitations of gestures, acts of pointing, pragmatic resources of all types, and to do so they had to be based on the notion that language has at least some independence from contextual and pragmatic effects. This does not mean that language cannot be modulated in context, but that in principle it must be possible to interpret sentences on the basis of what they conventionally say in virtue of conventional meanings. This independence is what guarantees that human beings follow a social path of interpretation when they proffer utterances. This independence is, furthermore, a model on which rules of language use are based. Pragmatics is not chaotic but essentially follows some rules or principles of language use and these are available to all users of the language and must be followed by all. It is only through the independence of language from pragmatics that language users could break out of the otherwise necessary lack of communication. The relative independence of semantics from pragmatics, furthermore, is what guarantees that pragmatics can have orderly effects on top of semantics, what guarantees that inputs and outputs are correlated through something which cannot alter completely the nature of the input. If it was possible to deform and freely alter the input, then nothing would prevent pragmatic effects from resulting into a chaotic magma in which the individual path of intentionality could prevail over the social path of intentionality, thus rendering communication impossible.⁸

6 On the Equivalence of Minimal Semantics and Encoded Meaning in RT

Wedgwood (2007) criticises C & L saying that, despite their attacks on RT, C & L's notion of semantic content is entirely equivalent to the notion of encoded meaning in RT. For the reader's convenience, I report an extract:

When C & L speak of truth conditions, they do not refer to the common (moderate contextualist) conception of 'what makes the proposition intuitively expressed by a given utterance true', but rather something that is indeed considerably more minimal. (1) is an example:

- (1) 'Steel isn't strong enough' expresses the proposition that steel isn't strong enough (C & L 2005: 61).

Though this kind of statements appears less than dazzlingly enlightening, C & L insist that this is the level at which truth conditions should be stated, arguing that it is the business of metaphysics, not semantics, to determine conditions on truth more precisely.

⁸ If am asked 'What did you mean by saying S in context C?', the set of my answers should be heavily restricted by the rules of language and the rules of use (the rules valid for the language game I am playing). I cannot reply by saying what might be convenient to me at that point of the interaction. The latitude in possible answers is severely restricted by the severe judgments of those who could reconstruct the meaning on the basis of what I said and the context in which I said that.

The point for current purposes is that (1) shows how C & L's strict semantic argumentation—in common with their essentially communication-based 'shared content' arguments—is committed to the idea that the semantic content of a given sentence is just whatever remains common to its meaning across all contexts of utterance. There is no other way to interpret statements like (1).

By this definition, C & L's minimal semantic content is entirely equivalent to the notion of encoded meaning in RT. This is necessarily the case, since encoded meaning is by definition just that which is conveyed by a given linguistic form, irrespective of context. (Wedgwood 2007: 664).

According to Wedgwood, to say that some meaning is encoded is entirely neutral with respect to questions such as whether it has other properties, such as being propositional. All logical considerations militate against the equation of sentential levels of meaning with propositional meanings.

If we take what Wedgwood says seriously, we should concede that

(2) The sentence 'She is happy' \leftrightarrow It is true that Mary Thornton is happy.⁹

The sentence 'She is happy' is equivalent to the proposition that Mary Thornton is happy. Presumably we have to assign the words 'entirely equivalent', as used by Wedgwood, a meaning which is radically different from what the words mean, say a speaker's meaning. Would this still be correct? I think it is not. There is no question that a sentence and a proposition can be equivalent. A sentence can serve to express a proposition in a given context, but, as pointed out by Strawson and by Frege (see Capone 2006), a sentence can be neither true nor false in itself, while a proposition is something that can be true or false. Sentences can be said to be equivalent in meaning only if they contain no indexical parts: alternatively we can say that two sentences are equivalent in meaning if, in case the indexical parts are saturated in exactly the same way by providing the same contextual value, uttering of one or the other sentence constitutes the same statement. The operation done by Wedgwood is very curious: the author subtracts important features from the programs of C & L and of Sperber and Wilson, and then says that by the phrase 'semantic content' they refer to more or less the same thing. The element subtracted from C & L's program is partial, local intrusion, at the level of deictic terms; the one subtracted from Sperber and Wilson is radical intrusion, at the level of enrichment of predicate meaning. After these elements are subtracted, he shows that the programs are similar. What is not captured is the spirit of Sperber and Wilson's program, which is to argue in favour of modulation, denying that verbs like, for example, 'open' can be assigned uniform truth-conditions. There is no denying that Sperber and Wilson's program is more radical.

One could argue that I misrepresented Wedgwood's point, which is that minimal content and encoded meaning are both assigned the role of being the content that is common across contexts, and that this means that SM will need to rely on

⁹ It was pointed out to me that there is no way to make this equivalence statement syntactically correct. But this is exactly the point of the example, there can be no equivalence statement of this sort.

similar processes as RT does when it comes to explaining how intuitive contents are arrived at in particular contexts. Wedgwood is not arguing that minimal content and encoded meanings are logically equivalent—but rather that the roles they are assigned by the two theories are structurally similar.

Now, acceptance of the above considerations, whether or not they represent Wedgwood's stance, would surely be easier than accepting the text a fragment of which I quoted. Of course, one would wonder what role structural similarity would play in stressing the compatibility of the two theories being compared, if, as one should concede, full propositionality of the minimal level of meaning is an obstacle in the way of reconciling the two theories.

The types of arguments I presented above might be replied to by saying that it is true that RT sees the predicate of an utterance as generally underdetermining the property it is used to express, but RT denies that the 'completion' of the predicate contributes to semantic content. To justify my position I should argue why the minimal proposition should be considered the semantic content.

If semantic content is used as expressing the idea of the skeletal semantic representations of linguistic meaning, then I agree that Relevance Theorists do not claim that completion processes contribute to semantics. The challenge to justify the position that the minimal proposition should be the semantic content, is a challenge for C & L, as I am simply imagining how they might defend themselves.

We should start with making connections between semantics and the idea that semantics deals with communication and transmission of factual information. On this idea, two statements have the same content, not if they have roughly the same semantic meaning, but if they report the same situations, if they connect with the world in such a way that both statements are about the same situation. While surely we can use two different sentences to express the same situation, and we could also use the same sentence to express different situations, one point is clear: if two statements consist of the same sentences and they are proffered in the same context (say in replicas of the same context) and with the same intentions, the two sentences will be taken to make a report on a certain state of the world. If reporting on the world is the basic point of communication, we should be able to choose between a semantics which more directly connects with the transmission of information about the world and a semantics in which this connection is less direct. Suppose we choose an indirect connection between semantics and the purpose of transmitting factual knowledge. This means that semantics mainly deals with schematic information, which produces factual information only when coupled with contextual information. Is such a view plausible? And which view is more basic? The one in which semantics has a direct connection with the purpose of transmitting information, or the one in which semantics is only indirectly connected with the purpose of transmitting factual information? Choosing the indirect connection is a way of committing oneself to a view in which semantics is directly connected with some purpose, other than communicating factual information. But what purpose could this be? The alternative that language serves merely or primarily to construct relationships does not surmount these difficulties, as

constructing relationships somehow relates to the expression of information about the world (about the way we feel about others).

And why should a looser relationship between semantics and communication of factual information be privileged? It is as if the theorist were to accept that language serves the purpose of providing information, but not information about the world, and that information about the world must be reconstructed on the basis of schematic, fragmentary information. But if this is the case, we must accept that this is the case even for prototypical utterances. A request for water as expressed by ‘I am thirsty’ provides or should provide bits of information, from which I reconstruct the puzzle of an intentional entity, the proposition that the speaker is thirsty.

The idea of a possible indirect link between language and communication makes it difficult to understand, for example, how the first intentional linguistic acts in the history of mankind were interpreted. If one starts with the idea of fragments of information being conveyed in semantic acts, one never really reconstructs the point of the utterance, as the utterance could be about ideas in general, rather than about facts. And yet, the intentional dimension of communication, as pointed out by Jaszczolt (1999) is a basis for all communication, even a default principle of communication. The very primitive acts of communication were interpreted as being about the world, and not about abstract ideas (and this made it possible for the hearers to utilize contextual clues for interpretation), but this could be possible only if there was perhaps an a priori principle, working in language (that is to say operative in the minds of language users) like the one produced by Jaszczolt, in favour of the intentionality of linguistic acts. It is true that in another book Jaszczolt (2009) opts for merger representations in which linguistic and non-linguistic representations merge, to form a propositional dimension, but this is not to say that prototypical, or primitive linguistic acts were devoid of intentionality or that in no cases the semantic representation corresponds to a proposition. I think it is clear that Jaszczolt is more interested in the complicated cases, but this is not like denying that there are simple cases.

7 A Problem for Semantic Minimalism: Indexicals

Wedgwood (2007) writes:

C & L’s self-imposed commitment to the propositionality of minimal semantic content forces them to make this content less minimal than it could be (...). It also introduces a difference between their conception of minimal semantic content and RT’s encoded meaning, as Carston (2007) stresses. In order to stick to what they recognize to be propositional meanings, C & L have to assume that minimal semantic content contains no unsaturated indexicals or ambiguities. The problem is that this immediately makes minimal semantic content something other than that which is shared across different contexts: the saturation of indexicals is a context-dependent process. (Wedgwood 2007: 666).

Let us see how this works in practice. Consider the following utterance:

(3) She is happy.

[(3) is uttered by Mary who points to Ann when she says ‘she’].

According to C & L we could report Mary’s utterance by saying (4):

(4) Mary said that she is happy.

Such reports can be uttered on the basis of the identity of the referent of the pronominal. Suppose this is a rule constraining indirect reports. The main question is whether there is identity of content between (3) and the reported part of (4), assuming identity of the referents of ‘she’ in the two utterances. If there is such an identity (and we suppose the answer is positive), then C & L’s theory is a theory of abstract semantic content. The fact that semantic content is propositional is no impediment to having a view of content that is context-invariant. Propositionality simply means that we keep fixed certain elements which are tied to context while we use the tests for context-invariance. If, by keeping fixed the contextual contribution of pronominals, we obtain an invariant interpretation of the sentence, the results of this analysis will contribute to semantics and not to pragmatics. The fact that the semantic content of the sentence is propositional (in that it incorporates the semantic interpretation of indexicals and pronouns) is hardly a reason for thinking that C & L reject minimal semantics or for thinking that RT is a more minimal theory of semantic content. Since the contribution of context at the level of pronominals is factored out in the attempt to show that the reported utterance in an indirect report has the same content as in the original utterance, the theory has the level of abstraction required. Furthermore, it is simply not true that RT is a more minimal semantic theory simply because it factors out the contextual interpretations of pronominals, given the radical claim that the predication in an utterance is largely underdetermined (e.g. ‘The car is red’). We do not grant that RT is more minimal at the level of pronominal interpretation, but even if we granted that, given that the predication act is largely underdetermined under RT, Wedgwood should be discouraged from arguing that RT is a more minimal semantic theory than C & L.

At this point the following objection by Wedgwood (p.c.) should be answered:

But why couldn’t RT just argue in parallel that the contextual contribution required by ‘red’ is ‘factored out’—you get an invariant interpretation of ‘that is red’ so long as you fix both the referent of ‘that’ and the way in which the object satisfies ‘red’. The point is that RT holds both pragmatic enrichments together, indexical saturation and predicate explication).

The question Wedgwood is pressing is why C & L treat the former as part of minimal semantics and the latter as part of pragmatic communication.

The answer to this question is very simple. C & L grant that pragmatic intrusion is indispensable when it comes to pronominals and similar expressions drawn from the ‘Basic Set of genuinely context-sensitive expressions’, while they do not grant that pragmatic intrusion occurs at the level of predicates such as, e.g., ‘red’. Given the asymmetric status of pronominals and predicates, it does not come as a surprise that the pragmatics of pronominals is somehow incorporated into minimal

semantics, while the pragmatics of predicates is relegated to the level of speech act interpretation. Presumably, the reasoning implicit in this asymmetric treatment is that pronominals and other deictic expressions do not contribute much to the thought expressed without pragmatic intervention, that is without relating the sentence to the context in which it is uttered. There is no thought at all, as Frege would say, without anchoring deictic elements to a given context. So, presumably in this conception there is an implicit equation between knowing the semantics of a sentence and knowing the thought it would express. A good question to ask is whether there could be genuine thoughts without anchoring deictic elements and pronominals to a context. The thought expressed by (5)

(5) She went to the cinema

cannot be the thought that a female person went to the cinema, since that would be expressible through (6) as well:

(6) A woman went to the cinema.

And (5) is different from (6).

A minimal semantics without provision of referents for pronominals would not describe thoughts, but would simply provide skeletal elements for the provision of thoughts. Instead, if my understanding of C & L is correct, they want to equate semantics with the vehicle for expressing thoughts, not just fragments of thoughts. Wedgwood writes:

Wedgwood writes:

It is particularly strange that C & L should include the referents of overt indexicals in that meaning that is infallibly shared, since this is one of the most obvious sources of misinterpretations in language use (given an utterance of *He's ready*, surely a response like *Who? Mark or Paul?* is at least as likely as *Ready for what?*: yet, according to C & L's position, the former question should be pre-empted by 'shared content' while the latter is not) (Wedgwood 2007: 666).

The point is exactly the opposite of the one made above. C & L admit that there is a tightly restricted class of genuine context-sensitive expressions, such as pronominals, which have variable interpretation in relation to the actual context of use and linguistic expressions that are not context-sensitive. They deny that 'ready' is a context-sensitive expression.

Yet, one might insist that Wedgwood is exploiting C & L's claim that 'ready' is not context sensitive but that 'he' is. Given C & L's view, the referent of 'he' makes it into the shared semantic content while the determination of 'ready' doesn't, yet as Wedgwood points out hearers often don't know who is being referred to by an utterance of 'he' (just as they may not know 'ready for what')—but C & L's position should rule this out—if we grasp shared content for C & L we should know who or what the referent is.

There is no doubt that some hearers may be in a position not to know the referent of a pronominal. Suppose I am persuaded that behind me there is a picture

of George Washington. It has been there for twenty years, and I expect it to be there while I am speaking. However, Mr Wedgwood to prove that C & L's theory is wrong suddenly removes the painting, while I say to my guests: 'He has a beautiful jacket'. When I turn round to point to the painting, I find out that George Washington is no longer there. My referential intention, as Wedgwood would want to say, cannot be understood correctly by my interlocutor. Is this situation the norm? Is this the way we ordinarily use pronominals and deictic expressions? I would say it is not. It is possible that the addition of a semantic clause restricting the use of pronominals to accompanying demonstrative gestures is too strong. However, surely there is a norm preventing speakers from using pronominals unless the hearer can identify the intended referent of the pronominal through contextual clues that are unambiguous. Suppose I am a general reviewing an army. There are thousands of soldiers before me. Could I legitimately use: 'He is clever', unless there are contextual clues that enable the hearer to identify the referent? Presumably I could not.

So, while I agree that unlike predicates, pronominals do not have a content that can be grasped without recourse to a context, there must at least be some basic agreement with C & L that pronominals belong to shared meanings, in the sense that given the semantic resources available to speakers/hearers and knowledge of rich contextual clues, one is able to understand the meaning of linguistic expressions in context. Presumably to work perfectly, C & L's view would have to be allied with a view of pronominals and indexical expressions in which semantics strictly regiment the content of a linguistic expression through a rule to the effect that pronominals and indexicals should be accompanied with unambiguous demonstrative gestures. Whether such a theory is desirable or terribly problematic is a topic for a different paper.

There is, obviously, the problem of how a context-sensitive expression can contribute a context-invariant meaning. But is this a real problem for semantics? While it is certainly true, as Wedgwood says, that the referent of a pronominal is provided by context, this is not to say that the semantics of pronominals cannot be furnished in a context-free manner. A very straightforward method is to let the pronominal refer to a thing *x* and to provide the semantics of the pronominal expression through a conditional:

(7) He is happy

is given the following semantics:

(8) If 'he' refers to *x*, then 'He is happy is true' just in case *x* is happy (Higginbotham, p.c.).

Following this reasoning, we can now understand how C & L can answer Wedgwood. The minimal semantics of a sentence includes the referents of pronominals and still remains context-invariant. To furnish the referent of a pronominal is just to furnish a specific value to 'x'. Replacing this value for *x* in (8), we obtain:

- (9) If ‘he’ refers to John, then ‘He is happy’ is true just in case John is happy. ‘He’ refers to John.

The second clause of (9) is provided through pragmatics, but this does not destroy the context-invariance of the minimal semantics of (8), since this is perfectly represented in the first clause in (9). All C & L should do is use context-invariant semantics and couple it with a specification of the value of the variable involved in the logical form of the minimal semantics.

8 Indirect Reports as a Test

Wedgwood discusses ‘Indirect reports’ as a way of testing intuitions about context-sensitivity. According to C & L indirect reports are a good test for context-sensitivity, as genuinely context-sensitive expressions, such as indexicals, as used in the original utterance, simply cannot be preserved in the indirect report. Here is an example:

- (10) John: I am happy;
 (11) John said that I was happy.

The sentence (11) cannot be a correct indirect report of (10), but something like (12) is required:

- (12) John said that he was happy.

On the contrary, if a lexical item (say one in the predication) is not context-sensitive, it can figure both in the original utterance and in the indirect report, as the new context of the indirect report does not change its interpretation. Here is an example from C & L (not a good one, actually):

- (13) John: Mary is ready.
 (14) John said that Mary was ready.

Furthermore, suppose that ‘ready’ was used in two (completely) different contexts, as in (15), (16), then it would still be correct to make the report as in (17):

- (15) John: Mary is ready (for the exam);
 (16) Ted: Mary is ready (to catch the train).
 (17) Both John and Ted said that Mary was ready.

This, however, should be understood as a quotational usage. The indirect reports promotes some common denominator interpretation, even if the two sentences were interpreted quite differently. I will return on the equivocation that allows C & L to obtain the wrong effects with the test in the case of ‘ready’.

However, Wedgwood (2007) writes:

C & L’s reasoning here rests on a crucial empirical question that they do not investigate: do indirect speech reports work consistently in this way? In particular, can we really make reliable

judgements of truth and falsity about such reports? Note that this consistency is crucial to C & L's arguments; if the answer to the question is no, we can no longer assume that the complement of 'said' in such cases is necessarily propositional. In that case, being the complement of 'said' simply is not a test for propositionality—with the consequence that indirect speech reports cannot be used to argue for the propositionality of encoded meaning that putatively distinguishes semantic minimalism from RT.

There are various ways in which indirect reports of the kind that C & L rely on can fail to report a single proposition. For example, being the complement of 'said' does not force disambiguation of homonyms. I can truthfully utter the sentence 'In C1 and C2, Nina said that John went to the bank' even if I know that Nina was referring in C1 to the financial bank and in C2 to the watercourse bank. (...) Significantly, one may easily do this with regard to C & L's example also: In C1 and C2, Nina said that John is ready—but in C1 she meant that he is prepared for the exam, while in C2 she meant he has his raincoat on. (Wedgwood 2007: 669).

It is legitimate to raise such doubts. In fact, if indirect reports merely allowed quotational interpretations, then their efficacy would be destroyed. As Wedgwood says, they would no longer be efficient tests for propositionality. However, the very examples he uses persuade us that indirect reports like (18)

(18) In both C1 and C2, Nina said that John went to the bank
made on the basis of utterances such as (19), (20)

(19) Nina: John went to the bank₁ (financial institution)

(20) Nina: John went to the bank₂ (river course bank)

are not licit. I think everyone agrees that (18) should be different from (21):

(21) In both C1 and C2 Nina uttered the words: 'John' 'went' 'to' 'the' 'bank' in this order.

It would not even be licit to utter (22):

(22) In both C1 and C2 Nina uttered the sentence 'John went to the bank'.

In fact, given the homonyms 'bank₁' and 'bank₂' there is not a single sentence, but there are two different sentences having distinct meanings.

Wedgwood (2007) also writes:

Similarly, C & L's test fails to disambiguate use/mention ambiguities.

My intuitions are completely different. So called use/mention ambiguities are not really cases of ambiguity in the written language, where inverted commas are conventionally used to make reference to mentions of a lexical item. Oral language uses different ways of marking inverted commas (...). Given these facts, it would simply be absurd to use indirect reports in ways that ignore the use/mention distinction (Wedgwood 2007, 669).

The attack on indirect reports as a test for a common meaning denominator is vaguely reminiscent of another attack on Cappelen and Lepore by Stanley (2005) Stanley's position is misguided, because it focuses on a specific problem in C & L's approach and draws generalizations about a more general point. Stanley, after noting C & L's position according to which

The semantic content of a sentence *S* is the content all utterances of *S* share (p. 143)

moves on to ponder on the consequences of this statement. He writes:

This passage is rather unclear. Is it the content that all assertions of *S* express, no matter how different their contexts of utterance? If so, “every bottle is in the fridge” has no semantic content relative to any context, since there is no one proposition that is asserted by every utterance of the sentence (and certainly not, as we have seen, the proposition that every bottle in the universe is in the fridge, since this is never asserted). If the common content of all utterances of a certain sentence is not the content of any genuine speech act, what is the motivation for thinking that common contents are always genuine propositions, rather than just Recanati’s “semantic schemata”? (Stanley 2005: 143).

However, this scepticism is simply defeated if one notices that C & L may well be wrong about quantifier domain restriction (a story like Stanley’s being more palatable), while being right in the general claim that the semantic content of a sentence is the content that all utterances of *S* share (Perhaps the addition is required that we should confine ourselves to serious, rather than non serious uses).

To go back to Wedgwood, one of the merits of C & L is to say that we should extrapolate some common denominator meanings from the various and disparate uses of some linguistic (sentential) expression. It is essential for this virtue of the theory to survive that both Wedgwood’s and Stanley’s contentions be silenced.

We should now reflect on the meaning of ‘ready’. Contextualists assume that ‘ready’ is more or less interpreted as ‘ready for that’. This would explain why in a context ‘John is ready’ means that he is ready for the exam, while in another context it means that John is ready for the tennis match. I claim that contextualism along these lines does not seriously jeopardise C & L’s work, as they, after all, admit that there is a restricted class of context-sensitive expressions, which they call ‘The Basic Set’. Pronominals belong to this class. But are there ways to make sense of C & L’s claim that ‘ready’ is not context-sensitive? It should at least be possible to discover that ‘ready’ is, semantically, not context-sensitive as contextualists say. I think it is useful to consult the Longman Dictionary of the English Language. The dictionary lists a number of senses, some of which appear to be more context-sensitive than the others. Let us start from the least context-sensitive:

- ‘spontaneously prompt’ (always has an answer);
- ‘Notably dexterous or skilled’ (he is very ready craftsman);
- ‘Immediately available’ (had little ready cash)
- ‘prepared for immediate use’ (dinner is ready’)

Then we have some context-sensitive senses:

- ‘forward or presumptuously eager’ (he is very ready with his criticism)
- ‘prepared mentally or physically for some experience or action’.

The real problem is to do with ‘ready’ what one should do with ‘open’, that is to say account for the various distinct uses of this word. Suppose we choose a sense such as ‘prompt’ (‘disposed to act as occasion demands’) as the basic sense from

the others can be derived pragmatically. At this point, we can easily explain how Intercontextual Disquotational Indirect Reports make this basic sense emerge as a context-invariant one.

I should clarify that I am not saying that ‘John is ready’ is context-independent, but I am arguing that there is a meaning of ‘ready’ which allows C & L to obtain the right results from the test of indirect reports.

Someone might criticise my views suggesting that by treating ‘ready’ and ‘enough’ as having hidden indexical components, I am actually rejecting the results of C & L’s tests. This suggestion is too quick, and to some extent resembles Stanley’s tactic to generalise from a local problem to a general problem. It should be admitted in principle that even if there is a reliable testing procedure, the person effecting the test may be wrong in executing it. We should remember that tests should be used judiciously, as they depend, after all, on intuitive judgements. In the case of ‘ready’, C & L strain their tests of context invariance. When we say that Julia and Mary are both ready, given that Julia is ready for her English exam and Mary is ready for her maths exams, we are certainly right **on the interpretation** that they are both well prepared for their respective exams, but saying that they are both ready in the sense that they are both ready for *x*, *x* being the (specific) thing each intends to do is surely false. It may very well be that there is an equivocation of meanings in the way C & L use this test, but this does not mean that all users of the test should be led away by things such as equivocation (of meaning).¹⁰

9 Justification for Semantic Minimalism

Last, I want to discuss Wedgwood’s following assertion:

Consider again the kinds of observation that C & L claim make shared context indispensable: co-ordinated action, collective deliberation, linguistic communication justifying beliefs, holding people responsible for what they say, and so on. Is minimal semantic content capable of supporting such things? If so, then an engineer who says ‘Steel is strong enough’ is held responsible just for that proposition (if we call it that) which is expressed by any and every utterance of this sentence. If the roof collapses, it should be considered reasonable of the engineer to point out that I just said steel is strong enough; I never said strong enough to support the roof. Such utterances would usually be described using words other than ‘reasonable’. (Wedgwood 2007: 670)

¹⁰ As Zwicky and Sadock (1975: 14) say, “If the semantic representation of certain sentences lack specification of some piece of meaning, then the applicability of transformation to them cannot possibly depend on whether or not this piece of meaning is present. If a sentence is unspecified with respect to some distinction, this lack of specification must be preserved by every transformational operation. But if a sentence is ambiguous, then it is possible for a transformation to apply in some but not all, of the cases, so that the effect of the transformation is to eliminate one or more understandings of the sentence”. In our case, the crossed reading can be obtained because the underspecified interpretations are preserved by the transformation of conjunction reduction. Since under-specification is preserved, the crossed readings can obtain.

While in principle we could agree with Wedgwood that more than minimal semantics is required for language to stimulate coordinated action, two things ought to be noted here:

C & L might reply that minimal semantics is like the pillars and beams of a house; the house can't exist without them. So communication is not possible without minimal semantics. Coordinated action as triggered by language implies a shared understanding of at least the basics of the utterances: that is to say their minimal semantics.

Second, in replying specifically to the point about the falling roof, an engineer who says 'Steel is strong enough' presumably means that 'steel satisfies one's needs. Then the pragmatics of his utterance will intervene in securing the interpretation that steel is strong enough for the roof. However, even if pragmatics did not intervene to secure such a specific interpretation, the engineer cannot be accused of incompetence or of having said something irrelevant.

10 On the Compatibility Between Radical Contextualism and Semantic Minimalism

Jaszczolt (forthcoming) claims that there is compatibility between Semantic Minimalism and Radical Contextualism. The compatibility in question mainly stems from the fact that (or as a consequence from the fact that) she abandons the syntactic direction principle, that is to say the assumption that the enriched propositions obtained through radical pragmatics should be considered developments of logical forms provided by sentential semantics. In other words, if I understand Jaszczolt well, in many cases, arriving at an enriched proposition is not a matter of taking a logical form as a point of departure and arriving at an enriched logical form, but it is a matter of going beyond what the initial logical form says, of bypassing that logical form. Consider the examples

(23) You are not going to die

(24) Child: Can I go punting?

Mother: You are too small

It seems implausible that the hearer goes from the literal logical form 'You are not going to die at t' to the non-literal 'You are not going to die from that cut' or that in (24) the hearer first recovers 'You are too small', whatever its context-invariant meaning ought to be, and then moves on to the contextualised meaning 'You are too small to go punting'.

There are also many other examples, from Capone (2009), where recovering an explicature is not just a question of developing a logical form, but is a question of bypassing it altogether:

(25) Sicily is a triangle

is one such case. In this case the enrichment is not of an augmentative type, but of a subtractive type, as it is implausible that Sicily is a literal triangle. So, it appears that the assumption that an explicature should be a development of a logical form in the sense that the logical form it starts with appears intact in the explicature must be abandoned.

I am aware that more than one approach will be available to this example, so I propose to use other examples, to show that the logical form of a sentence can be bypassed altogether in semantic interpretation. Cases of irony can prove that what is literally said need not be literally expressed and that, therefore, all cases in which a literal meaning is projected, heavy pragmatic processing is required.

The point by Jaszczolt that implicatures and explicatures can be calculated at the same time in the same merger representation is a bit more contentious. One of her best examples is the following:

(26) Everybody is going to Egypt this spring.

She claims that the main point of the assertion (the primary meaning) is that the interlocutors should consider going to Egypt this spring. She argues that primary meanings should be allocated a prominent place in a contextualist theory of meaning. That preoccupation, for Jaszczolt, seems to be incompatible with the syntactic direction principle. This point seems to me to be less well established, UNLESS we take her claim to be that a merger representation merges information deriving from explicatures and implicatures and that in many cases the literal interpretation is overridden by contextual considerations.

Let us return to the compatibility between semantic minimalism and radical contextualism. Once we agree with Jaszczolt that minimal semantics lacks psychological plausibility, a point that I think is conceded by C & L (2005), there are no problems in understanding how contextual considerations can override minimal semantics and how minimal semantics is absorbed and transformed into richer and often radically different propositional forms.

According to Jaszczolt, Semantic Minimalism is a project which is busy with truth-conditions at the sentential level, while radical pragmatics is busy with truth-conditions at the utterance level. In the same way in which semantic minimalism adheres to a compositionality principle, radical pragmatics adheres to a compositionality principle at a different level, given that the enriched propositions contain constituents which are not present at the level of logical form. It goes without saying that the compositionality principle employed in radical pragmatics is not a level of syntactic combination that belongs to surface structure (or to the logical form of the sentential semantics), but presupposes a level of syntactic combination that is (in principle) usable at the level of surface structure (or of the logical form of the sentential semantics).

According to Jaszczolt (2005), merger representations have the character of compositionality. This is to be accepted, as some implicit constituents, hence structure, is imposed on merger representations; we should suppose that the

combination of certain semantic interpretations with certain pragmatically supplied constituents should follow compositionality. Consider the following example:

(27) That piano is better.

In (27) there is a missing constituent, and this one must combine with ‘That piano is better’. Now, in examples like (27) the kind of compositionality which the merger representation displays is exactly the same which the full sentence would exhibit. (There are, of course, infinite discussions on whether (27) is a complete sentence or not; what is incontrovertible is that the full alternative to (27) has the same compositional structure as an utterance of (27)). However, Jaszczolt claims that in some cases there are some mismatches between the sentence and the utterance, and that while compositionality cannot be held at the level of the sentence, it can be held at the level of the merger representation. The best examples of this mismatch are belief reports. Jaszczolt’s ideas derive from Frege. Compositionality, at the level of the sentence, in the case of (28) breaks down.

(28) John believes that Mary is happy.

It breaks down because, by the compositional picture, when one replaces ‘Mary’ with a coreferential NP, one obtains (or may obtain) a false statement. So, a picture in which we compositionally build up the meaning of the sentence/statement by replacing each constituent with the appropriate value it has in context breaks down because such a compositional picture would require the possibility of substitution of identicals in slots associated with, say, proper names. At the level of the merger representation, further structure is provided, that quickly shows how it comes about that substitution of identicals is not always licit in intensional contexts. (Implicit modes of presentations are provided and integrated into the semantics, say as appositives, following (Capone 2008b). Belief reports are ideal candidates for the task of showing that compositionality obtains at the level of merger representations. I would add that ‘de se’ beliefs are even better for this purpose.¹¹ Although my views on belief reports do not exactly coincide with Jaszczolt, we agree that the so called scope ambiguities do not resolve the problem of compositionality. In fact, assuming scope ambiguities, it is possible to differentiate the ‘de re’ and the ‘de dicto’ reading of a belief report

(‘de re’ reading: John is such that Mary believes of him that he is clever)

(‘de dicto’ reading: Mary believes that John is clever).

¹¹ My paper on ‘de se’ attitudes (Capone 2010b) shows that ‘de se’ constructions are cases of intrusive constructions à la Levinson. ‘De se beliefs’ like ‘John remembers being in Oxford’ are beliefs about the self—they have first-personal readings which are truth-conditionally different from ‘de re’ readings. In my paper I argued that the mode of presentation ‘I’ is implicit in de se attitudes and furnished through pragmatic intrusion. Pragmatic Intrusion is also instantiated in the internal dimension of PRO in ‘de se’ constructions. See also Lewis (1979).

However, a problem to be noted is that the ‘de re’ reading entails the ‘de dicto’ reading, and thus scope ambiguities do not explain at all the problem of opacity (there would be some kind of circularity). It is best to accept, as Jaszczolt does, that compositionality is at the level of merger representation, where an implicit mode of presentation motivates opacity in the case where intentionality is dispersed (the strongest level of intentionality correlating with ‘de re’ readings, according to Jaszczolt). Does this amount to saying that the compositionality principle applies only at the level of merger representations? Should we reject the compositionality principle at the level of sentential semantics? (However, it should be born in mind that sentential semantics often consists of logical forms that do not match with the veneer of sentence structure). There is a sense in which the reply to this question is positive. If we do semantics in the merger representation, compositionality in the merger representation is all that is required (Jaszczolt 2005, p. 72). However, I also sympathize with more traditional semanticists, who claim that compositionality is an important characteristic of semantics. I can comfort them by saying that, after all, their semantics is done at a level of abstraction, thus it is methodologically correct to attribute compositionality to sentence structure, even if it has been abstracted away from the merger representation, for methodological reasons. The compositionality we find in the sentential semantics percolates to the semantics from the merger representation. However, since semantics is done, for methodological reasons, in abstraction from the merger representation, it appears as if the compositionality is there in the semantics. And in a sense, perhaps not the one by Jaszczolt, it is in the semantics.

11 A Final Note

In this section I will sketch a solution to some of the problems discussed, while I cannot go into the details of such a solution. It appears to me that Semantic Minimalism seeks to establish an asymmetry between subject and predicate positions. Subject positions are invariably subject to pragmatic intrusion, and there is nothing one can do about it. Furthermore, pragmatic intrusion provides a constituent of the proposition, whether minimal or at the level of speech act theory. Without such a constituent there can be no minimal proposition, so C & L cannot really provide minimal semantics without such constituents. Since subject positions are usually positions for reference assignment and reference assignment is pragmatic (having to resort to a number of contextual clues), there is no expectation that the subject position can provide a constituent of thought without referential resolution; and there is no expectation that there can be a minimal thought without referential resolution and the assignment of a constituent to the subject position. The predicate position, instead, has got a different status. It is true that a predicate is not immune to pragmatic enrichment; but it is not as dependent on pragmatic intrusion as the subject. Many cases have been provided to try to show that predicates cannot really furnish minimal truth-conditions. However, I am not

persuaded that these cases are really against Cappelen and Lepore. Consider the case of ‘The lemon is yellow’. This may well require pragmatic intrusion, but only at the level of the subject. We may enrich the proposition up to ‘The lemon’s peel is yellow’. Since the pragmatic intrusion is required at the level of the subject, the predicate is not affected. Could we extend a similar treatment to ‘John is ready’? Presumably, we need to transfer the pragmatic intrusion from the predicate to the subject. One way to do so is through an implicit apposition constituent: ‘John [who must take the exam] is ready’. We could extend this position further by positing a null prepositional phrase as sub-categorized by ‘ready’. We thus obtain ‘John [who must take the exam] is ready [for it]’. And now we have obtained a considerable advantage. While I previously said that ‘ready’ subcategorizes the constituent ‘for that’, which is fundamentally deictic, the prepositional phrase ‘for it’ is anaphoric. This means that the constituent [who must take the exam] is to all effects part of what is said, given that it is indispensable for anaphoric resolution. This could explain, presumably, why ‘Mary is ready’ cannot really mean that Mary is ready for that, for that and for that. The enrichment process is constrained and part of its constraints is that one should make sense of a sentence by enriching the subject first and one can enrich the predicate by anaphoric resolution through materials provided as apposition in the subject. Furthermore, one cannot say ‘Mary is ready and not ready’ meaning ‘Mary is ready for that1 and not ready for that1’. This presumably follows from the fact that the prepositional phrase ‘For it’ is anaphoric and thus it would be a contradiction if the first occurrence of ‘it’ and the second occurrence of it both referred back to the same constituent. ‘Mary [who must take an exam] is both ready [for it] and not ready [for it]’.

If I am right about these data, then we have bumped into the deep question of constraints to pragmatic enrichment and where they come from. But now we must ask the following questions. Do all predicates follow the same pattern as ‘ready’? Why is it that predicates can be enriched only through anaphoric resolution, while subjects can be enriched through provision of apposition constituents (that is to say in a really free way)? Why is it that indexicality cannot occur in predicates (unless use is made of explicit deictic pronominals), while anaphoric pronominals are tolerated? In reply to the first question, it appears that things are very much the same with ‘happy’. We cannot easily say ‘Mary was happy but unhappy’, meaning ‘Mary was happy about this but unhappy about that’, as we have the same anaphoric pattern as before: ‘Mary [who found her jewel] was happy about it but unhappy about it. It is also clear that somehow the anaphoric pronominal incorporates a reference to the event of finding Mary’s jewel, which compels the semanticist to incorporate the event in the apposition close through a device such as ‘Mary, who was such that there was an event of finding her jewel...’. Now I address the question of why pragmatic intrusion in the predicate appears to be subordinate to pragmatic intrusion in the subject and why it can occur through anaphora and not through deixis. A tentative answer has to do with a principle which has been brought to my attention by Jaszczolt (1999): Do not multiply levels of interpretation. This seems to correspond precisely with Cappelen and Lepore’s notion of semantic minimalism, in which intrusion is granted for subjects but not

or only minimally for predicates. In other words, Parsimony of levels of interpretation (POL) compels us to minimize the loci of pragmatic intrusion and to utilize an already existing and necessary locus. Since the subject is a necessary locus of pragmatic intrusion, reference being necessarily a pragmatic process and the subject being the locus for an NP that is also a referential position, pragmatic intrusion must occur primarily in the subject and if pragmatic intrusion occurs in a predicate, then this process must be subordinated syntactically to the subject position. Subordination allows us to keep the level of pragmatic intrusion in the predicate to a minimum (anaphoric resolution), while the pragmatic intrusion occurring in the subject is more radical.

Nelson Morales (personal communication) has thought of ways of falsifying my considerations. He thinks that loosening is such a case. He thinks that, in certain circumstances, a statement such as ‘Mary is ready and not ready’ is not contradictory (consider the interpretation ‘Mary [who must take the exam] is [physically] ready but [emotionally] not ready’). His remarks are certainly useful, even if his example does not necessarily require loosening. As should be clear to anyone, contradiction is a property of statements and NOT of sentences. Someone who says ‘Mary [who must take the exam] is [physically] ready and not [emotionally] ready [for it]’ introduces some modifiers and this kind of modifier introduction is one of the various manifestations of pragmatic intrusion, not necessarily loosening. Notice that the anaphoric phrase [for it] stays there all the time and even in his example, one is prohibited from interpreting: Mary [who must take an exam] is physically ready for this [the exam] and emotionally not ready for that [getting married].

His example is very interesting and useful, as I said. Loosening applies to predicates like ‘a triangle’, when one says things such as ‘Sicily is a triangle’. However, these cases are different because what is involved is a modification of the predicate to make the predicate fit the subject: Sicily has the rough shape of a triangle. Alternatively: Sicily’s shape resembles a triangle. Strictly speaking, it would not be true to say that Sicily is a triangle. And again, one could contrive the case in such a way, that the pragmatic intrusion is at the level of the subject and not of the predicate. One, in fact, could have the following result:

Sicily’s rough shape is a triangle.

This would very much be compatible with Jaszczolt’s POL.

Let us now consider the following objection by Burton-Roberts (p.c.):

In the rest of the paper you contrast indexicals (including pronominals) and predicates. However, in the final section you advert to a different notion of ‘predicate’. Here ‘predicate’ identifies a sentence constituent (‘THE predicate’) in contrast to the sentence constituent you refer to as ‘the subject’. On this latter use it corresponds to VP, assuming $S \rightarrow NP + VP$. It is this that leads you to refer to an asymmetry in C & L’s treatment of subjects and predicates and, on that assumption, to your proposal.

But this (VP) sense of predicate is not the sense of ‘predicate’ appealed to in the rest of the paper which I’ll call “pred-1”. Pred-1’s can and do occur in the subject (and indeed all over the place, it covers common nouns, verbs, adjectives and adverbs). Any *descriptive* subject expression includes a pred-1. So I don’t think C & L’s position can be said to result in an asymmetry between subject and predicate—and that perhaps you need to look again at your proposal in the light of that.

Answer: it is true that C & L speak of sub-constituent predicates, and of them they say that they are not to be modulated (their meanings are stable and are not in need of modulation). On the contrary for more radical treatments, predicates (at the sub-constituent level) require free enrichment because they cannot contribute to a full proposition; in other words, the contribution of the predicate to the proposition is sub-propositional. But then if my predicate roughly corresponds to a VP constituent, it is clear that the VP which contains as a sub-constituent a Predicate 1 also contributes a sub-propositional contribution; in other words, by applying that VP to the subject assuming that the subject can be used to identify a referent, one still obtains a sub-propositional element, that is to say an element which is not fully truth-evaluable. Since subjects are referential elements and reference always needs pragmatics means for being secured, it may be useful to transfer pragmatic intrusion from the level of the Predicate 1 and, subsequently, from the level of the next projection, VP to the subject, which, as I have already said, needs to secure its referent through pragmatic means.

Now consider the idea (correct, as far as I can understand) that after all Predicates of the type Pred1 can occur inside the subject (adjectival modification is an obvious illustration (e.g. ‘A tall girl’). As I said predicates of the type Pred 1 may be in need of pragmatic intrusion. But this is ok, since we hypothesed to start with that pragmatic intrusion is more parsimonious if it applied at the level of the subject.

Let us consider the next objection by Burton-Roberts (p.c.):

I don’t think you CAN “transfer the pragmatic intrusion to the subject” from the predicate, as you suggest, i.e. by interpolating a non restrictive (NR, appositive) relative clause. [1] Interpolating an NR relative does not actually enrich/modulate the referential semantics of the subject (unlike restrictive relative clauses); indeed, it has been argued that NR relative clauses aren’t true sentence constituents anyway; [2] there is no limit to the range of the content of such NR clauses you might add (anything you might happen to know the referent of the subject); [3] as I understand it, you want to restrict the content of the relative clause in a way that might be relevant to the modulation of the predicate. However, that could be argued to be circular, either empirically or methodologically: first you have to identify the subject in order to know what modulation of the predicate might be relevant but you are arguing that the subject must be modulated in a way relevant to how the predicate is modulated.

Answer: why should we secure modulation in the subject by taking into account possible ways of enriching the predicate? The predicate does not require any enrichment if we are ready to enrich the subject in the proper way. When I said that ‘Sicily is a triangle’ may correspond to the proposition ‘Sicily’s rough shape is a triangle’ I gave up completely the idea that the predicate or VP where it is

contained needed pragmatic enrichment. I said that I transferred enrichment from the predicate to the subject as a way of speaking, as a way of contrasting my analysis with those of radical contextualists, but the transferral of the enrichment from predicate to subject, in fact, may never occur (otherwise, it would be correct to say that it would be circular). All we need is to enrich the subject.

My opponent argues that Non-restrictive relative clauses may not be relevant to pragmatic enrichment. Yet, if one makes use of them in the context of a sentence such as ‘Mary is ready’, which contextualists argue to be context-sensitive, and it can be demonstrated that they play a role in modulating the meanings of such sentences, it goes without saying that they must be ancillary to pragmatic enrichment, whether or not, by themselves, they constitute cases of pragmatic intrusion. (In any case, see Capone 2008 on pragmatic intrusion and appositives). But then which are the reasons for doubting that the addition of non-restrictive relative clauses may amount to pragmatic intrusion? Presumably my opponent thinks that the addition of a non-restrictive relative clause will be indifferent to truth-conditional evaluation. Consider the case ‘John, who is the President, never accepts to be contradicted’ as a pragmatic enrichment of ‘John never accepts to be contradicted’. Surely this is an obvious case where the pragmatic intrusion does not intrude into truth-conditional meaning; so we may very well call it a pragmatic intrusion but not a truth-conditional pragmatic intrusion. And this is pretty uncontroversial. However, what about more complicated cases, such as Mary believes that [John does not accept to be contradicted] [APPosition]? In Capone (2008), I articulated the idea that apposition clauses serve to provide modes of presentation (sentential appositions being modes of presentation of the that-clause dependent on ‘believe’). But then there are at least some cases in which appositions (which is what non-restrictive relative clauses are similar to) result in truth-conditional pragmatic intrusion.

Finally, let us consider the charge of circularity: “first you have to identify the subject in order to know what modulation of the predicate might be relevant but you are arguing that the subject must be modulated in a way relevant to how the predicate is modulated”. But I said previously that my way of speaking ‘shifting pragmatic intrusion from the predicate to the subject’ is only a way of speaking and we need not think of it in this way. It is not the case that I need to identify the subject and the predicate etc. The predicate usually has slots one of which is occupied by the subject. So the subject can only occur in that slot. Thus I need not identify the subject, but the subject allows me to identify the object it refers to, and this can be done, usually, if the NP corresponding to the subject is expanded pragmatically. And most importantly I am not arguing that the subject must be modulated in a way relevant to how the predicate is modulated. In fact, what I am denying is that the predicate is being modulated. The modulation of the predicate is only a theoretical construction of other theorists—certainly a plausible construction until other possibilities are investigated. However, when other possibilities are investigated, one can easily see that there are syntactic means for gluing the pragmatic increments to subjects and thus explaining away what *prima facie* can be considered pragmatic enrichments of the predicates. What I am basically

claiming is that yes, there is a phenomenon of pragmatic intrusion, but the location of the intrusion is wrong. But then why is it that theorists are so adamant in claiming that predicates are loci of pragmatic intrusion? I will reply that theorists wear the eyeglasses of their own theories. We usually see what we want to see or what we can see through our ordinary glasses. Superficial syntax may also lead us astray and once we may consider other possible ways of gluing constituents to the existing ones, enrichments can be constructed in different ways. The merger representations in which literal and pragmatic meanings are combined are open to combinations which amount to different syntactic glues, which are different from the superficial syntax we can observe by considering literal meanings. The existence of merger representations à la Jaszczolt (2005) opens us up to the idea that there may be different modes of syntactic combinations other than those visible at the level of superficial syntax. We need not impute these non-standard modes of syntactic combination to logical form. We may simply impute them to the Merger Representations.

12 Conclusion

I confined myself to the discussion of Wedgwood's points from the perspective of C & L. In general, we should be taught some salutary lessons by C & L concerning minimal semantics. This is not to say that we should abandon a contextualist approach to language. But perhaps it is convenient to make such an approach as compatible as possible with C & L's assumption that the Basic Set of context-sensitive expressions should be maximally restricted. There are some interesting context-sensitive expressions mentioned by C & L, such as 'enemy', 'friend', 'foreigner', 'outsider' etc. (C & L p. 1), but these are almost never discussed in the literature. I think we should concentrate on cases such as these or such as those dealt with in Capone (2006), and Capone (2008b).

The ideas on the semantics/pragmatics debate are quite fluid for the time being, but I have a propensity to follow Jaszczolt in her radical pragmatic perspective on language, trying to make her picture compatible with more classical theories on semantics. The idea that semantic minimalism and contextualism should not be enemies has recently been voiced by Recanati (2010). Of course, to enhance this compatibility one would have to accept that at least in many cases, but not in all cases, the meanings of sentences are truth-conditional and that at least in basic cases knowledge of the language will allow one to say what must be true for an utterance of a given sentence to be true. Contextualism, in such a hybrid theory, would be considered as a tool for enhancing the semantic potential of a language, rather than for arguing that language in itself is value-less. I remember distinctly Higginbotham's (personal communication) words in this respect—pragmatics is a **path** that makes truth-conditional semantics work. It is very useful to think of semantics and pragmatics in this way: as a path and a vehicle. You cannot use your car unless you have a road; the road is what makes your car useful. The road

without the car is useless. Nevertheless, you can use your car in as many roads as you want to. This invariably involves a certain degree of autonomy.

Appendix 1

Abbreviations

In this paper, I make use of the following abbreviations:

SM	Semantic minimalism
RT	RT
RC	Radical contextualism
MC	Moderate contextualism
C & L	Cappelen and Lepore

Appendix 2

Wedgwood, in replying to my paper, was bothered by the fact that C & L allow pragmatic intrusion for indexicals but not for predicates. However, on my view that at least in some cases, predicates subcategorize for null prepositional phrases a limited form of pragmatic intrusion could be allowed for predicates as well. On my view, the asymmetry could be dissolved. Of course, this approach may be suitable for some cases; I am not saying that it should be suitable for all predicates. In fact, it is plausible that some forms of modulation à la Recanati should be accepted.

References

- Bach, Kent. 2006. The excluded middle: Semantic minimalism without minimal propositions. *Philosophy and Phenomenological Research* 73(2): 435–442.
- Blommaert, Jan. 2005. *Discourse. A critical introduction*. Cambridge: CUP.
- Borg, Emma. 2004. *Minimal semantics*. Oxford: OUP.
- Capone, Alessandro. 2004. 'I saw you' (towards a theory of the pragmemes). *RASK* 20: 27–44.
- Capone, Alessandro. 2005. Pragmemes *Journal of Pragmatics* 37: 1355–1371.
- Capone, Alessandro. 2006. On Grice's circle (a theory-internal problem in linguistic theories of the Gricean type). *Journal of Pragmatics* 38: 645–699.
- Capone, Alessandro. 2007. Review of Emma Borg's minimal semantics. www.semanticsarchive.net.
- Capone, Alessandro. 2008a. Review of Cappelen, H. & Lepore, E. Insensitive semantics. *Journal of Pragmatics* 40(9): 1626–1635.

- Capone, Alessandro. 2008b. Belief reports and pragmatic intrusion: The case of null appositives. *Journal of Pragmatics* 40: 1019–1040.
- Capone, Alessandro. 2009. Are explicatures cancellable? *Intercultural Pragmatics* 6(1): 55–83.
- Capone, Alessandro. 2010a. Pragmemes. *Journal of Pragmatics* 42: 2964–2977.
- Capone, Alessandro. 2010b. On the social practice of indirect reports. *Journal of Pragmatics* 42: 377–391.
- Capone, Alessandro. 2010c. Between Scylla and Charybdis”: The semantics and pragmatics of attitudes “de se”. *Intercultural Pragmatics* 7(3): 471–503.
- Capone, Alessandro. 2010d. What can modularity of mind tell us about the semantics/pragmatics debate. *Australian Journal of Linguistics* 30(4): 497–520.
- Capone, Alessandro. 2010e. Pragmemes revisited (conflicts and power within the class). In *Perspectives on language use and pragmatics*, ed. A. Capone, 169–180. Lincoln: Muenchen.
- Capone, Alessandro. 2010f. Default semantics and the architecture of the mind. *Journal of Pragmatics* 43: 1741–1754.
- Cappelen, Herman, and Ernest Lepore. 2005. *Insensitive semantics. A defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell.
- Cappelen, Herman. 2007. Semantics and pragmatics: some central issues. In *Content and context: essays on semantics and pragmatics*, eds. G. Preyer, and G. Peter, 3–22. Oxford: OUP.
- Cappelen, Herman, and J. Hawthorn. 2009. *Relativism and monadic truth*. Oxford: OUP.
- Carpintero, Manuel Garcia. 2001. Gricean rational reconstructions and the semantics/pragmatics distinction. *Synthese* 128: 93–131.
- Carston, Robyn. 1991. Implicature, explicature and truth-theoretic semantics. In *Pragmatics. A reader*, ed. S. Davis, 33–51. Oxford: OUP.
- Carston, Robyn. 2002. *Thoughts and utterances*. The pragmatics of explicit communication. Oxford: Blackwell.
- Carston, Robyn. 2007. Linguistic communication and the semantics/pragmatics distinction. *Synthese* 165(3): 321–345.
- Cohen, J. 1971. Some remarks on Grice’s view about logical particles of natural language. In *Pragmatics of natural language*, ed. Y. Bar-Hillel, 50–68. Dordrecht: Reidel.
- Giora, Rachel. 2003. *On our mind: salience, context and figurative language*. Oxford: OUP.
- Jaszczolt, Katarzyna. 1999. *Discourse, beliefs, and intentions*. Oxford: Elsevier.
- Jaszczolt, Katarzyna (ed.). 2000. *The pragmatics of propositional attitude reports*. Oxford: Elsevier.
- Jaszczolt, Katarzyna. 2005. *Default semantics. Foundations of a compositional theory of acts of communication*. Oxford: OUP.
- Jaszczolt, Katarzyna. 2009. *Representing time*. Oxford: OUP.
- Jaszczolt, Katarzyna. Forthcoming. Pragmatic compositionality, syntactic direction principle and the contextualism/minimalism debate. Proceedings from ILCLI.
- Lewis, David. 1979. Attitudes De Dicto and De Se. *The Philosophical Review* 88(4): 513–543.
- Montminy, Martin. 2006. Semantic content, truth-conditions and context. *Linguistics and Philosophy* 29: 1–26.
- Recanati, François. 2004. *Literal meaning*. Cambridge: CUP.
- Recanati, François. 2006. Crazy minimalism. *Mind and language* 21(1): 21–30.
- Recanati, François. 2010. *Truth-conditional pragmatics*. Cambridge: CUP.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Roeper, Thomas. 1987. Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18(2): 267–310.
- Saul, Jennifer. 2004. What is said and psychological reality: Grice’s project and relevance theorists’ criticisms. *Linguistics and Philosophy* 25(3): 347–372.
- Michel, Seymour. 2010. Speech act pluralism, minimal content and pragmemes. In *pragmemes*, ed. A. Capone, 2870–2880. Oxford: Elsevier. (Special issue of the journal of Pragmatics 42).
- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance*. Oxford: Blackwell.
- Stanley, Jason, and Zoltan Szabó. 2000. Quantifiers and context-dependence. *Mind and Language* 15(2): 219–261.

- Stanley, Jason. 2005. Review of F. Recanati's literal meaning. *Notre Dame Philosophical Reviews* 9.
- Stanley, Jason, and Timothy Williamson. 1995. Quantifiers and context dependence. *Analysis* 55: 291–295.
- Stati, Sorin. 2002. *Principi di analisi argomentativa*. Bologna: Patron.
- Heusinger, Von, Turner Klaus, and Ken (eds.). 2006. *Where semantics meets pragmatics*. Oxford: Elsevier.
- Wedgwood, Daniel. 2007. Shared assumptions: semantic minimalism and RT. *Journal of Linguistics* 43: 647–681.
- Wilson, Deirdre, and Dan Sperber. 2002. Truthfulness and relevance. *Mind* 111: 583–632.
- Wilson, Deirdre, and Dan Sperber. 2004. RT. In *The handbook of pragmatics*, eds. L. Horn, and G. Ward. Oxford: Blackwell.
- Zwicky, Arnold, and Jerry Sadock. 1975. Ambiguity tests and how to fail them. *Syntax and semantics* 4(1): 1–36.

Putting the Pragmatics of Belief to Work

Igor Douven

Abstract In earlier work, I argued that pragmatic considerations similar to those that Grice has shown to pertain to assertability pertain also to the epistemic notion of acceptability; in addition to the pragmatics of assertion, there is what I called “the pragmatics of belief.” In this paper, the pragmatics of belief is applied to a puzzle related to the so-called Lottery Paradox and is shown to help resolve that puzzle.

Consider these sentences:

- 1 (a) Joan was hungry, but they only had hamburgers.
(b) They only had hamburgers, but Joan was hungry.

Someone who believes (1a) is unlikely to believe (1b), and vice versa. Equally, someone who believes one of (2a) and (2b) will typically be unwilling to believe the other:

- 2 (a) Harry attended the workshop and therefore Fran attended as well.
(b) Although Harry attended the workshop, Fran attended as well.

This is so despite the fact that, according to standard semantics, the members of these pairs have the same truth conditions.

This observation might give reason to question the correctness of standard semantics. In earlier work (Douven 2010), I argued that it rather motivates a rethinking of some core epistemological ideas. Specifically, I argued that our beliefs are not just sets of truth conditions, but are individuated finely enough to reflect what are traditionally regarded as pragmatic differences between the members of the above pairs of sentences; beliefs might just be sentences, in a language of thought or even, as some think, in one’s native language. I further argued that, as a consequence, pragmatic factors also play a role in determining

I. Douven (✉)

Department of Philosophy, University of Groningen, Groningen, The Netherlands
e-mail: i.e.j.douven@rug.nl

what we find acceptable, and not just in determining what we find assertable. The second claim I dubbed “the pragmatics of belief.”

In the aforementioned paper, I presented the pragmatics of belief in some detail, pointing to several close parallels with Gricean pragmatics, but also to some noteworthy dissimilarities. The proposal was backed by noting the phenomenological point that we do not tend to find the members of pairs of sentences like (1a) and (1b) jointly credible at the same time—as we did above—but also, and more importantly, by normative epistemological considerations. Here, I want to extend my case for the pragmatics of belief by showing that the theory can do important theoretical work. I already made a beginning with this in Douven (2012a), in which I argued that appeal to the pragmatics of belief yields a solution to Kyburg’s (1961) vexed Lottery Paradox, a solution that is both non-ad hoc and does justice to our intuitions about the connection between high probability and categorical belief.¹ Building on that result, the current paper puts the pragmatics of belief to further use in solving an epistemological puzzle, originally presented in Douven (2008b), which seems to show that the currently dominant type of solution to the Lottery Paradox (to which the aforementioned solution also belongs) is at odds with standard formulations of our epistemic goal.

In the first part of this paper, I summarize the normative considerations underlying the idea that pragmatic considerations also pertain to the notion of acceptability. In the second part of the paper, I rehearse the epistemological puzzle that I believe the pragmatics of belief can be instrumental in solving. The third and final part then presents the solution to that puzzle.

1 Pragmatics: From Assertion to Acceptance

It is no exaggeration to say that Gricean pragmatics is one of the major achievements of twentieth-century philosophy. Grice’s core insight was simple but powerful, to wit, that by asserting a sentence we may convey more than just the semantic content of the sentence. Supposing a speaker wants to be helpful—which, Grice argued, is a default assumption in all communication—the fact that he or she asserts what he or she asserts in the circumstances in which he or she asserts it may generate implicatures that go beyond the semantic content of the asserted sentence. Because true sentences may, when asserted, generate false implicatures, true sentences may be unassertable. This explains a host of linguistic phenomena that appear puzzling from a purely semantic perspective.

¹ I also appealed to the idea of a pragmatics of belief in defending a particular account of the assertability/acceptability conditions for indicative conditionals, in Douven (2008a); see also Douven (2011). That may give additional indirect support to the pragmatics of belief hypothesis, at least for those who find the designated account of conditionals attractive. For yet another application of that hypothesis, see Capone (2012), where the hypothesis is deployed in the course of defending a radically pragmatic account of quotation.

In expounding his pragmatics, Grice's focus is on linguistic communication. Still, in his (1989) he does remark, even if only in passing, that much of what he has to say about linguistic communication pertains to cooperative behavior more generally. As he notes (p. 28), just as we expect a speaker to be relevant in his or her contributions to a conversation, we expect someone helping us make a cake to hand us ingredients rather than a good novel. However, Harman (1999) is to be credited for being the first to make the more specific observation that much of what Grice has to say about linguistic communication may also pertain to *thought*. For example, Harman notes that “[o]ne will normally want descriptions used in reasoning to relate events in an orderly way ...” (p. 230), just as we want to relate events in an orderly way in communicating them to others. Harman does not elaborate on this observation. I have tried to do so in my (2010), among other ways by suggesting an epistemological rationale for the occurrence of broadly Gricean phenomena in thought.

To see what the rationale is, suppose that presently I know that Harry is very fond of Fran. However, human memory is not fully dependable; often enough, we forget things we know. Finding among my beliefs one to the effect that, although Harry attended the workshop, Fran did so too, and having forgotten that Harry and Fran are good friends, I might come to think that the contrast suggested by the phrasing of my belief is real, and that, for instance, Fran dislikes Harry. After all—I may consider—I must have had some reason for storing this belief in its specific wording! The simple way to forestall making such a mistake is to accept (2a) but refrain from accepting (2b). In short, it is not only that Gricean considerations effectively do play a role in determining what we find acceptable; they *ought* to play a role in it.

It seems not too much of a simplification to think of storing and retrieving beliefs in retentive memory as a sort of communication between different selves, existing at different points in time; your current self puts beliefs in a belief box from which they may be picked up by your future selves. In this picture, the foregoing proposal is rather straightforwardly suggested by Grice's considerations about linguistic communication. Just as we assume our interlocutors to be cooperative when we interpret their assertions, we assume our former selves to have been cooperative when we retrieve a belief from retentive memory. In my (2010), this thought led me to propose the following principle:

Epistemic Hygienics (EH): Do not adopt as beliefs things that could mislead your future selves.

Needless to say, (EH) is meant to be analogous to the core recommendation of Gricean pragmatics not to assert things that might mislead our conversation partners.

It is to be noted that (EH) is only a slogan. As it stands, it might suggest that there is precious little, if anything, that we can adopt as a belief. For is not any belief *potentially* misleading? The response to this worry is that (EH) is to be read as stating that, while any belief is potentially misleading, we should not *increase* the risk of misleading a future self by adopting a belief that we now know, or have

excellent reason to believe, suggests something false, especially if the additional risk can be easily avoided at no additional cost. Notice that the recommendation not to assert things that might mislead one's conversation partners should be interpreted in much the same way, given that there is little that one can assert that is not *potentially* misleading in some way or other.

As mentioned in the introduction, the parallel between standard Gricean pragmatics and the pragmatics of belief is not perfect. For later purposes, one disanalogy in particular is worth mentioning. If one knows the bookshop on Elm Street to be closed, it would be misleading to assert

(3) There is a bookshop on Elm Street which sells maps.

in response to a passer-by's question where she can buy a map of the city. For in that situation, asserting (3) would generate the implicature that the shop is open. On the other hand, having (3) among one's beliefs is unlikely to put one at risk of inferring that the shop is open at times when it is not. Thus, even if one knows that the shop is currently closed, (EH) does not counsel against adopting (3) as a belief. Contrast (3) with (4):

(4) On May 1, 2011, Susan tried to swim a mile.

If one knows that, on May 1, 2011, Susan not only *tried* to swim a mile but actually *succeeded* in her attempt, then (4) is not just unassertable, it is also unacceptable. By committing the sentence to retentive memory, one runs the risk of misleading one's future selves who may have forgotten that Susan's attempt was successful. They may infer from (20) that Susan's attempt failed. (This is a risk that, again, is easily avoided by adopting as a belief that Susan swam a mile on May 1, 2011, but not (4).)²

² In Douven (2010), I conjectured that the just-described disanalogy between Gricean pragmatics and the pragmatics of belief may be explained in terms of Grice's typology of implicatures, which distinguishes *conventional* from *conversational* implicatures, and which further divides conversational implicatures into *generalized* and *particularized* conversational implicatures (or GCIs and PCIs, for short). As is well known, conventional implicatures arise due to the conventional meaning of certain words, whereas conversational implicatures derive from the fact that the speaker says what he or she says, assuming that he or she intends to be helpful. GCIs are implicatures that arises normally, or by default, *barring* special contextual circumstances, while PCIs arise only *in virtue of* such special circumstances. The suggestion made in Douven (2010) was that while a sentence may be unassertable because it has a false implicature, of whichever of the foregoing sorts, for the issue of acceptance only conventional implicatures and GCIs matter. After all—the thought went—when we store a belief, we realize that a future self who retrieves the belief will in general be unable to recall the specificities of the context in which we happen to be at the moment of storing the belief; and when or if we retrieve a belief, we will assume that the former self who stored it will not have assumed that we are able to recall the specificities of the context that self was in when the belief was stored (even if we sometimes *will* be able to recall those specificities). As a result, when or if we retrieve a belief from memory, we will not be tempted to infer any PCI from it, be it one that an utterance of that belief might have in the particular context in which we happen to be, be it one that an utterance of that belief might have had in the particular context in which we happened to be when we stored it.

2 A Puzzle About Lotteries

Many find the following principle at least *prima facie* plausible:

High Probability Principle (HPP) High probability—specifically, probability above a certain threshold value θ close to 1—is enough for rational credibility.

However, if it is assumed that rational credibility is closed under logical consequence, then (HPP) gives rise to the so-called Lottery Paradox. Take any fair n -ticket lottery that will have one winner and such that $1 - 1/n > \theta$. Then, prior to the drawing, it is rational to believe of each ticket in this lottery that it will lose, by (HPP). By the closure assumption, it is rational to believe that all tickets will lose. But this contradicts the publicly known fact that there will be a winning ticket.

For Kyburg, this paradox was a reason to reject the closure assumption. But few have concurred, and currently almost all philosophers take the Lottery Paradox to show that one cannot rationally believe of any ticket in a fair lottery that it will lose and thus that (HPP) cannot be maintained as it stands. In Douven (2008b), I challenged this near-orthodoxy by arguing that (i) from any theory of rational credibility that solves the Lottery Paradox by (explicitly or implicitly) ruling no lottery proposition to be rationally credible, a theory not of that kind can be derived that also solves the paradox but is more conducive than the former to our epistemic goal (as commonly conceived), and (ii) for all anyone has shown so far, the latter is thereby preferable to the former.³

At the time of writing the aforementioned paper, my sympathies were with the near-orthodoxy; it has always seemed intuitively wrong to me that I could rationally believe of any given ticket that it will lose if the drawing of the relevant lottery is still to occur (or the result still to be announced). Still, at that time I could not see what was wrong with the argument leading to the (in my view) counter-intuitive conclusion. I thus thought of the paper as presenting us with a puzzle about lotteries and our epistemic goal, rather than as making a positive case for a particular solution to the Lottery Paradox. In hindsight, I would say that I overlooked the epistemic role of pragmatic factors, as summarized in the previous section. At least that is what I intend to show in Sect. 3: that the pragmatics of belief helps to resolve the puzzle by suggesting a slight but crucial reformulation of our epistemic goal.

In the present section, I want to briefly recap the argument for the puzzling conclusion. The argument has two presuppositions: first, what is arguably the standard conception of our epistemic goal, and second, a seemingly uncontroversial principle connecting our epistemic goal to theories of rational credibility.

The former is canonically expressed in terms of truth maximization and falsity minimization, but the precise formulation is of lesser concern. For the purposes of

³ Here and elsewhere, by “lottery proposition” I mean a proposition to the effect that a given ticket will lose, where the ticket is part of a large, fair lottery whose drawing is yet to occur and that will have one winner.

the argument, Alston's (1985: 59) claim that our epistemic goal is "to amass a large body of beliefs with a favorable truth–falsity ratio" does as well as Lehrer's (1974: 202) formulation, according to which we ought to aim "at both believing only what is true and at believing all that is true."⁴ These formulations are equally good because, while strictly speaking they state somewhat different conceptions of our epistemic goal, *practically* speaking, they amount to the same thing: the best that ordinary mortals can achieve in aspiring to believe only what is true and all that is true, is to have a large body of beliefs with relatively few false ones, that is, to maximize the number of their true beliefs while trying to avoid picking up too many false beliefs in the process—and that, as will be seen shortly, is what matters to the argument.⁵

The other presupposition is this:

Connecting Principle (CP) If one theory of rational credibility, T , is more conducive to the realization of our epistemic goal than another theory of rational credibility, T' , then T is to be preferred to T' when all else is equal.

I will not say anything about this principle here apart from noting how exceedingly weak it is. Not only does (CP) refrain from making any absolute claims about the relationship between theories of rational credibility and our epistemic goal; it also contains a *ceteris paribus* clause, leaving open the possibility that, when it comes to adjudicating between theories of rational credibility, conduciveness to our epistemic goal is trumped by other properties such theories may have.⁶

To set the stage, we consider a solution to the Lottery Paradox suggested by Harman (1986: 70 f) that is different from both the Kyburg-type solution that denies closure for rational credibility and the currently dominant type of solution that rules lottery propositions to be not rationally credible. Harman says (p. 71):

To say one can infer ... of any ticket [that it will not be the winning ticket] is not to say one can infer it of all. Given that one has inferred ticket number 1 will not win, then one must suppose the odds against ticket number 2 are no longer 999,999 to 1, but only

⁴ Bonjour (1985: 8) and Foley (1992: 183) present formulations in the same vein as Lehrer's. Note that by "beliefs," Alston must mean "explicit beliefs." If believing ϕ were to imply believing $\phi \wedge \phi$, believing $(\phi \wedge \phi) \wedge \phi$, etc., as it plausibly would if the notion of belief at issue were that of *implicit* belief, then the idea of aiming at a set of beliefs with a favorable truth–falsity ratio would not really make sense: supposing one has at least one true belief and one false belief, one would automatically have infinitely many true beliefs and also infinitely many false ones. By contrast, it is reasonable to think that the notion of belief at issue in Lehrer's formulation of our epistemic goal is that of *implicit* belief. As another comment on Alston's formulation, I note that, as it stands, it might suggest that, should we ever have attained a large body of beliefs with a favorable truth–falsity ratio, we can "epistemically retire" in the sense that we can legitimately forego any opportunity to extend our set of true beliefs; after all, we have already achieved our epistemic goal! Clearly, though, what Alston will have meant to say is that we should try to acquire an ever larger set of beliefs with a favorable truth–falsity ratio.

⁵ By "practically speaking" I also mean *realistically* speaking: it may be possible to come up with a scenario in which the different formulations of the epistemic goal do not amount to the same thing, but I strongly doubt that such a scenario could sound anything close to realistic.

⁶ For more on (CP), see Douven (2008b), Sect. 3.

999,998 to 1. And after one infers ticket number 2 won't win, one must change the odds on ticket number 3 to 999,997 to 1, and so on. If one could get to ticket number 999,999, one would have to suppose the odds were even, 1 to 1, so at that point the hypothesis that this ticket will not win would be no better than the hypothesis that it will win, and one could infer no further. (Presumably one would have to have stopped before this point.)

This solution allows one to adhere to (HPP) provided one is willing to abandon the principle that rational credibility is closed under logical consequence. For given (HPP), one could judge it rationally credible that ticket number 1 will lose *without thereupon adopting as a belief that the ticket will lose*, so that the odds against ticket number 2, or any other ticket, would not change. More generally, one could judge rationally credible, though not adopt as a belief, that ticket number *i* will lose, for all tickets *i* in the lottery. If rational credibility were then closed under logical consequence, one could rationally believe that no ticket will win. One would still face a contradiction with one's background knowledge.

But perhaps abandoning the closure principle in the foresaid form is not much of a sacrifice, given that one could replace it with the principle that what we *rationally believe* (as opposed to what we *can* rationally believe, whether or not we believe it) is closed under logical consequence, or better perhaps, the principle that if we rationally believe each $\gamma \in \Gamma$, and Γ entails ϕ , then it is rational for us to believe ϕ .⁷ And there may be no good reason why, if one of these principles is available, one should want to keep to the principle that rational credibility is closed under logical consequence.⁸

The important point to notice now is that a theory of rational credibility that avoids the Lottery Paradox along the above lines does better *prima facie* from the viewpoint of achieving our epistemic goal than any theory that solves the paradox by prohibiting us from believing of even a single ticket of any given lottery that it is a loser (call theories of the latter kind "NRC theories," for "Not Rationally Credible"). To spell this out, suppose that $t = 0.999$. Then, proceeding as Harman suggests, we can come to justifiedly believe of 999,000 tickets of a one million ticket lottery that they will lose. The chance that we would thereby have added to our body of beliefs *only* truths is not so big; it is, to be precise,

$$\frac{\text{ways to select 999,000 tickets from 999,999 losers}}{\text{ways to select 999,000 tickets from all tickets in the lottery}} = \frac{\binom{999,999}{999,000}}{\binom{1,000,000}{999,000}} = \frac{1}{1000}.$$

⁷ The former principle may well be false as a matter of empirical fact, though much will depend here on how one views the explicit-implicit belief distinction, a matter that need not detain us here.

⁸ This gives Harman's solution a clear advantage over Kyburg's, which allows one to believe of each ticket of a given lottery that it will lose and thus must reject even the principle that what is rationally believed is closed under logical consequence.

However, with a chance of 1 we would have added *at least* 998,999 truths to our body of beliefs, and *at most* one falsehood. One would think that 998,999 : 1 is a pretty favorable truth–falsity ratio!

The point swiftly generalizes to other theories of rational credibility that block the Lottery Paradox by allowing us to believe of some, but not all, tickets that they are losers (call theories of this kind “RC theories”). In contrast to the NRC accounts, all RC accounts allow us to benefit from the opportunity to improve the truth–falsity ratio of our belief set that a “sufficiently large” lottery gives us; they allow us to benefit from such an opportunity to an extent proportional to the number of tickets of which they allow us to believe that they are losers.^{9,10}

It does not follow that any RC theory of rational credibility is to be preferred to every NRC theory, nor even that some RC theory must be preferable to every NRC theory. This is so for two reasons. First, nothing said so far excludes the possibility that, all things considered, an RC theory makes the achievement of our epistemic goal harder than any NRC theory does. The truths that an RC theory allows us to accept and an NRC theory does not will, after all, form but a small minority among the propositions that are candidates for acceptance. And who knows how many other truths are being discounted by an RC theory and are ruled rationally credible by (at least some) NRC theories, or how many falsehoods the (or at least some) NRC theories rule to be not rationally credible that RC theories allow us to accept?¹¹ Second, neither can it be excluded that RC theories do worse than NRC theories in ways relevant in the sense of (CP). They might be preferable, all else being equal; but perhaps not all else is equal, or even can be equal. For instance, while blocking the Lottery Paradox, they might engender new paradoxes.¹²

⁹ Up to all save one of the tickets, that is. An account that allows us to believe of all tickets that they are losers while sticking to the principle that rational credibility is closed under logical consequence (or to one of the similar principles mentioned in the text above) may lead us into inconsistency, and thus lead to our believing not only all truths but also all falsehoods. (Though here, too, one’s view on the explicit–implicit belief distinction will be relevant. An inconsistent belief state will not explicitly comprise all falsehoods (nor all truths). But depending on how one defines the notion of implicit belief, it may comprise them implicitly.)

¹⁰ Some might want to object that, because by believing of some tickets that they will lose, we run the risk of adopting a false belief, it is not evident that doing so serves our epistemic goal, at least not if this is taken to imply that we ought to believe *only* true beliefs, as on Lehrer’s conception it does. However, recall the point made earlier that we cannot hope to achieve the goal of believing all that is true and only what is true without coming to hold some false beliefs along the way. Indeed, the risk that this will happen may be presumed to be much greater in general than in the specific case of a lottery of some of whose tickets we believe that they will lose. As intimated earlier, from a practical viewpoint there is no difference between the epistemic goal as formulated by Alston and the epistemic goal as formulated by Lehrer.

¹¹ In the literature, “acceptance” is often used as a more or less technical term to indicate a propositional attitude different from belief; see, e.g., van Fraassen (1980). Here, it is used to mean the act of beginning to believe something, of adopting the thing as a belief.

¹² Indeed, in Douven (2012b) I showed that Harman’s solution to the Lottery Paradox gives rise to a new paradox—“the Sequential Lottery Paradox,” as I called it.

However, in response to the first of the above points—the point that an RC theory may *overall* still do worse when it comes to helping us achieve our epistemic goal—we may note that, for any NRC theory of rational credibility T , we can construct an account T' by simply adding to T a clause to the effect that, given a sufficiently large, fair lottery with one winner, it is rational to believe of all but one of the tickets that they will lose (which, patently, does not amount to sanctioning (HPP)). As far as I can see, such a “primed” theory will not discount any true proposition unrelated to lotteries that is not also discounted by the original theory without the additional clause.¹³ Might it countenance falsehoods that are not countenanced by an NRC theory? It might. If you believe falsely of your own ticket that it will lose and rational credibility is closed under logical consequence, then a primed theory may well render rationally credible many falsehoods that the NRC theory from which it was constructed does not render rationally credible. For instance, it may render rationally credible that you will be as poor in the coming year as you were in the past year, that you will not anytime soon be in a position to go on a long vacation or to quit your current job, and many other propositions that we may assume to be false if your ticket is the winner. However, it does not follow that you will end up adopting a collection of beliefs whose truth–falsity ratio is far from ideal, and at any rate worse than that of the collection of beliefs the NRC theory would have allowed you to adopt. After all, for every falsehood you can derive from the false proposition that your ticket will lose, you will be able to derive a truth from the proposition that ticket number i will lose, for each i such that i is unequal to the number of your ticket and you believe that ticket number i will lose—such as the truth that either the owner of the ticket is, or will soon become, rich independent of the outcome of the lottery or that he or she will not be rich anytime soon. Thus, even though a primed theory may countenance falsehoods that the NRC theory from which it comes does not countenance, the *ratio* between the truths and the falsehoods that you can rationally believe given the former theory still seems to be more favorable than that between the truths and the falsehoods that you can rationally believe given the latter.

This leaves us with the second point, the point that while RC theories may be more conducive to our epistemic goal than NRC theories, the latter may still be preferable, or at least some NRC theory may still be preferable, because they do (it does) better in some other respect that is relevant in the sense of (CP). As to this, recall that the argument is purported to present a challenge to, not to refute, the NRC theories. Specifically, the challenge is, as we can now see, to produce differences between the NRC theories and their primed variants that could be held to

¹³ Perhaps it is reasonable to suppose that, according to any RC theory, but according to no NRC theory, we cannot rationally believe that the right solution to the Lottery Paradox entails that we cannot rationally believe any lottery proposition. And perhaps that claim could be said to be related to the Lottery Paradox but not really to lotteries. Now, I am about to argue that the claim is true. So then there *may* be truths that are discounted by a primed theory that are not discounted by the NRC theory from which it is derived. However, to hold this against the primed accounts at the present juncture would manifestly beg the issue.

trump conduciveness to our epistemic goal in a comparison between an NRC theory and its primed counterpart.

In my (2008b), I did more than present this challenge; I sought to show that the challenge is not easily met. For instance, I claimed that it would be wrong to reject the primed RC theories for the reason that they are ad hoc, the way they modify NRC theories being motivated solely by the wish to solve the Lottery Paradox in a manner that makes the RC theories more conducive to our epistemic goal than the NRC theories from which we obtain them. This would be wrong already because it is presently far from obvious that there are any non-ad hoc solutions to the Lottery Paradox, and hence that there are any NRC theories that could be claimed to be non-ad hoc.¹⁴ More importantly—I argued in that paper—even granting that some extant NRC theories do provide non-ad hoc solutions to the Lottery Paradox, it is hard to see why, in applying (CP), non-ad hocness should weigh more heavily than, or even as heavily as, conduciveness to our epistemic goal. Furthermore, I pointed to the fact that the primed RC theories require us to make random epistemic choices—arguably, we shall have to pick randomly which tickets to believe will lose—in contrast to, we may assume, their NRC counterparts. But I also argued that there is currently no reason to believe that the primed RC theories' reliance on random choices should, in a comparison with their counterparts, outweigh their greater conduciveness to our epistemic goal. I thus concluded that, as far as we presently know, the primed RC accounts are preferable to the NRC originals.^{15,16}

As already mentioned, I find this conclusion hard to accept, and it has occurred to me that so do many others. Still, I have not heard anyone specify reasons for his

¹⁴ Pace Nelkin (2000); see Douven (2003).

¹⁵ In personal communication, Jonathan Weisberg pointed out to me that the issue of random choice is actually inessential to the discussion, given that proponents of NRC accounts will surely want to treat slightly biased lotteries the same as perfectly fair lotteries: even if some ticket has a slightly higher chance of winning than the others, they would still want to deny that one can be justified in believing that it is the winner and all the others are losers. However, in that case one could pick a ticket for believing of it that it will win in an arguably nonarbitrary way.

¹⁶ Some might have qualms about the doxastic voluntarism that the primed accounts presuppose. But note that the kind of voluntarism at issue is of a quite modest sort. The primed accounts do not require that all or at least most belief formation is under the control of our will, nor do they require one to believe something in the absence of any reasons for believing it (unless the NRC accounts from which they come already do so). For consider that the high probability that a ticket will lose certainly gives some reason to believe that it will lose. This is worth emphasizing, as critics of doxastic voluntarism do not always distinguish between the thesis that there can be an element of decision in believing, even if perhaps only for restricted classes of propositions or under particular circumstances, and the thesis that one can come to believe just about anything one likes. And I believe, when sensibly understood, doxastic voluntarism to be in a better shape than many epistemologists seem to suppose; see in particular Ginet (2001) for the presentation of what strikes me as a particularly compelling argument in favor of the position. For those unmoved by the foregoing remarks, let me add that, as here I aim to undercut the argument given in Douven (2008b) and summarized in this section, it will not hurt my present case to suppose as much as possible on behalf of that argument.

or her discontent that went beyond a blanket appeal to intuition, and I do not see how intuition alone could justify us in dismissing the RC accounts. In the following section, I hope to buttress the intuition that something is amiss with the above conclusion by appealing to the pragmatics of belief.

3 The Puzzle Solved

In Douven (2012a), I presented a new solution to the Lottery Paradox. The solution was of the NRC type, and in line with currently common approaches to the paradox that rule lottery propositions not rationally credible by adding a defeater to (HPP), that is, by declaring high probability to be *defeasibly* sufficient for rational credibility. As a defeater, I proposed that adopting as a belief the proposition whose credibility is at stake should not lead to a violation of (EH). Most of the work in that paper then went into showing that the assertion of a lottery proposition generates a false implicature—to wit, that the asserter has insider knowledge of the outcome of whichever lottery the proposition (implicitly) refers to—and that the implicature is of a type relevant to (EH); that is, that lottery propositions are, in the relevant sense, like (4) and unlike (3). Thus, lottery propositions are not rationally credible, despite their high probability of truth, and despite the fact that *normally* high probability suffices for rational credibility. I showed that this solution to the Lottery Paradox escapes Douven and Williamson (2006) critique of virtually all previous solutions to the Lottery Paradox.

In the following, I shall not rehearse my argument for the claim that by adopting a lottery proposition one would violate (EH); the interested reader is referred to (Douven 2012a, Sect. 1). What I want to argue here is that the claim by itself is not enough to solve the puzzle about lotteries from the previous section.

This might sound surprising. For—it might be thought—if lottery propositions generate false implicatures of the sort relevant to (EH), then by accepting of any tickets of a fair lottery with one winner that they will lose, one would be violating (EH); and that the NRC accounts of rational credibility do not countenance accepting lottery propositions anyway, whereas RC accounts (including the primed ones) do countenance this, makes for a relevant difference in the sense of (CP), which was what we were looking for.

But this is too quick, for it is by no means clear that violations of (EH) outweigh conduciveness to our epistemic goal. Compare: According to Grice's maxim of Quality, we should not say what we believe to be false. But, as has been frequently observed, this maxim can be overruled on many grounds. For instance, if you can save your life by lying, it is excusable to neglect the maxim of Quality. And, for all I have said in the previous section, (EH) is not an unexceptionable principle either.

Luckily, we do not have to decide which is to prevail if (EH) conflicts with our interest in achieving our epistemic goal. For, as I will now argue, the idea of a pragmatics of belief suggests a revision of our epistemic goal that makes such conflicts impossible.

If there is no contrast between the conjuncts of (2a) and (2b), then, we said, for reasons of epistemic hygienics it is best *not* to believe (2b). Thus, it appears, we should *not* believe all that is true, nor should we improve the truth–falsity ratio of our body of beliefs by adopting the—as per our supposition—true (2b), else we would be quite openly running the risk of coming to hold *false* beliefs, like for instance the false belief that Harry’s attending the workshop made it somehow surprising that Fran attended as well. (As already noted in section, if we try to achieve our epistemic goal, we inevitably risk adopting some false beliefs along the way. But we presumably should not run a risk if there is no point in doing so. And clearly there is nothing to be gained by believing (2b), at least not if we are free to believe instead the equally true but non-misleading sentence (2a).)

One might object that our epistemic goal, in either of the versions stated in Section 2, postulates an epistemic *ideal* to which our cognitive deficiencies, and whatever strategies we may want to implement to counteract these deficiencies, are orthogonal. It is a bit as with logic, it might be said; logic has normative force even though no ordinary mortal may be sufficiently equipped to follow the laws of logic unflinchingly.

The analogy is misleading, however. Derivational systems have been devised that can be successfully taught to freshmen in only one semester and that can help us to obey the laws of logic whenever we care to do so. Nothing similar seems to exist that could help us overcome the deficiencies that were said to motivate the idea of a pragmatics of belief. We can train our memories, and try as hard as possible not to forget things, but full success in these respects is illusory, as we know too well. Therefore, we would seem ill advised to pursue the epistemic goal as traditionally conceived.

This conclusion might seem to confront us with a new puzzle. For—one wonders—how could the community of epistemologists, almost in its entirety, be so mistaken about what our epistemic goal is? There may be an easy explanation for this, though, for I think that something at least superficially similar to the standard formulations of our epistemic goal is still correct. To explain what that is, I must first introduce the notion of *rightness*. Take the word “but,” which is commonly said to signal a contrast between conjuncts.¹⁷ Could we not build this contrast-signalling aspect into the truth table for “but,” perhaps as follows (Tr is the truth predicate)?

Tr(φ but ψ) iff

- (i) Tr(φ);
- (ii) Tr(ψ); and
- (iii) there is a contrast between φ and ψ .

¹⁷ Whether this is correct is immaterial; “but” merely serves for purposes of illustration here. (Incidentally, examples like, “He walks slowly, but he walks,” suggest that it may *not* be correct to say that “but” signals a contrast, at least not in general.)

It is generally said to be a problem for this idea that we do not comfortably declare a sentence of the form “ φ but ψ ” false if we believe (i) and (ii), but not (iii), to hold of it.¹⁸ On the other hand, we will have no difficulty admitting that, for instance, sentence (2b) is somehow wrong if there is no contrast between its conjuncts. I submit that the notion of wrongness involved in this verdict is at least approximately captured by defining a sentence to be *right*, and *wrong* otherwise, iff the sentence is true and in addition does not carry any misleading implicatures of the type relevant to (EH). So, for instance, if φ reports an event that occurred before the event reported by ψ , “ ψ and φ ” is wrong even if φ and ψ are both true; the same for “ φ because ψ ” if the truth of φ has got nothing to do with that of ψ .

Equipped with this notion of rightness, the plausible way to reformulate our epistemic goal would seem to go something like this:

Reformulated Epistemic Goal (REG) We should aim at believing only what is right and at believing all that is right.

I think that (REG) sounds similar enough to the better-known formulations of our epistemic goal to make it understandable how, especially in the absence of an articulated distinction between truth and rightness, those formulations could attain their current popularity.¹⁹

The simple but crucial point to note now is that the primed RC accounts of rational credibility are *not* more conducive to this newly formulated epistemic goal than the original NRC accounts from which they come. It is true that the former allow us to believe more *truths* than the latter, but truth is not all that matters to (REG); rightness also involves pragmatic factors, and lottery propositions, whether true or not, are not right, given that, we said, they carry the false implicature that the speaker has inside information about the outcome of the lottery. We thus can, in accordance with intuition, maintain that lottery propositions are not rationally credible. Note that it would be wrong to say that we have herewith *met* the challenge from Section 2. Rather, that challenge has been *undermined*; with (REG) in place of the standard conceptions of our epistemic goal, the challenge simply does not get off the ground.²⁰

References

- Alston, W. 1985. Concepts of Epistemic Justification. *Monist* 68: 57–89.
 Bonjour, L. 1985. *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.

¹⁸ See, e.g., Jackson (1979: 123).

¹⁹ The reformulation of Alston’s version of the epistemic goal would obviously read: “We should aim at amassing a large body of beliefs with a favorable rightness–wrongness ratio.”

²⁰ I am grateful to Jake Chandler, Richard Dietz, and Jonathan Weisberg for valuable comments on a previous version of this paper. Thanks also to Alessandro Capone for a helpful discussion of the pragmatics of belief.

- Capone, A. 2012 The Pragmatics of Quotation, Explicatures and Modularity of Mind, manuscript.
- Douven, I. 2003. Nelkin on the Lottery Paradox. *Philosophical Review* 112: 395–404.
- Douven, I. 2008a. The Evidential Support Theory of Conditionals. *Synthese* 164: 19–44.
- Douven, I. 2008b. The Lottery Paradox and Our Epistemic Goal. *Pacific Philosophical Quarterly* 89: 204–225.
- Douven, I. 2010. The Pragmatics of Belief. *Journal of Pragmatics* 42: 35–47.
- Douven, I. 2011. Indicative Conditionals. In *A Companion to Philosophical Logic*, eds. L. Horsten, and R. Pettigrew, 383–405. London: Continuum Press.
- Douven, I. 2012a The Lottery Paradox and the Pragmatics of Belief. *Dialectica* 66: 351–373.
- Douven, I. 2012b The Sequential Lottery Paradox. *Analysis* 72: 55–57.
- Douven, I., and T. Williamson. 2006. Generalizing the Lottery Paradox. *British Journal for the Philosophy of Science* 57: 755–779.
- Ginet, C. 2001. Deciding to Believe. In *Knowledge, Truth, and Duty*, ed. M. Steup, 63–76. Oxford: Oxford University Press.
- Grice, H.P. 1989. Logic and Conversation. In his *Studies in the Way of Words*, 22–40. Cambridge MA: Harvard University Press.
- Harman, G. 1986. *Change in View*. Cambridge MA: MIT Press.
- Harman, G. 1999. (Nonsolipsistic) Conceptual Role Semantics. In his *Reasoning, Meaning, and Mind*, 206–231. Oxford: Oxford University Press.
- Kyburg, H. 1961. *Probability and the Logic of Rational Belief*. Middletown CT: Wesleyan University Press.
- Lehrer, K. 1974. *Knowledge*. Oxford: Oxford University Press.
- Nelkin, D. 2000. The Lottery Paradox, Knowledge, and Rationality. *Philosophical Review* 109: 373–409.
- van Fraassen, B.C. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Contexts, Fiction, and Truth

Alberto Voltolini

Abstract In this chapter I want to hold that contextualism—the position according to which wide context, i.e., the concrete situation of discourse, may well have the semantic role of assigning truth-conditions to sentences in context—may well accommodate (along with some nowadays established theses about the semantics of proper names) three data about fiction, namely, the facts that as far as discourse involving fiction is concerned, (i) sentences about nothing are meaningful (ii) they may be true in fiction (iii) yet they may also be true in reality. Moreover, I also want to hold that such a contextualist accommodation does not depend on endorsing a realist stance about fictional entities, although such a stance gives the best truth-conditional explanation for such data, in particular (iii).

Keywords Traditionalism · Contextualism · Narrow context · Wide context · Fictional context · Real context · Fictional truth · Real truth · Fictional entities

1 Three Data About Fiction and the Traditionalism-Contextualism Debate

Discourse involving fiction presents us at least three interesting data we have to account for. Firstly, as Frege (1892) originally envisaged, when we are uttering sentences like:

(1) Werther is fallen in love with Charlotte

with reference to Goethe's tale of *The Sorrows of Young Werther*—typically, as an engaged comment on what's going on in the tale—by means of the proper names “Werther” and “Charlotte” we are talking about nothing at all—there really are no

A. Voltolini (✉)
University of Turin, Turin, Italy
e-mail: alberto.voltolini@unito.it

things such as Werther and Charlotte, as one might say—and yet we are not uttering something utterly meaningless, unlike the famous Carnapian sentence:

(2) Pirots carulize elatically

whose meaninglessness depends on the meaninglessness of its constituents, or even unlike the famous Chomskyan sentence:

(3) Colourless green ideas sleep furiously

which makes no sense as a whole, although its constituents are meaningful.

Secondly (and relatedly), although our previous utterance of (1) is not really endowed with a truth-value, its meaningfulness is somehow related with truthfulness. For, as Walton (1990) forcefully pointed out, such a comment is true in a game of make-believe we play with Goethe's tale. In such a game, we indeed tell that utterance of (1) from the utterance of its contradictory:

(4) Werther is not fallen in love with Charlotte

which is false in such a game.

Thirdly and finally, there are even cases in which when we utter sentences like (1), we utter something which is really endowed with a truth-value. For, if we utter (1) completely outside of any make-believe game, we say something really true. In this respect, if while being examined on Goethe's contribution to German literature instead of uttering (1) we uttered:

(5) Werther is fallen in love with Charlotte's husband

we would say something really false so as to fail the exam—we would be in the same predicament as if we uttered:

(6) $5 + 7 = 13$

in a test on elementary arithmetics.

As regards the distinction between semantics and pragmatics, two main options nowadays face each other: *traditionalism* and *contextualism*. As a matter of fact, both “traditionalism” and “contextualism” are a sort of umbrella terms that cover different positions.¹ Yet a way to cash out this very general distinction concerns the different conceptions those options have with respect to the so-called *wide context*, i.e., the concrete situation of discourse.

To put it in Perry's (1997) terms, according to traditionalists, wide context never has a *semantic* role, that is, it never fixes the truth-condition a sentence has (in context); typically, for them wide context has the *post-semantic*, genuinely pragmatic, role of determining whatever extra-truthconditional factor of significance is conveyed by a contextual use of a sentence. The only context that for a

¹ For instance, both so-called *semantic minimalism*—cf. e.g. Cappelen and Lepore (2005)—and a certain kind of *indexicalism*—e.g. the one defended by Predelli (2005)—fall under traditionalism, while both a *moderate* form of contextualism, i.e., *indexicalism à la Stanley* (2007) and a *radical* form of contextualism *à la Recanati* (2004a), are forms of contextualism.

traditionalist has a semantic role is *narrow context*, i.e., a set of a limited series of parameters—typically, agent, space, time, and world. For narrow context enables sentences that by themselves lack truth-conditions to receive—automatically, as it were—a truth-conditional assignment. The paradigmatic case of a such a semantic situation concerns indexical sentences. Granted, without a narrow context an indexical sentence has no truth-conditions. Yet once one such context is provided, that sentence obtains a certain truth-condition; alternatively put, only utterances of indexical sentences, i.e., (roughly) indexical sentences in (narrow) context, have truth-conditions. Moreover, one such sentence in narrow context obtains such a truth-condition automatically. For that sentence also has a *linguistic meaning*, or better a *character*, which is completely determined by the linguistic meanings, or better the characters, of its sub-sentential elements. Character is a function from narrow contexts to truth-conditional contributions. When one such sub-sentential element is indexical, its character points precisely towards a certain parameter in narrow context as such a truth-conditional contribution—for instance, the contextual agent in the case of the indexical “I”. Thus, that parameter automatically determines the truth-conditional contribution that sub-sentential element gives is that narrow context to the indexical sentence in which it figures.²

Whereas according to contextualists, over and above narrow context, wide context *may* well have the semantic role in question. That is, not only indexical sentences, but all sentences of a natural language, may be affected by contextuality in the sense that once one such sentence is linked with a wide context it obtains a certain truth-condition, or, which is the same, what has that truth-condition is a certain utterance of that sentence, i.e., (roughly) such a sentence in (wide) context. Properly speaking, in order for the contextualist to win the battle against his enemy, it is enough for her to show that contextualism is correct with respect to the purportedly semantic role of wide context in at least one case.³

Theoretically speaking, the first two above data can be accommodated within a traditionalist framework. Yet to account also for the third datum, a contextualist perspective is required. Or so I will claim.

2 How Contextualism may Account for All Such Data

If the problem we are focusing here were only how to account for the first two data, a traditionalist may easily deal with it. Let us see.

² For this whole picture on indexicals, cf. Kaplan (1989a, b); see also Predelli (2005). Recanati (2004b: 1) claims that the character of an indexical is given by a token-reflexive rule. He frames this idea precisely by saying that such a character automatically obtains a truth-conditional contribution for that indexical in a certain narrow context by pointing to a given parameter in that context.

³ As Recanati (2004a: 116) claims.

Consider two narrow contexts that differ for their worlds, the real world and a fictional world respectively—to remain to our example, the world where the events corresponding to Goethe’s tale of *The Sorrows of Young Werther* unfold. Let us call the first a *real* and the second a *fictional* context. Consider moreover (1) and imagine it as being linked to those contexts respectively, so as to mobilize two different utterances of that sentence. We may imagine (1) as uttered by someone either independently of her being caught in the narration of Goethe’s tale or as an engaged audience of that tale. Finally, suppose that the so-called new theory of reference stemming out of the works of Donnellan (1966), Kripke (1980), and Putnam (1975) is correct, as is nowadays a commonplace. According to this theory, there are directly referential linguistic devices, i.e., terms exhausting their truth-conditional contribution in their referent.⁴ Proper names are paradigmatic samples of such terms. As a consequence, empty proper names, i.e., directly referential devices lacking a referent, provide no truth-conditional contribution to the sentences they constitute.

Now, in the domain of the real world, there are no such human beings such as Werther and Charlotte. Thus, the names “Werther” and “Charlotte” refer to no such beings in the real world. Given the new theory of reference, as uttered in a real narrow context, (1) fails to have truth-conditions at all, hence a truth-value: it has neither *real* truth-conditions nor a *real* truth-value.

Yet in the domain of the fictional world of the tale of *The Sorrows of Young Werther*, there are human beings such as Werther and Charlotte. In that tale, Werther is an unlucky man who has precisely fallen in a desperate love with a married woman, Charlotte. Thus, the names “Werther” and “Charlotte” do refer to such beings in that world. Given the new theory of reference, once (1) is uttered in the fictional narrow context that has *that* fictional world as its ‘world’-parameter, that utterance of (1) does possess a truth-condition, notably a *fictional* truth-condition. Indeed it is fictionally true, i.e., it is true in the context that has that fictional world as its ‘world’-parameter, iff (the man) Werther falls in love with (the woman) Charlotte. Since things in the world of Goethe’s tale unfold precisely in that way—Werther is so fallen in love—that utterance is also fictionally true; unlike the corresponding utterance of the contradictory (2), which is fictionally false.

So, the first two data can be accounted for by the traditionalist.⁵ Firstly, although there are no such human beings as Werther and Charlotte, an utterance of (1) is meaningful insofar as it is linked with a fictional narrow context so that in (1) in that context has a fictional truth-condition. In this respect, really empty names like “Werther” and “Charlotte” yet having a designation, hence, a truth-

⁴ For such an account of direct referentiality, cf. e.g. Recanati (1993) and Marti (1995).

⁵ Someone would add that, in order for such a traditionalist explanation to work, a sentence like (1) has to be considered a (hidden) indexical sentence, in order for truth-conditions to be assigned automatically to utterances of it. In such a case, it would suffice that proper names like “Werther” and “Charlotte” were taken to be indexicals of some sort. I have strong inclinations for such a view (cf. my Voltolini 1995).

conditional contribution, in a fictional narrow context can be labeled *fictional* proper names.⁶ Secondly, since things in the fictional world of that context unfold as (1) in that context fictionally says, that utterance is fictionally true, i.e., it is true in the fictional world of that context.

Yet how is it that (1) can also be really, not merely fictionally, true? Appealing to narrow context here does not help. For if we link (1) with a real narrow context, we get that the corresponding utterance has no real truth-conditions, hence no real truth-value. Or, to put things better, we have to say that if we take two utterances of (1), one of which is really untrue while the other is really true, we cannot account for this real truth-value difference, hence for this real truth-conditional difference, between these utterances by merely linking (1) with a real narrow context. For such a link does not discriminate between the two cases.⁷

Here comes the contextualist. For the contextualist, the first utterance of (1) is truth-conditionally empty given the emptiness of the names there involved, insofar as the wide context in which such an utterance occurs foresees no relation at all with a fiction involving such names. Yet the second utterance of (1) has a real truth-condition insofar as (1) is uttered in a wide context including the backward assumption for the utterer of being related with a fiction involving such names yet by standing completely outside any make-believe game underlying such a fiction. In such a wide context, (1) undergoes a *free enrichment*, analogous to that affecting e.g.:

(7) It rains

when it is uttered in a certain wide context including a certain location—Turin, say—in order for its utterer to mean that *it rains in Turin* (rather than, say, in London or in Palermo). For the contextualist, this amounts to saying that the truth-condition of the sentence undergoing free enrichment in one such wide context contains an *unarticulated constituent*, namely, a constituent that is not expressed by the sentence itself—(1) or (7), in our examples. In the relevant case of (1), this constituent is in its turn constituted both by a two-place relation, namely, the relation of *belonging to* holding between propositions and *stories*, i.e., sets of propositions—in this case, the proposition to the effect that Werther is fallen in love with Charlotte and the propositional set that constitutes Goethe's fictional

⁶ Of course, if the new theory of reference were not correct, there might be other semantic explanations of this datum, tracing back either to Frege's (1892) 'sense-and-reference'-theory of proper names or to Russell (1956) descriptivist theory of such terms. Yet as I said in the text it is nowadays a commonplace that such theories have serious drawbacks.

⁷ This point might be alternatively formulated by claiming that one and the same utterance of (1) may count both as a pseudo-assertion, i.e., as a sentence having a fictional truth-condition in a fictional context, and as a genuine assertion, i.e., as a sentence having a real truth-condition in a real context [For this way of putting things, cf. Evans (1982: 363–364)]. Yet since for a traditionalist an utterance counts as a sentence in *one* context, she has troubles in accounting for such a phenomenon. I myself have put forward this line of criticism against traditionalism with respect to multistable puns, namely, funny utterances having ambiguous significance. Cf. Voltolini (2012).

work, *The Sorrows of Young Werther*—and by the left-hand side propositional *relatum* of that relation—in this case, the above proposition.⁸

Since the relation in question is what is normally expressed by the preposition “in” and the propositional set in question may well be designated by a singular term of the kind “the story S”, the contextualist’s claim is a new way of accounting for the so-called *ellipticity thesis*, namely for the thesis that (in the relevant wide context) a sentence like (1) is elliptical for a sentence of the form “in the story S, p”, i.e.:

(1') In *The Sorrows of Young Werther*, Werther is fallen in love with Charlotte.

This traditional formulation allows us to easily see why the relevant utterance of (1) is really true yet the relevant utterance of (5) is really false. In its turn, this second utterance indeed says the same as the false:

(5') In *The Sorrows of Young Werther*, Werther is fallen in love with Charlotte’s husband.

For while it is really true that the story or propositional set constituting *The Sorrows of Young Werther* contains the proposition to the effect that Werther is fallen in love with Charlotte, it is really false that such a set contains the proposition to the effect that Werther is fallen in love with Charlotte’s husband.

The ellipticity thesis is widely defended in the literature (cf. e.g. Lamarque-Olsen 1994; Orenstein 2003; Rorty 1982).⁹ What is important for my present purposes is that this thesis can be defended in a way which gives a further support to the contextualist position.

There is plenty of evidence that, if it has to have a real truth-value hence a real truth-condition, the relevant utterance of a simple fiction-involving sentence must be taken as saying the same as a complex sentence of the form “in the story S, p” embedding the previous one—an *internal metafictional sentence*, as I have labelled it.¹⁰ Consider e.g. the following sentential pair, related to the Paladins’ epic:

(8) Roland is wise

(9) Roland is mad.

Intuitively we do not want that, when these sentences are uttered in order to say something really true, they say contradictory things, namely, that Roland is wise and that he is not wise. This is easily accounted for if we take the relevant utterances of those sentences to say respectively the same as the definitely non-contradictory:

⁸ As the careful reader will have already noticed, I here draw a distinction between *tales*—sets of fictional contents, i.e., of fictional truth-conditions for the utterances they mobilize—and *stories*, sets of real propositions, i.e., of truth-conditional contributions to the real truth-conditions of the utterances they mobilize. Cf. my Voltolini (2006). More on this later.

⁹ Normally the locution is interpreted as expressing an intensional operator, but I want to refrain “in the story S, p” from such an interpretation. Cf. Voltolini (2006: Chap. 6).

¹⁰ Cf. Voltolini (2006).

(8') In *The Chanson de Roland*, Roland is wise

(9') In *Orlando Enraged*, Roland is mad.

Now, if the contextualist's claim is correct, to say that when so uttered the former sentences are elliptical for the latter sentences respectively makes it vivid that in the respective wide contexts, (8) and (9) are really true, respectively, iff the proposition to the effect that Roland is wise belongs to the set of propositions constituting Chretien de Troyes' *The Chanson de Roland*, and iff the proposition to the effect that Roland is mad belongs to the set of propositions constituting Ludovico Ariosto's *Orlando Enraged*. But this is precisely the case. For according to the contextualist, what the relevant utterances of (8) and (9) present is just another case of a pseudo-disagreement between speakers that is dissolved once the different unarticulated semantic content of the above utterances is explicitly expressed. Consider two speakers *X* and *Y*, respectively asserting and denying:

(10) Sugarless coffee is disgusting.

For the contextualist, there is no genuine disagreement between *X* and *Y*. For by means of their respective utterances of (10), they are respectively saying (with assent) that *for X*, sugarless coffee is disgusting, and saying (with dissent) that *for Y*, sugarless coffee is disgusting. A different truth-conditional content that is explicitly expressed by the respectively true and false:

(10X) For *X*, sugarless coffee is disgusting

(10Y) For *Y*, sugarless coffee is disgusting.

3 Contextualism and Realism about Fictional Entities

Ad advantage of this contextualist way of accommodating the above data is that this way is neutral between an ontologically realist and an ontologically antirealist stance about fictional entities, or in other terms, between people believing that there are fictional entities like Werther and Charlotte, as designated by the names "Werther" and "Charlotte" respectively, and people believing the opposite. On behalf of the contextualist, I have first of all said that an utterance of (1) has no truth-conditions when the wide context to which it is linked in which names like "Werther" and "Charlotte" are mobilized is utterly non-fiction-involving. To be sure, I have also said when (1) is uttered in a fictional narrow context, the names in question do refer to two distinct individuals. Yet again this claim is neutral between realism and antirealism about fictional entities. For such a controversy is a controversy as to whether *in the domain of the actual world*, there are *fictional things* like Werther and Charlotte. This controversy is untouched by the fact that, *in the domain of a make-believe world* such as the world of Goethe's tale of *The Sorrows of Young Werther*, those names refer to two distinct *human beings*, the man Werther and the woman Charlotte. In point of fact, the idea that in such a world of make-believe those names refer has been originally maintained by an

antirealist like Walton in order precisely to also claim that those names refer to nothing in the actual world. Finally, I have said that, in the relevant wide context, (1) has a real truth-condition and a real truth-value—the True. But I have not said *which* truth-condition (1) has in such a case. Antirealists hold that such truth-condition is *non-committal*, i.e., it does not commit us to the fact that in the general domain of what there *actually* is, there also are fictional entities like Werther and Charlotte. Realists instead maintain that such a truth-condition is *committal*, i.e., it commits us to such entities.

This said, I think that the realist truth-conditional approach of sentences like (1) when uttered as having real truth-conditions is preferable to the anti-realist approach. In this section, I will provide some criticisms against the antirealist approach.

To begin with, recall that I have assumed the correctness of the new theory of reference. Thus, I rule out antirealist descriptivist approaches that, while holding that when a sentence like (1) has a real truth-value (in this case, the True), it is elliptical for an internal metafictional sentence like (1'), also claim that the empty proper names “Werther” and “Charlotte” embedded in it are respectively synonymous with definite descriptions lacking an actual denotation. Even merely *partial* descriptivism, i.e., a theory holding that *only* fictional proper names, like the two above ones, are to be equated with definite descriptions lacking an actual denotation,¹¹ seem to me to have no independent justification.

Thus, if one wants to stick both to the semantic conviction that proper names are not synonymous with definite descriptions and to the ontological antirealist stance on fictional entities, there are two main options one can nowadays choose. First of all, one may appeal to Negative Free Logics and say that, although simple sentences involving fictional proper names, or better empty proper names in general, are false for want of reference for such names, complex sentences like internal metafictional sentences embedding such sentences may well be true.¹² Alternatively, one may assume that locutions such as “in the story, S” work not only as intensional operators shifting, as such operators normally do, the

¹¹ Cf. Currie (1990) and Garcia-Carpintero (2007). In point of fact, Currie’s theory defends this idea only for *embedded* fictional proper names. When such names occur unembedded in simple fiction-involving sentences that cannot be taken as elliptical for internal metafictional sentences, notably *external metafictional sentences* like “Werther is a fictional character”, for Currie they do refer to fictional entities, notably fictional roles. Given however that there are *mixed* contexts such as “Although he was fallen in love with Charlotte, Werther is a fictional character”, made out of both an internal and an external metafictional sentence, this position seems to me even more untenable.

¹² This option is envisaged by Sainsbury (2009). In point of fact, Sainsbury there and elsewhere (2011) prefers another account, according to which sentences like (1), which are really false because of the referentlessness of their names, are only true under a presupposition, i.e., accepted as true. Yet this account simply amounts to denying datum (iii). For it deals with the putative real truth of (1) in the same way as with the merely putative truth presented by datum (ii). Yet who utters (1) in an examination does not even shallowly pretend that (1) is true, as this account postulates.

circumstances of evaluation of the sentence “p” they embed, but also as context-shifting operators, shifting the context of interpretation of their embedded sentence. According to such a view, an internal metafictional sentence of the form “in the story S, p” is true iff the embedded sentence “p” *in a fictional context of interpretation*, that is, as if it were uttered in such a context, is true in the world of that context; in other terms, such an embedding sentence is true iff the embedded sentence has a *fictional* truth-condition rather than a real one. This option is positively endorsed by pretence antirealists. As we have seen before, the fact that a sentence like (1), which is exactly the sentence embedded in an internal metafictional sentence like (1'), is true when uttered in the fictional context of Goethe's tale entails that the names “Werther” and “Charlotte” do refer to concrete beings in the make-believe world of that tale, but does not entail that such names refer in the actual world.¹³

The first option is forced to maintain that *all* simple sentences containing fictional names are false for want of reference for such names. It thus accounts in this way for what happens for a sentence like (1) when an utterly non-fiction-involving real (wide) context is given to it. Instead of saying, as I did, that the sentence is neither true nor false for it has no truth-conditions since its proper names refer to nothing, this option prefers to saying that the sentence is false precisely for the very same reason, i.e., want of reference for such names. Yet intuitively speaking not all simple fiction-involving sentences containing fictional proper names are false. Consider simple yet fiction-involving sentences that however mobilize neither fictional tales or nor fictional stories, that is, *external metafictional sentences* like the following ones,

- (11) Werther and Charlotte are Goethe's creations
- (12) Werther has fascinated many youngsters
- (13) Charlotte is unbearable.

We clearly feel that such sentences are true. An intuition the realist easily accommodates, by saying that such sentences predicate something true of certain fictional entities—Werther and Charlotte, in this case.¹⁴

As is well known, the second option has the apparently unpleasant consequence that, taken as context-shifting operators, locutions such as “in the story S” are *monsters* in Kaplan's sense. For this option does not account for the fact that, when “in the story S, p”-sentences contain indexicals, no such context shift for the

¹³ For this option, which is implicit in Walton (1990), see Adams-Fuller-Stecker (1997).

¹⁴ Incidentally, even in his second approach to sentences involving fiction (cf. fn. 12) Sainsbury has somehow to acknowledge that external metafictional sentences *really* possess their intuitive truth-value. For, as he admits, they are not merely accepted as true, they are *believed* to be true. Cf. (2011:147). Furthermore, even a *prima facie* antirealist such as Braun, who defends a somewhat similar idea that simple sentences like (1) are false for they express gappy propositions, i.e., propositions whose ‘subject’-position is empty for the relevant proper names refer to nothing, is forced to say that external metafictional sentences are true for the names they contain refer to fictional entities. Cf. Braun (2005).

sentences they embed may take place.¹⁵ For instance, if on the behalf of the pretence antirealist by embedding Proust's *Recherche* inaugural sentence:

(14) For a long time I used to go to bed early

I utter:

(15) In *In Search of Lost Time*, for a long time I used to go to bed early

I cannot (truly) say, as the pretence antirealist wishes, that the *Recherche*'s narrator has such a habit, but I am forced to implausibly (and falsely) say that *I* have such a habit in the *Recherche*. For the embedded indexical "I" cannot undergo the context shift that would take the embedded sentence as if it were uttered in a fictional context, hence to have a fictional truth-condition, as the pretence antirealist requires.

To be sure, this problem can be circumvented. For there may well be cases precisely involving fiction in which simple sentences embedded in complex sentences generating intensional contexts, namely, sentences of the form "according to the fiction F, p", undergo such a context shift, so that locutions of the form "according to fiction F" are indeed Kaplanian monsters.¹⁶ The real problem is another. In general, the fact that a sentence is true when uttered in a given context is no sufficient condition for the fact that another sentence is true when uttered in another context. This is particularly the case in our situation. The fact that a sentence expressing (part of) the fictional truth-conditional content a tale has in the context of an unreal world is true in *that* world says nothing as to the fact that a sentence expressing or reporting in the context of the real world, as a contribution to its real truth-conditional content, (part of) the content of a story is true in *this* world.¹⁷

¹⁵ Cf. Kaplan (1989a: 510).

¹⁶ For such cases (that *pace* Kaplan typically involve indexicals, notably modal or temporal ones), cf. e.g. Bonomi (2008) and especially Predelli (2008). To be sure, in that paper Predelli defends a view that is more complicated than the 'context-shift'-view I am here presenting. For one thing, according to Predelli even a sentence like (1) when uttered as a comment to Goethe's tale has to be understood as elliptical for a complex sentence involving one such context shift for its embedded sentence, i.e., "according to Goethe's tale, Werther is fallen in love with Charlotte". But let me leave these complications aside. In any case, for me complex sentences of the form "according to the fiction F, p" are not sentences of the form "In the story S, p". For according to me locutions of the "in the story S"-form are not intensional operators. Cf. Voltolini (2006). (For a less radical distinction between the "according to the fiction"—and the "in the story"—locutions, which takes both as expressing different intensional operators, cf. Sainsbury (2009: 79fn. 12).

¹⁷ Incidentally, this is why the fact that, when uttered in the context of a possible world, a certain sentence is true in that world says nothing as to the fact that a corresponding modal sentence (presenting a normal intensional context), uttered in the context of the real world is true in this world. Consider a sentence like "Vulcan is a planet". There surely is a possible world in which people successfully refer to something by using the name "Vulcan" in uttering the above sentence, which so expresses a *possible* truth-conditional content that makes that sentence also true in that world, if that very something is a planet in that world. However, the fact that in the unreal context that has such a possible world as its 'world'-parameter that sentence is true does

All in all, we have that the best antirealist options that refrain from descriptivism do not satisfactorily account for the truth of internal metafictional sentences, hence for the truth of their embedded sentences when uttered in the appropriate wide contexts. Thus, in order to appeal to the datum of the real truth of such sentences when so uttered as an evidence in favor of contextualism, it is better to maintain that the proper names figuring in it refer to fictional entities and thereby defend a realist stance towards such entities.

References

- Adams, Fred, Gary Fuller, and Robert Stecker. 1997. The semantics of fictional names. *Pacific Philosophical Quarterly* 78: 128–148.
- Bonomi, Andred. 2008. Fictional contexts. In *Perspectives on contexts*, ed. P. Bouquet, et al., 215–250. Stanford: Csl Publications.
- Braun, David. 2005. Empty names, fictional names, mythical names. *Noûs* 39: 596–631.
- Cappellen, Herman, and Ernie Lepore. 2005. *Insensitive semantics*. Oxford: Blackwell.
- Currie, Gregory. 1990. *The nature of fiction*. Cambridge: Cambridge University Press.
- Donnellan, Keith. 1966. Reference and definite descriptions. *The Philosophical Review* 75: 281–304.
- Evans, Gareth. 1982. *The varieties of reference*. Oxford: Clarendon Press.
- Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50. In *Translations from the philosophical writings of Gottlob Frege*, eds. P.T. Geach and M. Black (trans: Black, M.) 56–78. Oxford: Blackwell, 1980.
- García-Carpintero, Manuel. 2007. Fiction-making as a Gricean illocutionary type. *The Journal of Aesthetics and Art Criticism* 65: 203–216.
- Kaplan, David. 1989a. Demonstratives. In *Themes from Kaplan*, ed. J. Almog, et al., 481–563. Oxford: Oxford University Press.
- Kaplan, David. 1989b. Afterthoughts. In *Themes from Kaplan*, ed. J. Almog, et al., 565–614. Oxford: Oxford University Press.
- Kripke, Saul. 1980. *Naming and necessity*. Oxford: Blackwell.
- Lamarque, Peter, and Stein H. Olsen. 1994. *Truth, fiction, and literature*. Oxford: Clarendon Press.
- Marti, Genoveva. 1995. The essence of genuine reference. *Journal of Philosophical Logic* 24: 275–289.
- Orenstein, Alex. 2003. Fiction, propositional attitudes, and some truths about falsehood. *Dialectica* 57: 177–190.
- Perry, John. 1997. Indexicals and demonstratives. In *Companion to the philosophy of language*, ed. R. Hale, and C. Wright, 586–612. Oxford: Blackwell.
- Predelli, Stefano. 2005. *Contexts*. Oxford: Oxford University Press.

(Footnote 17 continued)

not make the case that the modal sentence “It is possible that Vulcan is a planet” is true when uttered in a real context that has the real world as its ‘world’-parameter. For, in order for *this* sentence to be true in the real world, it must report the content of a possibility *for us*, people in the real world. But this cannot be the case if “Vulcan” refers to nothing in the context of the real world, so that (given the new theory of reference) it is just a *flatus vocis* giving no truth-conditional contribution to the sentence in which it figures (as if we had merely said “It is possible that Xhjtz is a planet”).

- Predelli, Stefano. 2008. Modal monsters and talk about fiction. *Journal of Philosophical Logic* 37: 277–297.
- Putnam, Hilary. 1975. The meaning of 'Meaning'. In *Mind, language and reality. philosophical papers*, vol. 2, 215–271. Cambridge: Cambridge University Press.
- Recanati, François. 1993. *Direct reference*. Oxford: Blackwell.
- Recanati, François. 2004a. *Literal meaning*. Cambridge: Cambridge University Press.
- Recanati, François. 2004b. Indexicality and context-shift. Paper presented at the *Workshop on Indexicals, Speech Acts and Logophors*, Harvard University, 20.11.2004. http://jeannicod.ccsd.cnrs.fr/docs/00/05/36/04/PDF/ijn_00000552_00.pdf. Accessed 23 Aug 2011.
- Rorty, Richard. 1982. Is there a problem about fictional discourse?. In *Consequences of pragmatism*, 110–138. Brighton: The Harvester Press.
- Russell, Bertrand. 1956. The philosophy of logical atomism. In *Logic and knowledge*, 177–281. London: Allen and Unwin.
- Sainsbury, Mark R. 2009. *Fiction and fictionalism*. London: Routledge.
- Sainsbury, Mark R. 2011. Fiction and acceptance-relative truth, belief and assertion. In *Truth in fiction*, ed. F. Lihoreau, 137–152. Munich: Ontos.
- Stanley, Jason. 2007. *Language in context*. Oxford: Clarendon Press.
- Voltolini, Alberto. 1995. Indexinames. In *Karlový vary studies in reference and meaning*, eds. J. Hill, and P. Kot'átko, 258–285. Prague: Φιλοσοφία-Filosofia Publications.
- Voltolini, Alberto. 2006. *How ficta follow fiction. A syncretistic account of fictional entities*. Dordrecht: Springer.
- Voltolini, Alberto. 2012. Puns for contextualists. *Humana mente* 23: 113–140.
- Walton, Kendall L. 1990. *Mimesis as make-believe*. Cambridge: Harvard University Press.

Pragmatics and Philosophy: Three Notes in Search of a Footing

Alec McHoul

Abstract In this chapter I turn to a seminal essay by Gilbert Ryle. Here Ryle explores the possible distinctions between ordinary everyday knowledge and recondite philosophical knowledge. He uses a number of metaphors in order to achieve this including those of “morning” vs “afternoon” types of questions, and the distinction between the villager’s and the cartographer’s knowledge of the locale. I extend these metaphors by drawing on Thomas Kasulis’s distinction between the potter’s and the geologist’s knowledge of clay and use this to reflect on lay as against professional knowledges of language. The idea of a “socio-logic” of ordinary talk, from Harvey Sacks, is alluded to as a way of seeing how such ordinary talk is grounded.

Keywords Everyday knowledge-in-action · Specialist philosophical knowledge · Logical geography · Language-in-use · Ryle · Wittgenstein · Sacks

1

In his paper “Abstractions” (1971), Gilbert Ryle attempts to draw a line of demarcation between ordinary everyday knowledge and recondite philosophical knowledge. He uses a number of metaphors in order to achieve this including those of *morning-* versus *afternoon-*types of questions, and the distinction between the *villager’s* and the *cartographer’s* knowledge of the locale. What concerns me in this paper is how the distinction might operate, if at all, with respect to ordinary everyday knowledge of *language* and the disciplines that congregate under the name of “pragmatics”.

A. McHoul (✉)

School of Arts, Murdoch University, Murdoch, WA 6150, Australia
e-mail: a.mchoul@gmail.com

Ryle begins his paper with a quotation from St Augustine that is also relied on by Wittgenstein for similar (though not identical) purposes: “When you do not ask me what Time is, I know perfectly well; but when you do ask me, I cannot think what to say” (p. 435). To tackle this pseudo-problem, Ryle’s first metaphor is itself temporal. He writes of “morning” as against “afternoon” questions. In the morning, our latter-day Augustine works perfectly well with matters of time. He knows that “what happened yesterday is more recent than what happened a month ago” (p. 435); just as we know perfectly well that a minute is shorter than an hour; or that, if you promise to meet me at 19:00 and arrive at 19:30, you are “late”. But, in the afternoon, as it were, it is no longer a matter of efficient coping. Now, we are faced with an abstraction, Time as such; and we are now engaged in wondering whether Time can ever stop; whether it flows regularly or irregularly, and a whole host of related imponderables.

In short, what is Time—is it a Thing or a Process or a Relation? Is it a sort of cosmic river, only one without any tangible water between its non-existing banks? One which flows out of no spring and pours into no ocean? (p. 435)

Before lunch, the familiar world exhibits no significant problems; at least none of the transcendental sort. And we go about our lives knowing perfectly well how to operate with watches and matching our movements to them. After lunch, thrown into contemplation of an abstraction, we are bewitched. Let me suggest, as a first approximation only, that the morning might—in a rather distinctive way—be the time for pragmatics while the afternoon might equally be the time for philosophy.

That would be well and good; except that pragmatics, as a discipline, as a form of inquiry, is not simply life-as-usual, going about its business. It is an attempt to capture (in some way—how?) what goes on in that business. Re-inspecting this sentence: “the morning might be the time *for* pragmatics while the afternoon might equally be the time *for* philosophy”, the two *fors* appear to have similar force. Describing philosophy’s (“afternoon”) knowledge of everyday (“morning”) knowledge, Ryle writes: “We now have to operate *upon* what we ordinarily operate readily and unquestioningly *with*” (p. 444). And is that not what pragmatics also attempts to do?

Yet it seems to me that the *for* and the *operate upon* in these two formulations have different implications for philosophy on the one hand and pragmatics on the other. Both appear to bear *some kind of relation* to everyday knowledge that is different from that knowledge as such. Yet the relations in question must be different. And this is why I parenthesise—“in a rather distinctive way”. Now my job is to say something about what is distinctive for us.

2

My initial attraction to the villager and the cartographer (as a metaphor or model) has a slight drawback. The villager, in Ryle’s characterisation at least, has purely intuitive knowledge of his way about. At times, he even appears to be less a

villager and more a village idiot. Yes, he manages perfectly well; but he *is* a rather artless creature. There's a whiff, the faintest of redolences, of what Garfinkel (1967) called "cultural doper" about him. If we continued along that path, it would not be long before we reached a point where the cartographer morphed into the speculative meta-ethnographer and the villager into the (negatively) unscientific native that Wittgenstein found in his "Remarks on Frazer's *Golden Bough*" (1971). Not that I'm accusing Ryle here. No one who has carried an official membership card of the Guild of Philosophers has done more to rehabilitate the significance of the ordinary, Austin and Wittgenstein notwithstanding. But none of these "ordinary language philosophers" have genuinely considered *that*—the use of ordinary language—to have the status of the artisanal. That is, as worth looking at in terms of the arts and crafts of its actual deployment via captured (live, recorded, etc.) specimens. To be sure, the attitude has changed but the method—making up examples from a sort of philosophical intuition about everyday language—remains the time-honoured procedure. One sense of "the ordinary" is, to be sure, admirably preserved. But another, the one that interests us most, the pragmatic itself, is not.

A better distinction might be that invoked by Thomas Kasulis in his forthcoming book *Engaging Japanese Philosophy*. I'm not entirely sure what he does with it—for I have only heard him talking on the radio about it—but his distinction is between the potter's and the geologist's knowledge of clay. There are several advantages here. The potter's status as craftsman, artisan, artist even, is without question. And we are explicitly dealing with knowledge; and of knowledge of some specific, if complex, substance—clay.

Both the geologist and the potter know clay. But they know it *as* different things. The geologist can identify its chemical composition by a reductive analysis; its physical properties, such as its mass and density. He can compare it with other samples drawn from locations at a distance from the original. He can speculate on the likelihood of its occurrence at as-yet-unknown sites on the basis of its co-occurrence with other mineral types. And so forth. The geologist has learned all of this in classrooms and laboratories and also "on the job".

The potter, by contrast, has a different kind of knowledge. He has learned about clay from his artisanal forebears, perhaps. Or from colleagues as well as from experience. And he has been out to the same sites as the geologist, perhaps, in search of the best specimens. After years of potting, he knows what he is looking for. He feels and rubs the clay between his fingers; the same fingers that will be used for the making of his pottery. He sniffs the clay and notes its colour. He has felt and sniffed clay like this before; or else not *quite* like this before. He is capable of making very fine distinctions that point to several ends: Will this work well? Will it fire satisfactorily? Will it take decoration and a glaze? Will it endure? And a host of other questions whose end is a practical and/or artful work.

Both the geologist and the potter have their *métiers*. But they are distinct and different. And yet their object of interest is, in some ways, the same. Clay. Ditto, then, language for the linguistic professional and the everyday speaker, hearer, writer, reader, calligrapher, sign-writer, conversationalist, poet, rhetor.... But now,

for our purposes, the analogy, as it must, comes to an end (see Hesse 1963). The model of the geologist will work perfectly well for all the other professionals of language and their recondite knowledge of language: the grammarians, the morphologists, the phonologists, the syntacticians, the lexicographers, the philologists and, perhaps sadly, even for the semanticists and semiologists. It matters not to them what the ordinary user (cf the potter) *knows* of what he does with the actual inspectable stuff, language (cf clay). But it matters to us in linguistic pragmatics. And it matters to us, accordingly, as our unique object.

3

There is a similar distinction in philosophy, mapped for example by the two phases of Wittgenstein's career. As Rhee (2003, p. 34) shows, Wittgenstein asked a single question across the course of his philosophical life: What does it mean to say something? But in the first (geological?) phase the question became: *What* is a logical proposition? In the second (quasi-potteric?) phase it turned into: *How* is ordinary language used? (Much is made of the shift though Rhee is more interested in the continuity.)

The linguistic fraternity, our grammarians, morphologists and so forth, it seems to me are stuck in the rut of a quasi-formal-logical kind of investigation. Only pragmatics, again it seems to me, has made something of a shift parallel to Wittgenstein's. Again, we are all asking: What does it mean to say, write, hear, read ... something? But we are asking it in different ways. Our take on the "clay"—which we can at least *assume* to be the same substance, postmodern nonsense about object-constancy notwithstanding—is different.

If language were a coin, then, on one face would be its formal logic. On the other would be the way people live with it. To say that the latter has a "logic" of its own may be in order. But now the word "logic" has completely altered its meaning. In the first sense it is nominal: the logic of English syntax would be a case in point. In the second case it is attributive: Ryle's "logical geography" come to mind here; or, just perhaps, Wittgenstein's "logical grammar". Sacks's (1992, *passim*) term "socio-logic" is perhaps even better.

Sacks is not attempting merely to describe what ordinary speakers ("potters") know about *how* to go about doing things with talk. He is attempting to find out what *must be the case* for them to know some things about *how* to go about doing things with talk. He is looking for the (socio-logical) underpinning of what they know. But, at the same time, this has to be utterly on all fours with the actual materials: recordings of talk in many of Sacks's cases though by no means confined to such things.

The grounding question, the footing, not often investigated in detail these days, and certainly not in a great deal of what passes for "conversation analysis" or "the analysis of talk in interaction" is: What anchors or grounds the ways people actually live with language?

References

- Garfinkel, Harold. 1967. *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
- Hesse, Mary. 1963. *Models and analogies in science*. London: Sheed and Ward.
- Kasulis, Thomas. (forthcoming). *Engaging Japanese Philosophy*. Honolulu: University of Hawaii Press.
- Rhees, Rush. 2003. *Wittgenstein's On Certainty: There—like our life*, ed. D.Z. Phillips. Oxford: Blackwell.
- Ryle, Gilbert. 1971. Abstractions. In *Collected papers, Volume 2*, 435–445. London: Hutchinson.
- Sacks, Harvey. 1992. *Lectures on conversation, Volumes 1 and 2*. ed. Gail Jefferson. Cambridge: Blackwell.
- Wittgenstein, Ludwig. 1971. Remarks on Frazer's *Golden Bough*. *The Human World*, 3(May): 31.

A Brief Essay on Slurs

Luvell Anderson and Ernie Lepore

Abstract Slurs target race, nationality, religion, gender, sexual orientation, politics, immigrant status, line of work and many other demographics. They offend their targets—some more than others. Children who blurt out slurs are scolded regardless of their ignorance; but not uncommonly, targets appropriate slurs for themselves. The main questions of this paper are by which mechanism do slurs derogate, why some more forcefully than others, and how do targeted members succeed in mollifying some slurs? A lacuna in the literature on slurs is the rarity with which their offenses are specified. We are told little more than that they derogate, belittle, disparage, or diminish, but never how. Explanatory options here are, however, limited. We know a lot about how words achieve efficacy; with that in mind, in what follows we will canvass alternatives in the hope of illuminating the offensive nature of slurs. Our view, in brief, is that slurs are banished words, not because of any linguistic feature they exhibit or any content they carry, but rather because they are on a list of prohibited words—the higher up the list the more offensive their uses. This leaves a slew of questions: what determines whether a word is on or off the list, its position on the list; and why is it sometimes appropriate to flout its prohibition? The backdrop for this paper is the role of linguistics and philosophy of language in the investigation of slurs. Our answer is none.

Keywords Indirect reports · Offense · Pragmatics · Prohibition · Racism · Slurs · Slurring · Semantics · Societal/discourse rules · Taboo

L. Anderson (✉)
University of Memphis, Memphis, USA
e-mail: lndrsn14@memphis.edu

E. Lepore
Rutgers University, New Jersey, USA

1 On Speech and Harm

Increasingly philosophers (and linguists) are turning their attention to the surprisingly little explored lexical category of slurs. These are expressions that target groups on the basis of race, nationality, religion, gender, sexual orientation, immigrant status and sundry other demographics. In academic discussion, it's easy to convince oneself (we confess on occasions we have) that particular uses of slurs are inoffensive. This discussion wouldn't exist otherwise. As a safeguard against inurement, you should always ask how a targeted member, perhaps overhearing you, would react to your usage. You may find that what seems suitable is most definitely not. As many celebrities, including recently the radio host Dr. Laura Schlessinger, have learned, almost no speakers can opt to use a slur non-offensively.

But why are their uses so offensive? Though xenophobes are not offended by slurs, others are—with some being more offensive than others. Calling an Asian businessman 'suit' will not rouse the same reaction as using calling him 'chink'. Even co-extensive slurs vary in intensity of contempt. Christopher Darden once branded the N-word the "filthiest, dirtiest, nastiest word in the English language"; we doubt anyone reacts as such to 'negro,' yet it too has become a slur. How can words fluctuate both in their status as slurs and in their power to offend? Targeted members themselves are not always offended by confrontations with slurs, for example, so-called appropriated or reclaimed uses (the camaraderie use of the N-word among African-Americans and 'queer' among homosexuals). These data focus the discussion around three questions:

Why are some confrontations with slurs offensive? Why do some impact audiences more forcefully than others? How do targeted members sometimes succeed in mollifying them?

The consensus answer among philosophers to the first question is that slurs, *as a matter of convention*, carry negative attitudes towards targeted groups. Those who pursue this answer are committed to the view that slurs get across offensive *content or meaning*; they disagree only over the mechanisms of implementation. An alternative proposal is that slurs are *prohibited* words *not* on account of any content they get across, but rather because of relevant edicts surrounding their prohibition. This latter proposal raises a few pertinent questions, including how do words become prohibited, what's the relationship between a prohibition and offense potential, and why is it sometimes appropriate to flout such prohibitions? Let's start with conventional meaning.

Does the N-word mean something different from its neutral counter-part 'African-American' or the less intense 'negro'? The generally reliable O.E.D, a catalogue of senses, says no more than that a slur is pejorative, derogatory, debasing, disparaging or that it belittles its target group. But this information fails to distinguish the N-word from the less potent 'spook.' (It's no accident Clint Eastwood, though portraying a racist in a recent film, chose to use the latter, and not the former, in confronting two young black men.) Still, it permits us to infer slurs *supplement* the meanings of their neutral counter-parts with something

offensive about whomever they reference. This information, however meager, suffices to isolate a flaw in trying to pin the offensiveness of a slur on its predicative content.

Anyone who wants to disagree with what ‘Binyamin is a Jewish person’ ascribes to Binyamin can do so with a denial. So if ‘Binyamin is a kike’ were offensive because of what it predicates of Binyamin, we should be able to reject its offense by denying it. But its denial is no less inflammatory than the original. Therefore, however a slur offends, it is not through what it predicates of its subject.

Another fascinating aspect of slurs that challenges the view that their meaning renders them offensive tracks their behavior in indirect speech. Normally, an utterance can be correctly indirectly reported by *re-using the very expressions being reported on*. What better insurance for accuracy can there be in reporting another than to re-use her words? But then Eric’s offensive use of ‘A bitch ran for President’ should be correctly reportable with ‘Eric said that a bitch ran for President’. Yet this report does not capture *Eric’s* offense, though, interestingly, it guarantees an offense by *whoever is reporting him*. For all that the report reveals, Eric may have uttered ‘A woman ran for President.’ What’s gone wrong? We expect indirect reports of others to be *of others*, not of ourselves. This limit on indirectly reporting slurs is significant: if a slur carries offensive content conventionally, why *can’t* it be captured by reporting another’s use! Is the offense of another’s slur ineluctable? Is it possible that we can recognize the offense, but not re-express it? How odd.

Is there someplace else to look for an account of the offensive nature of slurs? May it be a matter of tone? Unlike conventionalized content, tone is supposed to be subjective. Words can disagree in tone but share their content. Might tone distinguish slurs from neutral counter-parts? No one can deny the N-word *can* arouse subjective images and feelings in us that ‘African-American’ need not, but as an account of the difference in offensive punch it can’t be the whole story.

Consider a xenophobe who only uses slurs for picking out a target group. He may harbor no negative opinions towards its members; he may use slurs only among likeminded friends when intending to compliment or pronounce affection (‘I love wops’ or ‘Chinks are smarter than the rest of us’) but his uses are pertinently offensive. The difference between a slur and its neutral counter-part is *not a matter of subjective feel*.

A major problem with any account that tries to explain the offensive nature of a slur by invoking content is how it can explain the general exhortation against even mentioning slurs. A quoted occurrence of some slurs easily cause alarm and offense. Witness the widespread preference for ‘the N-word’, which describes a slur without using *or* mentioning it. This is a major problem for anyone who thinks that it’s their meaning that makes the occurrences of slurs offensive since quotation famously nullifies content; even though ‘bachelor’ and ‘unmarried man’ are synonyms, the sentence ‘‘bachelor’ has eight letters’ is true while ‘‘unmarried man’ has eight letters’ is not. Content inside quotation is inert; but why then are mentioned uses of the N-word offensive?

A more controversial challenge is raised by the offensive potential of *incidental* uses of slurs, as witnessed by the Washington D.C. official who wound up resigning his job over the outcry that his use of ‘niggardly’ provoked. In 1999, the head of the Office of Public Advocate in Washington, D.C. used it in a discussion with a black colleague. He was reported as saying, “I will have to be niggardly with this fund because it’s not going to be a lot of money.” Despite a similarity in spelling, his word has no semantic or etymological tie to the N-word; mere phonetic and orthographic overlap caused as much a stir as standard offensive language. This is not an accidental use of an ambiguous or unknown slur, but an incidental one.

What conclusions should we draw? One suggestion is that uses of slurs are offensive simply because they sometimes constitute violations on their very prohibition. Just as whoever violates a prohibition risks offending those who respect it, perhaps the fact that slurs are prohibited explains why we cannot escape the affect, hatred and negative association tied to them and why their occurrences in scholarly discussions and even within quotation marks can still inflict pain. Prohibited words are usually banished wherever they occur! This explains why bystanders (even when silent) are uncomfortable, often embarrassed, when confronted by a slur. Whatever offenses these confrontations exact, the audience risks complicity, as if the offense were thrust upon them, not because of its content, but because of a responsibility we all incur in ensuring certain violations are prevented; when they are not, they must be reported and possibly punished. Their occurrences taint us all. Slurs in this regard are like the swastika—a symbol with a benign history before the Nazis, but whose appropriation so contaminated it that even incidental tokenings are charged and offensive to many. In short, Lenny Bruce may have gotten it right when he declared “the suppression of the word gives it the power, the violence, the viciousness.” It is impossible to reform a slur until it has been removed from common use.

The prohibition of slurs accounts for why another’s slur is ineluctable. Indirect reports fail to attribute slurring to whomever they report since reporter’s offense “screens off” the offense of the reported. A use, mention, or interaction with a slur, *ceteris paribus*, constitutes an infraction. We cringe when confronted by slurring *because* its uses are never tolerable. But what is its etiology as a slur if not through the content it communicates?

Words become prohibited for all sorts of reasons, e.g., by a directive or edict of an authoritative figure; or because of a tainted history of associations, perhaps, though conjuring up past pernicious or injurious events. The history of its uses, combined with reasons of self-determination, is exactly how ‘colored’, once used by African-Americans self-referentially, became prohibited, and so, offensive. A slur may become prohibited because of *who* introduces or uses it. This is the sentiment of a high school student who objected to W.E.B. Dubois’ use of ‘Negro’ because it “is a white man’s word.”

What’s clear is that no matter what its history, no matter what it means or communicates, no matter who introduces it, regardless of past associations, *once relevant individuals declare a word a slur, it is one*. When the Reverend Jesse

Jackson proclaimed at the 1988 Democratic National Convention that from then on 'black' should not be used, his effort failed. Many African-Americans carried positive associations with the term ('Black Panthers', 'Black Power', 'I'm black and I'm proud') and so Jackson's attempt at prohibition did not stick.

Doesn't prohibition over time itself endow a slur with conventionalized offensive content? Though dictating or legislating usage can provoke meaning change, it needn't. When the victors banish the language of vanquished foes, their decree will modify behavior but it won't change what banished words mean or what they reference. Likewise, a word needn't change designation or meaning once deemed a slur. Its associations might change, but to infer they are responsible for its offense potential is to put the cart before the horse.

What about discrepancies in intensity? Why is it worse to slur African-Americans with the N-word than with some others, or Irishmen with 'mick'? Have sufficiently enduring campaigns against the uses of certain slurs ameliorated them? Targets with investments in seeing a word removed from circulation are more affronted by confrontations with some slurs than by others. In contrast, as fewer individuals acknowledge offense, it comes to pass, contingent on authority, station, or relevance, that offensive intensity is lost or diminished—the campaign against the word effectively ends.

How this diminution is achieved we doubt admits of a simple answer. We doubt the phenomena constitute a homogenous front, but the practice of appropriation illuminates the process in one sort of case. Why can a black man use the N-word on another black man without slurring? A natural response for a content theorist is to posit ambiguity; but that can't be correct, for then why can't non-members utilize this second sense?

One possible explanation of appropriation is that targeted members can opt to use a slur *without* violating its prohibition because membership provides a defeasible escape clause; most prohibitions include such clauses. Oil embargos permit exportation, just not importation. Sanctions invariably exclude medical supplies. Why shouldn't prohibitions against slurs exempt certain individuals under certain conditions for appropriating a banished word? African-Americans are can inoffensively use the N-word among themselves; gays 'queer', women 'bitches', and so it goes for most targeted groups. The NAACP continues to use "Colored" relatively prominently (on their letterhead, on their banners, etc.).

Once appropriation is sufficiently widespread, it might come to pass that a prohibition eases, permitting—under regulated circumstances—designated outside members access to an appropriated use. Should this practice become sufficiently widespread, the slur, as with 'suit' 'Tory' and 'limey', loses its intensity. How escape clauses are fashioned and what sustains them is a complex matter—one we cannot take up here. Still some clarifications are in order.

One clarification at the outset is that what we are offering is not a proposal aimed at reducing slurring, but rather an account of the offensiveness of slurs. Keep in mind that for us we are insisting on explaining at least two features of slurs. First, how can some be more offensive than others (including those targeting the same group)? And second, why can some of us use slurs without being

offensive while others of us cannot? Our proposal that slurs are expressions whose occurrences are prohibited purports to explain these features better than views that invoke the content of the word, i.e. their meaning. In fact, appeals to meaning don't seem to be capable of coherently accommodating either of these features.

This key point cannot be stressed enough, since several commentators on our work seem to conflate our aim with an entirely separate aim, namely, imposing standards of political correctness. This has never been our intention. Further, our proposal does not claim prohibition exhausts a slur's content. Rather, the suggestion is that prohibition explains the phenomenon of offense. Slurs may very well carry negative content (even though they need not), but Prohibitionism claims that content does not explain, for instance, the variation in offensive potency among slurs.

2 Responses to Criticism

Having made these points, we can now turn to worries we have encountered with regards to our particular proposal. There are three main ones we will focus on: One is the idea that it is the intention or attitude of the speaker that renders the word offensive. Next, is that what makes slurs offensive are the ideas that a slur's use seems to permit one to infer about the target. And finally, though Prohibitionism explains some of the phenomena surrounding our use of slurs, it can't be the whole story.

The first main worry is that we should explain the offense of slurs in terms of the hateful attitudes or intentions rather than prohibition. We have heard it said that it is probably less the word itself than the intent; or that the intent of the user of a slur most often is to castigate the recipient explicitly; or that the power to hurt in a slur resides in the hatred and contempt that underlies its usage; or that a large part of the offensiveness of slurs comes from the fact that whoever utters one has to deliberately choose to use a term he knows will cause hurt and offense instead of one that is neutral.

Comments such as these are widespread; they represent two distinct strands of thought; one suggests the expression's meaning depends on what the speaker intends to say and the other suggests that slurs signal something about the speaker's mental state. The first idea can be captured in the slogan: "A word is only derogatory if I mean it to be." Those who endorse this view cannot be saying a word means just what its user intends, for we lack that kind of control over the meanings of words. What we can mean by our speech is much more constrained. Those who favor this view must also be taking features of the context of the conversation to play a prominent role in determining word meaning as well, so that a word is derogatory given certain features of its conversational context and the intent of the speaker. But how could that be correct? It certainly didn't work as an explanation for former Senator George Allen of Virginia. Recall his campaign for senator was derailed by his "playful" reference to an Indian American campaign

worker for the opposition. His stated intention (sincere or not) wasn't enough to change perceptions of his chosen expression's significance.

A second strand of thought exhibited in the comments suggests slurs express a hateful attitude. When someone slurs another, the speaker expresses an adversarial attitude towards the target and the target's social group. That attitude is not reported in the statement: when one slurs another he or she does not literally say "I hate Xs", but rather displays hatred towards Xs. It is as if an exclamation mark ends the slurring statement or the speaker is scowling.

This way of explaining slurs' offensiveness corresponds with the thought expressed by some that slurring says more about the person using the terms than it does about the target. But this view fails for various reasons. First, using slurs flippantly or without any feelings of hatred towards the target is still offensive. In fact, many find uses of slurs by out-group members offensive even if it is clear the speaker does not possess hateful attitudes toward the target.

This is illustrated by a scene in the movie "Rush Hour," starring Chris Tucker and Jackie Chan. During a visit to a bar Carter (Tucker) greets the bartender with a slur that, when shared among in-group members, can be taken as a term of endearment. Lee (Chan) follows Carter and essentially mimics what was said (including tone). Needless to say the greeting was not received well. But how could this be if the offense consisted in the attitude underlying its use? It is clear Lee possesses no hateful attitudes, yet his statement is still offensive.

Explaining offense in terms of attitudes also fails to account for the variation of offense among slurs. Presumably, people who use slurs to refer to others exhibit the same hateful attitudes toward a particular group no matter which expression is used. It is also the case that slurs used by members of subordinate groups towards members of dominant groups generally lack the intensity of slurs used by members of dominant groups towards the subordinate ones. This difference is unlikely to be explained by a difference in hateful attitudes underlying the use of those expressions. For these and other reasons, appealing to the underlying attitudes in slur use is insufficient to explain their offense.

Next, another alternative view that warrants attention is evidenced in comments like: what's offensive about slurs is the way that they institute an inference that all members of a group exhibit some negative property. For example, if there's a slur "z—" directed against members of a group X and alleging exhibition of the negative property F, then any use of "z—" means "...is a member of X and THEREFORE exhibits F." The key to both the meaning and the harmfulness of the use of the slur here is the unarticulated major premise: "All members of X exhibit negative property F." Though unarticulated, this is still part of the content of the slur, and it is more or less obvious why this prejudicial assertion is found harmful.

Whatever its merits and insights, we should reject this view because it too fails to explain the relevant data. First, the view as stated claims the inference instituted by the slur says all members of the targeted group exhibit the negative property. But we can imagine a bigot who, after being presented with evidence that the victim of his slur does not exhibit the negative property associated with the term,

still applies the word to his target. And secondly, for any given slur, what is the specific assumption one is licensed to infer? Here is a homework assignment: take a slur and try to figure out what everyone who hears it would infer. It is doubtful you will come up with a single idea that everyone who knows the term shares.

Another problem for this view is how it could explain why uses of slurs among in-group members generally fail to give rise to negative inferences. Think of reclaimed or appropriated uses of slur terms. These uses do not provoke offense among in-group users. The fan of inferentialism will need to provide a credible story about the differences between these scenarios.

Another worry others have expressed to us about our view is that if we try to explain the offensiveness of slurs only in terms of prohibition we won't be able to distinguish the offensiveness of ethnic slurs from that of profanities. For example, some have said to us that the use of prohibited ethnic slurs would have the same effect as profanities which end in "-er" and characterize a person by an act, typically sexual, which in most cases there is an understanding that the other person actually is not. The upshot is that there is a profound difference between profane attribution and ethnic slur attribution, which prohibition alone cannot capture.

It is appropriate for us to reiterate a point we mentioned earlier: Prohibitionism does not say slurs are devoid of content, only that their content does not explain variation in offense or appropriated uses. Further, the proposal does not deny that content can be a reason for prohibiting an expression. However, content is not a necessary feature for prohibiting a word; witness the barroom brawl the meaningless "mook" caused in Martin Scorsese's early film "Mean Streets."

One final challenge for the advocate of a content-based explanation: paraphrases of what a slur supposedly means do not match in offense. But why this mismatch if content explains the offense? Further, what content does a slur for some groups carry that others lack, such that the first is more offensive than the latter? These are considerations any serious view about slurs must address. To the best of our knowledge, only Prohibitionism even purports to offer an account of them.

References

- Lepore, Ernest. 2011. Speech and harm. *New York Times*.
 Lepore, Ernest, and Anderson, Luvell. Forthcoming. Slurring words. *Nous*.

Viewing the Study of Argumentation as Normative Pragmatics

Frans H. van Eemeren and Bart Garssen

Abstract In this chapter we explain that the pragma-dialectical approach to argumentation involves at the same time a pragmatic and a critical treatment of argumentative discourse. Starting from the speech act dimension of pragmatics, we indicate for each of the five components of the pragma-dialectical research program what a normative pragmatic approach amounts to. First we discuss the critical conception of reasonableness developed in the philosophical component of the research program, second the model of a critical discussion designed in the theoretical component, third the resolution-oriented analysis of argumentative discourse carried out in the reconstruction component, fourth the qualitative and quantitative research of the pursuit of reasonableness and effectiveness in argumentative discourse conducted in the empirical component, and fifth the development of adequate instruments for the analysis, evaluation and production of argumentative discourse in all sorts of argumentative practices undertaken in the practical component.

1 Introduction

The study of argumentation is often considered to be part of the discipline called *logic*, more in particular of *informal logic*.¹ In our contribution we would like to make clear that the study of argumentation can also be constructively viewed as being part of *pragmatics*, more in particular of *normative pragmatics*. In doing so

¹ See, among others, Johnson and Blair (2005), Johnson (2000), and Freeman (1991).

F. H. van Eemeren (✉) · B. Garssen
University of Amsterdam and ILIAS, Amsterdam, The Netherlands
e-mail: F.H.vanEemeren@uva.nl

we start from the theoretical perspective on argumentation that is commonly known as the *pragma-dialectical* approach to argumentation.² For the sake of simplicity, in discussing the pragmatic character of the pragma-dialectical approach we focus in this contribution in the first place on the speech act dimension of pragmatics.³ Our central question can therefore be specified as: can it be made clear that the study of argumentation can be constructively viewed as being part of (normative) pragmatics by pointing out that the pragma-dialectical approach to argumentation benefits in various respects significantly from taking a speech act perspective?

Referring to the pragma-dialectical approach as a speech act perspective does not do justice to it. Not only because this label narrows down the pragmatic aspect of this approach to one of its dimensions, but even more so because it undervalues another aspect of a more comprehensive theoretical framework. Next to the pragmatic aspect, the dialectical aspect representing the critical angle is the second vital characteristic of the pragma-dialectical approach. It is crucial of this approach that the complex speech act of argumentation is examined as part of a ‘critical discussion’ aimed at resolving a difference of opinion on the merits. This means that justice is done to the pragma-dialectical approach to argumentation only when the pragmatic aspect and the dialectical aspect are taken together. It is only then that this approach can be characterized as being part of normative pragmatics. In addition because in recent years the pragma-dialectical approach has been extended by taking also the rhetorical aspect of effectiveness into account in the theorizing about argumentation (van Eemeren 2010), doing justice to this approach requires starting from this extended version.

In our opinion, the decisive advantage of a speech act perspective over rival theoretical perspectives developed in the study of argumentation is, firstly, its capacity to provide a comprehensive coverage of all the moves that are made in argumentative discourse, and, secondly, its capacity to take systematically into account that argumentation is carried out through the use of ordinary language. We will take a closer look at the study of argumentation by sketching the five components of the pragma-dialectical research program distinguished in van Eemeren’s ‘Argumentation Studies’ Five Estates’ (1987b). The first component is the *philosophical* component, where the question has to be answered as to when it is reasonable to regard argumentation as acceptable. The second component is the *theoretical* component, where the favored philosophical conception of reasonableness is molded into a particular model of argumentation. The third component is the *reconstruction* component, where the theoretical model is used to find ways to get a better grasp on argumentative discourse. The fourth component is the *empirical* component, where the pluriform empirical argumentative reality is

² See van Eemeren and Grootendorst (1984, 1992, 2004), van Eemeren et al. (1993), (2007), (2009), and van Eemeren (2010).

³ We confine ourselves here to a speech act perspective which starts from the standard theory developed by Searle (1969, 1971, 1979).

systematically explored. The fifth component is the *practical* component, where the findings of the other components are integrated and employed to analyze and improve argumentative practice in different contexts and settings.

In a comprehensive research programme each of the five components should be somehow represented, but for its viability, the reconstruction component is indeed crucial. Connecting the philosophical and the theoretical components to the empirical and practical components prevents philosophizing and theorizing about argumentation from getting unrealistic and ethereal, and empirical research and practical advice from being unsystematic and *ad hoc*.

What do the five components look like in our approach? This question can be answered by means of some catchwords. In the philosophical component we take the position of a critical rationalist, in the theoretical component we commit ourselves to pragma-dialectics, in the reconstruction component we opt for a resolution-oriented analysis of argumentative discourse, in the empirical component we prefer a convincingness-centered description, and in the practical component we are in favor of a reflection-minded practice.

2 The Speech Act Perspective in the Philosophical Component

The philosophical component is about reasonableness with regard to the acceptability of argumentation. By adhering to a *critical* conception of being reasonable, in the philosophy of reasonableness that is chosen as the starting point, the dialectical idea of having a regulated critical discussion is made the basic principle of reasonableness. By linking the reasonableness of argumentation to problem-valid and conventionally valid rules for a critical discussion,⁴ the concept of reasonable argumentation is at the same time externalized, functionalized, socialized, and dialectified. What this means can be explained by pointing out that the study of argumentation is dealing with a subject matter that is not just an internal state of mind, but the verbal expression of attitudes and opinions (*externalization*); that does not derive its acceptability from someone's personal authority or sacrosanct origin, but from its contribution towards the achievement of a specific objective (*functionalization*); that does not depend on certain characteristics of a single individual, but on the interaction between two or more individuals (*socialization*); and, finally, that displays a quality which cannot be reduced to extremely absolute or relative concepts, as happens in geometrical or anthropological conceptions respectively, but is systematically connected with pursuing the critical aim of resolving a difference of opinion on the merits (*dialectification*).

⁴ For the notions of problem-solving validity and conventional validity see Barth and Krabbe (1982, pp. 21–22).

How does a speech act perspective facilitate the realization of these four meta-theoretical objectives? On a philosophical level this can be expounded by pointing out the functional character of argumentation rules in a critical discussion. These rules are not metaphysically based, but rest on their suitability for doing the job for which they are designed: the resolution of differences of opinion on the merits. Because of this, the rules should in principle be optimally acceptable to those whose first aim is to resolve their differences of opinion. This means that the rationale for accepting these rules as conventionally valid is, philosophically speaking, pragmatic. Philosophical pragmatists judge the acceptability of rules on the extent to which they appear successful in solving the problems they wish to solve. In fact, to them a rule is a rule only if it performs a function in the achievement of the objective for which the rule was designed (van Eemeren and Grootendorst 1988).

Apart from the critical dialectical conception of reasonableness, one could also choose for an anthropological rhetorical conception of reasonableness. According to anthropological approaches, argumentation can be considered reasonable to the extent that it is acceptable for the audience or auditorium it is aimed at. In the notion of *strategic maneuvering* introduced in the extended pragma-dialectical theory the dialectical idea of reasonableness and the rhetorical idea of effectiveness are combined (van Eemeren 2010). However, taking anthropological insights into account in analyzing and evaluating argumentative discourse, as happens in the extended pragma-dialectical approach, does not result in a new conception of reasonableness.

Speech act theory is *in optima* forma a pragmatic approach to language use: it describes the conditions which make it useful to distinguish between certain forms of language use because of their potential practical consequences. In this way it becomes clear which specific commitments people accept in their verbal behavior, and this makes it possible to explain exactly what it means to conform to the philosophical principle that you should be prepared to undertake certain obligations in a discussion. The speech acts which are most useful to all who share a certain goal—in this case: resolving a difference of opinion on the merits—are those that are in accordance with rules which have a degree of problem validity, so that they may become conventionally valid.⁵ The merits of a speech act perspective for a critical rationalist approach to the reasonableness of argumentation can also be explained in a more specific and less philosophical way by referring to the definition of argumentation as a speech act designed to justify or refute a proposition by convincing another person, who acts reasonably, of the acceptability of a positive or negative standpoint with respect to this proposition.

⁵ To acquire conventional validity, it must be clear which speech acts should be performed when and by whom. This has been shown in our model of a critical discussion by characterizing and locating the different types of speech acts which can play a specific role in the resolution of a difference of opinion (van Eemeren and Grootendorst 1984, pp. 95–118, 151–175). In this way a dialectical approach motivated by critical rationalism combines with an utilitarian pragmatist approach to language use.

By concentrating on the performance of a speech act as a means of verbally expressing an opinion, externalization is guaranteed. By specifying the purpose of this speech act, functionalization is likewise ensured. By addressing the speech act to another person, socialization is on its way. And by requiring the acceptability of the speech act for a reasonable person, the road to dialectification is paved.⁶

Defining argumentation as a speech act just like that, however, creates in at least three respects some kind of clash with Searle's standard theory. Firstly, unlike in speech acts such as asserting, requesting, and promising, in argumentation always more than one proposition is involved. Sometimes it appears that there is only one proposition, but on closer inspection it always turns out that in such cases one or more other propositions—which are genuine parts of the argumentation—have been left unstated.⁷ Secondly, unlike Searle's prototypes, the utterance of argumentation as a speech act always has a dual illocutionary or, as we prefer to call it, *communicative function*: besides functioning as argumentation it is also an assertion, a question, a form of advice, a proposal, or whatever. Thirdly, unlike most of Searle's examples, argumentation cannot stand by itself, but is always in a particular way linked to another speech act expressing a standpoint. If this specific relation is absent, any allusion to argumentation is unjustified.⁸

In order to solve the problem caused by these three differences, the standard theory of speech acts needs to be modified in such a way that it becomes applicable to larger units than single speech acts. This can be achieved by distinguishing between communicative functions at the sentence level on the one hand and communicative functions at some "higher" textual level on the other (van Eemeren and Grootendorst 1984, pp. 33–35).⁹ It is only at the higher textual level that the utterance of a speech act can have the communicative function of argumentation, whereas the utterances of speech acts such as assertions, requests, and promises are situated at the sentence level. The difference between speech acts at the sentence level and at the higher textual level can be made clear by referring to

⁶ The contribution of a speech act perspective to socialization and dialectification is formulated more cautiously than its contribution to externalization and functionalization. The reason for this is that addressing another person does not necessarily imply a discussion in which both parties play an active role, since trying to convince this person does not automatically lead to the idea of resolving a difference of opinion, but from a critical-rationalist point of view both are certainly moves in the right direction. This becomes even clearer when one realizes that the subject of research is everyday argumentation: in daily life argumentation is often part of a conversational setting of a more or less discussion-like character.

⁷ The unstated parts of an argumentation can be referred to as *unexpressed premises* (van Eemeren and Grootendorst 1992, pp. 60–72).

⁸ Of course, there can be other candidates: explanations, amplifications, elucidations, but certainly not argumentation.

⁹ There are also authors who, like we do, believe that besides illocutionary acts there are larger units of language use to be distinguished, but who have a rather different view of these larger units. See the discussion of the views of van Dijk and Quasthoff in van Eemeren and Grootendorst (1984, pp. 37–39).

elementary speech acts in the former case and to complex speech acts in the latter.¹⁰ Clearly, argumentation belongs to the second category.

Argumentation is aimed at convincing another person of the acceptability of a standpoint, but what is meant by that? One possible interpretation would be: attempting to evoke in that person the “feeling” of being convinced. However, from a critical-rationalist point of view this interpretation is undesirable, because it threatens the externalization by perceiving ‘being convinced’ as an internal mental state. As a consequence, what is gained at one end of the process would be lost at the other.

Fortunately, it is possible to avoid the danger of “internalization” by defining *convinced* as being prepared to accept the standpoint supported by the argumentation. Since accepting a standpoint can be expressed by the performance of a speech act in which the acceptance is stated, externalization is guaranteed. And, as a consequence, acceptance can be made subject to dialectical rules, which is not the case with being convinced if this concept is not externalized. Treating the acceptance of a standpoint by another person as the purpose of performing the complex speech act of argumentation means that this acceptance is seen as a perlocutionary or, as we prefer to call it, *interactional* effect of this speech act. In general, acceptance can be seen as an effect aimed for universally by all speech acts.

Although the standard theory of speech acts has not much to say about perlocutionary acts and effects, one cannot escape from getting the impression that these constitute a category as diffuse as diverse. It seems that all kinds of possible consequences of speech acts fall under the general heading of perlocutions: opening a window, quitting smoking, getting frightened, et cetera.¹¹ In our opinion, it is necessary to distinguish different kinds of effects upon the listener or reader which can be brought out by speech acts. Firstly, consequences intended by the speaker or writer should be distinguished from unintended consequences. Secondly, consequences brought about on the basis of understanding the speech act should be distinguished from consequences where this is not the case. Thirdly, consequences brought about by way of a rational decision by the listener should be distinguished from consequences which take place without any control on his part.¹² Only the first elements of these three oppositions lend themselves to systematic research in a theoretical framework of speech acts. This also applies to the acceptance of argumentation. It is intended by the speaker, it requires recognition of the complex speech act as argumentation, and it depends on rational considerations of the listener. Undoubtedly, all kinds of consecutive consequences can also occur, but these are beyond the scope of speech act theory. The next question

¹⁰ For practical reasons we do not go into the exact relations between sentences and propositions now. We just equate speech acts which consist of the expression of one proposition with one sentence speech acts.

¹¹ See Searle’s description of perlocutionary acts (1980, p. 25).

¹² For a discussion of these views, see van Eemeren and Grootendorst (1984), pp. 23–28 and van Eemeren (2010, chapter 2).

is, of course: when exactly will it be reasonable for the listener to accept a standpoint on the basis of the speaker's argumentation? Answering this question requires the formulation of rules for critical discussion and leads to the next Component, in which the socialization and dialectification of our critical-rationalist philosophy of argumentation get a particular theoretical shape.

3 The Speech Act Perspective in the Theoretical Component

In the theoretical component our preference for a critical-rationalist philosophy is molded into a pragma-dialectical theoretical model for argumentation. The dialectical aspect consists in there being two parties who attempt to resolve a difference of opinion by means of a methodical exchange of moves in a discussion, whereas the pragmatic aspect is represented by the description of the moves in the discussion as speech acts. In order to resolve a difference of opinion on the merits by means of argumentation, the language users involved must observe a number of rules. If they jointly attempt to resolve the difference by engaging in an interaction of speech acts according to these rules, then their discourse can be referred to as a critical discussion. The model of a critical discussion van Eemeren and Grootendorst presented in *Speech Acts in Argumentative Discussions* (1984) and *A Systematic Theory of Argumentation* (2004) explains which rules apply to the distribution of speech acts in the various stages the resolution of a difference of opinion should pass through. As an ideal model the model of a critical discussion reproduces only aspects that are relevant to the resolution of a difference of opinion: it provides a set of instruments to grasp reality and to determine the extent to which practice corresponds with the requirements for the resolution. In this respect, the model not only links theory to practice, but also combines normative and descriptive aspects.

A critical discussion passes through four discussion stages: the confrontation, the opening, the argumentation, and the concluding stage. In the confrontation stage the difference of opinion is externalized: it becomes clear that there is a standpoint which meets with doubt or contradiction, so that a difference of opinion arises. In the opening stage agreements are reached concerning the manner in which the discussion is to be conducted; the parties have to determine whether there is sufficient common ground to serve as a starting point. Only if there is such a common point of departure, it makes sense to undertake an attempt to remove the difference of opinion by means of argumentation. In the argumentation stage (as is obvious from the terminology) argumentation is advanced and reacted to. By definition, the purpose of adducing argumentation is to overcome possible doubts about the standpoint. And by reacting to the argumentation put forward, it can be made clear that this attempt has not yet succeeded in a fully satisfactory way. Finally, in the concluding stage the result of the discussion is established. It is only

if both parties agree on the result that the difference of opinion can be regarded as having been resolved. The model specifies which speech acts the participants in a critical discussion have to perform at the four stages in order to contribute to the resolution of the difference of opinion. The model is elaborated in rules for critical discussion that specify who may perform what type of speech act with what role at what stage of the discussion. These rules prescribe when the discussants are entitled, or indeed obliged, to perform a particular speech act.

Starting from Searle's taxonomy of speech acts (1979, pp. 1–29), it can be said that all kinds of assertives can be used to express standpoints and argumentation, and to establish the results of the discussion. The use of directives is restricted to challenging somebody to defend his standpoint and requesting him to put forward argumentation in support of his standpoint. Commissives are used to accept (or not accept) a standpoint or argumentation and to agree upon the division of dialectical roles in the discussion and upon the discussion rules. *Language usage declaratives*, finally, can be helpful in avoiding all kinds of misunderstandings.¹³ It should be noted that no other types of declaratives and no types of expressives are listed in the model, because they do not contribute in a direct way to the process of resolving a dispute.¹⁴

The starting point underlying the model is that the participants in a critical discussion have the intention of jointly resolving the difference of opinion. That is, the discussant whose arguments do not prove strong enough must be prepared to abandon his position, and the discussant whose doubts or objections have been overcome by the argumentation must be prepared to drop them. This is the basis of the reasonable attitude that is a prerequisite to the conduct of a critical discussion. The rules of the ideal model tell us what such an attitude amounts to and constitute a code of conduct for rational discussants who want to act reasonably.¹⁵

Apart from incorporating the critical-rationalist ideal into a coherent theoretical framework, the model serves both descriptive and normative purposes. With regard to the former, more will be said in the Reconstruction Component and the Empirical Component, with regard to the latter in the Practical Component. In the Theoretical Component, the importance of the model is shown for the analysis of the fallacies which traditionally pose theoretical problems, especially for the so-called informal fallacies for which the logically based approach does not provide a satisfactory account.¹⁶

¹³ The notion of language usage declaratives, which comprises speech acts such as defining, precisating, amplifying, explicating, and explicitizing, is introduced in van Eemeren and Grootendorst (1984, pp. 109–110).

¹⁴ The distribution of the various types of speech acts in the stages of a critical discussion is discussed in van Eemeren and Grootendorst (1984, pp. 95–118) and van Eemeren and Grootendorst (2004, pp. 62–68).

¹⁵ The rules of the code of conduct are discussed in van Eemeren and Grootendorst (2004, pp. 187–196).

¹⁶ Cf. Hamblin (1970), van Eemeren and Grootendorst (1992, pp. 102–106).

Fallacies can be analyzed as violations of the rules of the model. Any violation of the rules may have the consequence that the difference will not be resolved.¹⁷ Although the consequences of violating the rules may vary in their seriousness, every violation is a potential threat to the successful conclusion of the discussion, regardless of which party is responsible and regardless of the stage of the discussion at which it occurs. Seen in this perspective, all violations of the rules are incorrect moves in a critical discussion. These incorrect moves correspond roughly to the various kinds of traditional fallacies. In our pragma-dialectical theoretical conception, the term *fallacy* is reserved for speech acts which are derailments of strategic maneuvering because they hinder in any way the resolution of a difference of opinion by means of a critical discussion by violating a rule for critical discussion (van Eemeren 2010, pp. 193–196). Thus, fallacies are systematically connected with the rules for critical discussion, and our treatment of the fallacies is linked to a particular theoretical approach. In this conception, committing a fallacy is not necessarily tantamount to unethical conduct, but it is wrong in the sense that it frustrates efforts to arrive at the resolution of a difference of opinion.

How does a speech act perspective facilitate the development of a pragma-dialectical model for critical discussion? The main point is that by formulating the rules like this, the normative rules for carrying out speech acts in a critical discussion are linked in a natural way to the descriptive conditions of performing elementary and complex speech acts. In turn, these are closely connected with all kinds of more general descriptive rules for conducting everyday discourse and conversations, such as the rules referred to in Grice's maxims, and rules for turn-taking as described by conversation-analysts.¹⁸ To put it bluntly, our normative discussion rules can partly be seen as dialectical regulations of rules that already apply in ordinary conversations. Of course, this is a simplification but it draws attention to the fact that proposing normative rules for critical discussions has closer ties with reality than some people think. The conditions for the performance of speech acts as described by Searle serve as a link, so to speak, between the ideal and the real. In order to be able to fulfill this task of linking optimally, it is necessary, however, for the speech act conditions to be formulated in a way that makes it possible to reveal the similarities with conversational rules such as Grice's maxims. van Eemeren and Grootendorst integrated these two sets of conditions and rules in *Argumentation, Communication and Fallacies* (1992).

First they replaced the Gricean Co-operative Principle by a general Principle of Communication, requiring language users to be clear, honest, efficient, and to the point. Starting from this rule they formulated, as an alternative to the Gricean maxims of Manner, Quality, Quantity and Relation, the following general rules of communication: (1) "Perform no incomprehensible speech acts", (2) "Perform no

¹⁷ Having the required attitude and observing all the rules of the model do not guarantee that the participants in a discussion will actually bring their difference of opinion to a successful resolution, only that a number of preconditions for doing this have been met.

¹⁸ The conversational maxims are introduced in Grice (1975), the rules for turn-taking are discussed in Levinson (1983) and Edmondson (1981).

insincere speech acts”, (3) “Perform no unnecessary speech acts”, (4) “Perform no pointless speech acts”, and (5) “Perform no new speech acts that are not an appropriate sequel or reaction to preceding speech acts”. The rules of communication correspond to a large extent to Grice’s maxims. The advantage of this approach is that the maxims are now formulated as rules for the performance of speech acts.

The first rule of communication roughly corresponds to Searle’s propositional content condition and the essential condition. The second rule corresponds to his sincerity condition, the third and fourth rules correspond to his preparatory condition, whereas the fifth rule does not have a counterpart in his conditions. By integrating Gricean maxims with Searlean speech act conditions both are enhanced. Compared to the maxims, these rules are more specific as a consequence of their connection with the Searlean conditions, and they are more general because they are no longer restricted to assertions, as they are with Grice. The speech act conditions also profit from it, because it is now shown that the conditions for different speech acts are in fact specifications of more general rules of communication.¹⁹

This synthesis of Searlean and Gricean insights reveals the heterogeneous character of the original set of speech act conditions. Although Searle makes no difference in their importance, in the revised version it becomes clear that a distinction should be made between the propositional content and essential conditions on the one hand, and the sincerity and preparatory conditions on the other. This can be demonstrated by looking at the consequences of violations of the various conditions. In the case of violation of the first two, no recognizable speech act has been performed, whereas in the case of violation of the second two, though the performance of the speech act is not quite successful or “happy” in the full sense, a recognizable speech act *is* performed. This crucial difference can be accounted for by realizing that there is a correspondence between the propositional content condition and the essential condition on the one hand, and Grice’s maxim of Manner (“Be perspicuous”) and our first communication rule (“Perform no incomprehensible speech acts”) on the other. Violating these two conditions damages the recognizability of a speech act, whereas violating one of the two others affects its correctness by causing insincerity, inefficiency or irrelevancy. In order to express this difference terminologically, we refer to the first two as *identity conditions*, and to the second two as *correctness conditions* (van Eemeren and Grootendorst 1984, p. 41).

By integrating Searlean and Gricean insights in our rules of communication, we think an important step has been made towards a comprehensive theory of everyday communication and interaction. Of course, much remains to be done. For example, all useful kinds of concepts from conversation analysis have to be

¹⁹ Searle does not believe that all speech act conditions are specifications of Grice’s maxims, because some of them (such as the essential condition and the sincerity condition) are internal to specific kinds of speech acts (1980, pp. 22–23). We do not agree with him on this point.

incorporated in the theoretical framework. Up to now, many conversation analysts have shown some reluctance to make use of speech act theory—or any other theoretical framework at all for that matter. As a consequence, conversation analysis lacks a firm theoretical foundation. This lends most of the results of the analyses an *ad hoc* character and makes them less interesting. It also makes it more difficult to carry out the required integration. The important notion of ‘adjacency pair’, for instance, could gain from taking theoretical insights from speech act theory into account. Not only should speech act theory become more conversation-oriented but conversation analysis should also become more speech act-oriented.²⁰

People engaged in argumentative discourse are characteristically oriented towards resolving a difference of opinion on the merits and may be regarded as committed to norms instrumental in achieving this purpose, maintaining certain critical standards of reasonableness and expecting others to comply with the same standards. At the same time, however, these people are also, and perhaps even primarily, interested in resolving the difference of opinion effectively in favor of their case, i.e. in agreement with their own standpoint or the position of those they represent. In examining actual cases of argumentative discourse, the conceptualization of argumentation as a communicative and interactional (speech) act complex aimed at resolving a difference of opinion for a reasonable judge by advancing a constellation of reasons the arguer can be held accountable for as justifying the acceptability of the standpoint(s) at issue therefore needs to be complemented by taking, besides the presumption of reasonableness involved in argumentation, also the quest for effectiveness into account.²¹

In all stages in every argumentative move that is made the aims of maintaining reasonableness and achieving effectiveness go together. It is the simultaneous pursuit of these two aims that takes place in strategic maneuvering. Strategic maneuvering refers to the efforts made in all moves carried out in argumentative discourse to keep the balance between reasonableness and effectiveness (van Eemeren 2010).

In argumentative discourse, strategic maneuvering manifests itself in three aspects of the moves that are made, which can be distinguished only analytically: ‘topical choice’, ‘audience adaptation’, and ‘presentational design’. Topical choice refers to the specific selection that is made from the topical potential—the set of dialectical options—available at a certain point of the discussion, audience adaptation involves framing a move in a perspective that agrees with the intended audience, and presentational design concerns the selection the speaker or writer

²⁰ Searle admits that conversation analysis may be useful to the analysis of conversations, but he also emphasizes the crucial differences between speech act theory and conversation analysis (1986, pp. 7–19).

²¹ It should be noted that “effectiveness” is not completely synonymous with “persuasiveness.” For one thing, aiming for effectiveness is not limited (as is persuasiveness) to those parts of the argumentative discourse (arguments) that can be reconstructed as belonging to the argumentation stage but applies also to the parts of the discourse that belong to the other stages (van Eemeren 2010).

makes from the existing repertoire of presentational devices. In their strategic maneuvering aimed at steering the argumentative discourse their own way without violating any critical standards in the process, both parties may be considered to be out to make the most convenient topical selection, to appeal in the strongest way to their audience, and to adopt the most appealing presentation.

A clearer understanding of strategic maneuvering in argumentative discourse can be gained by examining how in argumentative practice the rhetorical opportunities available in the dialectical situation are exploited. Each of the four stages in the process of resolving a difference of opinion on the merits is characterized by having a specific dialectical objective. Because, as a matter of course, the parties want to realize these objectives to the best advantage of the position they have adopted, every dialectical objective has its rhetorical analogue. Because in each discussion stage the parties are out to achieve the dialectical results that serve their rhetorical purposes best, in each stage the rhetorical goals of the participants in the discourse will be dependent on—and therefore run parallel with—their dialectical goals. As a consequence, the specifications of the rhetorical aims that may be attributed to the participants must take place according to dialectical stage. This is the methodological reason why the study of strategic maneuvering that we propose boils down to a systematic integration of rhetorical insight in a dialectical framework of analysis.

4 The Speech Act Perspective in the Reconstruction Component

In the reconstruction component the theoretical model is used to get a better grasp of the empirical reality of argumentative discourse. Following on from our speech act perspective, we opt in the reconstruction component for an orientation towards the resolution of differences of opinion. A resolution-oriented reconstruction means that one tries to get a clearer picture of everyday argumentation by reconstructing argumentative discourse in terms of a critical discussion. Since the reconstruction starts from an ideal model, it can be called a *normative reconstruction*. The main task of such a reconstruction is to connect the real to the ideal by confronting empirical reality with a theoretical model. The reconstruction component is determined by the theoretical component, which in turn is determined by the philosophical component, and is instrumental in molding both the empirical and the practical components. Therefore, the reconstruction component has a pivotal function which is really vital to the study of argumentation.

In a resolution-oriented normative reconstruction, everyday discourse is treated as a discussion which solely aims at resolving a difference of opinion. This does not mean that every piece of discourse is automatically regarded *in toto* as a critical discussion. We first check to what extent it is possible to carry out the analysis as if the discourse were a critical discussion. The normative

reconstruction represents the angle of approach which is, seen from a pragma-dialectical theoretical perspective, the most relevant and the most illuminating.²²

A resolution-oriented normative reconstruction entails some specific operations which amount to carrying out four dialectical transformations. The first transformation is called dialectical *deletion*, and involves a selection from the discourse as it occurred in reality. Elements that are relevant for the resolution process are included, others are left out. The second transformation is called dialectical *addition*, and involves a process of completion. This is a matter of making implicit elements explicit: supplying unexpressed premises, et cetera. The third transformation is called dialectical *permutation*, and involves an ordering and rearranging of elements in the discourse. In contrast to a purely descriptive record, a normative reconstruction need not necessarily follow the order of events in time, or the linear presentation. It rearranges the elements in such a way that the process of resolving a difference of opinion as described in our ideal model, is most directly reflected. The fourth transformation is called dialectical *substitution*, and involves a clear and uniform notation of the elements which fulfill the same dialectical function in the discourse. Ambiguities and vagueness are replaced by unambiguous and clear standard formulations, whereas different formulations of elements with the same function are reduced to one single standard formulation.²³

Normative reconstruction of ordinary discourse from a resolution-oriented perspective depends to a large extent on the model for a critical discussion, but a speech act perspective can contribute substantially to the carrying out of these dialectical transformations. It is of special significance to the transformations of addition and substitution. With regard to the addition of implicit elements, such as unexpressed premises, a speech act perspective has already proven its usefulness.²⁴ Therefore, it may be more interesting to look closer at the transformation of substitution. This pertains, for example, to the replacement of a statement in which the communicative function of a standpoint remains implicit by an explicit standard formulation. In speech act theory it is a recognized fact that in ordinary discourse the communicative function—or, as Searle calls it, *illocutionary force*—of a speech act is, as a rule, not expressed explicitly. In practice, however, this does not normally present much of a problem. Sometimes verbal indicators give an unambiguous hint towards the desired interpretation. In the absence of such indicators, the verbal and non-verbal context, together with logical and pragmatic inferences that can be made and background information of various kinds, provide

²² Naturally, other angles of approach are also possible. A Freudian psychological analysis would undoubtedly be able to produce yet other interesting results—but there again the same sort of restriction would apply. Things that appear as relevant from one angle remain out of sight when regarded from another. However, one angle of approach need not necessarily preclude another. The same conversation can very well be examined and analyzed from different angles, though it is a good idea to make sure not to confuse the different perspectives.

²³ For the pragma-dialectical method of reconstruction see van Eemeren (1986, 1987a), van Eemeren and Grootendorst (1990), van Eemeren et al. (1993), and van Eemeren et al. (2012).

²⁴ For the reconstruction of unexpressed premises see van Eemeren and Grootendorst (1992).

sufficient clues.²⁵ So, a speech act perspective points to a phenomenon which is inherent to everyday language use, and it also makes it clear that a critical discussion is no exception.

All kinds of indirectness, however, can pose a problem. To demonstrate this, the following piece of discourse can serve as an example:

Let's take a cab. You don't want to be late for the show, do you?

Carrying out a resolution-oriented normative reconstruction, we would say that we have argumentation here all right, but where is the standpoint and what constitutes the argumentation? The standpoint is to be found in the first sentence, the argumentation in the second. However, the first sentence clearly has the communicative function of a *proposal*, and the second of a *question*, and a normative reconstruction should take this into account. How can we justify the attribution of the dialectical function of a standpoint to the first sentence, and that of an argumentation to the second?

Performing a proposal presupposes that the speaker himself believes it to be a *good* proposal. According to the preparatory conditions for the performance of a proposal, the speaker wants it to be accepted by the listener otherwise his proposal would be pointless. One way to get the proposal accepted by the listener, would be to show that it is in his interest. By asking rhetorically whether the listener wants to be late for the show, the speaker indirectly provides a reason that is possibly conclusive for the listener: he surely knows very well that the listener does *not* want to be late (and it is understood that not taking a cab would cause this unwanted effect).²⁶ By adding the rhetorical question to his proposal, the speaker tries to resolve a potential difference of opinion with the listener in advance. This explains how the speaker's proposal can be transformed into the standpoint that it is wise to take a cab, and his rhetorical question into the argumentation that otherwise they will be late for the show (which is undesirable). Although more could be said about this reconstruction, it suffices for showing the merits of a speech act perspective in helping to get the transformation of substitution carried out properly. Without speech act theory, no satisfactory explanation can be given.

An important advantage of a speech act perspective is that it emphasizes that argumentation in ordinary discourse is not always about assertive statements, nor does argumentation always consist of assertive statements. Argumentation can pertain to all kinds of speech acts and it can be presented by means of all kinds of speech acts. This is really a big step forwards compared to the many argumentation theorists who restrict argumentation to assertions which can be considered true or false. A speech act perspective not only shows that this is incorrect, but it can also

²⁵ For indicators of argumentative moves see van Eemeren et al. (2007).

²⁶ The sentence in parentheses refers (by convention) to the unexpressed premise in the argumentation.

systematically account for the diversity by turning to the correctness conditions of the speech acts involved.²⁷

The example also confirms that it is necessary to modify the standard theory of speech acts in several respects. Firstly, in the standard theory both identity and correctness conditions are formulated, though not always consistently, from the speaker's perspective, whereas it is clear from the example that it is necessary to differentiate between the correctness of a speech act from the speaker's point of view and the correctness from the listener's point of view. Seen from the first perspective, it is sufficient that the speaker believes that his proposition is in the interest of the listener, but seen from the second perspective it is also required that the listener thinks likewise. Only if the latter is actually the case, can the proposal be acceptable to him. So the correctness conditions have to be formulated from both a speaker's perspective and a listener's perspective (and the same applies to the identity conditions).

Secondly, the Communication Principle underlying the rules of communication as specified by the speech act conditions, has to be complemented by an Interaction Principle ("Perform no speech acts which are not correct and which are not acceptable to the listener") which can account for the fact that when performing a speech act the speaker is not only supposed to believe that this speech act is correct from his own point of view, but also that it is acceptable to the listener to whom it is addressed. Otherwise, it would remain unexplained what it means to give a reason in support of a proposal, as happens in the example. As a rule, every speech act presupposes its own acceptability. This is accounted for in the Interaction Principle, which enables us also to explain why speakers, even when the listener does not ask for it, take the effort to establish the acceptability of their speech acts by putting forward direct or indirect arguments.²⁸

The ideal model and the speech act perspective provide together the basis for an adequate a normative reconstruction of argumentative discourse. The reconstruction is greatly helped by the inclusion of all kinds of other factors such as knowledge of the word. Taking the strategic maneuvering that takes place in the argumentative discourse into account is also helpful to reach a fully justified reconstruction of all speech acts involved. A reconstruction of the strategic maneuvering lays bare the functional design of the discourse in aiming for

²⁷ See Jackson (1985) and van Eemeren (1986). As a matter of fact, one of Austin's reasons to develop a theory of speech acts refers to the simplistic dichotomy between truth-functional statements (which make sense) and non-truth-functional statements (which are considered to be metaphysical) made by logical positivists. According to Austin, performative utterances ("I hereby declare..." et cetera) are both non-truth-functional and non-metaphysical. Of course, it would be a serious philosophical setback to return to the age of hard-boiled positivism.

²⁸ The Interaction Principle reflects in its formulation Sperber and Wilson's Principle of Relevance: "Every act of ostensive communication communicates the presumption of its own optimal relevance" (1986, p. 158). There are important differences, however. Firstly, the Interaction Principle is more general than the Principle of Relevance, secondly, it is formulated in terms of speech act theory, whereas Sperber and Wilson reject speech act theory completely (1986, pp. 243–254).

effectiveness within the boundaries of reasonableness. Therefore, in some cases it may provide decisive reasons for going for the one reconstruction rather than for the other. In addition, a reconstruction of the strategic maneuvering can strengthen the justification of the analysis because it may provide extra support for the likelihood of a certain analysis.

5 The Speech Act Perspective in the Empirical Component

In the empirical component the various manifestations of argumentative reality are explored systematically by examining to what extent the insights gained in the philosophical, theoretical, and reconstruction components are reflected in the argumentative moves that are carried out in everyday argumentative discourse. Starting from a critical-rationalist philosophy, a pragma-dialectical theory, and a resolution-oriented reconstruction, a convincingness-centered description of the distinctive properties of everyday argumentative discourse must be carried out. This means that pragma-dialectical empirical research should be aimed at answering questions concerning how the way in which in ordinary situations ordinary arguers go about resolving their differences of opinion relates to the procedure for resolving a difference of opinion on the merits portrayed in the model of a critical discussion.²⁹

In order to resolve a difference of opinion on the merits it must be clear to the people involved that there is a difference, and what the difference exactly amounts to. They also have to recognize the communicative function of the speech acts that are performed in argumentative discourse as specific argumentative moves—expressions of doubt, criticism, argumentation etc. (van Eemeren et al. 1989). In principle, the way in which they evaluate the reasonableness of the performance of these argumentative moves also needs to correspond with the norms incorporated in the rules for critical discussion (van Eemeren et al. 2009). There are, of course, a great many other possibilities for empirical research, but these topics seem of primary importance. The way in which they are tackled can illustrate the important role the speech act perspective can play in this component.

How differences of opinion come into being and develop in everyday situations is subject to conversation-analytical research. What kind of speech acts do ordinary arguers use in expressing their differences of opinion and dealing with them? How do they make clear that there is a difference of opinion and how do they try to avoid getting involved in a dispute? How do they try to resolve or settle their differences? How does the way they go about compare with the ideal model of a critical discussion and the rules prescribing in what way exactly a difference

²⁹ In the pragma-dialectical empirical research it has to be made sure that the preliminary 'higher order' conditions for conducting a critical discussion have been fulfilled, so that the psychological state of the arguers, socio-political power relations between them etc. do not allow for alternative explanations of the research results. See van Eemeren and Grootendorst (2004, pp. 189–190).

of opinion should start and should be resolved? How can the differences that are found be explained by pragmatic knowledge and insight about argumentative reality? According to the model of a critical discussion, for instance, in the confrontation stage an antagonist of a standpoint must state his doubts clearly and unambiguously but in an informal situation doing this can be 'face-threatening' for both parties. It also creates a potential violation of the 'preference for agreement' governing normal conversation.

The general question is: how do ordinary language users try to combine meeting the dialectical requirements of reasonableness formulated in the ideal model with meeting the rhetorical requirements of effectiveness imposed by the argumentative reality in which they operate? Though much more could be said about this, we content ourselves here for the time being with the answer that in the strategic maneuvers that are made more often than not implicit and indirect speech acts are performed. Qualitative empirical research should make clear how exactly these maneuvers are carried out in practice and what precisely the strategic function of implicitness and indirectness is. In this kind of empirical research a speech act perspective is most helpful. The same applies to the quantitative empirical research aimed at tracing the factors influencing the recognition of argumentation. A crucial question then is whether ordinary language users are capable of recognizing argumentation which is not explicitly presented in that way. In this research it is interesting to look in particular at the various kinds of implicitness and indirectness in presenting argumentation, because this sort of argumentation is expected to be more difficult to recognize than its explicit and direct counterparts. The experiments we have carried out so far suggest that it is indeed the case that the presence of verbal indicators does significantly facilitate the ease of recognition of argumentation and in the absence of sufficient contextual or other clues indirect presentations do indeed pose more problems.³⁰

The main point we want to make about empirical research in referring to these examples is that starting from a speech act perspective makes it possible, as we have done, to start a coherent set of research projects covering the most pertinent empirical problems. A speech act perspective provides not only the necessary link between the various components of the research program, but also selects the empirical research problems and determines the order of priorities.

For a full and systematic description of the distinctive properties of argumentative reality it is necessary to take account of the possibilities of strategic maneuvering in the specific macro-contexts in which argumentation occurs. Due to the context-dependency of communicative practices, the possibilities for strategic maneuvering are in argumentative reality to some extent determined by the institutional preconditions prevailing in the communicative practices concerned. This makes it necessary to situate the empirical research of strategic maneuvering in the macro-context of the 'communicative activity type' in which the argumentative discourse takes place.

³⁰ See van Eemeren et al. (1989).

Communicative activity types are conventionalized practices whose conventionalization serves, through the implementation of certain ‘genres’ of communicative activity, the institutional needs prevailing in a certain domain of communicative activity. Communicative activity types belong to specific communicative domains, such as the political domain (e.g. plenary debate in parliament, political interview, Prime Minister’s Question time), the legal domain (e.g. court proceedings, arbitration), the scientific domain (e.g. peer review, conference paper) and the medical domain (medical advertisement, health brochure, medical consult). In principle, each of these communicative activity types has its own institutional goals and conventions, which are instrumental in realizing the institutional point of the communicative activity type concerned (van Eemeren 2010, pp. 140–143). It should be examined empirically which preconditions these institutional goals and conventions impose on the performance of speech acts in the various communicative activity types.

The extended pragma-dialectical theory provides the analyst with the tools to give a more refined, accurate and comprehensive analytic and evaluative account of argumentative reality than could have been given using only the dialectical tools of standard pragma-dialectics. With the help of the notion of strategic maneuvering it becomes possible to research argumentative discourse empirically in such a way that not only the dialectical dimension pertaining to its reasonableness is taken into account, but also the rhetorical dimension pertaining to its effectiveness (van Eemeren 2010). It should be clear, however, that extended pragma-dialectics does not provide an empirical model of the various ways in which in real-life argumentative discourse ordinary arguers try to make effective argumentative moves within the boundaries of dialectical reasonableness.³¹ This is because the notion of strategic maneuvering is incorporated in a theoretical model with a normative character, which is not a tool for describing the argumentative behavior of ordinary arguers and their pursuit of effectiveness empirically. Strictly speaking, a normative model cannot be put to the test empirically, since the model can neither be falsified nor confirmed by means of empirical data. This does by no means mean, however, that from an empirical point of view the model is useless. On the contrary: the extended model makes it possible to derive theoretically motivated hypotheses about how the speech acts performed in ordinary argumentative practice can contribute to the rational persuasiveness of argumentative discourse that is usually called *convincingness*.

6 The Speech Act Perspective in the Practical Component

In the practical component, the various findings of the other components are integrated and employed to value and improve argumentative practices. Philosophical and theoretical principles are necessary to formulate an ideal, the results

³¹ For the differences between *effectiveness* and *persuasiveness* see van Eemeren (2010), p. 39.

of empirical research tell us to what extent reality is distinct from this ideal, and normative reconstruction help us in appreciating and bridging the differences. A critical-rationalist philosophy and a pragma-dialectical theory, combined with a resolution-centered reconstruction, stimulate a reflection-minded practice. This amounts to the promotion of critical discussion and the reflection on all sorts of argumentative practices by supplying adequate instruments to analyze, evaluate and produce them.

Reflection should result in the first place in more sophisticated insight in the problems in specific argumentative practices. Such insight can only come about if all complications which are inextricably connected with the kinds of argumentative realities concerned are acknowledged right from the outset and incorporated in the reflection process. Again, a speech act perspective turns out to be very useful. This can be demonstrated by comparing the analysis and evaluation of argumentation resulting from a speech act perspective with the analysis and evaluation resulting from a formal logical perspective. Then it transpires immediately that although logic can play an important role in carrying out analytical as well as evaluative tasks its role should not be exaggerated, as a great many practical textbooks tend to do (van Eemeren and Grootendorst 1990).

The main advantage of a speech act perspective is that certain crucial characteristics of ordinary language use, such as implicitness and indirectness, are duly taken into account. In a logical perspective, they are either completely ignored or treated as incidental pitfalls which have to be disposed of as soon as possible. At best, some useful *ad hoc* observations are made. Of course, this does not come as a surprise, since one hardly expects to solve practical problems if they are not first recognized theoretically.

By bringing in pragmatic considerations, a speech act perspective does justice to the functionality of language, but introduces also a considerable dose of uncertainty. The analysis and evaluation of argumentation in ordinary discourse always have an open and conditional character and no ready-made procedures are available. The alternative to having such procedures boils down to attempting to gain as much insight as possible in the communicative and interactional factors that play a part in argumentative discourse. On this basis guidelines can be drawn up to facilitate and improve the performance of analytical and evaluative tasks.

The difference between a pragmatic speech act perspective on argumentation analysis and evaluation and a logical perspective can be illustrated by taking the identification of fallacies as an example. As far as the analysis of fallacies is concerned which do not fit the standard definition of fallacies as invalid arguments which seem valid, the logical approach can only offer an *ad hoc* explanation which has in fact nothing to do with logical validity. In addition, a logically oriented method for detecting fallacies is in practice usually very hard to apply.³² Our pragma-dialectical approach to fallacies brings all kinds of violations of discussion

³² See Hamblin (1970), van Eemeren and Grootendorst (1987), Grootendorst (1987), and van Eemeren et al. (2011).

rules into the analysis, not just the ‘logical’ errors. By defining fallacies as possible threats to the resolution of a difference of opinion on the merits, it can in a great many more cases be explained why a certain argumentative move is fallacious.

Of course, the pragma-dialectical discussion rules do not provide a set of simple tricks that only have to be learnt by heart for a successful practice. The rules can only be applied to discussions which are designed to resolve differences of opinion. If that is not the case, there are no fallacies in the dialectical sense. However, borderline cases do occur. This is the reason why the identification of fallacies is always conditional: fallacies are only fallacies if they occur in a discourse which can be regarded part of a critical discussion.³³ Because the language used in discussions is usually implicit, it is not always possible to say with certainty whether or not any of the discussion rules has been broken. This has nothing to do with the analytical instruments or the way they are used but is a direct consequence of the nature of the subject under investigation. In a speech act perspective, the openness of the analysis and evaluation is not treated as a disadvantage, but as a fact of life which should receive the attention it deserves by paying careful attention to the way in which the speech acts performed are to be reconstructed.

In pragma-dialectics the rules for critical discussion provide general norms for reasonableness. Any move that is an infringement of any of these rules, whichever party performs it and at whatever stage in the discussion, is a possible threat to the resolution of a difference of opinion on the merits and must therefore—and in this particular sense—be regarded as fallacious. In the pragma-dialectical approach a fallacy is thus a hindrance or impediment to resolving a difference of opinion on the merits and the specific nature of a particular fallacy depends on the way in which it interferes with the resolution process. This theory of fallacies proves to be theoretically and practically adequate, but what is still lacking are the specific *criteria* by which argumentation as it occurs in practice should be evaluated. We think that the pragma-dialectical theory of argumentation can be remedied in this respect by enriching it by the inclusion of insight from rhetoric, as has happened in extended pragma-dialectics.

Although in strategic maneuvering the pursuit of dialectical objectives can well go together with the realization of rhetorical aims, this does not automatically mean that in practice the two objectives will always be in perfect balance. If a party allows his commitment to a critical exchange of argumentative moves to be overruled by the aim of persuading the opponent, we say that the strategic maneuvering has got ‘derailed’. Such derailments occur when a rule for critical discussion has been violated. Because derailments of strategic maneuvering always involve violating a rule for critical discussion, they can be identified with the wrong moves in argumentative discourse designated as *fallacies*. Viewed from

³³ In case of irresolvable doubt, the ‘strategy of maximal reasonable interpretation’ is recommended (van Eemeren 1987a). Fortunately, taking the strategic maneuvering into account makes it often unnecessary to take refuge to such a ‘maximal’ strategy because including the strategic maneuvering that takes place in the considerations leads to a definitive solution.

this perspective, fallacies are derailments of strategic maneuvering that involve violations of critical discussion rules.

Each mode of strategic maneuvering has, as it were, its own continuum of sound and fallacious acting and the boundaries between legitimate and illegitimate argumentative acting are not in all cases immediately crystal clear (van Eemeren 2010, ch. 7). More often than not, fallacy judgments are in the end contextual judgments that depend on the specific circumstances of situated argumentative acting. The criteria for determining whether or not a certain norm for critical discussion has been violated may be dependent on how the argumentative discourse is disciplined in a particular case, i.e. on the institutional conventions of the communicative activity type concerned. This does not automatically mean that there are no clear criteria for determining whether the strategic maneuvering has gone astray, but only that these criteria take a shape that may vary to some extent from the one communicative activity type to the other. Depending on the institutional requirements pertaining to the activity type, it may, for instance, vary who or what counts as authoritative, so that a certain appeal to authority may be legitimate in the one case but not in the other. In a civil law case, for example, referring to precedent may be a perfectly legitimate appeal to authority while in a criminal law case it may in some legal systems not be allowed—let alone in a scientific discussion. This variation is one of the factors that explain why it may not always be immediately apparent to all concerned that a fallacy has been committed (van Eemeren 2010, p. 203).

References

- Barth, E.M., and E.C.W. Krabbe. 1982. *From axiom to dialogue. A philosophical study of logic and argumentation*. Berlin: Walter de Gruyter.
- Edmondson, W. 1981. *Spoken discourse. A model for analysis*, vol. 27. Longman linguistics library. London: Longman.
- van Eemeren, F.H. 1986. Dialectical analysis as a normative reconstruction of argumentative discourse. *Text* 6(1): 1–16.
- van Eemeren, F.H. 1987a. For reason's sake: Maximal argumentative analysis of discourse. In *Argumentation: Across the lines of discipline. Proceedings of the conference on argumentation 1986*, eds. F.H. van Eemeren, R. Grootendorst, J.A. Blair, and Ch.A. Willard, 201–216. Dordrecht: Foris. [PDA 3].
- van Eemeren, F.H. 1987b. Argumentation studies' five estates. In *Argument and critical practices: Proceedings of the 5th SCA/AFA conference on argumentation*, ed. J.W. Wenzel, 9–24. Annandale, Virginia: Speech Communication Association.
- van Eemeren, F.H. 2010. *Strategic maneuvering in argumentative discourse*. Amsterdam: Benjamins.
- van Eemeren, F.H., B. Garssen, and B. Meuffels. 2009. *Fallacies and judgments of reasonableness. Empirical research concerning the pragma-dialectical discussion rules*. Dordrecht: Springer.
- van Eemeren, F.H., B. Garssen, and J.H.M. Wagemans. 2012. *The pragma-dialectical method of analysis and evaluation*. In *Reasoned argument and social change. Selected papers from the seventeenth biennial conference on argumentation sponsored by the National Communication*

- Association and the American Forensic Association*, ed. R.C. Rowland, 25–47. Washington, DC: National Communication Association.
- van Eemeren, F.H., and R. Grootendorst. 1984. *Speech acts in argumentative discussions. A theoretical model for the analysis of discussions directed towards solving conflicts of opinion*. Dordrecht/Cinnaminson: Foris Publications.
- van Eemeren, F.H., and R. Grootendorst. 1987. Fallacies in pragma-dialectical perspective. *Argumentation* 1(3): 283–301.
- van Eemeren, F.H., and R. Grootendorst. 1988. Rationale for a pragma-dialectical perspective. *Argumentation* 2(2): 271–291. [Special issue on questions about philosophical argumentation, guest editor H.W. Johnstone, Jr].
- van Eemeren, F.H., and R. Grootendorst. 1990. Analyzing argumentative discourse. In *Perspectives on argumentation. Essays in honor of Wayne Brockriede*, ed. J.W. Wenzel, 86–106. Prospect Heights: Waveland.
- van Eemeren, F.H., and R. Grootendorst. 1992. *Argumentation, communication, and fallacies. A pragma-dialectical perspective*. Hillsdale: Lawrence Erlbaum.
- van Eemeren, F.H., and R. Grootendorst. 2004. *A systematic theory of argumentation. The pragma-dialectical approach*. Cambridge: Cambridge University Press.
- van Eemeren, F.H., R. Grootendorst, S. Jackson, and S. Jacobs. 1993. *Reconstructing argumentative discourse*. Tuscaloosa: The University of Alabama Press.
- van Eemeren, F.H., P. Houtlosser, and A.F.S. Henkemans. 2007. *Argumentative indicators in discourse. A pragma-dialectical study*. Dordrecht: Springer. Argumentation Library.
- van Eemeren, F.H., R. Grootendorst, and B. Meuffels. 1989. The skill of identifying argumentation. *Journal of the American Forensic Association* 25(4): 239–245.
- Freeman, J.B. 1991. *Dialectics and the macrostructure of arguments*. Berlin: Walter de Gruyter.
- Grice, H.P. 1975. Logic and conversation. In *Syntax and semantics 3: Speech acts*, eds. P. Cole, and J.L. Morgan, 41–58. New York: Academic.
- Grootendorst, R. 1987. Some fallacies about fallacies. In *Argumentation: across the lines of discipline, Proceedings of the conference on argumentation 1986* (vol. 1), eds. F.H. van Eemeren, R. Grootendorst, J.A. Blair, and Ch.A. Willard, 331–342. Dordrecht: Foris. [PDA 3, Vol. 1 of 3].
- Hamblin, C.L. 1970. *Fallacies*. London: Methuen.
- Jackson, S. 1985. What speech acts do for argumentation theory? In *Advances in argumentation theory and research*, eds. J.R. Cox, M.O. Sillars, and G.B. Walker, 127–138. Annandale: Speech Communication Association.
- Johnson, R.H. 2000. *Manifest rationality: A pragmatic theory of argument*. Mahwah: Lawrence Erlbaum Associates.
- Johnson, R.H., and J.A. Blair. 2005. *Logical self-defense*. New York: IDEA.
- Levinson, S.C. 1983. *Pragmatics*. Cambridge: Cambridge University Press. Cambridge Textbooks in Linguistics.
- Searle, J.R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, J.R. 1971. What is a speech act? In *The philosophy of language*, ed. J.R. Searle, 39–53. London: Cambridge University Press.
- Searle, J.R. 1979. Literal meaning. In *Expression and meaning*, ed. J.R. Searle. Cambridge: Cambridge University Press.
- Searle, J.R. 1980. An interview. *Versus* 26: 17–27.
- Searle, J.R. 1986. Notes on conversation. In *Contemporary issues in language and discourse processes*, ed. D.G. Ellis, and W.A. Donahue, 7–19. Hillsdale: Lawrence Erlbaum.
- Sperber, D., and D. Wilson. 1986. *Relevance. Communication and cognition*. Oxford: Blackwell.

Rhetoric and Pragmatics: Suggestions for a Fruitful Dialogue

Francesca Piazza

Abstract The aim of this paper is to show that classical rhetoric can provide valuable insights in the contemporary debate in pragmatics. This is especially true for Aristotelian rhetoric, due to its philosophical approach. In the first part of the paper, I discuss the conditions under which ancient rhetoric can be a real partner of current pragmatics: (1) rhetoric must be understood as a type of knowledge (a *techne*) and not as a “jumbles of techniques”; (2) we need to consider persuasion as an anthropological feature and not only as a specific case of communication; (3) we should not exclude truth from the rhetorical field. The second part of the paper focuses on what can be considered the basic insight of classical rhetoric (stated in a well-known passage of Aristotle’s *Rhetoric*, 1358a 37-b1): *speakers and listeners are inside and not outside discourse*. If adequately interpreted, this statement has important consequences that can be interesting for current pragmatics. These consequences can be schematically summarized as follows: (1) a broader conception of persuasion, useful to overcome the opposition informative/persuasive; (2) the consideration of speaker and listener as internal component of discourse (with the latter in a key position); (3) the consideration of the stylistic elements of a discourse from a cognitive point of view; (4) the overcoming of the sharp opposition cognitive/emotional due to the consideration of *ethos* and *pathos* as discursive elements.

1 The Marriage Between Pragmatics and Rhetoric

In 1999 M. Dascal and A. Gross published a paper entitled “The marriage of Pragmatics and Rhetoric”. More specifically, the marriage they are referring to is the marriage between Aristotelian rhetoric and Gricean pragmatics. Although the

F. Piazza (✉)
University of Palermo, Palermo, Italy
e-mail: francesca.piazza@unipa.it

authors acknowledge that «it is a union with problems on both sides of the aisle» (Dascal and Gross 1999: 108), they think that their reconstruction attempts to combine the strengths of Gricean pragmatics and Aristotelian rhetoric are possible and fruitful. «It does so by giving one strand of the rhetorical tradition a cognitive reading, one that suggests the compatibility of two otherwise separate understandings of linguistic communication and interaction. We believe that this reconstruction benefits both pragmatics and rhetoric» (Dascal and Gross 1999: 128).

A few years later, Larrazabal and Korta, following Dascal and Gross, propose a new perspective called *Pragma-Rhetoric* that is «a pragmatic and rhetorical view in discourse analysis, combining both disciplines in order to explain the intentional phenomena that occur in most communicative uses of language, namely the communicative intention and the intention of persuading» (Larrazabal and Korta 2002: 1). Even Larrazabal and Korta are aware that «it is quite difficult to ‘marry’ such an ancient discipline as rhetoric with such a new discipline as pragmatics, if we do not put both in the same “register level”, i.e. in the level of intentionality». (Larrazabal and Korta 2002: 1). Therefore, their efforts are mainly directed towards proposing a reading of ancient rhetoric that places at its center the Gricean concept of communicative intention.

In this paper, I will discuss this proposal with the aim of showing that rhetoric brings a rich dowry to this marriage; classical rhetoric can provide valuable insights on the contemporary debate in pragmatics. Doubtless, the *Pragma-Rhetoric* approach is attractive and theoretically productive, but it is possible to go one step further.

In my view, the main advantage of this perspective is that it considers ancient rhetoric (especially the one developed by Aristotle) as a real partner and not only as an ancestor. Indeed, it should be noted that the comparison between rhetoric and pragmatics often recognizes merely generic similarities between the two disciplines. Furthermore, in many cases these comparisons are based on the assumption that pragmatics, due to its modernity, is more scientific than ancient rhetoric. But, in this way, the dialogue between the two disciplines becomes unfruitful and rhetoric is interesting only from a historical point of view.¹

Instead, the *Pragma-Rhetoric* approach takes very seriously the possibility that ancient rhetoric still has something to tell us. Both Dascal and Gross and Larrazabal and Korta offer a cognitive reading of the main concepts of Aristotelian

¹ I refer, for example, to Sperber and Wilson who state that «if relevance theory is right, then it offers a solution to the rhetorician’s dilemma, a way of being precise about vagueness, of making literal claims about metaphors and ironies, without abandoning any of the Romantics’ intuitions. However, rhetoricians could not adopt this solution without jeopardizing the very foundations of rhetoric. For what this solution implies is that metaphor and irony are ordinary exploitations of basic processes of verbal communication, rather than devices based on codified departures from the ordinary use of language. Moreover metaphor and irony exploit quite different basic processes and are more closely related, the former to loose talk, the latter to a variety of echoic uses, than to one another. The very notion of a trope is better dispensed with. *If so, then rhetoric has no subject matter to study, or to teach*» (Sperber e Wilson 1990, italics mine).

rhetoric, not just those traditionally included in *inventio* but also those relating to *elocutio* and *dispositio*.² From their point of view, this cognitive reading is the only way to put both rhetoric and pragmatics «in the same register level» and by doing so establish a fruitful dialogue between the two disciplines. But this seems to suggest that ancient rhetoric needs modern pragmatics to develop its theoretical potential, showing that a narrow view of classical rhetoric remains in use. In fact, when concluding their paper, Dascal and Gross state that the main advantage that rhetoric derives from its marriage with pragmatics is «the realization that its practices are not a jumbles of techniques, but are, instead, a coherent theory in the cognitive class» (Dascal and Gross 1999: 128). This statement seems to imply that, according to the authors, before its union with pragmatics, rhetoric was only a “jumbles of techniques” and not a “coherent theory”. But I believe that this is not true, at least not for Aristotelian rhetoric. Therefore, I will follow a different path. A path that sees ancient rhetoric as a theory and not merely as a collection of practical advices requiring a cognitive re-elaboration.

Furthermore, a comparison between rhetoric and pragmatics is more interesting if it is able to show not only the similarities between the two, but also (and above all) their differences. In others words, a comparison should aim to show whether ancient rhetoric can still give us a (at least partially) different perspective on language and speech.

To achieve this aim, I will focus on what I perceive to be the basic insight of classical rhetoric. For the moment, we can schematically summarize this basic idea as follows: *speakers and listeners are essential components of speech and not only its external users, they are inside and not outside discourse (logos)*. I believe that this idea has not been properly valued by contemporary scholars, although it is consistent with many perspectives in current pragmatics. Therefore, a fruitful comparison between these two disciplines should take this basic idea into account. However, to avoid misunderstandings, I will start with some general remarks, then move to a more detailed analysis.

1.1 What Kind of Rhetoric?

First of all, it should be noted that I will be referring to Aristotelian rhetoric. Indeed, due its philosophical approach, it is the best choice for a worthwhile dialogue with current pragmatics. Our background assumption is that Aristotle's *Rhetoric* is a philosophic work. This means that, as Garver says, it is «a piece of philosophic inquiry» and therefore it can be read «in the hopes of learning something about contemporary philosophic issues» (Garver 1994: 3).³

² In ancient rhetoric, *inventio* was the part concerning the discovering of arguments, *dispositio* that relating to the arrangement of these arguments, while *elocutio* mainly dealt with stylistic elements of a discourse.

³ See also Grimaldi (1972) and Piazza (2008).

For our purposes, the Aristotelian definition of rhetoric as «an ability, in each [particular] case, to see (*theorein*) the available means of persuasion (*to endechomenon pithanon*)» (Arist. *Rhet.* 1355b 26–27, transl. Kennedy 1991) is particularly significant. This definition clearly reveals a theoretical point of view on rhetoric and is based on the belief that persuasion plays a crucial role in human life. Indeed, at the very beginning of his work Aristotle says that

rhetoric is an *antistrophos*⁴ to dialectic; for both are concerned with such things as are, to a certain extent, within the knowledge of all people and belong to no separately defined science. A result is that all people, in some way, share in both; for all, to some extent, try both to test and maintain an argument [as in dialectic] and to defend themselves and attack [others, as in rhetoric] (1354 a 1–6, transl. Kennedy 1991).

Therefore, to use language in order to argue and persuade is a skill that all people possess to some degree, even if

some do these things randomly and others through an ability acquired by habit, but since both ways are possible, it is clear that it would also be possible to do the same by [following] a path; *for it is possible to observe [theorein] the cause why some succeed by habit and others accidentally, and all would at once agree that such observation is the activity of an art [techne]* (1354a 6–11, transl. Kennedy 1991).

Thus, according to Aristotle, both dialectic and rhetoric are arts (*technai*) and this means that it is possible to try to discover *why* a speech may be persuasive («the available means of persuasion»). In fact, in the Aristotelian perspective a *techne* is not a collection of practical advice but instead is a type of knowledge. Principally, two aspects differentiate it from science: first, it has a productive purpose (while science only has a theoretical aim); second, it deals with things that «can be otherwise» (while science deals with «what cannot be otherwise»). However, it is still a knowledge that is able to discover the causes (Arist. *Met.* 981a 5–28; EN 1140 a 1–13). This means that, as a *techne*, rhetoric also has theoretical value (it is the ability «to see the available means of persuasion») and it is just one of the reasons why the Aristotelian perspective can still be interesting.

Moreover, it is also worth noting that since rhetoric, as a *techne*, deals with «what can be otherwise» it cannot guarantee success, and is always fallible. To stress this aspect, Aristotle states that the specific function (*ergon*) of rhetoric is «not to persuade but to see (*idein*) what can be persuasive (*ta hyparchonta pithana*) in each case» (*Rhet.* 1355b 10–11, the translation is mine). Therefore, the aim of rhetoric is to show the steep and slippery path toward what can be persuasive. It is a very difficult task because, as Aristotle clearly says, «the persuasive

⁴ Concerning the word *antistrophos*, commonly translated “counterpart”, see Kennedy (1991: 28, n.1).

is *persuasive to someone*» but «since no art examines the particular—for example, the art of medicine does not specify what is healthful for Socrates or Callias but for persons of a certain sort (this is artistic, while particulars are limitless and not knowable)—neither does rhetoric theorize about each opinion—what may seem so to Socrates or Hippias—but about what seems true to people of a certain sort, as is also true with dialectic» (*Rhet.* 1356b 27–36, transl. Kennedy 1991). We can see a tension here between the practical purpose of persuading *someone* (that is always a particular individual) and the theoretical impossibility of knowing particulars. On the one hand, to be effective the speaker cannot be satisfied with just persuading «people of a certain sort» because he has to persuade Socrates or Hippias. On the other hand, he knows that rhetoric can only tell him what can be persuasive for «people of a certain sort» (*Rhet.* 1356b 36). The gap can only be bridged by natural skill, experience and, last but not least, by good luck. In my view, this is another interesting aspect to current pragmatics. In fact, if we look carefully, pragmatics seems to be in a similar position, since it looks at language from the standpoint of a real interaction between speakers and listeners.

1.2 Persuasion as Anthropological Feature

Another aspect that makes Aristotle’s rhetoric particularly suitable to interact with modern pragmatics is the Aristotelian conception of persuasion. As we have already seen, for Aristotle persuasion is not just a matter for specialists but a natural skill which all human beings possess to some degree. To come to a better understanding of the place of persuasion in Aristotle’s philosophy, we should bear in mind that the meaning of the Greek words we usually translate with “persuasion” (*peitho*; *pistis*), “to persuade” (*peitho*), “to be persuaded” (*peithomai*) or “persuasive” (*pithanos*, *pistos*) is much broader than ours. These words can be used to mean a large variety of things that we see as different from persuasion, such as erotic seduction, political obedience, trust or faith. Without going into detail, we can say that what links all of these cases is a sort of “assent”, “to say yes” to someone (or, on the other hand, to try to get someone to say yes).

This broader definition of “persuasion” leads to a broader conception of rhetoric. Moreover, reading *Rhetoric* in the context of the other Aristotelian works makes Aristotle’s view of persuasion, or better *pistis* in its all meanings (*persuasion*, *conviction*, *proof*, *belief*, *trust*, *faith*)⁵ clear; it plays a crucial role in understanding human nature. This is also confirmed by many passages (see, for example, DA, 428 a 19–24; EN 1098 a 3–6) in which Aristotle explains humanity’s specificity (the possession of *logos*) in terms of our capacity to persuade each other or, more generally, to pay heed—or not—to what other people say.⁶ According to Aristotle, persuasion—seen as our power to act on other people

⁵ Concerning the meaning of the word *pistis* in Aristotle see (Grimaldi 1972: 57–68), Garver (1994: 142) and Piazza (2008: 45–47).

⁶ See Piazza (2008: 13–29).

through language—is an inner and constitutive component of a human being, the only animal that has *logos* as well as a very complex form of social life (*polis*).⁷ This means that persuasion is an anthropological feature and not only a particular case of communication. This is exactly the reason why seeing «in each case, the available means of persuasion» is a very ambitious task, certainly much more ambitious than teaching how to make discourses more elegant and effective.

It should be noted, however, that saying that persuasion is an anthropological feature is not the same as upholding that all speech acts always have (or that they should have) an explicit persuasive purpose. Instead, it is a way to put persuasion at the heart of linguistic activity. In this way, persuasion can be a good observation point from which our speech act can be better understood. Indeed persuasion as an anthropological feature can show the links between language and other aspects of human life (cognitive, social and emotional). For this reason, this conception of persuasion is a preparatory step for making rhetoric a real partner of pragmatics, instead of only one of its ancestors.

1.3 *Persuasion Versus Communication?*

A larger conception of persuasion can also help us to overcome the sharp opposition between “informative” and “persuasive”. It is a distinction still active in many current perspectives but it can obstruct the dialogue between rhetoric and pragmatics. Moreover, it risks suggesting the idea that persuasion is only a form of manipulative communication.

Dascal and Gross, for example, consider the distinction between informative and persuasive goals to be one of the discrepancies (albeit negotiable) between rhetoric and pragmatics. They state that this discrepancy «is an artifact of the paradigmatic case of talk exchanges on which Grice bases his system: the exchange of information. Rhetorical exchanges are another special case of talk exchanges; they have a different purpose: persuasion. In persuasion, the initial agreement of H to be open to influence is *in itself* an instance of the Cooperative Principle. The introductory remarks of S in the typical classical *exordium* or introduction—remarks designed to obtain the attention, interest and trust of the audience—are meant to carry this process one step further: to establish a beachhead of the Cooperative Principle as the foundation of continuing influence. In information exchanges, cooperation is presupposed; in persuasive exchanges, it must be established, strengthened, and maintained» (Dascal and Gross 1999: 109–110).

I suspect that this distinction between “information exchanges” and “persuasive exchanges” is too sharp and can produce misunderstandings. First, are we really sure that the «exchange of information» is «the paradigmatic case of talk

⁷ Arist. *Pol.* 1252b 30–1253a 35. Concerning the link between *logos* and *polis* see Lo Piparo 2003: 3–33).

exchanges»? As Dascal and Gross themselves state, persuasion «often has as its goal, or interim goal, not a proposition but an alteration in emotional state or in the disposition to approve or disapprove of conduct or character» (Dascal and Gross 1999: 109). If this is true, we can say that this is not only a special case, but the goal of most of our talk exchanges. Perhaps we can differ from Grice's view, and consider persuasion to be more «paradigmatic» than the «exchange of information». The effort to «obtain the attention, interest and trust of the audience» does not only concern the typical situation of rhetorical speeches, but is present in ordinary conversation. Similarly, the talk exchanges in which the Cooperative Principle «must be established, strengthened, and maintained» are much more frequent than they seem to be at a first glance. By contrast, the cases in which the aim of our utterances is only an information exchange are actually quite rare. Indeed, often the information exchange also has the goal to induce in our listeners an emotional state, a disposition or even stimulate behavior. In addition, information is a necessary step of a persuasive effort.⁸

With this, I am not altogether denying the distinction between informative and persuasive aspects. I only want to argue that it could be more productive to see both the informative and the persuasive dimension as components (more or less) present in all of our talk exchanges.

1.4 Rhetoric and Truth

The sharp distinction Dascal and Gross make between informative/persuasive discourse leads them to see another discrepancy between rhetoric and pragmatics concerning the conversational maxim prescribing truthfulness. Indeed, according

⁸ Concerning this distinction, Larrazabal and Korta (2002) sustain a partially different position. They distinguish “communicative” and “persuasive” intention and claim that «it is very clear that these two intentions are in different levels. We need first the fulfillment of communicative intentions in order to make possible then the fulfillment of persuasive intention (particularly, the intention to convince in argumentative discourse). Both in monological discourse and in dialogical (or multilogical discourse)—in what we are more interested—the unit of analysis is a unique speech act, where by means of the satisfaction of the communicative intention one can get the satisfaction of a persuasive intention (we are speaking of course of persuasive communication)» (p. 7). In their view, the main distinction between the two types of intentions is that «persuasive intention in general is not an overt intention. It can be an overt intention as in the case of intention to convince (by arguments) or as in particular kinds of persuasive intentions in special discourse contexts. But it can clearly also be a covert intention: think, for example, about a situation where the speaker intends to persuade the hearers hiding the real persuasive intention behind her discourse behaviour, because this is just the way of getting her goal in that particular situation» (p. 8). According to Larrazabal and Korta, the focus on persuasive intentions allows rhetoric to also include «perlocutory effects on the audience, intended or not intended by the speaker (...). usually ignored by pragmatics studies. This is where Rhetoric can make its contribution. Persuasive as well as convincing and other kinds of perlocutory intentions seem to constitute the basis of rhetorical studies of linguistic use» (p. 5).

to Dascal and Gross this maxim can be applied «to rhetorical exchanges only as an injunction that its *appearance* must be maintained» (Dascal and Gross 1999: 110). It is certainly true that persuasion can be deceptive, but this is not only a prerogative of rhetorical exchanges. This extension of Grice's maxim is interesting and useful provided that it does not imply that rhetoric can only deal with an apparent truth.

I cannot address the complex issue concerning the relationship between rhetorical discourse and truth in detail here.⁹ However, there is an important aspect we should note before continuing. One should note that there is a widespread misunderstanding impeding the dialogue between rhetoric and pragmatics. I am referring to the conviction that rhetoric is not interested in truth, an idea often also ascribed to Aristotle. The evidence used to defend this attribution is a passage of *On Interpretation* (17a 1–6) where one finds the distinction between the so-called *apophantic* and *semantic* speech. In this passage, Aristotle distinguishes sentences—called *apophantic* (constative, declarative)—that can be judged true or false from sentences (such as prayers) that, albeit meaningful, cannot be true or false. Immediately afterwards, he says: «let us therefore dismiss all other types of sentences but the *apophantic*, for this last concerns our present inquiry, whereas the investigation of the others is more appropriate to the study of rhetoric or of poetry» (17a 5–6). Unfortunately, this statement has been generally understood as an exclusion of truth from the realm of rhetoric and poetics. But, looking more in depth, here Aristotle is saying something different. Indeed, stating that the study of sentences that are neither true or false is more appropriate to rhetoric and poetics does not mean that they *only* deal with sentences of this type.

On the contrary, rhetoric (we leave poetics out here because it would make our discussion too long) is also interested in sentences that claim something about the world and therefore can be true or false. If we cast aside the common underestimation of rhetoric (especially by philosophers) we see that rhetorical speech constantly involves truth (real or apparent) in its difficult but inescapable link with persuasiveness. This link is expressly stated by Aristotle when he says that: «persuasion (*pistis*) occurs through the *logos* when we show the truth or the apparent truth from whatever is persuasive in each case» (*Rhet.* 1356a 19–20, transl. Kennedy 1991).

It is enough to show that Aristotle does not leave the true/false couple out of the realm of rhetoric. It is true that rhetoric also deals with *non-apophantic* sentences, like prayers, orders, oaths, etc. However, this does not mean that it is not interested in truthfulness. If rhetoric is an ability «to see the available means of persuasion», then it is interested in all speech acts that can concur to achieve the *pistis* (in the sense, as seen above, of *persuasion, conviction, belief, trust, faith*). It would be very surprising if it did not also care about sentences that can be true or false. Doubtless, this is not a warrant of truthfulness, but it isn't a disregard for truth either.

⁹ See Piazza (2004).

In short, we can say that rhetoric can be a real partner of pragmatics only under certain conditions. First, it must be understood as a type of knowledge (a *techné* in the sense seen above) and not as a “jumbles of techniques”; second, we need to consider persuasion an anthropological feature and not only a specific case of communication; third, we should not exclude truth from the rhetorical field.

2 The Dowry of Rhetoric

Against this background, we can return to the basic insight of classical rhetoric as seen above: the idea according to which *speakers and listeners are inside and not outside discourse*. This insight is explicitly stated in a well-known passage of Aristotle’s *Rhetoric*:

logos is constituted by three elements, by the *person speaking*, by the *topic spoken about* and by the *person spoken to*, and the objective (*telos*) of the speech relates to the last, (I mean the hearer) (Arist. *Rhet.* 1358a 37-b1, transl. Kennedy 1991, partially modified).

It is a simple statement, yet rich with theoretical consequences. It expresses a leading principle of Aristotle’s *Rhetoric* and it is also useful to understand the Aristotelian conception of language and speech. Let us analyze it in detail.

First of all, we should clarify an important point. Unfortunately, this statement has often been interpreted as an anticipation of the code-based model of communication (or even of Jakobson’s communicative functions).¹⁰ According to this view, Aristotle shares the idea of language as a *code* used by an *addresser* to send a *message* to an *addressee*. I believe that this is a misreading of the Aristotelian point of view. Indeed, if we read this passage trying to forget the modern terminology, we see that we do not have here *addressers* or *addressees*, nor *codes* or *messages* that would be transferred from one mind to another. In the quoted passage we only have the *person speaking*, the *person spoken to* and the *topic spoken about*. It is not just a terminological difference, but a different point of view on language and communication.

Moreover, to better understand the Aristotelian statement, it should be emphasized that it does not concern *communication* in general but *discourse* (*logos*). The *person speaking*, the *person spoken to* and the *topic spoken about* are not the constitutive elements of the communicative system, as many scholars sustain.¹¹ More exactly, they are *the components of speech* (*logos*). The difference is not superficial. Indeed, in the first case, speakers and listeners, even though they are essential to the communicative process, are external users. They remain

¹⁰ See Kennedy (1991: 47, n. 75) and Venier (2008: 14).

¹¹ See Mortara Garavelli (2006: 3).

outside the *logos* (that they *use* to communicate). In this way, the discourse seems to only coincide with one of its elements, the “message” or “what is said”. Instead, here Aristotle is saying something different: *logos* is not only “what is said” but it is made up (this is the meaning of the Greek verb *synkeitai*, often used by Aristotle to mean the material composition of something) of all of the three elements, *speaker*, *listener* and *topic*. Each of these plays a crucial role in building the speech that gains its real consistency only thanks to the relationship between these three elements. In this way, speakers and listeners can be considered as internal components and not only as external users.

To avoid misunderstandings, the meaning of the expression used above, “what is said”, should be clarified. In the Aristotelian passage, it does not mean the “informative content” of a speech, the “message” that the speaker sends to the listener. Furthermore, it is not even “what is said” in the sense of current pragmatics, as opposed to “what is meant”. In the Aristotelian passage, “what is said” is the topic spoken about, the subject matter (in ancient Greek, *pragma*) or, perhaps better, the issue at stake. Therefore, it is both what is explicitly said and what is implied and, at least in a sense, it also includes the “context”, the “common ground” thanks to which we can understand each other.

2.1 *The Primacy of the Listener*

Another important consequence of this perspective concerns the relationship between speaker and listener or, better, the primacy of the latter. As we have seen, concluding the passage quoted above, Aristotle stresses that the aim (*telos*) of the speech relates to the listener.

To better understand this statement, it is useful to bear in mind that in the Aristotelian perspective, the concept of *telos* has both logic and ontological priority and it plays a crucial role in understanding both natural and artificial processes. Therefore, the Aristotelian specification clearly means that one of the three components of *logos*, the listener, occupies a key position in the discursive relationship.

This leading role of the listener is a distinctive feature of the rhetorical point of view and it is one of the most important contributions that rhetoric can make to the contemporary debate on pragmatics. As we have anticipated above, according to Aristotle, «the persuasive is *persuasive to someone*». Even if we don't have full knowledge of the individual, we should never forget that our speech is always addressed to a listener. But Aristotle goes a step further. Broadly speaking, the listener is always a *judge* (because he/she must pass judgment on the speaker's speech, see Aristotle, *Rhetoric*, 1391b 8–23). It is for this reason that he/she is not simply an addressee but the *telos* of the whole process. This means that, unlike the main perspectives in current pragmatics, the starting point is no longer the speaker, but the listener.

To be more precise, the fundamental position of the listener does not imply an underestimation of the other two elements. Indeed, as was said above, each component is essential to the building of *logos*. Moreover, we should not forget that speakers and listeners in our ordinary talking continually exchange roles. Hence, to say that the *telos* of the speech relates to the listener is not only a matter of hierarchy. Instead, it involves seeing our talk exchanges from the point of view of listening and not only from that of producing. It is not a trivial difference but a fruitful shift in the way we see our talk exchanges.

2.2 The Triad Ethos-Pathos-Logos

The Aristotelian statement according to which «*logos* is constituted by three elements, by the *person speaking*, by the *topic spoken about* and by the *person spoken to*, and the objective (*telos*) of the speech relates to the last, (I mean the hearer)» is not only a general consideration. It also represents one of the organizing principles of Aristotle's *Rhetoric*.¹² Indeed, the triad *speaker*, *listener*, *topic* is also the basis of the well-known triple-partition of *entechnoi pisteis* (means of persuasion embodied in art or "artistic proofs"): *ethos*, *pathos* and *logos*. They are the rhetorical means of persuasion (*pisteis*) provided "by us" and "through speech" opposed to the so-called *atechnoi pisteis* ("non-artistic" means of persuasion) that are not provided by us, but instead are pre-existing.¹³

Each of the three *entechnoi pisteis* corresponds to one of the three elements of the triad. Indeed, *ethos* is the proof based «on the character of the *speaker*», *pathos* consists «in disposing the *hearer* in some way» and *logos* is the «argument itself, by showing or seeming to show something» (Arist. *Rhet.* 1356a 2–5, transl. Kennedy 1991). According to Aristotle, the three *entechnoi pisteis* are closely connected and, if the speech is effective, persuasion can be seen as the result of the interaction between *ethos*, *pathos* and *logos*.¹⁴

As is generally recognized, this triple-partition represents the theoretical frame of the whole of Aristotle's *Rhetoric*. It plays a role not only in *inventio*, but also in *dispositio* and *elocutio*. We can say that, if adequately interpreted, it is the most vital part of the rhetorical tradition. Then, a fruitful dialogue with this tradition

¹² It is not possible here to deal in detail with all of the consequences of this idea on the entire Aristotelian rhetorical system. I limit myself to only indicating that the classical distinction between the three species of rhetoric: deliberative, judicial, and epideictic (cfr. *Rhet.* 158b 1 sgg.) is also based on it.

¹³ Arist. *Rhet.* 1355b 37-40: «I call *atechnic* those that are not provided by "us" [i.e. the potential speaker] but are preexisting: for example, witnesses, testimony of slaves taken under torture, contracts, and such like; and *artistic* whatever can be prepared by method and by "us"; thus one must *use* the former and *invent* the latter» (transl. Kennedy 1991).

¹⁴ See Grimaldi (1972), Fortenbaugh (2002), Garver (1994) and Dow (2007).

should take account of it.¹⁵ Without entering into strictly philological questions, let us look more closely at this triple-partition.

First of all, we should stress that, coherently with what was said above concerning Aristotle's conception of *logos*, each element must be considered in its relationship with the others. More exactly, due to the fact that speakers and listeners are inside and not outside the *logos*, *ethos* and *pathos* can be seen as internal components of the persuasive process and not only as external (or, worse, "irrational") expedients.

As Aristotle himself says, *ethos* and *pathos* have an intimate connection with language because they are provided through speech and thus they should not be separated from *logos*. In the case of the third *pistis*, based on *logos* itself, this connection is easy to see. Instead, the case of the other two *pisteis* seems more troublesome to some scholars. According to Dascal and Gross, for example, the inclusion of *ethos* and *pathos* in their project of a cognitive rhetoric could seem to be a «counterintuitive move» (Dascal and Gross 1999: 108). Although they think that this inclusion is possible, they observe that «it is not immediately obvious, however, that Grice's notion of inference can be applied to so-called emotional "proofs" and "proofs" from character, surely, "proof" applied to emotion and character is a metaphoric extension that obscures more than it explains» (Dascal and Gross 1999: 111).

I disagree with this statement. For Aristotle, saying that *ethos* and *pathos* are *pisteis* is not a metaphoric extension. The word "proof" applied to emotion and character seems to be a metaphor only from our modern perspective. Indeed, as we have already seen, the Greek word *pistis* does not mean exactly the same thing as the English word "proof" (although in many cases "proof" can be a good translation of *pistis*). The extent of the semantic field of "pistis" allows Aristotle to use this term to indicate *ethos* and *pathos* as well, without metaphoric extension. They are *pisteis* in the sense of "means of persuasion". Moreover, *qua* artful (*entechnoi*) *pisteis*, *ethos* and *pathos* are means of persuasion provided through *logos*. For this reason, it is quite possible to use Grice's notion of inference in the case of *ethos* and *pathos* as well without forcing Aristotle's position in any way.

2.2.1 Ethos and Logos

The intimate relationship of *ethos* and *pathos* with language emerges clearly from many passages of Aristotle's *Rhetoric*. Let us begin with *ethos*. Aristotle describes it in this way:

[There is persuasion] (*pistis*) through character (*ethos*) whenever the speech is spoken in such a way as to make the speaker worthy of credence (*axiopiston*); for

¹⁵ Also Dascal and Gross (1999) and Larrazabal and Korta (2002) consider the triad *ethos-pathos-logos* still useful for the modern pragmatics.

we believe (*pisteuomen*) fair-minded (*epieikes*) people to a greater extent and more quickly [then we do others] on all subjects in general and completely so in cases where there is not exact knowledge (*akribes*) but room for doubt (*amphidoxein*). And this should result from the speech, not from a previous opinion than the speaker is a certain kind of person; for it is not the case, as some of the technical writers propose in their treatment of the art, that fair-mindedness (*epieikeia*) on the part of the speaker makes no contribution to persuasiveness; rather, character (*ethos*) is almost, so to speak, the controlling factor in persuasion (Aristotle, *Rhetoric*, 1365 a 4–13, transl. Kennedy 1991).

First of all, it should be observed that in the quoted passage, Aristotle is not saying that the speaker's credibility makes the speech credible. Instead, he is stating that *the way in which the speech is spoken* makes the speaker trustworthy.¹⁶ It is for this reason that the *ethos* Aristotle is referring to is not the pre-existing reputation of the speaker, but a character constructed "through *logos*". This does not mean that it is an apparent character. Indeed, I agree with Garver that *ethos* is artificial but not for that reason, unreal (Garver 1994: 197).¹⁷ More exactly, as it is explicitly claimed by Aristotle, to be an artful *pistis*, *ethos* "should result from speech". Yet, it does not necessarily imply that it is deceptive. Instead, this means that the speaker has to construct his/her character due his/her words, namely *inside* the speech. It is the reason why «more sophisticated, politically wise, audiences trust a speaker *qua* speaker, but not therefore *qua* man» (Garver 1994: 196). This is only possible because the speaker is an internal element and not only an external user of the speech. Hence, persuasion through *ethos* is not something added from outside to reinforce a reasoning. Instead, it is something that the audience can infer from the speech itself.

As Garver has rightly pointed out, in Aristotle's *Rhetoric*, «*logos* and *ethos* are integrated. [...] There is a mutual implication and interdependence between the two terms. We trust a speaker, and impute *arete* and *phronesis* to him, when he presents us with a cogent and intelligent argument. We infer from *logos* to *ethos*. But rhetorical arguments are about indeterminate matters. So, in order to regard an argument as cogent and intelligent, we have to trust the speaker. [...] We need

¹⁶ I do not agree with Kennedy that here «Aristotle is not thinking of style and delivery but of thought and content» (Kennedy 1991: 38, n. 41). I believe that we have no textual reasons for thinking so. On the contrary, the use of the verb *legein* seems to indicate that Aristotle is referring to speech as a whole, therefore also including style and delivery.

¹⁷ See also Garver (1994: 196): «the *ethos* which the audience trusts, then, is the artificial *ethos* identified with argument. It is not some real *ethos* the speaker may or may not possess. It is an *ethos* not necessarily tied to past experiences of the speaker, not an *ethos* acquired through performing similar actions in the past. It may be *likely* that the good speaker is able to deliberate intelligently because of past experiences, but it is not *ethos qua* product of past experience that the audience trusts, but the *ethos* as exercised in some particular argument. Insofar as audiences do actually draw that further inference about moral character, demonstrating its superfluosity should be seen as a contribution the *Rhetoric* makes to political science and legislation» (Garver 1994: 196).

trust to perceive the argument correctly. We infer from *ethos* to *logos*. It is the interdependence of *ethos* and *logos* that prevents such inferences from becoming a vicious circle» (Garver 1994: 191–192).

It is not possible to deal with all of the questions concerning the interpretation of Aristotle's *ethos* here. But I hope that what was said is enough to show that, far from being a counterintuitive move, the application of Grice's notion of inference to *ethos* is a straightforwardly Aristotelian move.

2.2.2 *Pathos* and *logos*

Something similar can be said concerning *pathos*. As we have already seen, it is an artful *pistis* too, therefore it is provided through *logos*. Indeed, Aristotle says:

[There is persuasion (*pistis*)] through the hearers when they are lead to feel emotion (*pathos*) by the speech, for we do not give the same judgment when grieved and rejoicing or when being friendly and hostile (1356a 14–15, transl. Kennedy 1991).

Like in the case of *ethos*, the *pathos* Aristotle is speaking about in this passage is a *pathos* achieved by the speech itself. It can only be an artful *pistis* under this condition. So intended, *pathos* is not an external device that is only useful to seduce or deceive the audience. Instead, it is an essential part of the persuasive speech. Again, this is possible thanks to the consideration of the listener as an internal component of *logos* or, better yet, as the *telos* of the whole process.

To come to a better understanding of the role of *pathos* in persuasive speech, one should bear in mind that persuasive speech generally aims at a practical choice and that, in Aristotle's perspective, practical reason is intrinsically connected to desire and motivation. This means that there is a strong connection between *orexis* (desire) and *krisis* (judgment). Thanks to this connection, emotions have an important cognitive component. This clearly emerges from the definition of *pathe* in the Book II of *Rhetoric*:

The emotions (*pathe*) are those things through which, by undergoing change, people come to differ in their judgments and which are accompanied by pain and pleasure, for example anger, pity, fear, and other such things and their opposites (1378a 19–22, transl. Kennedy 1991).

As we can see, emotion and judgment are so strictly connected to the extent that Aristotle considers *pathe* to be a crucial factor of our differences in judgment. This strict connection between *pathos* and judgment is also pointed out by Aristotle in the definition of each *pathos*, but we cannot consider it in detail here (Rhet. II, 2-1).

The importance of the cognitive component in Aristotle's theory of the emotions is generally recognized by contemporary scholars.¹⁸ Dascal and Gross themselves claim that «Aristotle's theory of the emotion is compatible with Grice in that it is straightforwardly cognitive» (Dascal and Gross 1999: 120). This is certainly true. But I think that Aristotle's perspective, although compatible with that of Grice, is more exhaustive. Indeed, according to Aristotle, the relationship between the emotional and cognitive aspect is one of interdependence. Aristotle focuses not only on the cognitive component of the emotional states but is above all interested in the connection between the two. It is precisely for this reason that Aristotle's perspective is still interesting.

For Aristotle, in the case of practical reason, our judgments can never be independent from our desires (See Nussbaum 1996). Indeed, «thought (*logos*) by itself moves nothing» but «what moves us is thought aiming at some goal and concerned with action» (EN, 1139 a 35–36). To move a living being, it is necessary that he feels pleasure or pain. Human beings are no exception. Certainly, in the case of human beings the action is not only a matter of pleasure and pain, even if these are necessary to act. Human action (*praxis*) is possible only thanks to the strict connection between desire and judgment (see *De Anima*, 432b 27–33b 30). Or, to put things in a better way, an action is *human* in that it is a result of this connection. This is well expressed in the famous Aristotelian description of the human being as “desiderative reason” (*orektikos nous*) or “rational desire” (*orexis dianoetike*) (EN, 1139b 4–5).

If we return to the definition of *pathe* cited above, we note that judgment is only one of the components of our emotional state. Another essential feature is the presence of pleasure and pain: *pathe* are always *accompanied by pain and pleasure*. This presence, together with the cognitive component, makes *pathos* one of the motives of our action. Naturally, it is possible that emotion can cause mistakes and deception but this can also happen with reason. Therefore, one can agree with Garver that it is not true that «emotional appeals are necessary only because audiences are weak and corrupt. (...) But the *Rhetoric* will instead exhibit a constitutive role for the emotions that derives from its constitutive role in good *praxis* and good character» (Garver 1994: 108).

If we remember that persuasive speech generally aims at a practical decision, we can understand the reasons why *pathos*, integrated with *ethos* and *logos*, is an internal factor of the persuasive process.

This means that, in order to understand persuasive communication, one cannot overlook either considerations relating to the truth of “what is said” (the issue at stake) or considerations relating to the participants in the talk-exchanges (“the speaker” and “the person spoken to”). There is not a “rational” component opposed to an “irrational” one. Instead, we have a complex interplay of elements, all necessary for the achievement of persuasion (if it is achieved).

¹⁸ See Fortenbaugh (2002), Leighton (1982), Conley (1982), Cooper (1996) and Striker (1996).

3 The Pleasure of the Words

The intimate connection between the three elements of speech, as well as the crucial role of *ethos* and *pathos*, also emerges in the case of another important concept of classical rhetoric, that of *prepon* (appropriateness, lat. *aptum*). It is the capacity to adapt words both to the topic and to the speakers/listeners. It is not a superficial trait, like a matter of “cosmetics”. On the contrary, it plays an important role in persuasive processes. Indeed, choosing the “right” words can be a decisive move. Aristotle discusses this notion in the section of Book III devoted to *lexis* (style, *elocutio*)

He says that:

The *lexis* will be appropriate (*prepon*) if it expresses *emotion* and *character* and is proportional to the *subject matter*. Proportion exists if there is neither discussion of weighty matters in a casual way nor shoddy things solemnly and if ornament is not attached to a shoddy word. (...). Emotion is expressed if the *lexis*, in the case of insolence (*hybris*), is that of an angry man; in the case of impious and shameful things, if it is that of one is indignant and reluctant even to say the words (...). Proof from signs is expressive of character, because there is an appropriate *lexis* for each genus and moral state (Rhet. 1408 a 10–27, transl. Kennedy 1991, italics mine).

Again, here we have the three constitutive components of the speech: the speaker, the listener and the topic spoken about. To be appropriate, a *lexis* should be able to express each of these components.

In order to understand this passage well, it is important to clarify the meaning of the word *lexis*. It means something more general than our “style”. As Aristotle himself says in the *Poetics* (1450b 13), with *lexis* he means “expression of the meaning in words”. I agree with F. Lo Piparo (1999) that *lexis* is the word Aristotle uses to indicate the actions we perform through language. Hence, it is something very similar to Wittgenstein’s *language game* or Austin’s *speech act* (Lo Piparo 1999, p. 122). In this way, the rhetorical notion of *prepon* becomes more interesting for modern pragmatics. The appropriateness, so intended, is not only a matter of professional writers or poets but also concerns our ordinary talk exchanges. The ability to adapt our words to the context, the issue and the listener is an essential part of our linguistic competence.

This larger conception of “style” goes in the same direction indicated by Dascal and Gross. They state that «in both Aristotle and Grice stylistic devices involve inference for their comprehension» and that «style often has a persuasive as distinct from an aesthetic purpose (Dascal and Gross 1999: 111). I agree with them that Aristotle’s theory of “style” is a “cognitive theory” in that tropes (like metaphor, metonymy, hyperbole etc.) and schemes (like antithesis) involve inference to be understood (Dascal and Gross 1999: 121). But, again, I think that even in this case the cognitive component, although essential, is not everything.

According to Aristotle, what make a trope or a scheme persuasive is not a merely intellectual factor. An expression (*lexis*) persuades because it is also

pleasant. In fact, in Aristotle's view, pleasure and understanding are strictly connected and this connection is one of the guiding principles of *Rhetoric*: «(...) let the following be our first principle. To learn easily is naturally pleasant to all people, and words signify something, so whatever words create knowledge in us are the pleasantest» (1410b 10–13, transl. Kennedy 1991).

In other words, this means that the more a *lexis* is able to produce a quick and pleasant knowledge, the more persuasive it is. This guiding principle is very often pointed out by Aristotle when he explains the reasons why a particular type of reasoning or *lexis* is better than another. For example, concerning the enthymemes (the rhetorical syllogisms), he says that

«in the case of all syllogistic argument, both refutative and demonstrative, those are most applauded that [hearers] foresee from the beginning, but not because they are superficial (at the same time, too, people are pleased with themselves when anticipating [the conclusion], and [they like] those that they are slower to apprehend to the extent that they understand when these have been stated» (1400b 31–35, transl. Kennedy 1991).¹⁹

A good metaphor is also based on this strict connection of pleasure and knowledge. To be persuasive, a metaphor should have “clarity” (*saphes*) “sweetness” (*edus*) and “strangeness” (*xenikon*) (1405a 8). The simultaneous presence of these qualities produces easy and pleasant learning in the listener. For this reason, appropriate (and therefore persuasive) metaphors «should be transferred from things that are related but not obviously so, as in philosophy, too, it is characteristic of a well-directed mind to observe the likeness even in things very different» (1412a 10–15, transl. Kennedy 1991). As this last quotation clearly shows, the ability to create this effect of pleasant knowledge in the listener is not a trivial one and it is also difficult (if not impossible) to learn (1410b 7–10).

We also see the same principle applied to the type of *lexis* called *periodos*:

I call period an expression (*lexis*) having a beginning and end in itself and a magnitude easily taken in at a glance. *This is pleasant and easily understood*, pleasant because opposed to the unlimited and because the hearer always thinks he has hold of something, in that it is always limited by itself, whereas to have nothing to foresee or attain is unpleasant (1409b 1–5, transl. Kennedy 1991).

Similarly, also the persuasive power of antithesis depends on the ability to produce an easy and pleasant understanding: «such a *lexis* is pleasing because opposites are most knowable and more knowable when put beside each other and

¹⁹ See also 1410b 20–27: «Those things are necessarily urbane, both in *lexis* and in the enthymemes, which create quick learning in our minds. That is why superficial enthymemes are not popular [with audiences] (by superficial I mean those that are altogether clear and which there is no need to ponder) nor [are] those which, when stated, are unintelligible, but those of which there is either immediate understanding when they are spoken, even if that was not previously existing, or the thought follows soon after; for [then] some kind of learning takes places, but in neither of the other case» (transl. Kennedy 1991).

because they are like a syllogism, for refutation [*elenchos*] is a bringing together of contraries» (1410 a 20–24, transl. Kennedy 1991).

This effect is reinforced if the antithesis is also concisely expressed. The reason is always the same: «the cause is that knowledge results more from contrast but is quicker in brief form» (1412 b 22–24, transl. Kennedy 1991).

Moreover, we also find this principle in the case of another aspect of style that has been often overlooked because it is wrongly considered not really cognitive. I am speaking about *rhythm*, an aspect Aristotle considers to be very important in a persuasive speech. A persuasive *lexis*, says Aristotle, «should be neither metrical nor unrhythmical». The reason is always the same: a metrical *lexis* diverts attention, while an unrhythmical one is «unpleasant and unknowable» (see 1408b 21–30, transl. Kennedy 1991). Then, even in the case of rhythm we have a strict connection between cognitive and emotional factors. A suitable rhythm can be persuasive precisely thanks to this connection.

It is worth noting that the use of this guiding principle is not limited to *lexis* but also concerns *taxis* (*arrangement*). In their paper, Dascal and Gross rightly include arrangement in their project of a cognitive rhetoric. They say that «arrangement—the organization of a speech—may seem equally remote from the inferential process, which apparently functions only *within* arguments. But we shall show that arrangement can also be a form of proof» (Dascal and Gross 1999: 124). I agree with Dascal and Gross provided that we do not forget that, even in the case of arrangement, the cognitive aspect cannot be separated from the ethical and emotional ones. Indeed, if we read the section of the *Rhetoric* devoted to *taxis* (III, 13–19) we encounter the same guiding principle: the speaker should be able to produce an easy and pleasant knowledge in the listener. This is stressed again by Aristotle, for example, when he says that in judicial speeches and in epic the introduction (*prooimion*) has the function of guiding the listener so that he can understand the speech without undue effort:

in [judicial] speeches and in epic there is a sample of the argument in order that [the audience] may know what the speech is about and [their] thought not be left hanging. The unlimited leads astray; he who gives, as it were, the beginning into the hand [of the hearer] allows him, by holding on, to follow the speech (1415a 12–16, transl. Kennedy 1991).

Before concluding, one last remark. Here, like in all of the other passages quoted above, the key position of the listener emerges again (the *telos* of speech). Or, to be more precise, these passages confirm that speakers and listeners are internal elements of speech. Hence, the evaluation of an argument can never be separate from the relationship between speakers and listeners. This means that the triad *person speaking-person spoken to-topic spoken about* (and the correspondent one *ethos-pathos-logos*) is always active in our talk exchanges.

To clarify my view, I quote Dascal and Gross again: «the *logos* of actual arguments is never separable from their *ethos* and *pathos*, nor is the logical aspect of argument separable from its expressive and contextual aspects. In other words,

we must adhere to the methodological dictum that, when we talk about argument, we *must* go beyond *the logos*; we must talk about emotional proofs and proofs from character; we must talk about, style and arrangement (Dascal and Gross 1999: 116–117)». What I have tried to show is that, if we accept the basic insight of Aristotelian rhetoric, talking about emotional proofs and proofs from character, as well as style and arrangement, is not going *beyond logos*. If Aristotle was right, proofs from *ethos* and *pathos*, style and arrangement, are *inside* the *logos*. I believe that this is the richest dowry ancient rhetoric can bring to its marriage with modern pragmatics.

References

- Conley, Thomas. 1982. *Pathé and pisteis*: Aristotle *Rhet.* II, 2-11. *Hermes*. CX: 300–315.
- Cooper, John M. 1996. An aristotelian theory of the emotions. In *Aristotle's rhetoric*, ed. A. Oksenberg Rorty, 238–257. Berkeley: University of California Press.
- Dascal, Marcelo, and Alan G. Gross. 1999. The marriage of pragmatics and rhetoric. *Philosophy and Rhetoric* 32(2): 107–130.
- Dow, Jamie. 2007. A supposed contradiction about emotion-arousal in Aristotle's *Rhetoric*. *Phronesis* 52: 38–402.
- Fortenbaugh, William W. 2002. *Aristotle on emotion*. London: Duckworth. (1st edn. 1975).
- Garver, Eugene. 1994. *Aristotle's Rhetoric: An art of character*. Chicago: University of Chicago Press.
- Grimaldi, William. 1972. *Studies in the philosophy of Aristotle's Rhetoric*. Wiesbaden: Steiner.
- Kennedy, George A. 1991. *Translation, introduction, notes and appendices to: Aristotle, on Rhetoric. A theory of civic discourse*. Oxford: Oxford University Press.
- Larrazabal, Jesus M., and Kepa Korta. 2002. Pragmatics and Rhetoric for Discourse Analysis: Some conceptual remarks. In: Dascal's Festschrift. *Manuscrito* 25(2): 33–48.
- Leighton, Stephen R. 1982. Aristotle and the emotions. *Phronesis* 27: 143–174.
- Lo Piparo, Franco. 1999. Il corpo vivente della *lexis* e le sue parti. Annotazioni sulla linguistica di Aristotele. *Histoire Épistémologie Langage* 21(1): 119–123.
- Lo Piparo, Franco. 2003. *Aristotele e il linguaggio. Cosa fa di una lingua una lingua*. Laterza: Roma.
- Mortara Garavelli, Bice. 2006. La via linguistica alla retorica negli studi italiani. Frammentazioni attuali del dominio retorico, legami e intersezioni con altri settori delle scienze del linguaggio. In *Retorica: Ordnung und Brüche. Beiträge des Tübinger Italianistentag*, eds. R. Franceschini, R. Stillers, M. Moog-Grünewald, F. Penzenstadler, N. Becker, H. Martin, Tübingen: Gunter Narr Verlag.
- Nussbaum, Craven Marta. 1996. Aristotle on emotions and rational persuasion. In *Aristotle's Rhetoric*, ed. A. Oksenberg Rorty, 303–323, Berkeley: University of California Press.
- Piazza, Francesca. 2004. *Linguaggio, Persuasione, Verità*. Roma: Carocci.
- Piazza, Francesca. 2008. *La Retorica di Aristotele. Introduzione alla lettura*. Roma: Carocci.
- Sperber, Dan and Wilson, Deirdre. 1990. Rhetoric and relevance. In *The ends of rhetoric: history, theory, practice*, eds. D. Wellbery and J. Bender, 140–155, Stanford: Stanford University Press.
- Striker, Gisela. 1996. Emotions in context: Aristotle's treatment of the passions in the *Rhetoric* and his moral psychology. In *Aristotle's Rhetoric*, ed. A. Oksenberg Rorty, 286–302 Berkeley: University of California Press.
- Venier, Federica. 2008. *Il potere del discorso. Retorica e pragmatica linguistica*. Roma: Carocci.

Debating with Myself: Towards the Psycho-Pragmatics and Onto-Pragmatics of the Dialectical Self

Marcelo Dascal

Abstract The above experiences were indeed experienced by me while writing this piece. Yet, none of them is idiosyncratic to Marcelo Dascal. Anyone who has engaged in writing has certainly experienced them, albeit perhaps without acknowledging and naming them as I suggested above. Nor are these phenomena typical of this particular kind of intellectual activity.

As I set out to establish the plan of this essay, after much reading, consultation and reflection, I consider whether its purported scope isn't perhaps too broad, whether certain issues that seemed to me at one point or another relevant are indeed relevant to its purported topic, whether certain arguments support the claims I want to put forth and whether certain objections I can foresee are important enough to be discussed, whether one or another order of presentation will be more efficient for achieving my purpose, whether the terminology to be employed should be technical or not, whether I should include summaries of earlier work I make use of or rather refer the reader to such work, and a host of similar questions. In deliberating about these questions, I am in fact debating with myself. Reasons for and against the alternatives raised are confronted in my mind and weighed, and the result is the plan I finally adopt, which spells out the 'prior intention' (cf. Searle 1983)—or rather, the set of prior intentions—that I will try to implement in the intellectual activity I am about to engage in.

We all struggle against each other. And there is always something in us which struggles against something else in us [Quoted in P. Dews, "The nouvelle philosophie and Foucault". *Economy and Society* 8: 164–165 (1979)].
Michel Foucault.

M. Dascal (✉)
Department of Philosophy, Professor Emeritus, Tel Aviv University, Ramat Aviv,
Tel Aviv 6978, Israel
e-mail: dascal@post.tau.ac.il

As I set out to implement these prior intentions, by translating them into ‘intentions-in-action’, i.e., by actually writing each section, sub-section, paragraph, sentence, and footnote, I am not satisfied with some of my formulations, which I reject in favor of others. Here too I am debating with myself, exercising self-criticism so to speak ‘in the making’.

When I had set up the plan, I had decided, for reasons of conciseness and lack of time, to avoid footnotes and digressions. But at a certain point of my writing I simply cannot resist the temptation to mention or discuss some tangentially relevant work or alternative view. I thus make an exception to my self-imposed rule. Conflicting tendencies within my mind confronted each other in a debate of sorts, and one of them triumphed. I have been the victim of akrasia or weakness of will.

As I am about to finish the work, to put the final period in the last sentence, I feel I should go on a bit and explain myself better. I re-read what I have written and realize that several of my arguments are problematic and need more work. But I am exhausted and somewhat bored by dwelling so long and intensively on this; my body aches and I badly need well-deserved rest; an important soccer match is about to begin. The motivation for shutting down the computer and dispatching the essay to the publisher without further ado and the desire to turn on the TV overcomes my best judgment. I cheat myself into believing that the essay is O.K. as is, that it is perfectly clear, sound, stylish, provocative, persuasive. I give into self-deception or wishful thinking—the result of yet another kind of inner debate.

1 Introduction

The above experiences were indeed experienced by me while writing this piece. Yet, none of them is idiosyncratic to Marcelo Dascal. Anyone who has engaged in writing has certainly experienced them, albeit perhaps without acknowledging and naming them as I suggested above. Nor are these phenomena typical of this particular kind of intellectual activity. Self-debate is as common as inter-personal debate, in all spheres of life. But, unfolding sometimes in the innermost inarticulate recesses of the mind, self-debate is of course much more difficult to observe directly. So, why not try to understand it through its external counterpart, which is a well documented and easily observable—although rather neglected—phenomenon? Why not put to use in the study of self-debate whatever is known and theorized about inter-subjective debate? This is precisely the strategy of the present essay. As obvious as it is, it has not been adopted—as far as I know—by any of the many studies of the most typical kinds of self-debate. I hope to show that this strategy is rewarding, both in complementing and systematizing such studies and in suggesting solutions for some of their problems. The tactics I will employ for implementing this strategy is methodologically unorthodox, for it stresses mainly the analogies between debating with oneself and debating with others. I count on the critics for the next step: pointing out the differences and thus

leading to further elucidation of the psycho-pragmatics and onto-pragmatics of the puzzling notion of self-debate.¹

I begin by enriching our database with illustrations of self-debate drawn from different sources; these illustrations show the ubiquitous nature and varieties of the phenomenon and the kind of mechanisms invoked in attempts to account for it (Sect. 2). I then discuss the two main kinds of relation—metonymic and metaphoric—connecting intra-personal and inter-personal debates (Sects. 3 and 4). The possibility of systematizing the structural analogies observed is then discussed in terms of a typology of inner debates parallel to the typology of external debates I have previously developed. Reconstructions of the *prima facie* similar typologies of external and internal debates in Aristotle’s rhetoric and ethics serve as a starting point for the discussion (Sect. 5). After an interlude on self-deception—as treated in the recent philosophical literature—and its ontological difficulties (Sect. 6), I return to the task of elaborating a typology of self-debates, introducing and making use of a distinction between ‘soft’ and ‘hard’ rationality (Sect. 7). Finally, in a further look at the onto-pragmatic dimension, I explore the possibility of considering the self-debating multiple self as real enough, provided one conceives of its partition as soft rather than hard (Sect. 8).

2 Varieties and Sources of Self-Debate

Every time we have to make some choice between alternatives of any sort, we are likely to be torn between the available options and to engage in a decision process in which we must adjudicate between their advantages and disadvantages. It is not far fetched to depict this process as an inner trial in which, as it were, each alternative ‘appeals’ to different interests or motives inside ourselves that endeavor to prevail and thus to determine our choice. Whether this process takes place consciously or unconsciously, through careful deliberation or on the spur of the moment, rationally or instinctively—something akin to a debate seems to be under way between at least two contenders within our mind (whatever they may actually ‘be’) defending opposite positions. And this happens dozens of times a day in our lives. Sometimes the inner debate is prompted by external circumstances—quite informal (a student question to be answered, a road to be chosen when in a crossroads, what to wear in my daughter’s wedding) or more formal (deciding what electrical appliance to buy, how to vote, whether to sign a contract, what to answer to the priest’s crucial question in your wedding ceremony); sometimes it is a purely internal process (does this argument indeed support this thesis?);

¹ I wish to thank Varda Dascal, Yaron Senderowicz, Ilana Arbel, Adelino Cardoso, and Mark Glouberman for their suggestions and especially for being the ‘others’ in a debate that helped me in conducting the self-debate that this essay reflects.

sometimes we are acutely aware of the debate that is taking place within us (as in some of the formal situations just mentioned) and sometimes totally unaware of it.

Whatever their form, these episodes are far from marginal in our mental life, for their outcome determines to a large extent our behavior, the course of our thoughts, our beliefs and desires, our projects—in trivial as well as in crucial matters. Yet, although much has been written about the insidious nature of some of them and the consequent difficulties we have in controlling them, little is known about their *modus operandi*. Let us consider a few examples of such difficulties.

Adam Smith's *Theory of Moral Sentiments* devotes a chapter to self-criticism and its difficulties:

There are two different occasions upon which we examine our own conduct, and endeavor to view it in the light in which the impartial spectator would view it: first, when we are about to act; and secondly, after we have acted. Our views are apt to be very partial in both cases; but they are apt to be most partial when it is of most importance that they should be otherwise (Smith 1976: 157).

According to Smith, in such moments it is impossible to detach oneself from one's undertaking and to evaluate one's action as an equitable judge. This is particularly difficult "when we are about to act", for "the eagerness of passion will seldom allow us to consider what we are doing, with the candor of an indifferent person" (ibid.). Following Malebranche, he argues that the reason for this is that, in such situations, the passions all justify themselves and seem reasonable and proportioned to their objects. After the action, however, we cool down and become almost as indifferent as an indifferent spectator. We can then "identify ourselves, as it were, with the ideal man within the breast" and view our conduct "with the severe eyes of the most impartial spectator". But, besides being now almost useless, our judgments even then are not "entirely candid", for self-love makes it "simply too disagreeable to think ill of ourselves" (Smith 1976: 158). He concludes that "this self-deceit, this fatal weakness of mankind, is the source of half the disorders of human life" (ibid.).

About one century before Smith, the Jansenist moralist Pierre Nicole is much more critical of self-love and, consequently, much less optimistic than Smith about the prospects of overcoming its deleterious effects. Whereas Smith attributes to the self-deceit induced by self-love, which damages our capacity of self-criticism, the responsibility for *half* our troubles, Nicole sees in self-love's impairment of our capacity for self-knowledge the cause of *all* vices (Nicole 1999: 331). He begins his essay "On self-knowledge" by observing that, although there is universal agreement about the duty to fulfill the Socratic precept *know thyself*, "nothing is more hateful to men than the light that uncovers them unto their own eyes and forces them to see themselves as they are". This tendency—which is not just a bad habit of some, but a "general inclination of corrupted human nature"—leads men to do all they can in order to avoid self-knowledge.² This seems to conflict with the opposed tendency, equally natural, men have to contemplate with pleasure the

² Nicole (1999: 310, 311).

image they have of themselves. But the conflict is merely apparent and its examination reveals the core of the strategy of self-deception employed to satisfy simultaneously both tendencies—namely, suppressing one’s defects so as to retain only a pleasing self-image. As a result, while a man “contemplates himself continuously, he never sees himself truly, for instead of himself he sees only the vain phantom he has created of himself” (p. 312).

One of the devices made use of for this purpose is to hold a confused—rather than distinct and detailed—idea of the “I”, since, according to Nicole, were this vague idea of the *Moi* unfolded, one would hardly find anything lovable in it. But this is not sufficient, for my self-image is not immune to what I observe around me. For example, the defects and misery of others and the opinions others have of me keep reminding me of my own imperfections. In order to prevent such things from tarnishing my cherished self-image, more sophisticated self-deception devices are required. Regarding the variety of defects we observe and criticize in others (including the often voiced ridicule of those who deceive themselves), my mind—“helped by self-love” (p. 319)—develops a technique of not performing the seemingly mandatory double-step inference consisting in (a) generalizing these observations and (b) concluding that, if all men I know have such defects, then I am also likely to have them. This seems to be grounded on an ingrained belief in the singularity of one’s self vis-à-vis all the others, which exempts one from applying oneself even elementary logic. Whatever its source, its result is a sort of blindness: “[Men] see constantly the image of their own defects in those of all other persons, and they don’t want to recognize them in those images” (p. 320).

As for the particular judgments of others about me, the mechanism of self-deception takes different forms. The simplest one is to exercise ‘selective perception’—simply to ignore those who are against us and to pay attention to those who are favorable to us (p. 318). Things become more complicated when we face the rare charitable and well-intentioned people who “try to rescue us from the illusion in which we live regarding ourselves” (p. 321). The way self-love gets rid of this embarrassment consists in accusing such people of acting on self-interest, ill faith, or similar motives, according to the principle that “nothing can be condemned in us by a concern for equity and justice” (ibid.). Since to be accused of such motives, especially when he believes to be acting altruistically, hurts the critic’s self-image, next time he will think twice before volunteering his criticism, so that in the long range the species of well-intentioned critics will tend to vanish. Instead, people will keep for themselves their criticism or will vent it before everyone else except its target, i.e., the person who might have benefited most from them. In so doing, however, they make it easier for one to rank them among the cohorts of slanderers and gossipers; this, in its turn, further justifies one’s discarding their criticism.

Another problem is that one of the greatest defects we do not want to acknowledge in us is that of not wanting to see the truth about us. This, of course, is incompatible with our constant effort not to admit any criticism. The solution is to “grant something to the one as well as to the other” of these passions (p. 323). Self-love’s subtle maneuver for achieving this feat is to be carefully asymmetrical

in the partition: While criticism of our essential defects is ruled out, the proof that we do want to know truth about ourselves is our readiness to admit at least some criticism, provided it is only of “minor defects that do not disfigure the image we have of ourselves, preserving all its beauty” (p. 324).

That we create so many successful stratagems to hide the truth in a matter of such importance as the knowledge of ourselves indicates, for Nicole, that our presumed ‘natural love of truth’ is far from actually guiding our behavior. In fact the main usage of this presumption is “to persuade ourselves that what we love is true”. In other words, the idea that we love truth is the very principle of wishful thinking: “[W]e do not love things because they are true, but we believe they are true because we love them” (p. 326).

Both Smith and Nicole believe there are ways to combat self-deception. One of these ways is by means of rules of conduct. For Smith, these are the “general rules of morality”, stemming from our “continual observations upon the conduct of others”, which “insensibly lead us to form to ourselves certain general rules concerning what is fit and proper either to be done or to be avoided” (Smith 1976: 159). He insists that these rules are “founded upon experience” and stresses that, although once established they function as standards of judgment of an action, we should not be misled into thinking that “the original judgments of mankind with regard to right and wrong, were formed like the decisions of a court of judicatory, by considering first the general rule, and then, secondly, whether the particular action under consideration fell properly within its comprehension” (p. 160). Nevertheless, “when they have been fixed in our mind by habitual reflection, [they] are of great use in correcting the misrepresentations of self-love concerning what is fit and proper to be done in our particular situation” (ibid.).

Nicole’s position is radically opposed to Smith’s moral inductivism. For him, given our corrupt nature, moral rules must stem not from experience but directly from God. Even so, they are not immune to wishful thinking and thus may not suffice to overcome self-deception, for we can attempt to deflect the rule in order to adjust it to our corrupt inclinations: “[Men] do not want only to follow their interests and their passions; they want also to be approved when following their interests and their passions” (Nicole 1999: 338). We can nevertheless adopt practical, rather than moral, rules that may be of some help in preventing self-deception. A remarkable example of such a rule, which can enable us to steer a correct course between criticism and flattery, is to realize that the language of both is “peculiar”—the former a language of moderation, the latter a language of exaggeration. In both cases we have to learn to discern between the “precise meaning of the expressions and the thoughts they allow us to read in the mind of those who use them” (p. 374). The rule that formulates this accurate lesson in pragmatics is that, in order to assign the appropriate import to criticism and flattery, we must expand what is only half-said in the former and add weight to it, while we must look out for what the latter silences and subtract weight from what it actually says. Self-love, of course, leads us to do exactly the opposite of what this rule prescribes, as we have seen.

Another Jansenist thinker, Pascal, whose influence upon Nicole is noticeable, was also concerned with the inner battle between the ‘lower’ and the ‘higher’ parts of the soul. I want to single out here his attack on the ‘imagination’ as a key factor in this battle, because of the particularly insidious character he attributes to it. In a well-known section of the *Pensées*,³ Pascal describes the deceptive activity of the imagination as follows:

It is this deceptive part of man, this master of error and falsity, and the more perfidious because it is not always so; for it would have been an infallible rule of truth if it were an infallible rule of lie. But, although it is for the most part false, it gives no mark of its quality, marking with the same sign the true and the false.

...

[The imagination] is a superb power, enemy of reason, which pleases itself in controlling and dominating it. ... It makes [one] believe, doubt, deny reason; it suspends the senses, it makes them sense. ... (Pascal 1960: 95).

Imaginative ability grants self-assurance, hence self-satisfaction and assertiveness. Persons endowed with it regard people imperiously; they dispute daringly and confidently ... and their joyful face often gives them the upper hand in the opinion of the audience, since imaginary wise men are favored by judges of the same nature (ibid.).

While imagination makes its practitioners happy, “reason cannot but render its friends miserable; the former covers [its friends] with glory, the latter with shame” (p. 96).

The overwhelming superiority of imagination over reason ultimately stems from the latter’s prudence, which does not allow reason’s “friends” the same unbridled self-assurance afforded by imagination to its. In the camp of reason, there is no decisive and unquestionable principle of truth, whereas imagination provides “many and excellent” principles of falsity. Furthermore, the “two principles of truths—reason and the senses—, besides the fact that they lack sincerity, abuse each other reciprocally” (p. 99). It is within this division of the camp of truth that imagination has a place too, not as a fiend but as a friend. Pascal the scientist is well aware of the positive role of the imagination in geometrical constructions, in conceiving ideal experimental conditions à la Galileo, in the value of analogies for theory construction à la Kepler.⁴ But, like the senses and reason itself, imagination—even when properly used—does not provide certainty. Furthermore, the fact that the boundaries of its legitimate uses in the pursuit of truth cannot be precisely drawn prevents them to be distinguished sharply from their illegitimate counterparts—which form the majority of its uses—and grants the ‘negative’ imagination so much power of deception. What causes the self to be prey to the many kinds of deception Pascal ascribes to imagination is that the deceptive agent

³ Paragraph 82 of the Brunschvicg edition, as published in Pascal (1960).

⁴ Recall, for example, his experiments on the vacuum and his essay *De l’esprit géométrique*. Pascal (1985) contains, along with this essay, the draft of a Preface for a planned *Traité du vide*, as well as other posthumously published texts, notably the *Ecrits sur la Grâce*—a judicious juxtaposition of writings, which highlights the dual role Pascal assigns to the imagination.

cannot be identified as such and ruled out once and for all by some visible ‘mark’; any such attempt would risk to throw the baby with the bath water, since the very same agent may be, in important cases, non-deceptive.

Thus, the negative ‘force of the imagination’—a theme that emerges forcefully in renaissance literature about the relations between the parts of the soul (Piro 1999)—does not stem—for Pascal—from the imagination’s *belonging* to the lower part or being *inherently* unreliable. It is because this is not the case, i.e., because the imagination can eventually have also a positive ‘higher’ role, that it can be a persuasive *lawyer* of the passions and deceive the mind (‘esprit’). Unlike Pascal, other thinkers of the period attribute the force of the imagination to an inherent, closer connection with the sensitive soul and with the will, upon which the passions act:

...these passions originate in the senses, they are always on their side, the imagination never represents them to the mind without speaking in their favor; with such a good advocate they corrupt their master and are victorious in every cause; the mind listens to them, examines their reasons, considers their inclinations, and in order not to sorrow them pronounces itself very often to their advantage; the mind thus betrays the will whose prime minister it is; it cheats this blind queen and, disguising the truth, presents her with unfaithful reports; ... the imagination, filled with the species it receives from all the senses, solicits the passion and shows to it the beauties or the ugliness of the objects that can move it; ... [and] the appetite follows blindly all the objects that the imagination proposes to it (Senault 1664: 15–17).⁵

Whatever the precise ‘location’ of the imagination in the map of the soul, whatever the source of its power, one must beware of it, because it can be—and more often than not actually is—appropriated by the ‘wrong’ party, thus becoming its deception-engendering representative in the self-debate going on within the soul.

As a last illustration of the phenomenon of inner debate, let us turn now to the twentieth century, focusing on the phenomenon that came to be known as ‘cognitive dissonance’ in the wake of the psychological theory (Festinger 1957, 1964) that first proposed a systematic account of its nature and consequences. According to this theory, a cognitive dissonance occurs when there is a “nonfitting relation” among one’s cognitions—understood as “any knowledge, opinion, or belief about the environment, about oneself, or about one’s behavior” (Festinger 1957: 3). For example, one may think children should be quiet in the presence of adult guests and yet be pleased when one’s child noisily captures their attention; or one may be a convinced anti-racist and yet withdraw one’s children from school as soon as Arabs enroll in it. Dissonance is, thus, an inconsistency within an individual’s set of cognitions in a given domain. It is more rigorously defined as follows: two “elements of cognition” x and y are dissonant if not- x follows from y (p. 13). One should not be misled, however, by the formal character of this definition. Festinger himself points out that he introduced the term ‘dissonance’ instead of ‘inconsistency’ because the former “has less of a logical connotation” (p. 2), and stresses that ‘follows from’ should not be invariably interpreted as a strict logical relation

⁵ Quoted in Piro (1999: 168).

(p. 14). Logical inconsistency is but one of the possible sources of dissonance; others include a variety of contextual elements, such as motivation and desired consequences, cultural conventions, past experience, etc.

Assuming that “an individual strives toward consistency within himself”—an assumption amply supported by “study after study”—the theory claims that dissonance is “psychologically uncomfortable”; consequently, it is a “motivating factor in its own right”, which pressures the individual to act in order “to try to reduce the dissonance and achieve consonance”, as well as “to avoid situations and information which would likely increase the dissonance” (pp. 1–3). In this respect, cognitive dissonance is analogous to drives such as hunger. The strength of the pressure to reduce dissonance is claimed to be a function of the degree or magnitude of dissonance, which in turn is a function of the number and importance for the individual of the dissonant elements between two dissonant clusters of cognitions. These factors are also determining—along with other factors—of the resistance experienced in attempting to reduce or avoid dissonance, be it by trying to change a “behavioral cognitive element” (e.g., modifying one’s behavior or one’s perception thereof) or an “environmental cognitive element” (e.g., changing one’s opinion about someone by relying on the support of others for the new opinion), or by suppressing potentially dissonant elements or adding consonant ones, thereby reducing the weight of the dissonant elements in the cluster.

Several of these notions can be straightforwardly applied to the examples of inner debate discussed above. For example, the clash between one’s pleasant self-image and one’s unpleasant image as reflected in others’ criticism clearly involves cognitive dissonance, and the various maneuvers described by Nicole amount to attempts to reducing it through one or more of the methods described by Festinger. Generalizing a bit, consider the theory’s view of the resistance to the reduction of dissonance or to the avoidance of its increase as what explains the persistence of certain kinds of dissonance as opposed to cases in which a belief is easily changed in the light of counter-evidence. Persistent dissonance could be seen as an indicator of a sort of inner debate. Its resistance to reduction might be due to the fact that it arises between cognitive elements belonging to a core of basic beliefs, feelings, or behaviors of the self, or else to a set of nodal cognitions, whose change would require corresponding changes in a vast array of cognitions. In either case, the dissonant elements would be ‘entrenched’ components of the individual’s mental set (cf. Dascal and Dascal 2004a, b)—in Nicole’s examples, self-love and the desire for self-knowledge; these components somehow co-exist in the self, in spite of their latent inconsistency; nevertheless, none of them can be easily given up, even when their inconsistency becomes apparent and causes the discomfort characteristic of cognitive dissonance.

Research in cognitive dissonance theory has devoted particular attention to the kind of dissonance that arises after a decision between mutually exclusive options has been made. Experimental data (cf. Festinger (ed) 1964; Brehm 1956) show that immediately after decision a process of ‘rationalization’ of the decision begins, whereby the “attractiveness” of the non-chosen option is lessened and the attractiveness of the chosen alternative is increased. This “spreading apart” of the

alternatives is achieved by lowering the evaluation of the “good” features of the rejected alternative and raising that of the adopted one, as well as by lowering the evaluation of the “bad” features of the latter and raising that of the former—thus reducing the dissonance between the choice made and those cognitive elements in the options considered that are against it.

Notice that the dissonance the person is under pressure to reduce after having made the decision arises not between the two alternatives, but between her “knowledge of the direction in which (s)he is going” after having decided and the “cognitive cluster of elements” that “accumulated during the process of coming to a decision” (Festinger 1957: 35) and runs counter what she decided. Obviously these elements must remain cognitively available after the decision—otherwise they could not cause post-decision dissonance. Post-decisional dissonance, therefore, is a rather surprising phenomenon, at least insofar as it does not conform to the intuitive idea that a decision ‘puts an end’ to doubt and inner debate by picking up one of the alternatives and removing the others, along with their supporting arguments, from the decider’s cognitive horizon. Instead, it suggests that, as a rule, the pre-decision debate goes on beyond the decision point—albeit in a slightly different and somewhat attenuated form.

It is perhaps in order to preserve at least part of the intuition in question that Festinger emphasizes the singularity of the moment in which a decision “is made”, which, according to him, creates a discontinuity in the cognitive process between what precedes and what follows it. Although he admits that the evidence about what happens immediately prior to a decision is both scant and problematic (cf. Festinger (ed) 1964: 6, 9), he posits a sharp distinction between pre-decisional ‘conflict’ and post-decisional ‘dissonance’:

The person is in a conflict situation before making the decision. After having made the decision he is no longer in conflict; he has made his choice; he has, so to speak, resolved the conflict. He is no longer being pushed in two or more directions simultaneously. He is now committed to the chosen course of action. It is only here that dissonance exists, and the pressure to reduce this dissonance is *not* pushing the person in two directions simultaneously (Festinger 1957: 39).

On this picture, self-debate at the stage of cognitive deliberation toward a choice is like a fierce battle between opposed forces trying to win the person’s acquiescence, whereas at the stage of dissonance reduction after the choice is made it is a rather quiet process of suppressing the remaining resistance and regrouping and redeploying the victorious forces after the battle. These metaphors suggest quite different types of self-debate and a sharp gap between them. Nevertheless, the robust experimental results on post-decisional dissonance rather suggest a continuity between them.

3 Intra-Personal and Inter-Personal Debates: Metonymic Relations

In what follows, I will assume that the phenomena presented in Sect. 2 and in the opening paragraphs of this essay illustrate mental processes that can be properly termed ‘self-debate’, for they (a) take place within the mind and (b) are significantly similar to the debates we are most familiar with, namely inter-personal agonistic confrontations.⁶ Let me now begin to implement the strategy of trying to learn more about self-debate through the glasses of inter-personal debate.

The first step will be to examine the kinds of relations obtaining between the two phenomena. For this purpose, I will rely on the general classification of relations proposed by Leibniz, who distinguished between two principal types of relation—*connexio* and *comparatio*.⁷ The former involves a (spatial or temporal) contiguity between the relata and some sort of direct influx or link between them; part-whole, part-part, order, possessor-possessed, and causal relations, among others, belong to this type. The second type is characterized by a ‘distance’ between the relata, a lack of direct ‘contact’ (hence of direct influx) between them; this type comprises relations of similitude, difference, opposition, super-ordination, exemplification, and analogy, among others. Jakobson (1962: 64) considers this dichotomy—which he designates with the terms *metonymic* and *metaphoric*—as essential for understanding not only verbal behavior, but for human behavior in general. I retain here his terminological suggestion.

Let us consider first the metonymic relations. Obviously, there are causal relations between an external debate and several mental activities directly involved in it. In addition to the mental processes normally involved in understanding, interpreting, evaluating, and reacting to any interlocutor’s intervention in a non-agonistic conversation for example, engaging in a debate with someone requires also more specific mental activities such as anticipating what might be the opponent’s replica to a possible reply to his present objection I am considering, and what I might respond to his anticipated replica. These and similar mental activities, however, are part and parcel of the communicative chores of taking part in the *social* activity “debate”, which takes place between an opponent and a self who interact in a public arena. Although these are mental moves that take place *in foro interno*, they are entirely subservient to what is currently going on *in foro externo*. As such, they do not configure per se the quite distinct phenomenon of a debate the self engages with itself, not necessarily in the presence of an external opponent or prompted by the need to react to his objections. Insofar as this

⁶ In my previous work on inter-personal debates (e.g., Dascal 1998, 2000), I have employed the expression ‘polemical exchange’ as the general term for their various types. In the present context, I prefer the more neutral term ‘debate’.

⁷ Leibniz (1903: 355). He also employs the terms *conjunctio* for the former and *convenientia* for the latter. On this classification of relations and their counterparts in 19th–20th century thought, see Dascal (1978: 106–115).

phenomenon enjoys a relative autonomy from the social environment, its study belongs properly to *psycho-pragmatics* (at least to the extent it involves the use of linguistic/semiotic or quasi-linguistic/quasi-semiotic means), along with other ‘solitary’ mental processes such as meditating, reasoning, daydreaming, etc. In contradistinction, the investigation of the mental contributions to inter-personal debate belongs straightforwardly to another branch of pragmatics—*socio-pragmatics*.⁸

Genuine self-debate can, however, bear a metonymic relation of the causal type to external debate. Criticism by others may engender, along with a public polemical exchange, an inner process of self-criticism or at least of self-examination—which may or may not end up with the modification or one’s self-perception. This will depend—in Nicole’s terms—on whether the self-love or the self-knowledge drive prevails, or—in cognitive dissonance theory’s terms—on the comparative resistance to dissonance reduction of one’s self-concept vis-à-vis the critical challenge to it (cf. Thibodeau and Aronson 1992), or on several other factors (e.g., personality traits such as a tendency to be self-critical or not, the credibility of the critic in the eyes of the criticized, etc.). Although triggered by a move that usually leads to external debate, the self-debate in question often pursues its own course and has its own intensity, quite independently of the external debate, which can even abort right at the outset.

The impact of suddenly being in the vortex of an intellectual controversy, where “one’s ingrained, taken-for-granted sense of how *certain* things are ... is suddenly and insistently confronted by something very much at odds with it”, is powerfully described by Smith (1997: xiv–xv):

[A]n impression of inescapable noise or acute disorder, a rush of adrenalin, sensations of alarm, a sense of unbalance or chaos, residual feelings of nausea and anxiety. ...

Thus a sense of intolerable wrongness in some journalist’s description or fellow academic’s analysis can set the mind’s teeth on edge and produce a frenzy of corrective intellectual and textual activity: letters to the editor, exposures, rebuttals, and sometimes tomes and treatises. ... [These activities are especially energetic when the wrongness is experienced as] bearing on one’s personal safety, dignity, or even identity (social, professional, and so forth), so that a response seems summoned and obligatory.

Although she stresses the public aspect of the response, the inner response she describes to what she views as an example of “the familiar but by no means simple phenomenon of *cognitive dissonance*” (ibid.) is certainly no less intense, and in all likelihood consists in engaging in self-debate. Confronted with external objections

⁸ For the distinction between socio- and psycho-pragmatics, see Dascal (1983: 42–52). Viewing ‘psycho-pragmatics’ as a branch of pragmatics or the theory of language use is justified in the present case insofar as one takes into account that language is actually ‘used’ in mental processes not only by way of employing in them words or a ‘mental lexicon’ (Aitchinson 1994), but also through the presence in such processes of underlying structures that are originally or typically linguistic. One such structure is the dialectic structure of inter-personal debate (cf. Dascal 2002).

that question the very grounds of my basic beliefs, my data, my research aims, and my methodology, I can hardly content myself with a public reaction. Even if I adopt what Peirce called the ‘method of tenacity’ (i.e., strict adherence to my position, disregarding the onset of doubt due to the criticism and avoiding further encounters with such critics), I must somehow reassure myself that I am right and “they” are wrong. If I view myself as justly tenacious, i.e., as having justifiable grounds for my position, then I must discuss with myself, at the very least, whether I have made no mistakes in my assumptions and in my inferences from them. And if I go as far maintaining a methodology that calls for ‘symmetry’ between “those beliefs currently seen as absurd or wrong as well as those now generally accepted as true” (Smith 1997: xvi), when facing systematic misrepresentations of my position and radical objections to it, I certainly must engage in a thorough inner search of where I might have gone wrong, perhaps by not really living up to the principle I profess. Any of these forms of self-examination is likely to involve an inner debate, independently of whether I also engage in an external debate or not.

The effect of external debate on inner debate may be less direct, in the sense of being more sensitive to the style of argumentation and attitude of an opponent than to the content of the positions he attacks and defends. A beautiful description of this kind of effect is given by Proust:

And when Bergotte’s opinion was thus contrary to mine, he in no way reduced me to silence, to the impossibility of providing any reply, as M. de Norpois’ would have done. This does not prove that Bergotte’s opinions were less valid than the Ambassador’s; on the contrary. A powerful idea communicates some of its power to the person who contradicts it. Partaking of the universal community of minds, it penetrates, grafts itself onto the mind of him whom it refutes, among adjacent ideas, with the aid of which, counter-attacking, he complements and corrects it; so that the final verdict is always somewhat the work of both discussants. It is to ideas which are not, strictly speaking, ideas at all, to ideas which, based on nothing, can find no foothold, no fraternal support in the mind of the adversary, that the latter, grappling with thin air, finds nothing answer. M. de Norpois’s arguments (on art) were unanswerable because they were devoid of reality.⁹

It seems that in genuine cases of self-debate prompted by or influenced by external debate, the voice of the external objector is internalized by the self-debating self into an ‘inner voice’ that sort of becomes part of the self, thus distancing itself from the objector’s voice and acquiring sufficient autonomy for conducting a debate with other ‘voices’ present in the self. Even in the formal situations that require a decision of sufficient importance to be preceded by a deliberative process, the evidence of post-decision dissonance strongly suggests that the inner debate follows its own course quite autonomously, for it does not cease once it supplies the decision required by the formal situation.

⁹ Proust (1954: 562).

4 Inter-Personal and Intra-Personal Debates: Metaphorical Relations

Metaphorical relations between mental activity and external, bodily activities abound in the literature. Sometimes these relations are (or are taken to be) ‘etymologically’ grounded and as such as providing solid evidence for the nature of mental activities. Locke, for example, invokes the sensory origin of words such as ‘apprehend’, ‘comprehend’, ‘conceive’, ‘imagine’, commonly used to describe mental operations, in support of his thesis that “notions quite removed from sense have their rise thence, and from obvious sensible ideas are ... made to stand for ideas that come not under the cognizance of our senses” (*Essay* 3.1.5). Horne Tooke, an influential and controversial figure in English ‘philology’ in the 18th century goes a step further and views the role of language in thought as straightforwardly metonymic, by claiming—in the footsteps of Hobbes—that “what are called [the mind’s] operations are merely the operations of language”—a thesis he also based on ‘etymological’ evidence.¹⁰

In general, language figures prominently in the metaphorical conceptualization of mental processes. Recall, for example, the notions of a ‘mental lexicon’ (Aitchinson), of ‘voiceless speech’ as what thinking amounts to (Watson), of ‘inner speech’ (Vygotsky), of the polyphony of inner ‘voices’ alluded to above, of mental ‘dialogues’ (Bakhtin). It seems all too natural to add to this series the notion of ‘inner debate’ and to explore how far it can take us in conceptualizing adequately the phenomena it is supposed to refer to.

In this exploration, we should bear in mind the tendency—illustrated above—to forget that, *prima facie* at least, a metaphoric, unlike a metonymic relation, relates a target and a source that belong to distinct domains or categories: “A relationship of the form *A is B* (for example AN ARGUMENT IS A FIGHT) will be a ... clear metaphor if *A* and *B* are clearly different kinds of things or activities” (Lakoff and Johnson 1980: 85). The category-distance between target and source cannot be obliterated if a relation is to count as metaphorical. If one’s concept “fight” includes, in addition to physical aggression and the consequent pain, also psychological aggression and pain, for instance, then ‘an argument is a fight’ would come closer to being just a literal predication.¹¹ What preserves the distance is the selectiveness of the properties metaphors project from one domain onto the other, which requires from their users and interpreters to be attentive to which properties of the source domain are expected to have counterparts in the target domain and which not. The more ‘central’ or ‘essential’ are the properties of the one transferred to the other, the less the relation remains metaphoric.

In our case, we should ask what—if any—are the counterparts in the target “self-debate” of essential features of the source “external debate”, such as

¹⁰ John Horne Tooke, *Diversions of Purley* (1786), II, 51b (quoted in Aarsleff 1967: 13).

¹¹ For further analysis of the metaphoric and metonymic relations in ARGUMENT IS WAR, see Dascal (2004a, b).

discursiveness or the fact that external debates take place between clearly demarcated and ontologically independent entities. If we go as far as positing as a condition for speaking of self-debate the splitting of the self into ontologically distinct components such as the Freudian id, ego, and superego, each acting intentionally in its combat with the other (as in external debates), then we are very close to interpreting the expression ‘self-debate’ not metaphorically but as providing a literal theoretical account of the phenomena it refers to. This, however, would require the theorist to go further and answer such questions as: What exactly is the ‘discursive’ medium or ‘inner language’ in which self-debate is conducted? In what sense can intentions—a typically conscious phenomenon in external discourse—be assigned to an *unconscious* part of the self? And—last but not least—if several selves indeed co-exist within ‘the breast’ (as Adam Smith would put it), how to account for the feeling of unity of ‘the’ self as well as for properties attributed to ‘the’ person (e.g., responsibility) that seem to presuppose some level of identity at which the ‘several selves’ must become nothing but aspects of the ‘real’, unitary self?

No doubt these are important and heavy questions, having to do mostly with ontological matters. Presumably, it is the difficulty of providing satisfactory answers to these and similar questions suggested by the metaphor ‘self-debate’ that ultimately has led some scholars to disparage the ‘scientific’ status of ideas such as those of dichotomous and tripartite selves, on the grounds that “we ought not to take the notion of ‘several selves’ very literally” (Elster 1986a, b: 30). Elster concludes his analysis of several theories that appeal in one way or another to this notion by claiming that, in the phenomena these theories discuss, “in general, we are dealing with exactly *one* person—neither more nor less” (ibid.). It is the job of *this* person, this unitary self, to deal with *his* problems of cognitive coordination and motivational conflicts, i.e., there is no need, in accounting for the observed phenomena, for additional explanatory constructs such as “several selves”. Particularly misleading, in Elster’s eyes, is the metaphor of inner debate, for motivational conflicts and cognitive coordination problems of an individual “do not sort themselves out in an inner arena where several homunculi struggle to get the upper hand” (pp. 30–31).

Regarding Elster’s substantive conclusion, I will try to show later on that it is more profitable to discuss these ontological issues in terms of the ‘onto-pragmatics’ (Dascal 1992) of self-debate—a perspective that may help to dissolve the apparently unsolvable dilemmas they seem to involve, just as the ‘psycho-pragmatic’ perspective permits to deal more productively with the traditional issue of whether language is constitutive of thought or not (Dascal 1995). But, in line with my belief in the need of metaphoric conceptualization even in science, I beg to disagree with Elster as to the value of pursuing the attempt to understand self-debate via its relation with external debate—even if this relation is “nothing but metaphoric”. I will substantiate this value by considering in the remaining of this section and in Sects. 5 and 7 what I take to be significant *structural* analogies between intra-personal and inter-personal debates, regardless of their eventual ontological implications—which I will address in Sects. 6 and 8.

The metaphor ARGUMENT IS WAR certainly captures the agonistic aspect of both inner and external debates. It permits, in both, to speak of ‘battles’ between ‘contenders’ who have opposed aims and targets, of ‘forces’ that are ‘mobilized’ in order to achieve these aims, of ‘positions’ that are ‘attacked’ and ‘defended’, of ‘strategic’ and ‘tactical’ moves that may be successful or not, of ‘defeat’ and ‘victory’. Structurally, this metaphorical conceptualization raises such questions as: Who are the contenders in the ‘war’? What are the aims? What are the criteria of partial success and of complete victory? What are the weapons and moves typically used or permitted?

In both inner and external debates, the answers to these questions—i.e., the way one fills in the appropriate slots of the general structure—will vary with the particular type of debate and its circumstances. And the fact that one is ‘inner’ and the other ‘external’ constrains the nature of the candidates for each slot. Thus, in inner debate the contenders may be different ‘faculties of the mind’, e.g., the ‘will’ and the ‘intellect’ (as in Pascal and Senault), or opposing components or elements of the same faculty (as the drives “self-love” and “self-knowledge” in Nicole or the dissonant cognitions in the theory of cognitive dissonance). Accordingly, each of the specified contenders will use in the combat ‘his’ characteristic and most effective weapons and moves, pitting—respectively—passions and emotions against reasons, or passions against passions, or reasons against reasons. Each will also attempt to ‘neutralize’ (i.e., rule out or disqualify) the adversary’s weapons: Paraphrasing Pascal, the ‘heart’ will acknowledge reasons that Reason doesn’t accept, and vice versa. The aims of each side may range from leading the mind into a certain mental state (the adoption of a desire over another, of a belief over another, of an intention over another) to influencing the self’s behavior (looking for additional information in order to reach a decision, action taken in the wake of a decision, acting on an intention). Success or failure in achieving these aims may be total (complete self-persuasion vs. suspension of judgment; implementing a decision vs. not acting on it) or partial (some degree of self-persuasion; partial, hesitating action).

5 Towards a Typology of Intra-Personal Debate: The Aristotelian Clue

Taking now into account the different types of debates in each of the two domains, let us try to see whether the structural analogy extends to the level of a parallel between the typology of external debates and that of inner debates. I refer the reader to my publications on the former (e.g., Dascal 1998). For our purposes here, the following schematic summary should suffice.

Traditionally, two main types of debates are distinguished. Systematically, they can be described as ‘ideal types’ and characterized by their differences in aim, extension, procedures of decision, preferred moves, form of conclusion, and

cognitive gains. One is taken to be paradigmatic of what a scientific, rational debate should be, while the other is typical of conflicts of interest, where emotion and power struggles prevail. I have chosen to call them, respectively, ‘discussion’ and ‘dispute’:

Discussion

Aim: determining the true position

Extension: a well-defined problem or issue

Decision procedure: application of a method of decision the contenders accept

Preferred move: proof

Possible closing: solution

Possible cognitive gain: elimination of mistaken belief

Dispute

Aim: victory over the adversary

Extension: a well-defined content and attitudinal divergence

Decision procedure: no “internal” method of decision agreed upon by contenders

Preferred move: stratagem

Possible closing: dissolution by intervention of “external” factors

Possible cognitive gain: discovery of irreconcilable positions or attitudes

The discussion-dispute dichotomy is usually perceived as exhaustive, so that contenders (as well as analysts) tend to view a particular dispute either as belonging fundamentally to one of these types or to the other, and to emphasize one or more of the following oppositions between them:

<u>Discussion</u>	<u>Dispute</u>
• <i>The truth</i>	• <i>My truth</i>
• The issue can be <i>decided</i>	• The divergence <i>cannot</i> be decided
• Logic	• Rhetoric
• Rational	• Irrational
• Is about content	• Is about attitudes
• Leads to change of opinion	• Does not lead to change of opinion

One of the contributions of recent research on inter-personal debates has been to propose a third, intermediary ideal type, which mitigates to some extent the dichotomous character of these oppositions. It has been labeled ‘controversy’ and has the following features:

Controversy

Aim: persuasion of the adversary or of a competent audience

Extension: begins with well defined issue and quickly extends both to ‘horizontally’ and ‘vertically’ related problematizations

Decision procedure: opponent’s methods are questioned

Preferred move: argument

Possible closing: resolution

Possible cognitive gains: clarification of divergence; conciliation of opposites; emergence of innovative ideas

In a controversy, unlike in a dispute, each contender does not assume a priori that the adversary is entirely wrong while he is entirely right, thus abandoning from the outset any hope of rationally persuading the other to change his mind. External intervention (e.g., by a tribunal) can dissolve the dispute, but usually does not change the contender's confidence in the correctness and justice of their positions. On the other hand, controversy differs from discussion in that, while it is predicated upon the possibility of rational persuasion, it does not assume that this can only be achieved through the acceptance by the contenders of the unquestionable results of the application of a method both unconditionally accept. In controversy, questioning of basic assumptions, methodology, and the modularity of issues is the rule. This leads to a wide span of disagreements that can be quite radical—including doubts about the alleged certainty of the decision procedures. Hence, rational persuasion in controversy, unlike in discussion, is not usually comparable to the dramatic revelation of the truth and compelling full conversion.

Our task now is to attempt to spell out a parallel typology for self-debates and to find out whether it is adequate for accounting for the phenomena this expression is supposed to refer to. In other words, we must characterize the presumed inner analogues of discussion, dispute and controversy as described above (let us call them 'self-discussion', 'self-dispute' and 'self-controversy'), to check how far the analogy goes, and to show the descriptive and explanatory power of these theoretical constructs.

We can begin by following a clue provided by Aristotle. He applies the same term, *deliberation*, to the external debate that takes place in an assembly in the course of a collective decision and to the inner debate that takes place within an individual who seeks to determine the proper ethical course of action for himself, in his particular circumstances.¹² Aristotle's terminological choice is not casual, for it reflects a deep analogy between his conceptions of the two domains, and there is significant overlapping between the two texts.¹³ Deliberative

¹² Hampshire, who elaborates the Aristotelian notion of deliberation as the cornerstone of his moral theory, puts to use the external-internal debate analogy in this vivid passage: "The picture of the mind that gives substance to the notion of practical reason is a picture of a council chamber, in which the agent's contrary interests are represented around the table, each speaking for itself. The chairman, who represents the will, weighs the arguments and the intensity of the feeling conveyed by the arguments, and then issues an order to be acted on. The order is a decision and an intention, to be followed by its execution. This policy is the outcome of the debate in the council chamber" (Hampshire 1991: 51).

¹³ Aristotle considered the notion of deliberation sufficiently important on its own to write a special treatise devoted to it (*Peri Symboulías*, which is listed in Diogenes Laertius' catalogue). On possible traces of this treatise in the *Rhetoric*, see the Spanish translation and commentary by Q. Racionero, pp. 99, 214, 222.

discourse—one of the three kinds of oratory dealt with in the *Rhetoric*—gives advice to individuals (in private counseling) or to an assembly about taking a certain course of action. Deliberation—as dealt with in the *Nicomachean Ethics*—is the mental process through which an individual establishes his preferences and decides how to do. In both domains, deliberation’s aim is to lead to rational persuasion in favor of one of the options—in the external domain, other-persuasion; in the inner domain, self-persuasion. In both domains, deliberation concerns the choice of a future action among a set of possibilities—none of which is certain to occur. Therefore, it deals not with the necessary—which is the object of science—but with the probable. Nevertheless it is a rational procedure of evaluating the available options through comparison of their advantages and disadvantages. Consequently, deliberating well comports a form of correctness, although it does not yield knowledge or truth. It is a “correctness of thinking” which differs from the correctness of ‘demonstrating’ characteristic of science, but only of ‘calculating’, characteristic of deliberation.¹⁴

These shared properties alone, I submit, clearly suggest that ‘deliberation’ occupies, in Aristotle’s scheme—both in its inter-personal and intra-personal versions—the slot occupied by ‘controversy’ in my inter-personal typology. This would justify (at least as a weighty *ad verecundiam* argument) taking intra-personal deliberation, as described by Aristotle, as a prime example of a type of self-debate that may be called ‘self-controversy’. It remains to be seen whether a similar analogical argument applies to the other components of the Aristotelian typology.

As we have seen, Aristotle emphasizes the difference between prudence or practical wisdom, the ‘intellectual virtue’ that relies upon deliberation, and science. Unlike the former, where the ‘calculating’ deliberative procedure evaluates the best alternative but does not *determine* it with certainty, self-debates in scientific matters are those capable of being solved by rigorous proof. In this respect, they correspond quite straightforwardly to what I have provisionally termed ‘self-discussion’. So, the *Nicomachean Ethics* suggests, for the sphere of self-debate, a plausible analogue for my ‘discussion’ in the inter-personal debates. A certain kind of ‘self-criticism’ might be taken as a paradigmatic example of self-discussion in this sense. It would consist in examining (e.g., in a Cartesian or in a Popperian way) one’s reasoning or explanation by applying to them a precise ‘method’ (e.g., methodical doubt or the logic of refutation, respectively) that would be capable of determining whether we should reject them as mistaken or be assured of their correctness.

Does the *Rhetoric* support this second analogy? Of the two genres of discourse that flank the deliberative genre, the most likely Aristotelian candidate for the inter-personal counterpart of self-discussion seems to be the forensic. One might think that the juridical system provides the judge, who has to decide whether an

¹⁴ In this paragraph, I refer to, paraphrase and extrapolate statements that can be found in *Rhet.* 1357a–b, 1358b, 1359a; *Nic. Eth.* 1112a–1113a, 1139a–b, 1141a, 1142b.

injustice or wrongdoing (defined as an action that violates the law and inflicts damage; *Rhet* 1368b8) has been committed by the accused, with the methodical tools necessary to decide with certainty. Ideally, he would have to accept nothing but the ‘complete proofs’, rather than ‘probabilities’ or ‘signs’ (*Rhet*. 1359a7). If this were the case, there would be no room for ‘deliberation’, and the judge’s decision would be certain if it resulted from the correct application of the method. It turns out, however, that none of the five kinds of juridical ‘proofs’ examined by Aristotle (*Rhet*. 1375a23-1377b14)—laws, witnesses, contracts, tortures, oaths—are exempt of qualifications and variations. This prevents their mechanical application and requires the judge to deliberate about each and to attribute to them an appropriate relative weight, regardless of the conviction displayed by the counsels who base their case on them. We are thus in presence of a process that resembles more self-controversy than self-discussion.

If we turn now to the remaining rhetorical genre, the epideictic—the kind used in ceremonial oratory for praise or censure of some person, action or event—it seems to be quite removed from anything like debate, and even less of ‘dispute’, since no agonistic element is present in it at all. The audience attending this kind of discourse functions as a sympathetic spectator, and as such is not addressed in order to deliberate or decide about the merits of the eulogized person; only the quality of the eulogy and the capacity of the orator in moving the audience is of subject to the audience’s evaluation (*Rhet*. 1358b8). In this respect, epideictic discourse is similar to the work of an artist or artisan, and corresponds somehow to the third kind of intellectual virtue—‘art’—that the *Nicomachean Ethics* contrasts with deliberation (*Nic. Eth.* 1140a). The only resemblance I discern with dispute in either case lies in the deep self-conviction that animates the orator and the artist, which cannot be shaken by reasons of any sort, internal or external.

So, in spite of a promising beginning, with a clear support for ‘self-controversy’ as a distinct type of self-debate and for the role of deliberation therein, and in spite of the triadic structure of Aristotle’s typology in both the internal and the external spheres of debate, his authority only partially supports my proposal. Let us turn, then, to other venues (including Aristotle himself) in order to show the plausibility and usefulness of this proposal.

6 Self-Deception and the Splitting of the Self

Self-deception is a mental phenomenon of a clearly agonistic nature, and thus a *prima facie* example of self-debate. It seems also obvious that, *qua* self-debate, it must belong to the category of ‘self-dispute’: Recall how, in Nicole’s description, ‘self-love’, engaged in a ruthless battle with reason, is not at all embarrassed in employing a rich assortment of devious maneuvers in order to induce one to hold beliefs about oneself which are contrary to those one’s best judgment would recommend. Although easy to classify, if one grants it is a form of self-debate, it is not evident that self-deception actually can be viewed as such. For, under closer

analysis, self-deception consists not only in a situation in which somehow a person believes at one and the same time that a proposition is true and that the same proposition is false; it also is such that—in its most acute cases—one of the contradictory beliefs is a causal factor in bringing about the other and even causally ‘sustains’ it (Davidson 1986: 89). Self-deception thus displays such a flagrant form of inconsistency, such aberrant irrationality, that one must ask whether *one* ‘self’ can bear it, and if so, what kind of self it is. No wonder that recent philosophical accounts of self-deception (e.g., Davidson 1986; Rorty 1986; Pears 1986; Searle 2001: 235–236) have reached the conclusion that, in order for self-deception to be possible, it is necessary to posit some sort of splitting of the self. This section will examine this ontological issue—a detour from the typological enterprise pursued in the previous section, to which we will return in the next section.

Let us assume, for the sake of the argument, that the self is split in at least two subsystems—be they quite permanent (as for Pears and Searle), evolving and variable in their configuration (as for Rorty), or having their borderline traced according to the case at hand (as for Davidson). Once one splits the self in any of these ways, one lands squarely in a situation basically equivalent to that of external debate. If the contradictory beliefs *p* and *not-p* of a self-deceiver are each held by one of *two* contenders, there is no logical inconsistency in any of them per se, for each holds only one of these beliefs. Metonymically, the awareness by one contender that his opponent holds a belief that is inconsistent with his own, poses, of course, an inner problem for him: is he right and the opponent wrong or vice versa? But this problem does not consist in discovering an inconsistency *within himself*, for the situation is not that he, John, believes that *p* and *not-p* nor that John believes that *p* and John believes that *not-p*. Just as in the inter-personal case each of the beliefs in conflict is held by one of the contenders, John or Tom, so too in the intra-personal case each is held by one of John’s subsystems, *J*₁ and *J*₂, none of which can be blamed for contradicting *itself*, nor—a *fortiori*—for self-deceiving itself. Self-contradiction—and *self*-deception—only arises if the two subsystems are co-present and interact at some ‘integrative’ level of the self, where the contradiction is both explicitly represented and must be dealt with. As long as the subsystems are independent ‘mini-selves’ operating strictly apart and not actually integrated in a common self, as seems to be the case in Dissociative Identity Disorder (previously known as Multiple Personality Disorder), no self-contradiction arises. Self-deception can thus exist without challenging the self’s consistency because its logically clashing elements are, so to speak, distributed among different constituents of the self.

But this solution seems to raise more problems than it solves. First, if the self is no longer unitary or integrated, then what is the point of preserving ‘its’ consistency or ‘its’ rationality? By giving up its integrative character, it seems that one has *ipso facto* given up the requirement that the self should possess any of the family of ‘coherence’ properties to which certainly consistency belongs. So, perhaps what is to be preserved is the consistency of some privileged component of the self. Which one? To be sure, the consistency of each of the contending

components of the split self is preserved thanks to the distribution of the logically clashing elements across them. But surely the aim of the exercise was not to preserve these posited mini-selves' consistency. Maybe the aim is then 'consciousness'? But since this presumed higher order component of the self integrates the conflicting inputs of the mini-selves and 'is aware of' them, it explicitly contains the clashing cognitive elements, so that the inconsistency that was to be avoided by the distribution maneuver raises its head once more. What this shows is that the proposed solution fails as an account of self-deception because it is achieved at the expense of depriving the prefix 'self-' in 'self-deception' of significance.

For the same reason, this solution rules out the possibility of considering self-deception as a genuine example of *self*-debate, i.e., as a debate that takes place between one entity and itself, one entity that argues with itself, playing the alternating roles of proponent and opponent. Given what we know about cases of self-deception, were it a genuine self-debate, it would certainly be a particularly good example of the type 'self-dispute'. Yet, on the solution discussed above for the problem of its irrationality, self-deception is in fact an 'inter-subsystems' dispute, completely equivalent to an inter-personal dispute, for which the prefix 'self-' does not add anything. We may indeed imagine self-deception as a particularly ruthless dispute between two virtually autonomous entities (Dr. Jekyll and Mr. Hyde?), each absolutely convinced of being right and consequently as allowed to fight without constraint for its own interests by means of whatever deceptive stratagem available in order to achieve victory, at any price, over its competitor. The 'self' is exempted from inconsistency by playing no role in this dispute other than that of a thin membrane enveloping the contending subsystems and of an arena where the fight takes place.

7 Towards a Typology of Self-Debate: 'Hard' and 'Soft' Rationality

Obviously, the role of the self should be more substantial than that of a tenuous container in order to account for genuine '*self*-debates'. This 'substantiality' has, as we have seen, an onto-pragmatic side; but it has also a psycho-pragmatic side. The later pertains to the conduct of the debate, i.e., to the kind of interaction that takes place between the 'debaters' in a self-debate (be they subsystems of the self or the self as a whole) and its underlying norms: What are—if any—the explicit or implicit rules of the debate? What are the aims of the debate and its criteria of success? Does each of the participants have a distinct role, e.g., of accuser or defender, of mere spectator or third party, of a judge that must decide the issue? What kinds of moves are permitted and what kinds of moves debaters fulfilling specific roles are expected to perform at each stage? What kind of rationality—if any—is presupposed by each of these moves? It is the answers to questions such as

these that should differentiate between distinct ideal types of self-debate. Though each of these questions deserves full and detailed attention, for reasons that will become apparent I will focus here mainly on the last one.

A comparison between weakness of will and self-deception is a good starting point. Recent philosophical analyses of weakness of will and self-deception highlight a profound difference between these two phenomena (e.g., Davidson 1969, 1986: 80–81; Searle 2001: 234–236, 2005; and also Rorty 1986, although to a lesser extent). On the face of it, both involve some sort of irrationality. Yet, the irrationality characteristic of the former is somewhat ‘milder’ than the hard irrationality that characterizes the latter. For, strictly speaking, while in a typical case of weakness of will or incontinence (e.g., having decided to do something and doing something else instead) there is no formal inconsistency, self-deception, on the contrary, involves the simultaneous holding by someone of the belief that p and the belief that *not* p , which is logically unacceptable. This difference in the kind of irrationality involved goes a long way in explaining why weakness of will is ‘less of a problem for philosophical psychology’ than self-deception (Davidson), why the former is ‘less dangerous’ than the latter (Rorty), and why it is a ‘common and natural form of irrationality’, stemming from the very nature of rational action (Searle). The reason for that, in my opinion, is that, being ‘milder’, the irrationality of *akrasia* can be handled in terms of a plausible conception of rationality as reasonableness, whereas the hard irrationality attributed to self-deception implies either the impossibility of this phenomenon or else some implausible splitting of the self into independent or at most very loosely connected subsystems in order to render it possible.

In the preceding Sect. 1 have discussed the ontological difficulties caused by such a splitting and, in the next section, I will try to show how the distinction between hard and soft rationality can help to mitigate—at least onto-pragmatically—these difficulties. Before this, in the present Sect. 1 make use of this distinction as the psycho-pragmatic basis for my proposed tripartite typology of self-debate.

Let me first explain the distinction I have in mind.¹⁵ By ‘hard’ rationality I understand a conception of rationality that has standard logic and its application as its fundamental model. This conception views logical inconsistency as the paradigmatic expression of irrationality and regards certainty as the principal aim and sign of knowledge. Since mathematics is the most successful implementation of this ideal of rationality, hard rationality privileges what it takes to be the basic reasons of this success. Accordingly, it considers, as conditions of rational thinking and praxis or as their preferred manifestations, such parameters as: uncompromising obedience to the principle of contradiction; precise definitions formulated in terms of necessary and sufficient conditions; conclusive argumentation modeled upon deduction; formalization of this procedure by means of a symbolic notation; quantification and computability; axiomatization of domains of knowledge; and the like.

¹⁵ For details, see Dascal (2005).

By ‘soft’ rationality I understand, broadly speaking, a conception of rationality that seeks to account for and develop the means to cope with the host of situations—theoretical as well as practical—where uncertainty and imprecision are the rule. Although acknowledging the applicability and usefulness of the high standards of hard rationality in certain fields, it rejects the identification as ‘irrational’ of all that falls short of them. It deals with the vast area of the ‘reasonable’, which lies between the hard rational and the irrational. The model underlying the idea of soft rationality is that of a balance where reasons in favor and against (a position, a theory, a course of action, etc.) are put in the scales and weighed. But there is a deep difference between ‘weighing’ reasons and ‘computing’ them.¹⁶ For, except in a handful of cases, the weights of reasons are not precisely quantifiable and context-independent. Consequently, weighing them does not yield conclusive results whose negation would imply contradiction. The balance of reason, unlike deduction, “inclines without necessitating”—in Leibniz’s felicitous phrase.¹⁷ Even so, if the weighing is properly performed, the resulting inclination toward one of the plates provides reasonable guidance in decision-making.¹⁸ Soft rationality’s logic is, thus, non-monotonic and cannot be reduced to standard deductive logic. It is the logic of presumptions that rationally justify conclusions without actually proving them, of the heuristics for problem-solving and for hypothesis generation, of pragmatic interpretation, of negotiation, and of countless other procedures we make use of in most spheres of our lives.

It is also the logic of deliberation, a process that is usually depicted in terms of the balance model.¹⁹ It is because its rationality is of the soft variety that deliberation can be associated with ‘self-controversy’, in which no hard decision procedure is available. Let us now try to position ‘self-dispute’ and ‘self-discussion’

¹⁶ “Rationes non esse numerandas sed ponderandas” (reasons are not to be counted but weighed), stresses Leibniz on several occasions (this quote, e.g., is from his letter to Gabriel Wagner, 27.2.1697; in Leibniz 1875, vol. 7: 521).

¹⁷ Again, a phrase often employed by Leibniz. The *locus classicus* is the *Discours de Metaphysique* [1686], par. 13 (in Leibniz 1999: 1546).

¹⁸ It should be noted that, being based upon non-conclusive inferences, soft rationality provides only a weak warrant for its recommendations, if compared with the strong warrant provided by hard rationality. Nevertheless, this does not mean that it is chronically guilty of a cognitive error analogous to weakness of will, which Davidson calls ‘weakness of the warrant’ and defines as follows: “a person has evidence both for and against a hypothesis; the person judges that relative to all the evidence available to him, the hypothesis is more probable than not; yet he does not accept the hypothesis” (Davidson 1986: 81). A person using soft rationality may, of course, incur in this error, but only if he fails to take into account the reasons he considers relevant, to assign them proper weights, to compare them properly, and to admit that the result may be overridden by the eventual presentation of ‘heavier’ reasons.

¹⁹ Although certainly relevant for the soft vs. hard rationality distinction, Hampshire’s (1991: 52) emphasis on the argument-dependence of a deliberative decision (as opposed to the argument-independence of the rightness or wrongness of an arithmetical calculation) focuses on the fact that a moral or political decision is a decision to act *for certain reasons*, and not on the soft character of the inference, which is what I am stressing here.

in the typology of self-debate in terms of how they distance themselves from deliberation. We begin with a detour via weakness of will.

Aristotle, as we have seen, considers deliberation the cornerstone of rational decision-making and rational action. It is through deliberation that one establishes one's preferences, according to which one should act—if one is to act rationally. Failing to do so is to be guilty of *akrasia*. This sin is, thus, a sin against rationality. Remember, though, that deliberation for Aristotle is not demonstration. The latter implies necessity, whereas the former precludes necessity, since where things occur necessarily, there is no choice, and therefore no room for deliberation.²⁰ So, weakness of will can be blamed at most for using soft, not hard irrationality. This makes it easier to account for than the presumed hard irrationality of self-deception, for two reasons. First, since the negation or rejection of conclusions of soft inferences does not imply contradiction, it is no great scandal if such a negation or rejection happens. Rather, it is a matter of routine, which does not challenge (unless it happens massively) its soft rationality's 'rules' of inference. Second, since the main reason for splitting the self was to prevent logical inconsistency, such a need does not arise in the case of soft irrationality, which at most displays less acute forms of cognitive dissonance. These, the self is used to handle without disintegrating, just as it is used to assess opposed reasons in the course of a deliberation without being literally torn apart by them.

John Searle's (2001, 2005) explanation for the routine—hence rather unproblematic from the point of view of philosophical psychology—character of *akrasia* takes a *prima facie* quite different route. It is based on his deeply anti-deterministic concern for the freedom of choice characteristic of voluntary human action. As such, human action cannot be, in his view, *compelled* by anything—not even by the most strictly rational reason for action. This concern leads him to posit a series of 'gaps'—between the process of deliberation and the formation of an intention or decision to act according to its result, and between the intention and its actual execution (Searle 2001: 14–15). These gaps are required for the human agent to be able to exercise her freedom. For, were she not free to decide otherwise than 'recommended' by her best deliberative judgment and to act otherwise than she decided to act, the judgment would be 'causally sufficient' to produce the decision and the decision, to undertake the action, and the agent could just sit back and passively watch the causal chain deliberation-decision-action to unfold without his intervention. That we often experience such a smooth flow should not, however, mislead us into believing that there is no gap, nor that the agent's active 'doing' in every step of the chain is absent. Fortunately *akrasia* is there to remind us of this—it is "but a symptom" of what is going on each time we act voluntarily (Searle 2005: 76). *Akrasia* happens routinely and is easy to account for because all it does is to break the appearance of smoothness, by bringing to light what is always there

²⁰ In the same vein, Leibniz writes: "The nature of the will requires freedom, which consists in that a voluntary action is spontaneous and deliberate, i.e., that it excludes necessity, which suppresses deliberation" (*Causa Dei* [1710], in Leibniz 1875–1890, vol. 6: 441).

anyhow, albeit hidden: the gaps and free agency, which is supposed to be possible only thanks to them.

Searle's mode of argumentation erects mighty barriers between a "Classical Model" (as he calls it) that he considers to be entirely mistaken and his own theory, wherein gaps are excavated and establish no less mighty abysses between mental processes. In his search for a formal model of practical reason, he considers exclusively *deductive* models, and—on this basis alone—concludes: "it is impossible to get a formal logic of practical reasons which is adequate to the facts of the philosophical psychology" (Searle 2005: 56). All this suggests a puzzling reliance on hard rationality as the only available model and tool for the account of rational action, even when this model seems to be patently inadequate for Searle's own purposes. It is to some aspects of his position that might benefit significantly if handled in terms of a soft rationality approach that we now turn.

Let us ponder first about Searle's treatment of *akrasia*. He gives the following formula summarizing its basic form:

1. It is best to do A and I have decided to do A, but I am voluntarily and intentionally doing B (Searle 2001: 235).

It should not come as a surprise that this is directly connected to the gap theory. It provides direct support to it: That there is *akrasia* is a decisive counter-example against those who do not accept the thesis that the same freedom not to do what one has decided occurs in *all* cases of voluntary action. Accordingly, there should be a parallel formula for non-akratic rational action:

2. It is best to do A and I have decided to do A and I am voluntarily and intentionally doing A.

Notice, however, the telltale difference between the two formulae: *but* in the former vs. *and* in the latter. The interpretation of the adversative conjunction *but* in general requires a delicate pragmatic operation, for it rarely—if ever—refers to a semantic or logical opposition between the two conjuncts (see Dascal 2003: chapter [Implicature, Inference and Cancellability](#)). To make a long story short: Cases of *akrasia* are no doubt knock-out counter-examples to the thesis that deliberation and decision are causally sufficient for the corresponding action and thus necessarily bring it about; but per se they do not eliminate the possibility that, having made his mind after deliberation, the agent is more likely than not to act upon his decision; in other words, (2) rather than (1) is the default rule; i.e., one is expected to act according to (2), though not forced to do so, as shown by (1); this explains the *but* in (1)—its function being to indicate that the performance of B is a strong enough reason to cancel, in a particular case, the expectation that A would be performed generated by the default rule (2).

The notion of default is a soft notion, and what has been sketched above is a plausible soft rationality account of the feeling of irrationality *akratic* action arouses. Similarly, the feelings of having an infinite range of possibilities each

time we make a choice and the freedom of choosing any of them, which Searle describes as “the way in which *akrasia* characteristically arises”,²¹ can be tracked to the fact that deliberation is not conclusive, among other things because of its context-dependence. Each possible choice is associated with its own set of relevant reasons and their weights, so that the simple consideration of its possibility is likely to modify the balance’s inclination, by bringing into focus ‘other grounds’ for the decision.²² In this sense, *akrasia*—and possibly the very idea of the gap—stems from the fact that deliberation is an open-ended process that always goes on, an idea corroborated by the phenomenon of post-decision dissonance discussed in Sect. 2.

Going metaphysical, one might ask whether these and many other ways in which soft-rational factors *incline* us toward an action without *compelling* us to do it should lead to viewing them as an entirely different type of causation, ‘intentional causation’, thus transforming what is observed as “a certain amount of slack between deliberation and formation of intention ... and between the intention and the actual undertaking” (Searle 2001: 231) into a metaphysical gap rather than the normal way in which soft reasons exercise their causal powers. Having so many motives for going soft-rational,²³ perhaps Searle does not do so simply because he himself obliterates the distinction soft vs. hard rationality by arguing that even the latter has no compelling power: An inference may be valid, “but there is nothing that forces any actual human being to make that inference” (p. 21). Like Descartes’s God whose freedom was above the rules of logic, the freedom of Searle’s rational agent seems to be absolutely unbounded, i.e., neither compelled nor inclined by anything.

Returning to earthly matters, consider again the *akratic* in the fraction of a second between he concludes his deliberation with a decision and acts contrary to this decision. At precisely this moment, he is only a *potential akratic*. Some reason against his decision, which he had not put in the balance, pops up in his mind, perhaps due to an objection coming from an external critic. This is when the *potential akratic*’s self-debate is most acute. *To do or not to do?* He had decided to stop smoking; but he didn’t take into account the weight of the painful urge to puff his pipe precisely in moments like these, when he needs all his mental powers focused on the essay he must by all means finish *today*. This is a strong reason not

²¹ “[A]s a result of deliberation we form an intention. But when the moment comes there is an indefinite range of choices open to us and several of those choices are attractive or motivated on other grounds. For many of the actions we do for a reason, there are conflicting reasons for doing not that action but something else. Sometimes we act on those reasons and not on our original intention. The solution to the problem of *akrasia* is as simple as that” (Searle 2005: 76–77).

²² This might explain Hampshire’s notion of argument-dependence of a deliberative decision (see note 20), while at the same time not restricting it to the case in which the decision is that of *acting for certain reasons*.

²³ From the choice for the jacket of *Rationality in Action* of Vermeer’s masterpiece *Woman Holding a Balance*, which depicts a situation of extremely delicate weighing of jewels, to the explicit use of the balance model—the emblem of soft rationality—there would be but a small step, were it not for the gap.

so much for reconsidering his sensible decision to stop smoking, but for admitting an exception, only this one, please. He has been told to be especially careful against such “only this one” pleadings and to resist them by sticking to the decision no matter how persuasive they are. Obviously, this person has already engaged, in this very fraction of a second, in a new round of deliberation, which may yield a different decision and action. If it does, then he will have committed *akrasia*.

How extensive and careful the new round of deliberation will be, and how this may influence its results, depends on such factors as how solid is my previous conviction, how serious I take the objection to my previous decision to be, how important the issue is for me, how urgent it is to act on it, etc. It also depends upon character and style. In a letter to Leibniz, Thomas Burnett of Kemeny describes an extreme deliberative attitude: “Mr. Locke is known here as a man who meditates and thinks more than anyone at what he writes, especially in philosophy. He considers and reconsiders, he examines an idea front, back and across (so to speak), and he sees a thing through all possible points of view, so that he will always have an advantage over a man that does not think so much as he does”.²⁴ The more you examine an issue from all sides, the more likely you are to question entrenched assumptions, methods, and solutions—yours or others’—i.e., the more controversy-like the self-debate is. This may of course lead to indefinite suspension of judgment, and will be justified only regarding important issues that are not pressing.²⁵ More often, however, the result of the re-deliberation can be the *re-confirmation* of one’s previous position, the *conversion* to the opposite position, or a new position emerging from some *reconciliation* between the contending positions that accommodates the new configuration of weights in the balance. In each case, the balance’s inclination may become stronger or weaker in the wake of the re-deliberation process.

But there are other ways of reacting to the (inner or outer) objection and to the ‘new reason’ it brings to the fore. If I am very strongly persuaded of my decision, feeling certain about it, an objection—however, well grounded—is unlikely to make me doubt that decision. I may then entirely overlook it, and avoid engaging in a costly re-deliberation process. In this case, no self-controversy will arise, and there is no danger of *akrasia*. Now, my virtually absolute confidence in my decision and its consequent virtual immunization vis-à-vis objections may have its origin in one of the following scenarios:

1. There was no deliberation and I have a firm intuition, a gut feeling that this is what I have to do.
2. There was deliberation and—as a matter of principle or of habit—I do not change my mind after I make a decision.

²⁴ Letter from Burnett to Leibniz, 3 May 1697 (in Leibniz 1875, vol. 3: 198).

²⁵ “This is how the Areopagites would in effect absolve this man whose case was too difficult to be decided; they would postpone it for a long time and grant themselves one hundred years to reflect about it” (Leibniz, *Nouveaux essais sur l’entendement humain* [1705], 2.21.23; in Leibniz 1875, vol. 5: 167).

3. There was deliberation and I am sure the decision is correct because it was reached by the application of a faultless method.

In scenario (1), there was no deliberation before the decision because I never had any doubt about it being the right decision. My confidence is based on immediate ‘evidence’, which stems from my privileged self-knowledge. Just as it does not need justification (hence, is not a matter of deliberation) it does not admit objections (hence, is unquestionable). Since others do not have the privileged access I have to the source of my evidence, they are a priori disqualified to question my certainty. If their objections to my choice surreptitiously ‘infect’ my mind, I must beware of this treacherous ‘voice’ disguised as my own and not listen to it at all. If I appear to argue against it inside myself, it is not because I admit there is room for a real debate about my decision; I do so only for eristic purposes, for reassuring myself through ‘winning’ this pseudo-debate of what I anyhow know is right. I am just fighting back the siren’s voice, the stratagem designed to divert me from what I am sure I have to do, with a counter-stratagem. I may also avert the danger by Peirce’s already mentioned ‘method of tenacity’, namely, refusing to expose myself to those who might challenge my self-assurance and thus “systematically keeping out of view all that might cause a change in [my] opinions”.²⁶ In both cases, my maneuvers are examples of self-deception. Aristotle calls the kind of person described in this scenario ‘impulsive’ or ‘excitable’ and classifies her as someone who has no self-control, i.e., as a potential *akratic* (*Nic. Eth.* 1141a20). But her decisions are far from volatile, for they will change only as the result of a clash with a more powerful contrary gut feeling. In the absence of such a clash, she will cling tenaciously to her decision no matter what.

In scenario (2), the person resembles rather the continent man who is in full self-control. He has deliberated and reached a decision to which he is committed and which he abides by. The difference between them lies in the fact that a person of the first kind is “hard to convince, and not easily persuaded to change their convictions”, whereas the continent man is “ready on occasion to yield to persuasion”. Whereas the latter stands his ground against passion and yields to reason, the former holds her ground against reason. Aristotle calls such persons ‘obstinate’, and lists among the motives of their stubbornness “the agreeable sense of victory in not being persuaded to change their mind and the annoyance of having the decrees of their sovereign will and desire annulled”.²⁷ For these people, the decision point marks the end of the deliberative process. Reopening this process and accepting any modification of its outcome would mark for them self-defeat, in the sense that the self would thereby admit serious imperfections in both its reason and will—in the performance of rational choice and in the capacity to stand for it.

²⁶ “The fixation of belief”. In Peirce (1931–1958, vol. 5: 235; first published in *Popular Science Monthly* 12: 1–15, 1877). For a critical discussion of this article by Peirce, see Dascal and Dascal (2004).

²⁷ The quotes in this paragraph are from *Nic. Eth.* 1151b4–15.

In scenario (3), as opposed to (2), the basis of the person's conviction is not one's personal pride, eager to defend *post factum* the decision made even if one realizes that it clashes with reason. It is, rather, a deep belief that the deliberation method was appropriate to the issue at hand, correctly applied, and rigorously conclusive, so that the decision reached admits of no revision whatsoever on rational grounds. The person can eventually admit that, although having carefully checked the results, both by herself and with the help of competent colleagues, some minimal possibility of mistake remains. This may be a human mistake or a 'bug' in the procedure. A doubt about the results, provoked by internal or external questioning, is a good occasion for verifying the correctness of the procedure, her way of applying it, or similar possible problems. This gives rise to what might look like re-deliberation. But the only way this kind of person admits of persuading her to abandon her conviction is to show that some *mistake* in the strict sense was made—i.e., a mistake that could be corrected by re-running the procedure (hopefully through independent researchers). Once the truth is determined through such a re-calculation, no reason for doubt can remain, and she, on her own accord, is compelled to accept the result (whatever it is) and have in it the same degree of confidence she had in her previous—whether to be preserved or abandoned—result. Unlike in scenario (2), to abandon in this way her previous conviction is not perceived in scenario (3) as a defeat, but rather as a victory of the method one has adopted and as a further proof one is in the right track.

Each of the sources of conviction in (1), (2), and (3) differ in their positioning vis-à-vis deliberation; but they all share their repugnance for it. Each has its own reasons for dispensing with deliberation in self-debate, but all are persuaded that they can handle better than deliberation. Deliberation is repugnant to them because of the time and energy consuming its appropriate execution demands. As against this, conviction of the three sorts is supposed to provide a finite, mental energy and time saving way of deciding any self-debate. Furthermore, deliberation is repugnant to conviction-based strategies because of its open-ended nature. One can begin deliberating about a well-delimited issue and may end up discussing the whole universe, due to the eagerness not to overlook in the weighing process any of the relevant reasons. As against this, convictions keep the discussion within watertight compartments, for which they are most effective, and thus avoid the dangerous spreading of doubt to broad philosophical and methodological issues. And, most of all, the great enemy of conviction is—rightly—perceived to be the soft rationality characteristic of deliberation. Conviction believes to be able to overcome this enemy either by adherence to a hard rationality method (scenario 3), or by the adamant refusal to call into question a decision made, even though it was made by a soft-rational deliberation whose results can in principle be questioned (scenario 2), or else by giving up altogether the assumption that reliable decisions and beliefs are those based on rational (soft or hard) decision making processes and replacing such presumed grounds by unexplainable intuitions, feelings and impulses (scenario 1).

Table 1 Types of self-debate

Type	Subtype	Phase	Begins with	Procedure
Self-controversy			Doubt	Deliberation
Self-discussion			Doubt	Demonstration
Self-debate	Stubborn	Pre-decision	Doubt	Deliberation
		Post-decision	Conviction	Tenacity
	Impulsive	Pre-decision	Intuition	Tenacity
		Post-decision	Conviction	

If we put together what we have been belaboring in this and previous sections, we come up with a typology of self-debate that is roughly parallel to that of interpersonal debate, which I present below in a self-explanatory schema (see Table 1).

8 A Soft-Partitioned Self?

To conclude, consider again self-deception, a typical tactic in self-dispute, which has provided some of the strongest philosophical arguments in favor of the self's partition. Maybe Davidson is right in claiming that self-deception cannot achieve its objective of inducing in oneself a belief contrary to one's best judgment through the kind of causality by which *reasons* operate, for "nothing can be viewed as a good reason for failing to reason according to one's best standard of rationality" (Davidson 1986: 92).²⁸ If so, "one's best standard of rationality" functions as a strong demarcating criterion, which segregates 'reasons' from whatever else that is capable to influence one's beliefs. The emphasis on the *inconsistency* that, according to Davidson, is what makes self-deception irrational suggests that the criterion in question is a hard criterion. That is to say, it excludes soft, non-monotonic influences, especially those of a non-cognitive character, such as one's habits, tastes, and inclinations. Admittedly, on Davidson's category distinction between reasons and causes, such influences are causes, not reasons. But all they need in order to become members of the privileged club of reasons is to cross the thin threshold of awareness and become "represented as internal reasons" as Searle (2001: 115) puts it, i.e., accessible to consciousness. We know, however, that mental representations can be vague, obscure, and comport a lot of implicit elements, and that they move easily in and out of consciousness. This threshold seems, therefore, to be too thin to establish a hard demarcation, not to mention a gap.

There are other reasons for softening the self's presumed partition. Indeed, the fact that Davidson attributes a provisional and ad hoc character to the division of the self in his account of self-deception suggests that the demarcation is, for him,

²⁸ On the distinction between 'reasons' and 'causes', see Davidson (1963).

quite soft in nature. Consider, for instance, the role the ‘principle of total evidence’, a central piece of Davidson’s ‘best standard of rationality’, plays in the demarcation. Suppose this principle is softened, for who in fact acts on *total available evidence* and how should availability be defined—as Davidson himself (1986: 81n) asks? As a matter of fact, Davidson argues that, in order for the irrationality of self-deception to be possible, there must be a *causal* (i.e., not involving reasons) step in its generation, and proposes that this step consists in the “temporary exile or isolation” (from the domain of reasons) of the requirement of total evidence. But isn’t a principle that can be temporarily suspended best viewed as a *presumption*, whose occasional violation is to be expected and does not amount to an *inconsistency*? Obviously, such routine violations do not create any permanent boundaries or gaps. At most they reveal the need of pragmatically motivated, variable soft lines one draws within the self in order to make possible the understanding of phenomena such as self-deception that otherwise not even the most charitable version of the principle of charity could rescue from hard irrationality.

Davidson is not alone in hinting at the possibility of mitigating the self’s partition. Others are more explicit about this possibility and less concerned about the partition’s eventual failure in performing its alleged explanatory job if the softening goes too far. A case in point is the work of George Ainslie (1982, 1986), which suggests a way to soften what is perhaps the hardest ever proposed partition of the normal self—the Freudian triad id, ego, and superego. Ainslie distinguishes between two different functions of what is usually lumped together as ‘mechanisms of ego defense’: “(1) the avoidance of unpleasant perceptions per se” and “(2) the avoidance of impulses, that is, reward-seeking behaviors which could lead to unpleasant perceptions” (Ainslie 1982: 737). The first task has to do with current perceptions whose avoidance gives immediate reward; the second is concerned with long run interests of the ego and requires some mechanism capable of giving up an immediate reward that might lead to future unpleasantness for the sake of a delayed reward. Mechanisms of this kind include ‘precommitting devices’, which are “farsighted but must stay in control of behavior over the long run if they are to obtain the rewards that are their *raison d’être*” (p. 746). Ainslie analyzes a wide variety of such devices, but the one which is most interesting for our purposes here is the ‘bunching of choices’. This consists in grouping choices according to their diagnostic value for the chooser, i.e., according to whether each of the choices is perceived as a “precedent for future choices in the same set” or bunch (p. 772). In our terminology, this perception endows a choice under consideration with extra weight in the deliberation process. Suppose I have decided, after careful deliberation, to stop smoking. But I feel an intense urge to smoke right now. If I give in and violate my self-imposed rule, this will signal to myself that I may do the same in the future. This additional weight must be put in the balance when deciding whether to satisfy my urge, because ignoring it may lead, against my best judgment, to the crumbling of the rule I adopted following my best judgment. The ‘bunching of choices’ thus functions as a reasonable self-control mechanism that may help to prevent incontinence. However, it may be hardened

into compulsive behavior that forbids any exception to the particular rule in question, as well as to any other rule—which is precisely the superego’s strategy (Elster 1986a, b: 22). Just as it has to defend itself from the immediate pleasure-seeking id, the ego has thus to defend itself from the imposing of hard rules by the superego. The way to do this, in Ainslie’s theory, is a sort of inner bargaining that loosens the bunching of choices in rigid sets, while at the same time not allowing the person “to shift boundaries repeatedly in order to make exceptions” at will, by setting up ‘bright lines’ that tell him when the circumstances are such that an exception is justified (Ainslie 1982: 772). Whether or not this account of the ego-superego conflict obviates the need to posit them as separate autonomous entities, it illustrates well how this conflict involves several of the types of self-debate discussed in this essay, as well as how soft rationality and the context-sensitive onto-pragmatic perspective are directly relevant to this issue.

In any case, if we grant some of the above ways of softening the self’s partition, then the posited parts of the self are, in fact, not so separated as the personalities “Dr. Jekyll” and “Mr. Hyde”. Furthermore, the self would thus acquire some substance of its own. First, because each of its presumed parts would no longer be constituted by a fixed set of elements, since what is at one point in one of them can routinely become a member of another; these elements, thus, would become multi-functional entities that ‘belong’ *de facto* to the self, not to any one of its supposed parts; they are elements that the self can *use* for different purposes, just like words can be used for thinking as well as for communicating; the partition of the self would turn out to be, therefore, not ontological but onto-pragmatic. Second, because even those who insist on a strict boundary separating the domain of rationality from the causal domain, like Davidson and Searle, ultimately allow these parts of the self to interact *causally*, “since reason has no jurisdiction across the boundary” (Davidson 1986: 92). Together, the mobility of the elements and the causal interaction between the self’s posited parts requires enough shared ground for seeing them as belonging to one and the same substantial ‘personality’.²⁹

Does the softening of the partition lines in the ‘multiple self’ make it, ultimately, a ‘mere metaphor’, a mere ‘way of speaking’, a mere ‘pragmatic device’ because it does not correspond to something ‘stable’, ‘embodied’, ‘real’ in the person? To answer this question would require us to embark in a new inquiry about what might qualify the self’s subsystems (not to mention the self itself) as ‘real’:

²⁹ Furthermore, the across boundary possibility of causal interaction may be all that is required for the intervention, in a self-dispute, of an ‘external’ factor not directly involved in it, which can however, be causally effective in ‘dissolving’ it or reducing its level of conflict. Consider some sort of self-regulating mechanism analogous to an analyst who interprets the fighting parties’ aims and moves in the hope of dissolving or reducing the unrest or the dysfunctions caused by the conflict, or the inner analogue of Kant’s ‘umpire’ or ‘legislator’ (similar to Smith’s ‘impartial observer in the breast’), who does not take sides in the metaphysical disputes between ‘dogmatic sects’ but only draws from their combats lessons about the limits of pure reason (Kant 1971: Aviii, A423, B452; Dascal 2000, for further details). These speculations suggests that the quarreling parts of the ‘self’ share, after all, a common interest in preserving a functioning condition, a well being that regulates the wildness of self-disputes.

For example, would their functional identity suffice or should one also find in the brain separate modules in charge of performing these functions? Provisionally, we can stop here and suggest that the answer to this question depends on whether one considers the literal vs. metaphorical boundary a *hard* or a *soft* boundary.

References

- Ainslie, George. 1982. A behavioral economic approach to the defense mechanisms: Freud's energy theory revisited. *Social Science Information* 21(6): 735–779.
- Ainslie, George. 1986. Beyond microeconomics. Conflict among interests in a multiple self as a determinant of value. In ed. Elster, 133–175.
- Aitchinson, Jean. 1994. *Words in the mind: An introduction to the mental lexicon*, 2nd ed. Oxford: Blackwell.
- Aarsleff, Hans. 1967. *The study of language in England, 1780–1860*. Princeton, NJ: Princeton University Press.
- Aristotle. 1996. *Nicomachean Ethics*, (trans: Rackham, H., introd. S. Watt. Ware). Hertfordshire: Wordsworth.
- Aristotle. 1990. *Retórica*, (trans: introd. and notes by Q. Racionero). Madrid: Gredos.
- Bakhtin, Mikhail.M. 1984. *Problems of Dostoevsky's Poetics*. ed. Emerson, C. 1984. (trans: Emerson, C.). Minneapolis, MN: University of Minnesota Press.
- Brehm, Jack W. 1956. Post-decision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology* 52: 384–389.
- Dascal, Marcelo. 1978. *La sémiologie de Leibniz*. Paris: Aubier-Montaigne.
- Dascal, Marcelo. 1983. *Pragmatics and the philosophy of mind*, vol. 1. Amsterdam: John Benjamins.
- Dascal, Marcelo. 1992. Why does language matter to artificial intelligence? *Minds and Machines* 2(2): 145–174.
- Dascal, Marcelo. 1995. The dispute on the primacy of thinking or speaking. In *Philosophy of language—An international handbook of contemporary research*, vol. 2, eds. Dascal, M., D. Gerhardus, K. Lorenz, and G. Meggle, 1024–1041. Berlin/New York: De Gruyter.
- Dascal, Marcelo. 1998. Types of polemics and types of polemical moves. In *Dialogue analysis VI* (= Proceedings of the 6th IADA conference, Prague 1996), vol. 1, eds. Čmejrková, S., J. Hoffmannová, O. Müllerová, and J. Světlá, 15–33. Tübingen: Max Niemeyer.
- Dascal, Marcelo. 2000. Controversies and epistemology. In *Philosophy of science* (=Vol. 10 of *proceedings of the twentieth world congress of philosophy*), ed. Tian Yu Cao. 159–192. Philadelphia: Philosophers Index Inc.
- Dascal, Marcelo. 2002. Language as a cognitive technology. *International Journal of Cognition and Technology* 1: 35–61.
- Dascal, Marcelo. 2003. *Interpretation and understanding*. Amsterdam: John Benjamins.
- Dascal, Marcelo. 2004. Argument, war and the role of the media in conflict management. In *Jews, muslims and mass media*, ed. Parfitt, T. with Y. Egorova. 228–248. London: Routledge.
- Dascal, Marcelo. 2005. The balance of reason. In *Logic, thought and action*, ed. D. Vanderveken, 27–47. Dordrecht: Springer.
- Dascal, Marcelo, and Dascal, Varda. 2004. A des-fixação da creença. In *O Processo da Crença*, eds. Gil, F., P. Livet, and J.P. Cabral. 321–353. Lisboa: Gradiva.
- Davidson, Donald. 1963. Actions, reasons, and causes. In Davidson, D. 2001. 3–19.
- Davidson, Donald. 1969. How weakness of the will is possible?. In Davidson, D. 2001. 21–42.
- Davidson, Donald. 1986. Deception and division. In ed. Elster, J. 79–92.
- Davidson, Donald. 2001. *Essays on actions and events*, 2nd ed. New York: Oxford University Press.

- Elster, Jon. 1986a. Introduction. In ed. Elster, J. 1–34.
- Elster, Jon (ed.). 1986b. *The multiple self*. Cambridge: Cambridge University Press.
- Festinger, Leon. 1957. *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Festinger, Leon (ed.). 1964. *Conflict, decision, and dissonance*. London: Tavistock.
- Hampshire, Stuart. 1991. *Innocence and experience*. Cambridge, MA: Harvard University Press.
- Jakobson, Roman. 1962. Deux aspects du langage et deux types d'aphasie. *Les Temps Modernes* 188.
- Jakobson, Roman. 1964. Deux aspects du langage et deux types d'aphasie. In *Essais de Linguistique Générale*, ed. R. Jakobson, 43–67. Paris: Minuit.
- Kant, Immanuel. 1971. *Kritik der reinen Vernunft*. Leipzig: Reclam (A = 1st ed., 1781; B = 2nd ed., 1787).
- Lakoff, George, and Mark Johnson. 1980. *Metaphors we live by*. Chicago: The University of Chicago Press.
- Leibniz, Gottfried W. 1875–1890. *Die Philosophischen Schriften*, ed. Gerhardt, C.J. Berlin [repr. Hildesheim, Georg Olms, 1965].
- Leibniz, Gottfried W. 1903. *Opusculs et fragments inédits de Leibniz*, ed. Couturat, L. Paris: Presses Universitaires de France
- Leibniz, Gottfried W. 1999. *Sämtliche Schriften und Briefe*, series 6, vol. 4, ed. Leibniz-Forschungstelle der Universität Münster. Berlin: Akademie Verlag.
- Locke, John. 1690. *An Essay Concerning Human Understanding*. ed. Yolton, J.W. London and New York: Everyman's Library (1967).
- Nicole, Pierre. 1999. De la connaissance de soi-même. In *Essais de morale*, ed. Thirouin, L. 309–379. Paris: Presses Universitaires de France (1st ed. 1675).
- Pascal, Blaise. 1960. *Pensées*. Text of the Brunschvicg edition, with introduction and notes by Ch.-M. des Granges. Paris: Garnier Frères [1st, posthumous ed. 1669].
- Pascal, Blaise. 1985. *De l'esprit géométrique, Ecrits sur la Grâce et autres textes*, ed. Clair, A. Paris: Flammarion.
- Pears, David. 1986. The goals and strategies of self-deception. In ed. Elster, J. 59–77.
- Peirce, Charles S. 1931–1958. *Collected papers of Charles Sanders Peirce*, ed. Hartshorne, C. Cambridge, MA: Harvard University Press.
- Piro, Francesco. 1999. *Il Retore Interno: Immaginazione e passione all'alba dell'età moderna*. Napoli: La Città del Sole.
- Proust, Marcel. 1954. *A l'ombre des jeunes filles en fleurs*, 1^{ère} partie. In M. Proust, *A la recherche du temps perdu*, vol. 1, ed. Clarac, P. and A. Ferré. Paris: Gallimard.
- Rorty, Amelie O. 1986. Self-deception, *akrasia* and irrationality. In ed. Elster 115–131.
- Searle, John R. 1983. *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Searle, John R. 2001. *Rationality in action*. Cambridge, MA: The MIT Press.
- Searle, John R. 2005. Desire, deliberation and action. In *Logic, thought and action*, ed. D. Vanderveken, 49–78. Dordrecht: Springer.
- Senault, Jean F. 1664. *De l'usage des passions*. Paris: Le Petit [1st ed. 1641].
- Smith, Adam. 1976. *The Theory of Moral Sentiments*, ed. Raphael, D.D. and A.L. Macfie. Oxford: Clarendon Press [6th ed. 1790].
- Smith, Barbara H. 1997. *Belief and resistance: Dynamics of contemporary intellectual controversy*. Cambridge, MA: Harvard University Press.
- Thibodeau, Ruth, and Elliot Aronson. 1992. Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin* 18: 591–602.
- Vygotsky, Lev S. 1962. *Thought and language* (trans. Hanfmann, E. and G. Vakar). Cambridge, MA: The MIT Press.
- Watson, John B. 1913. Psychology as the behaviorist views it. *The Psychological Review* 20: 158–177.

Truth, Negation and Meaning

Franco Lo Piparo

Abstract ‘True’ and ‘False’ are defined through a linguistic rule requiring the negation operator. This is the elaboration of an idea proffered for the first time by the Stoics on the basis of some remarks by Aristotle and then in modern times by Frege and Wittgenstein. Another thesis of this essay is the following: the true/false rule is a sort of UR-Regel underlying all linguistic practices (including prayers and commands) and all human cultures. Reinterpreting the notion of Spielraum put forward by Wittgenstein in 4.463 of the Tractatus, I will present an implicational pragmatic theory of a true proposition. Jokes and reductio ad absurdum are explained as examples of Spielraum.

Keywords True · False · Spielraum · Aristotle · Stoics · Frege · Wittgenstein

1

Society and human cultures that are not governed by the truth rule do not exist. I will clarify this with an example. No court can condemn a witch to be burnt at the stake if she does not prove, after checks, to be a *true* witch. If a woman that proved not to be a witch was burnt as a witch, representatives of the culture and the society to which the court belongs would maintain that the judges of that court have condemned a person to be burnt at the stake on the basis of a false judgment: they would say that it *is not true that the woman X is a witch*. Naturally, the judges of the court would reply that that judgment of falsehood is false and that it is *true that the woman X is a witch*. The present writer and the readers know that witches do not exist and therefore they are convinced that a condemnation to be burnt at the stake for witchcraft rests in any case

F. Lo Piparo (✉)
University of Palermo, Palermo, Italy
e-mail: lopiparo@unipa.it

on *false judgments*, and therefore they maintain that, *independently of the facts ascertained, it is not true that the woman X is a witch*.

The situation described prompts two considerations.

1.1

Both the person condemning a person to be burnt at the stake for witchcraft, and those who, though believing in the existence of witches, believe that that particular woman *X* condemned to be burnt was not a witch, as well as anyone who is convinced that no woman is a witch, to express and uphold their judgments and evaluations, have to resort to the rule of truth. The different evaluations presuppose sharing the rule of truth. «Men—Aristotle observed—have a sufficient natural disposition to truth and in most cases attain truth: for this reason it is part of the same disposition to tend through conjectures (*stokhastikōs*) towards shared opinions (*éndoxa*) or towards the truth» (*Rhet.* 1355a 15–18).

The truth rule, as a natural disposition, has the theoretical statute of instinct: it is not acquired through training, but every human animal finds it in his own natural cognitive baggage. From the observation that the truth rule unifies cultures and societies that are very distant from one another we draw the following axiom:

(A1) If a culture or a society is human, then it is governed by the truth rule.

1.2

The three different evaluations are not only governed by truth but by the rule of justice: *it is right/it is not right to condemn the woman X to be burnt*. The three truth judgments can be reformulated in terms of justice:

- (a) *It is true* that *X* is a witch, and therefore it is *right* to condemn *X* to be burnt.
- (b) *It is not true* that *X* is a witch, and therefore it is *not right* to condemn *X* to be burnt.
- (c) *It is not true* that witches exist, therefore it is *not right* to condemn *X* to be burnt.

To them we need to add a fourth judgment with a more complex structure:

- (d) *Even if it were true* that *X* is a witch, it *would not be right* to condemn *X* to be burnt.

The logical-argumentative structure of the four judgments allows us to formulate a second axiom:

(A2) Only animals governed by the truth rule are able to consider the problem of the rules with which to distinguish what is right from what is not right. Or: no rules of justice without a truth rule.

Axioms (A1) and (A2) could be rephrased through an even more general axiom able to encompass them both:

(A3) If a culture and a society are human, then truth is the Ur-Regel from which the other rules originate and without which the other rules could not work.

2

Let us examine a case that is apparently different from that of the court that affirms that *X is a witch*. In the middle of a conversation my interlocutor exclaims: *I have toothache*. I suspect that it is a pretext to interrupt the discussion. The truth or falsehood of the affirmation will be the object of my evaluations and it will necessarily guide my action.

In comparison to *Mary is a witch* or *The table is red* or *The table is a metre high*, the truth/falsehood of *I have toothache* raises a new problem. The first three utterances are affirmations on facts that, being external to the person making the affirmation, are also public, that is to say accessible in principle to everybody. *I have toothache*, instead, is an account of a private experience to which only the person making the assertion in principle has direct access. Is the rule of truth always the same? Not only is the answer positive, but it is also possible to maintain that the semantics of utterances like *I have toothache* is a good model explaining the way in which utterances like *Mary is a witch* or *The table is one metre high* are related to the world.

We begin by recording the fact that a person who able is to transform a pain, personally experienced, into the utterance *I have toothache* knows, at the moment when s/he projects the pain into the linguistic representation of the pain, if s/he has toothache or not. At a later moment s/he could for instance say: «I lied. It was not true that I had toothache when I said *I have toothache*». A person who says *I have toothache*, at the moment of saying it submits her/his own experience to the truth rule, which is not private. In the terminology most frequently used by Wittgenstein, the utterance *I have toothache* is not a private image (*Vorstellung*) but a public image (*Darstellung*).

From the previous axioms the following corollary derives:

(C1) If a culture and a society are human, then the experienced and private world of its members is also governed by the public truth rule.

3

What does the truth rule do? All humans know it. Indeed, one of the conditions for an individual to belong to the human race is knowing what the truth rule does. Here we are naturally talking about operational knowledge, not explicit theoretical knowledge. The clear and simple definition given by Aristotle states the minimum operations of the rule: «Falsehood consists of *saying* (*légein*) that what is, is not, or that what is not, is; truth consists instead of *saying* (*légein*) that what is, is, or that what is not, is not» (*Met.* 1011b 26–27). In more recent times, Tarski has re-phrased in almost the same terms the operational field of the rule: «a true sentence is one which says that the state of affairs is so and so, and the state of affairs indeed is so and so» (1956, p. 155).

The simple and necessary (but not sufficient) operations stated by Aristotle and Tarski lead us to observe that the attribution of truth/falsehood has to do with language: «Truth and falsehood consist of *saying that...*» (Aristotle); «a true sentence is one which *says that...*» (Tarski). What in Aristotle was operational description was transformed by the Stoics into an abstract theoretical utterance: «for something to be true or false it is first of all necessary for it to be *sayable* (*lektón*)» and therefore «truth and falsehood lie in the *sayable* (*lektón*)» (Arnim 1903–1905, vol. II, p. 187). Wittgenstein's reflections are set in the same tradition of thought: «It is what human beings *say* that is true and false; and they agree in the *language* they use. That is not agreement in opinions but in form of life (*Lebensform*)» (*PI*, I, § 241).

The arguments with which Frege maintained the linguistic nature of the attribution of truth/falsehood to images is useful to us for the continuation of our discourse. An image in itself—the German logician maintained—is neither true nor false. For the question to be asked on its truth or falsehood, it is essential to recognize in the image the intention to represent something. But this is only possible if the image is correlated to a thought expressible through a proposition of the type: «This image represents (is an image of) this and this other». It is of this thought-proposition that precisely truth or falsehood is predicated.

When it is stated of an image (*Bild*) that it is true, the intention is not really to ascribe to it a property that would belong to it in complete isolation from other things, and one has something totally different in mind: the intention is to say that that image somehow corresponds to this thing. 'My representation (*Vorstellung*) corresponds to Cologne Cathedral' is a proposition (*Satz*): now we are concerned with the truth of this proposition. What in a wholly improper way is called the truth of images and representations is thus linked to the truth of propositions (Frege 1966, pp. 32–33).

That the attribution of truth or falsehood to an image is correlated with a linguistic operation was noticed by Aristotle in the passage in *Poetica* in which he observes that «men, looking at images (*eikóna*), feel pleasure because it happens to them that, examining them, they learn and argue (*sullogízēsrai*) on what everything is, for instance that this is» (*Poet.*, 1448b 15–18).

Axioms (A1)–(A3) could be rephrased as follows:

(A4) Truth and falsehood have to do with speaking and therefore only linguistic animals can practise the truth rule.

«Why—Wittgenstein asks—can't a dog simulate pain? Is he too honest? Could one teach a dog to simulate pain?» Naturally if the dog does not lie it is not out of an excess of honesty but because lying is a language game and therefore only linguistic animals can learn to play it. And why, Wittgenstein also asks, are we not prepared to believe that «the smile of an unweaned infant is a pretence?» (*PI*, I, § 249–250). The answer is identical: the unweaned infant does not simulate because it does not yet speak. The learning of mendacious behaviour will coincide with the learning of language.

Axioms (A1)–(A4) can be rewritten through a corollary that sums up their most salient features:

(C2) Language and truth equally arise with human animality and they characterize its species-specific Lebensform.

Where there is language, that is to say a proposition, there is the true/false rule, and where there is the true/false rule there is a proposition. The affirmation that «what can be true or false» is a proposition must not be seen to mean «as if we had a concept of true and false, which we could use to determine what is and what is not a proposition». In this case the concept of truth would play the role of a cogwheel with which to distinguish propositions from non-propositions: «What *engages* with the concept of truth (as with a cogwheel), is a proposition. *But this is a bad picture*». Bad because it suggests a before and an after: the proposition, on one side, and the true/false attribution that «fits (*paßt*)» to it, on the other. As in the game of chess there is not on one side the piece called 'king' and then the rule that says: «The king in chess is *the* piece that one can check». There are not two entities (the king on one side, the rule of checkmating on the other) that from outside someone combines together, but something like a Möbius strip that articulates itself in the rule that defines a figure (the king) as the one that one can check. The same goes for the 'proposition': «a *proposition* can be true or false can say no more than that we only predicate 'true' and 'false' of what we call a proposition». The use of the words 'true' and 'false' is a «constitutive part (*Bestandteil*) of the linguistic game and if so it *belongs* to our concept 'proposition' but *does not 'fit' (paßt) it*» (Wittgenstein, *PI*, I, § 136).

Frege defended the same thesis: Truth «is distinguished from other predicates through the fact that it is implicitly affirmed every time something is said» (Frege 1979, p. 128).

4

From axioms (A1)–(A4) and from the relevant corollaries there follows a theoretical affirmation that must be tested:

(T1) No historic-natural language can speak of the world (and a language that does not speak of the world is not a historic-natural language) if it does not contain a vocabulary and a syntax with which to practise the operations of the truth rule.

Let us return to *I have toothache*. What makes the proposition a highly species-specific event is not its communicating or manifesting to another a pain personally experienced (this result could also be achieved, and better, through inarticulate cries) but the fact that, in principle, the proposition may not correspond to a toothache really experienced: not having toothache and saying *I have toothache*. The same consideration applies to the example from which we started: *Mary is a witch* has a meaning if a person that understands it knows that it is possible to affirm the proposition (*Mary is not a witch*) that denies it.

Between the content of the proposition *I have toothache* and the experience of the toothache and, in general, between what a proposition says and the facts of which the proposition speaks there is a connection interwoven with discontinuity. Negation is one of the operators rendering possible the discontinuous connection that links language and reality. «Proposition is everything that can be true or false»—observes Wittgenstein—means the same thing as ‘Proposition is everything that can be denied’» (*PG*, § 79). «Any affirmation can be negated: if it has sense to say *p* it also has a meaning to say *not p*. (...) A proposition therefore is any expression which can be significantly negated» (*KL*, p. 22).

We can further specify the theoretical statement (T1):

(T2) If a language does not contain a vocabulary and a syntax with which to use the negation operator, in it and with it the true/false rule cannot be practised.

The thesis that we are upholding here is that negation is a specific operator of historic-natural languages and, only subordinately, of the expressive systems that are transformations or variants of historic-natural languages. Transformations and/or variants of historic-natural languages are mathematics, formal logic, and so-called artificial languages.

5

The true/false rule and the negation operator that makes it work duplicate the world: (a) the world of the *Sinne* or thoughts that live in and with language; (b) the world that, though connected with the former, lies outside verbal thoughts. Language speaks of the world that is outside words but it is not the secondary effect of the nonverbal world. «As it is possible—Aristotle maintained—to show in and with language (*apophainesthai*) that what exists does not exist and that what does not exist does exist, just as it is possible to show in and with language as not existing what does not exist and as existing what exists (...), it is possible as well to deny everything that someone has affirmed and to affirm everything that someone has denied» (*Int.*, 17a 26–31). Wittgenstein’s idea is consistent with that of Aristotle: «It would be odd to say: (...) ‘A red patch looks different when it is

there from when it isn't there—but language abstracts from this difference, for it speaks of a red patch whether it is there or not» (*PI*, I, § 446). And again: “if I say falsely that something is *red*, then, for all that, it isn't *red*. And if I want to explain the word ‘red’ to someone in the proposition ‘This is not red’, I do it by pointing to something red” (*PI*, I, § 429). The working of the negation operator in the *Tractatus Logico-Philosophicus* had been explained in the same terms:

But it is important that the signs ‘*p*’ and ‘*not p*’ can say the same thing. For it shows that nothing in reality corresponds to the sign ‘*not*’. (...). The propositions ‘*p*’ and ‘*not p*’ have opposite senses (*Sinn*), but to there corresponds to them one and the same reality (*TLP* 4.0621).

This happens because the sense of the proposition is not an effectively existing state of things but a possible state of things. In other words, the proposition is not the image of an effectively existing state of things but the image of a possibility. I quote from the *Tractatus*:

The image (*Bild*) represent-depicts (*darstellt*) a possible situation in logical space (2.202). The image contains the possibility of the situation that it represents (2.203). The image represents (*darstellt*), independently of its own truth or falsehood, by means of its pictorial form (*die Form der Abbildung*) (2.22). What the image represent-depicts (*darstellt*) is its sense (2.221).

Let us return to *I have toothache* and *Mary is a witch*. The persons that say and/or listen to the two propositions can distance themselves from them because their meanings do not represent what effectively happens but what may happen: the person that says *I have toothache* may lie and the dentist may explain to him or her in turn that his/her teeth are in a good state; a judge declares that *Mary is a witch* and others affirm that *Mary is not a witch*.

6

Let us imagine a scene in which Mr. Brown, after saying *I have a bad toothache*, happily bites on a walnut and breaks it. I cannot access Mr. Brown's internal experience but I can legitimately suspect that it is false to say that Mr. Brown has toothache. The reasons for the plausibility of this judgment of falsehood are explained in propositions 4.022 and following of *Tractatus*:

The proposition *shows* its sense. The proposition *shows* what relationship there is between things, *if* it is true (*Der Satz zeigt, wie es sich verhält, wenn er wahr ist*). And it *says that* things are this way (*Und er sagt, daß es sich so verhält*) (*T* 4.022).

To understand a proposition means to know what is the case (*was der Fall ist*) if it is true. One can understand it, therefore, without knowing whether it is true (4.024).

Let us explain the thesis that Wittgenstein is upholding here by examining the semantics of the proposition:

I went to Rome by train.

To understand the proposition it is not necessary to know if the person uttering it is saying something true or false: «a proposition can be understood without knowing whether it is true». However, understanding it means “knowing what is the case if it is true”, that is what the world would be like and what it would not be like if that proposition was true. A person that understands the proposition, in fact, cannot but know (in a non-psychologicistic sense of ‘knowing’) the potentially endless set of facts entailed by that very proposition if it is true. He/she knows, for instance, that if the proposition is true, then the person uttering it went to Rome and *not* to Paris or to Palermo, knows that to go there s/he did not use the airplane or the car or the bicycle, knows that the trip took place at a previous time and will *not* take place after the time at which the proposition is uttered, and knows a great number of other things. A person that knows what *the case is if a proposition is true* does not only know the fact the proposition states but knows above all the vast set of facts that are compatible and incompatible with the fact stated by the proposition: he/she knows what *happens* but also what does not *happen* (“S/he did not go to Rome by plane”), and *what could or could not happen* if the proposition were true.

Here ‘knowing’ does not mean instantly being able to master all the infinite implications derivable from the proposition but being able to recognize the logical connection between the fact stated by the proposition and the facts implied by it. This type of knowledge requires the capacity to use linguistic negation: «if I went to Rome by train then I did not go there by plane», etc.

Negation and the true/false rule are constitutive of the specific cognitive nature of propositions. Thanks to them the cognitive horizon of the human animal broadens beyond the confines of what effectively exists:

Positive and negative propositions are on a single plane. If I use a ruler (*Maßstab*), I do not only know the length of a thing, but I also know the length that it does not have. If I verify the positive proposition, in so doing I also falsify the negative proposition. At the instant when I know that the azalea is red, I also know that it is not blue. The two things are inseparable. The conditions for the truth of a proposition imply the conditions for its falsehood and vice versa (Wittgenstein, *WWK*, p. 87).

The most innocent and simplest proposition (for instance, Tarski’s *Snow is white*) does not describe a fact but shows a network of compatible facts (positive facts) and incompatible ones (negative facts). And since facts are representable through propositions, the same concept can be formulated in this way: «The proposition affirms (*bejaht*) every proposition that follows from it» (Wittgenstein, *TLP*, 5.124). This logical-syntactic aspect of understanding a sentence and therefore of the semantics of languages is discussed by Wittgenstein immediately before the often quoted 4.024:

The proposition constructs a world with the help of a logical scaffolding, so that one can actually see from the proposition how everything stands logically *if* it is true. One can *draw inferences* from a false proposition. (*TLP*, 4.023)

The final gloss (*One can draw inferences from a false proposition*) clarifies the thesis even better: in understanding, what counts is not the real truth of propositions but what is involved (*was der Fall ist*) in the assumption, not necessarily

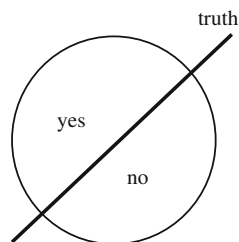
explicit, of their truth, and this is possible because «from a false proposition conclusions can be drawn» and «a proposition can be understood without knowing whether it is true». The network of facts generated by the proposition is independent of the real truth or falsehood of the proposition. The *Tagebücher* of the years 1914–1916, from whose notes the *Tractatus* was distilled, added to the observations faithfully transcribed in 4.023, by way of an explanatory comment: «So I can see that, if (x, ϕ) . ϕx were true, this sentence would be in contradiction with a sentence ψa » (Wittgenstein, *TB*, p. 104).

7

What is expounded in the *Tractatus* is an original syntactic theory of truth and meaning. In 4.024 Wittgenstein does not maintain that in order to understand the meaning of a proposition one needs to know what the conditions or the circumstances are that make it true (a thesis referable to some version of verificationism or semantic referentialism) but something very different: understanding the meaning of a proposition is equivalent to knowing how to recognize all the facts that the proposition, if it were true, should imply or exclude in conformity with the general laws of logic. A person that understands a proposition is submitted to the same constraints of logical coherence as God creating the world: «If a god creates a world in which certain propositions are true, then by that very act he also creates a world in which all the propositions that follow from them come true. And similarly he could not create a world in which the proposition “p” is true without creating all its objects» (Wittgenstein, *TLP*, 5.123).

A proposition—and this is the thesis that we are formulating with Wittgenstein’s help—produces two complementary sets of facts: the set of facts compatible with the proposition and the set of facts incompatible with the proposition, if the proposition is true. *The border that divides and generates the two sets is the meaning of the proposition.* Hence understanding the proposition is equivalent to grasping, «with the help of its logical scaffolding», the border that divides the two sets generated by the proposition. Starting from this original theoretical intuition, another affirmation by Wittgenstein that otherwise appears very obscure becomes comprehensible: «A proposition must restic reality to two alternatives: yes or no» (4.023). The proposition, precisely, fixes and divides reality into two parts: the part that is compatible («yes») and the one that is incompatible («no») with the fact shown by the proposition. «If two propositions contradict one another, their structure shows it; likewise, if one follows from the other. And so forth» (4.1211). «That the truth of a proposition follows from the truth of other propositions is seen from the structure of the propositions» (5.13). «If the truth of a proposition follows from the truth of others, this is expressed through relations in which the forms of those relations go with one another. It is not necessary for us to set them in those relations, connecting them with one another in a proposition. Those relations are internal and immediately exist when and because those propositions exist» (5.131).

Fig. 1 Sense or *spielraum* of the proposition



Every proposition, through the very fact of its existing, generates around itself what in the *Tractatus* is called a *Spielraum*, a structured space in which all the facts find an orderly position and can play out their movements. «The truth-conditions determine the game space (*Spielraum*) granted to the facts through the proposition. The proposition, the image, the model are, in a negative sense, like a solid body that restricts freedom of movement of others; in a positive sense, like the space, limited by a solid substance, where a body has its place» (4.463). Two types of sentence are exceptions: tautology and contradiction. «A tautology leaves open to reality the whole—the infinite whole—of logical space; a contradiction fills the whole logical space leaving no point of it for reality. Thus neither of them can determine reality in any way» (*Ibid.*).

We can give a second definition of the sense of a proposition equivalent to the first: *the meaning of a proposition is the Spielraum that its truth value generates, and understanding a sentence is nothing but the capacity to install oneself in its Spielraum* (Fig. 1).

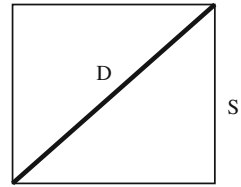
We are better able to understand the often quoted 4.022: «The proposition shows its sense. The proposition shows what relationship there is between things, if it is true (*Der Satz zeigt, wie es sich verhält, wenn er wahr ist*). And it says that things are this way (*Und er sagt, daß es sich so verhält*)».

8

The truth value of every proposition, tracing out the *Spielraum* within which the proposition is related to an indefinite multiplicity of other propositions, has a systemic effect. If the proposition *Snow is white* is true, then not only is snow effectively white but at the same time utterances like *When it snows the mountains are tinged with red* or *Last night's snowfall coloured the roof of the house green*, etc., must be false. Wittgenstein's explanation of understanding a proposition—founded on the systemic effects of the truth value—shows what a joke and a mathematical demonstration have in common.

That the truth/falsehood of a scientific proposition is not a private quality of the single proposition is too obvious a fact to devote much space to it. An example, taken from mathematics in classical Greece, will allow us at all events to clarify the concept.

Fig. 2 The demonstration of D and S



8.1

The work of a demonstration consists in the discovery of hidden spaces of the *Spielraum* of an utterance. Greek mathematicians, with a subtle and clever procedure, found the reasons for the incommensurability of the side and diagonal of the square in the hidden folds of the *Spielraum* generated by the truth of the proposition that affirms that an integer is either even or odd. The demonstration, to which Aristotle too makes reference (Aristotle, *AP*, 41a 23–27; 50a 37–38), is given in Appendix 27 of Book X of the critical edition of Euclid’s *Elementa* edited by Heiberg. Here I will give my own adaptation of it.

Theorem *If it is true that an integer is either even or odd, then it is also true that the side and diagonal of the square are incommensurable. (Incommensurable means that their relationship is not expressible through a ratio between two integers).*

The demonstration consists in the exploration of the *Spielraum* of the proposition *An integer is either even or odd*. Let us see (Fig. 2).

If D and S were commensurable then their relationship would be representable through the ratio between two integers (x/y) and we could write the proportional ratio:

$$D : S = x : y \tag{1}$$

(In words: the diagonal D is to the side S as the integer x is to the integer y).

Before proceeding further we need to bear in mind that Greek mathematicians knew that a ratio between any two numbers can be transformed into an equal ratio between two numbers that are prime to one another (a demonstration of this is found in Book VII of *Elementa*). Examples: $8/16 = 1/2$; $9/12 = 3/4$; $10/25 = 2/5$. Hence in (1) x and y are supposed to be numbers that are prime to one another.

From Pythagoras’ theorem we know that the square that has side D is equal to two squares that have side S . In a formula:

$$D^2 = 2S^2 \tag{2}$$

From (1) and (2) it follows that:

$$x^2 = 2y^2 \tag{3}$$

Equivalence (3) has the following characteristics: (a) $2y^2$ is an even number since any number multiplied by 2 is even; (b) if $x^2 = 2y^2$ then x^2 is also an even

number; (c) since only even numbers multiplied by themselves result in even numbers ($2^2 = 4$, $4^2 = 16$ but $3^2 = 9$, $5^2 = 25$, etc.), x is an even number.

A *crucial passage in the demonstration*. If x is even, y is odd since x and y are prime to one another: if y too was even, x and y would not be prime to one another because they would have the number 2 as a common divisor, as well as unity.

If x is even, then it is divisible by 2 and the following equivalences hold:

$$2w = x \quad (4)$$

$$4w^2 = x^2 = 2y^2 \quad (5)$$

From (5) there follows:

$$2w^2 = y^2 \quad (6)$$

We have seen that in Eq. (3) ($x^2 = 2y^2$), y necessarily has to be an odd number and x an even number. If we apply the same reasoning to Eq. (6) ($2w^2 = y^2$), we are forced to conclude that y cannot not be even.

Conclusion: if D and S were commensurable they would be measured by numbers simultaneously even and odd. But since a number is either even or odd, D and S are incommensurable magnitudes. The *Spielraum* generated by the truth of the proposition *Integers are even or odd* contains the reasons for the incommensurability of the side and diagonal of the square.

8.2

Demonstrations are possible because, in Wittgenstein's words, «there is no proposition that is all by itself» (*PG*, § 124) and «every proposition affirms (*bejaht*) every proposition that follows from it» (*TLP* 5.124). That is to say: «To understand a proposition means to know what is the case (*was der Fall ist*) if it is true. One can understand it, therefore, without knowing whether it is true» (4.024).

A very similar thesis was upheld by Aristotle when he noticed that speaking is a *sullogizesthai* or cohesive and non-casual concatenation of utterances so that all verbal expressions are consequences and premises of other verbal expressions. Aristotle called the internal concatenation of propositions (*lógoi*) *sylogism*. «A *sylogism* is a discourse in which, given some ⟨meanings⟩, necessarily—that is to say because of the fact that these meanings are ⟨set⟩—we must simultaneously have (*sumbaínei*) a different meaning from the meanings set» (*PA*, 24b 18–20; *T*, 100a 25–27). The *sullogizesthai* is a common element of all linguistic games, the rigorously logical ones of scientific reasoning but also the rhetorical ones with which we persuade each other and we seek with approximations the way things really are. To syllogisms called rhetorical he gave the name *enthymemes*: «Given the existence of some ⟨meanings⟩ (*tinōn óntwn*), the fact that, because of them (*dià taúta*), ⟨a meaning⟩ which is different (*éteron ti*) and new (*parà taúta*) co-occurs (*sumbaínein*) as a consequence of the fact that these meanings are (*tauta eínai*)—

both universally and mainly—there {in *Topics*} is called syllogism, and here {in *Rhetoric*} *enthymeme*» (Aristotle, *Rhet.* 1356b 16–18).

In both syllogisms—syllogism-syllogisms and *ō*syllogism-enthymemes—the truth rule is operative: «Knowing the truth and knowing what is similar to the truth belong to the same capacity and besides men have a sufficient natural disposition to truth and in most cases attain truth: for this reason tending through conjectures towards shared opinions [*éndoxa*] or towards the truth belongs to the same disposition» (Aristotle, *Rhet.* 1355a 14–18).

9

And jokes? Is the notion of *Spielraum* or systemic effect of truth also valid for explaining understanding of a joke? Let us check the explicative power of the theory on a very circumstantial example.

A dialogue taken from a comic strip by Chiappori:

A: Berlusconi complains. He says that in government his hands are tied ... tied by the President, by Parliament, by the Northern League, by the oppositions, by the trade unions....

B: And so he insists on governing with his feet! [The idiomatic expression *to do something with one's feet* means 'doing something in a slapdash way']

If the meaning of the sentence *Berlusconi says that in government his hands are tied* were limited to the fact that the sentence describes and makes it true or false—in other words, if the sense were, so to speak, a private matter of the single sentence—then every discussion developing from it would be a mysterious event and would fall outside the explicative range of theoretical semantics. What happens daily and always (every sentence uttered is the start of a discussion) could not be explained by the theory of meaning. But if the sentence with its truth conditions determines the *Spielraum* within which the facts of the world can be placed, the riposte (*And so he insists on governing with his feet*) is nothing but one of the possible logical consequences of the truth of the initial sentence. The quip was inscribed in the logical space determined by the sentence *Berlusconi says that in government his hands are tied* and it makes us laugh because it was so to speak before everyone's eyes though only the brilliant comedian succeeded in perceiving it. We also laugh because we feel that we too could have seen that connection.

Jokes do not create anything new but disclose logical consequences that everyone could have seen: «The proposition affirms (*bejaht*) every proposition that follows from it» (Wittgenstein, *TLP*, 5.124). Exactly like in demonstrations of a geometrical proposition: finding the *Spielraum* within which logical coherence is found by the truth values of the sentences *Berlusconi says that in government his hands are tied* and *Berlusconi insists on governing with his feet* is not very dissimilar from being able to explore the *Spielraum* within which the

incommensurability of side and diagonal is grasped as a consequence of the proposition that affirms that an integer is either even or odd.

10

The thesis that the truth rule governs the semantics of historic-natural languages is opposite to the one enounced, I believe for the first time, by Aristotle in *De Interpretatione*. «Every utterance (*logos*)—writes the philosopher of Stagira—has meaning (...) but not all ⟨meaningful⟩ utterances are apophantic ones [in post-Aristotelian terminology: assertors]. Utterances are apophantic if in them truth and falsehood exist and this characteristic is not found in all ⟨meaningful⟩ utterances: prayer, for instance, is a ⟨meaningful⟩ utterances but it is neither true nor false» (*Int.*, 16b 33–17a 4). For other examples of meaningful utterances that are not apophantic ones the work refers the reader to *Rhetoric* and *Poetics*.

Despite this passage in *De Interpretatione*, the Wittgensteinian version (and, I believe, also the Aristotelian version) of alethic semantics shows great explicative power. Let us examine the case of prayer. Is one prepared to recognize a prayer to have efficacy if at the same time one does not recognize as true all the utterances in which the thoughts contained in the prayer are expressible? *Our father in Heaven, hallowed be your Name, your Kingdom come, your Will be done on Earth as it is in Heaven ...*—is it possible to say the prayer meaningfully without believing in the truth of the thought contents (the *Sinne*) expressed by the prayer? It would be a strange kind of prayer if the person praying did not recognize as true the utterances that affirm the existence of a God-father that lives (even if metaphorically) in Heaven, is indistinguishable from his own Name, will one day enact a new spiritual order, and one whose Will governs the world, et cetera.

Another example. A child is trained to pray to a God that does not love blacks and homosexuals (*God protect me from the wickedness of blacks and the impurities of homosexuals*). Having become an adult, the person checks the affirmations implicitly contained in the prayer (what he or she checks them with may be a secondary issue), and arrives at the conviction that those thoughts-affirmations are false and changes prayers and/or religions. Why should the true/false rule be extraneous to prayer?

Another example that seems appropriate in this context: orders. Someone gives me the order: *Close the window!* and I realize that in the room we are in there are no windows. The context in which the command should be carried out forces me to move onto a level that precedes that of deciding whether to obey or not obey the order.

The moral, if a moral can be arrived at in these discussions: meaning, in linguistic animals, unlike in other animals, is tightly interwoven with the truth rule.

References

- Arnim, H. 1903–1905. *Stoicorum Veterum Fragmenta*. Leipzig.
- Heiberg, I.L. 1972. *Euclid Elementa* X ed., vol. 3. Leipzig: Teubner.
- Frege, G. 1966. Der Gedanke. Eine logische Untersuchung. In *Frege, Logische Untersuchungen*, 30–53. Göttingen: Vandenhoeck & Ruprecht.
- Frege, G. 1979. *Posthumous Writings*. Blackwell: Basil.
- Lo Piparo, F. 2000. Le théorème de Pythagore dans la linguistique grecque. *Histoire Épistémologie Langage* 22/1:51–67.
- Lo Piparo, F. 2003. *Aristotele e il linguaggio*. Roma-Bari: Laterza.
- Tarski, A. 1956. The concept of truth in formalized languages. In *Logic, Semantics, Meta-mathematics*, ed. A. Tarski, 152–278. Oxford: Clarendon Press.

Works of Aristotle

- AP. Analytica Priora*, ed. W. D. Ross, At Clarendon Press, Oxford 1964.
- Int. De Interpretatione*, ed. L. Minio-Paluello, At Clarendon Press, Oxford 1949.
- Met. Metaphysica*, ed. W. Jaeger, At Clarendon Press, Oxford 1957.
- Poet. De Arte Poetica*, ed. R. Kassel, At Clarendon Press, Oxford 1965.
- Rhet. Ars Rhetorica*, ed. W. D. Ross, At Clarendon Press, Oxford 1959.
- T. Topica*, ed. W. D. Ross, At Clarendon Press, Oxford 1958.

Works of Ludwig Wittgenstein

- KL. Wittgenstein's Lectures, Cambridge, 1930–1932*, From the Notes of J. King and D. Lee, Blackwell, Oxford 1980.
- PG. Philosophical Grammar*, ed. Rush Rhees, Oxford: Blackwell, 1969.
- PI. Philosophical Investigations*, eds. G.E.M. Anscombe and R. Rhees, Oxford: Blackwell, 1953, 2001³ (ed. Hacker-Schulte, 2009).
- TLP. Tractatus Logico-Philosophicus* (1921), translated by D. F. Pears and B. Mc Guinness, London: Routledge, 1975. Critical edition: *Logisch-philosophische abhandlung/ Tractatus Logico-Philosophicus*, ed. By B. Mc Guinness and J. Schulte, Frankfurt: Surkamp, 2001.
- TB. Tagebücher 1914–1916*, Blackwell, Oxford 1961.
- WWK. Wittgenstein und der Wiener Kreis (1929–1932)*, Blackwell, Oxford 1967.

Habermas' *Universal Pragmatics*: Theory of Language and Social Theory

Alessandra Pandolfo

Abstract Although Habermas' universal pragmatics has played a marginal role in studies on pragmatics, it can still make an important and meaningful contribution, precisely because it highlights the system of validity claims that lie in speech acts. This type of analysis allows one to consider the dialogic dynamics that engage speakers in the activity of reciprocal giving and asking for reasons for saying and doing things. The reasons why a speaker knows he or she can say what he or she says in the presence of other speakers constitute an essential element of the production of meaning; and reciprocally identification of the speaker's reasons by the listener is an indispensable condition of the activity of understanding of meaning. These factors allow one to render explicit the dialogue dynamics on which the quality of the utterance and understanding, like the quality of social relationships, depend.

Keywords Social theory · Validity claims · Speech acts · Reasons · Agreement

1 Introduction

Since the fifties and sixties the so-called pragmatic turn brought about by authors like Wittgenstein, Austin, Grice and Searle has enormously broadened the sphere of linguistic investigation. These authors, though with different ways, have expressed the same need to overcome the interpretative limits of semantic and syntactic analysis separating propositions from the contexts uttered in them. In order to explain the activity of linguistic utterance/understanding, it is necessary to recall the friction caused by the introduction of propositions in semantic theory.

A. Pandolfo (✉)
University of Palermo, Palermo, Italy
e-mail: alessandra.pandolfo@tiscali.it

This change of paradigm made it epistemologically urgent to redefine the confines between syntax, semantics and pragmatics. The most significant results of this operation, in my opinion, are two.

First of all, there is the overcoming of the negative prejudice according to which pragmatics would be a *waste-basket*¹ which contains linguistic facts considered minor or marginal. This nullifies the hierarchy between primacy of analysis of propositional meaning and almost accessory observation of the effects of meaning produced by the use of the proposition in the context of utterance. For linguistic pragmatics the effects of meaning are not added extrinsically to the literal meaning of the proposition. Rather, utterance activity is the correct locus where one can find the factors contributing to the genesis of meaning, without which the propositional sense, purified of the contaminations given by use, becomes a mere abstraction.

Another effect of the pragmatic turn lies in the overcoming of the rigid demarcation between linguistic and non-linguistic components. If meaning is constituted in contexts of utterance, it is necessary to take into account all those psychological, cognitive, environmental, cultural and social factors that are interwoven in the fabric of experience. Understanding a proposition is as immediate an operation in daily life as it is complex from the point of view of theoretical analysis, precisely because it brings all experience into play. Certainly at this point there is the risk that the notion of context of utterance can prove so dense and opaque as to resemble the famous “dark night in which all cows are black”. A contribution has been made to dispelling this danger in the last few decades by the rich proliferation of studies aiming to eliminate this opacity, and to show how the psychological, cognitive, cultural and social factors of the activity of uttering/understanding can be experienced in propositions, now implicitly presupposed, now explicitly signalled, now implied in conversational dynamics. Through these analyses, pragmatics shows the nature and dimensions of the interstices which are present in the linguistic space and in which there nests the richness of human experience.

In this context the universal pragmatics worked out by Habermas has occupied a less central position, prejudicially deemed too abstract and philosophical. Starting from these general considerations, my contribution intends to re-examine the validity of these prejudices in order to credit the Habermasian perspective with the merit of having highlighted fundamental aspects of the activity of utterance/understanding, otherwise investigated in a superficial way. These are the stages of the argument:

1. First of all I will try to understand the reasons for the “success” obtained by Habermas’ model among pragmatics scholars. They lie, in my opinion, in the peculiar nature of the pathway taken by Habermas, who aims to understand not only language, but more widely the social nature of man. But precisely for this reason, from my point of view, Habermas’ approach appears anything but

¹ See Bar-Hillel (1971, 405). See also Bianchi (2003, 7).

insubstantial and is all the more illuminating in that it succeeds in rendering explicit the normative potential of language on which the genesis of social institutions depends.

2. Secondly, I will dwell on the stages that lead Habermas to delineate the space of linguistic intersubjectivity as the locus of the mechanism of social coordination.
3. In the third section, after describing the model of universal pragmatics hinging on the system of validity claims, I will endeavour to stress its value. I will show that something truly fundamental in the activity of utterance/understanding is the dimension of the commitment of speakers, who in a more or less explicit way link their utterance to a weft of reasons. And the presence of these reasons places any speech act within a dialogic dynamics.
4. Lastly, I will expound my conclusions, showing the fertility of the synthesis between theory of language and theory of society effected by universal pragmatics. This synthesis renders accessible the dialogic dynamics of sociality and the relative political and ethical implications.

Before proceeding, a clarification of a methodological type² is necessary. I have preferred to keep in the background the vast critical literature on universal pragmatics, addressing which would have required much ampler discussion. It is well known that the most recurrent issues concern comparison between universal pragmatics and the *transcendental pragmatics* of Apel;³ reduction of the notion of truth to that of consensus;⁴ the criticisms of the normative character of pragmatics and its claims to universality, which would not take into account plurality of contexts;⁵ criticism of the prevalence of communicative action over strategic action;⁶ criticism of the notion of “ideal speech situation”.⁷ I will not expressly address such observations, but I propose a reading that tries to throw a different light on universal pragmatics aiming to valorise its strong points, with the desirable effect of attenuating the impact of some of these criticisms. Many critical observations are grounded if we attribute a point of view to Habermas that is either foundational, or descriptive; neither the one, nor the other. I believe that the effectiveness of universal pragmatics lies in the reconstructive point of view⁸ that says neither “how it should be” nor “how it is” but shows the explicative

² See Dews (1999), Honneth et al. (1989), Outhwaite (1994), Petrucciani (2000), Wellmer (1989).

³ See Apel (1989); according to Apel, Habermas, foregoing the transcendental point of view, enacts an idealisation of empirical reality, to which he remains bound. See also Kuhlmann (1985), Pedroni (1999, 286–287).

⁴ See Hentig (1973), Ilting (1976), Ferrara (1987).

⁵ See Rorty (1998, 293). To the contextualist Rorty Habermas' approach appears excessively abstract, as it is based on an ideal of universal validity.

⁶ See Rasmussen (1990).

⁷ *Ibidem*.

⁸ On the reconstructive character of science, see Habermas (1984, 363).

linguistic conditions of all aspects of sociality, both communicative ones and strategic and competitive ones. In this sense theory of language is inserted in broader investigation of the nature of sociality.

2 From Social Theory to Theory of Language

Within the literature on linguistic pragmatics there are not abundant references to universal pragmatics.⁹ This is due to the singularity of this model, which it is difficult to inscribe in the disciplinary confines of linguistics. Habermas' interest in the analysis of language is closely connected to analysis of the nature of social phenomena, and to the possibility of practising social criticism in the best Frankfurt tradition.¹⁰ But, from my point of view, the peculiarity of this pathway does not constitute a limit; rather, it provides precious hints for enriching our understanding of what happens when we speak.

How are social constraints and bonds produced and what is their nature? What relationship is there between society and social actors? What is the logic of social change? These are just some of the questions that lead Habermas to take the pathway of the so-called *sprachtheoretische Grundlegung*¹¹ of social theory; that is to say, it is a matter of identifying the normative matrix of sociality in the space of linguistic intersubjectivity.

To understand the value of this move, reference has to be made, though briefly, to the models of social research Habermas faces up to, denouncing their interpretative insufficiencies. This complex critical process leads to the definition of the *Theorie des kommunikativen Handelns*.¹² Although here it is not possible to reconstruct all the developments of the question, I would like at least to show what is at stake in establishing a correct relationship between theory of language and social theory. I am therefore concerned to go into Habermas' analyses of what happens when social theory either ignores language, or when, though considering language, it uses a "bad" linguistic theory incapable of accounting for the normative potential of linguistic action.

Wanting to link the genesis of social structures (institutions, norms, traditions, customs, values) to the linguistic actions of social actors, Habermas has reservations on the point of view of the theory of action, which often disregards the

⁹ Greater attention is paid to Habermas' position by Schlieben Lange (1975).

¹⁰ On the link between Habermas and the Frankfurt School see the recent De Simone (2010).

¹¹ Habermas uses this expression in "Vorlesungen zu einer Sprachtheoretischen Grundlegung der Soziologie" (1970–1971), published in 1984 in *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns* (from here on referred to as VE). This volume includes a big series of preparatory studies for *Theorie des kommunikativen Handelns*, which are important for understanding the shift in the philosophy of consciousness (*Bewusstseinsphilosophie*) toward linguistic pragmatics (*Sprachpragmatik*), as Habermas states in VE (1984, 7).

¹² Habermas (1981) (1988 edition). From here on this work will be referred to as TKH.

language dimension. The profile of the social actors has to integrate the acting subject with *speaking and acting subjects*.¹³ Language is the space in which there occurs the production of structures endowed with meaning that link speaking and acting subjects within a common normative fabric. This theoretical choice appears decisive, because it frees social sciences from the sterile choice between two opposed possibilities, aiming to trace the matrix of meaning either “inside” the mind of social subjects or in the devices of the social system that “from the outside” would impose normative constraints on single agents. The two perspectives, that of atomism and methodological individualism on one side and the functionalistic characterizing the systemic theory on the other, ignore the space of linguistic interaction, which instead for Habermas is decisive for defining the nature of sociality. The absence of this level of analysis jeopardizes the interpretative force of the theories, which can be traced respectively to the models of Max Weber and Talcott Parsons. In [Chaps. 2 and 7](#) of TKH, Habermas first expounds the perspective of the two famous sociologists, and agrees with their demands and analysis methodology; lastly, he brings out the interpretative deficit of the theories, which can be overcome precisely through recourse to the linguistic notion of *communicative action*.

The crucial question that orients the comparison with Weber and Parsons concerns the genesis and nature of the social bond, and more precisely the constitution of that set of rules, norms, values and traditions that orient and coordinate the action of individuals in general. But, in addition to establishing how this common background of indicative criteria is constituted, it is also necessary to take into account the tension between social action and individual action, between the horizon of meaning that constrains a group of subjects and the possibility of bending common sense in particular directions. It is also necessary to explore the dynamics of this tension, on which the richness of human society depends. Only in this way is it possible to understand to what extent the social organization tolerates the tensions, which give rise to zones of greater or lesser compactness and cohesion, possibly identifying the break-up points at which the compactness of the social fabric is lacerated.

With respect to such issues Weber's theoretical choices are mostly oriented toward the pole of the individual, as if he wished to reconstruct the genesis of the social orientations of meaning starting from the indicative criteria presiding over the action of the latter. By contrast, Parsons' theoretical choices are mostly concerned with clarifying the way in which the social system succeeds in ensuring that manifold individual orientations converge towards a single direction of meaning.

Let us begin to examine the critical observations that Habermas makes regarding Weber. Habermas appreciates Weber's attempt to describe the complexity of social action, through different modalities of action that bring different types of reason into play: *instrumentally rational action*, *value-rational action*, *affectual action* and *traditional action*. By this variegated typology Weber intends

¹³ Speaking and acting subject is an indivisible concept recurring in all Habermas' work.

to overcome the limits of the atomistic conception, valorising the normative components of social action, that is to say that set of values, constraints, customs and traditions that orient the choices of single individuals. But despite his interest in normative components, according to Habermas Weber would continue to have a preference for the point of view of the unrelated agent and to opt for the intentionalist perspective. In this connection, Weber confines himself to exploring the connection between the singular intention of the subject regarding meaning and the possible forms of action, thus failing to clarify the conditions on the basis of which the agent is able to plan different forms of action “inside” his/her mind, now orienting him/her towards the satisfaction of selfish needs and now towards the recognition of shared values. This analysis model has a predilection for investigation of the retrieval of adequate means with a view to the action objectives chosen by the individual. But it is difficult, in these terms, to establish the logic of the selection of purposes, on the basis of which an acting subject chooses to operate with a view to individual needs or in compliance with social constraints. It would seem that the matrix of individual needs and the matrix of social constraints and bonds operate as two parallel and unrelated circuits, without it being possible to understand what activates now the one and now the other. How is it possible, then, that the individual is attracted by common values and aims, which drive him or her beyond his or her individual needs? How does it come about that values and traditions exert constraint on a plurality of individuals? The possibility of accounting for the components of agreement or conflict inside social action depends on the answer to such questions. But the monologic genesis of intentions of meaning rules out this level of analysis. The perspective is broadened, nevertheless, if the intentions of the agent regarding meaning are considered not only as the unjustified principle from which the action originates, but also as the result of a dialogic process with which the subject agent defines his/her own objectives measuring individual needs against the values of the community he or she belongs to: «Dabei interpretieren sich Ziele und Werte wechselseitig».¹⁴ This process, however, requires a mind model implying a close relationship between social mind and linguistic mind going well beyond the intentionalist perspective.

Hence according to Habermas’ diagnosis, the explanatory deficit is due to the fact that Weber, though having set the notion of *meaning* at the basis of the theory of action, «erläutert ‘Sinn’ nicht anhand des Modells sprachlicher Bedeutungen und bezieht ‘Sinn’ nicht auf das sprachliche Medium möglicher Verständigung, sondern auf Meinungen und Absichten eines zunächst isoliert vorgestellten Handlungssubjekt».¹⁵ Habermas imputes the responsibility for the interpretative failure to the mentalist (and not linguistic) category of meaning. Weber would have realised the complexity of social life, without succeeding in identifying its possibility conditions, since the conscientialist (and not linguistic) mind model does not explain how social coordination comes about among actions planned

¹⁴ VE (1984, 318).

¹⁵ TKH (1981) (1988 edition), Band 1, 377.

“inside the mind” of single individuals. How can one coordinate individual objectives through common orientations of action? The analysis of the genesis of individual meaning does not make it possible to grasp the genesis of social meaning that has the power to constrain a community of individuals. But according to Habermas this is precisely the task of the social theory.

In this respect more adequate answers would seem to be provided by the systemic social theory with a functionalist orientation. Habermas, nevertheless, does not find these solutions wholly convincing, although Parsons proposed them as a correction to methodological individualism. According to Parsons, coordination among social actors is determined by the heritage of values, norms, rules and shared customs that express the general orientations of action of a collectivity. In this sense, he shifts his investigation from the intentions of the single actor to the connection between motivations and value orientations. The onus of the mediation between individual needs and general norms is on the social system and its mechanisms which are able to fulfil the integrative functions that give stability to the system. Society, on the basis of the systemic theory, works as a true cybernetic machine able to regulate itself in response to environmental variations and in view of its own self-preservation.¹⁶ According to Parsons, four functions are essential to the organization of the system and can be summed up with the abbreviation AGIL:

- *A (Adaptive)* indicates functions of environmental adaptation, enacted by the economic subsystem that guarantees the necessary resources;
- *G (Goal attainment)* indicates the functions performed by the political institutions that mobilize all the social energies toward priority aims;
- *I (Integrative)* indicates the functions of integration presided over by the legal subsystem that checks the effective validity of the norms, contemplating sanctionative mechanisms against their infringement;
- *L (Latent pattern maintenance)* indicates the functions of maintenance of the latent models pursued by the culture subsystem, that is to say by the set of organisms responsible for the socialization of the orientations of action.

Habermas, on the one hand, appreciates the global approach to the social phenomenon in a direction that integrates institutional, cultural, ethical-moral and economic aspects. He also appreciates the space that systemic sociology devotes to the notion of meaning, when it accepts that norms, values, institutions and traditions perform the function of integrating actions. On the other hand, Habermas makes some critical observations. The fact is that Parsons confines himself to distinguishing between the objectives pursued by each of the four functions of the system, without however, clarifying the mode of operation of each device. There is a certain levelling of the four functions on the cybernetic model that clarifies the integration role without specifying the modality. This jeopardizes understanding of the culture subsystem, since the dynamics of the working out of meaning that constitutes the peculiar normative identity of a society is not clarified.

¹⁶ See TKH (1981) (1988 edition), Band 2, 338–339.

The functionalist theory aims to highlight the force with which normative constraints absorb the action of individuals, because this proves to be integrated and not to disintegrate. In this sense, we can only wonder about the effectiveness with which such force is exerted in view of maintenance of the cohesion and equilibrium of the system. However, this renders the possibility of appraising the quality of the normative bonds entirely marginal and does not address issues that are instead crucial for a social theory aiming to integrate the descriptive aspect and critical aspects: whether normative constraints are imposed by authority or are the result of processes of sharing; whether adhesion to the normative constraints is produced by recognition of the validity of the constraints or exclusively derives from the use of sanctionative mechanisms; to what extent the aims of the social system can ignore the demands of social actors.

And so, assuming that the social system is integrated in virtue of the good operation of its devices, the solid compactness of the social system does not tell us a great deal about the possible life and action choices of the single social actors. This happens because Parsons interprets the validity of cultural values in the cybernetic sense of control functions.¹⁷ In this case too, Habermas attributes the weakness of this approach to the lack of correct consideration for language. Instead, Habermas attaches great value to the action of social coordination pursued by the systemic devices that are imperiously imposed on the social actors, and to the action of coordination through language that enables individuals to participate in working out the social fabric of meaning. But here we see the big gap between Parsons and Habermas: «Auf der einen Seite kann man [...] die Analyse der Sprache auf der Ebene einer Theorie des kommunikativen Handelns ansetzen. Dann kann man an allgemeine Linguistik und Sprachphilosophie anknüpfen, aber auch an die soziologischen Handlungstheorien, welche Interpretation und Verständigung als Mechanismus der Handlungskoordination untersuchen. Man begibt sich dieser Möglichkeit, wenn man andererseits die Ebene sprach- und handlungstheoretischer Untersuchungen *systemtheoretisch unterläuft* und den Mechanismus sprachlicher Verständigung für die Gesellschaftstheorie von vornherein nur unter dem funktionalistischen Gesichtspunkt der Systembildung fruchtbar macht. Auf diese Weise werden die rekonstruktiv gewonnenen Merkmale kommunikativen Handelns, [...] durch Elemente ersetzt, in denen sich die abstrakten Bestimmungen von allgemeinen Systembildungsprozessen *lediglich wiederholen*».¹⁸

We will conclude this section by stressing the theoretical aspects which Habermas tries to organize differently. Careful reading of the Weber and Parsons model convinces Habermas that the social theory must re-calibrate its scope, so as to highlight the dynamics of the social space produced by the polarity between individual demands and general orientations. This space is not accessible either to the too close gaze of methodological individualism, which if anything yields an

¹⁷ See TKH (1981) (1988 edition), Band 2, 370–371.

¹⁸ TKH (1981) (1988 edition), Band 2, 390.

action subject that is *untersozialisiert*,¹⁹ or to the excessively distant gaze of functionalism, which yields a global systemic configuration, losing sight of the role of action subjects that are *übersozialisiert*.²⁰ In both cases a harmful element is the presupposition of an insufficiently problematized notion of language that ignores the peculiar linguistic-normative nature of the social understanding/agreement. This leads Habermas to opt for the *sprachtheoretische Grundlegung*, which identifies the matrix of human sociality in the space of linguistic intersubjectivity. Grounded in the polar space of linguistic intersubjectivity, the social theory can at last account for the relationship between individual and society without either of the two poles entirely absorbing the other.

3 The Polar Constitution of the Space of Linguistic Intersubjectivity

Habermas' interest in theory of language is prompted by a search for the conditions of the tension between general and particular that constitutes social action. In this connection, the mechanism of social coordination creates constraints and bonds that govern the action of a society of individuals, but without eliminating the possibility of setting going subjective initiatives in the common space. For this reason, according to Habermas, the nature of sociality is only comprehensible on the basis of careful consideration of the linguistic constitution of human life.

The smallest nucleus of sociality involves the sharing of forms of organization of experience that establish a common space of action. The matrix of such forms of organization of experience has normative power, in that the set of meanings, rules, customs, norms and traditions exerts a binding power on the social actors; the latter, belonging to a social dimension, in point of fact find themselves following the rules that coordinate the actions of individuals. Now, as we have seen, the matrix of normative constraints and bonds does not exclusively lie either "inside the mind" or in the social devices that extrinsically dictate the conditions for understanding/agreement. The mind of the single individual would have normative power that is too weak and, to be aware of this weakness, one need only read Wittgenstein's observations on the fact that it is impossible to *follow a rule privatim*.²¹ Besides, the social devices would have such strong normative power as to pass "over the heads"²² of the social actors, taking their action in a single direction of meaning. The normative matrix of constraints and bonds therefore lies in the space of linguistic intersubjectivity. We will now endeavour to understand how this space is structured and what its peculiarities are.

¹⁹ VE (1984, 581).

²⁰ VE (1984, 581).

²¹ See Wittgenstein (1953, § 202).

²² Habermas often uses this expression.

The speaking and acting subject operates within a dense fabric of rules, norms and constraints and it is not by chance that Habermas thinks that the Wittgensteinian notion of *language game* is indispensable to describe the status of the rules that inhabit the space of linguistic intersubjectivity. Wittgenstein, as is well known, acutely analyzed the properties of human activity of “following a rule”²³ that I would sum up in the following way. First of all, criticizing the mentalist tradition, he stresses the public character of the practices in which the rules are rooted. The rules constitute a background of shared certainties, enabling the activity of coordination among the movements of the actors present on the stage of the interaction. Such sharing mainly occurs through forms of immediate and unintentional learning; there is a form of immediateness in the socialization of the practices of a collectivity, precisely because the rules inscribed in the practices are automatically internalized by the agents through their very actions. The background knowledge that is sedimented in the co-acting of the social actors constitutes a storehouse of meaning on which they draw to orient themselves in reality.

With an anti-conscientialist attitude Wittgenstein places the matrix of the constraints that coordinate action in the space *among* the speaking and acting subjects. This approach accounts above all for the sedimentation of the praxis and language rules in shared customs. According to Habermas, this is an extremely significant aspect of the game of social coordination; it is the aspect whereby the rules, stabilized in social practices, acquire validity. In this sense, the matrix of the validity of the rule lies precisely in accredited use. However, this is merely factual validity. When we say that a rule is factually valid, we mean that the rule is “ongoing” and exerts an effective imperative constraint on a group of agents. There we observe the concrete existence of a general interpretative criterion to formulate judgments of correctness/incorrectness on the single applications of the rule. We can thus identify both the mandatory character of the rule that is imposed on speaking and acting subjects and its unproblematic character; as Peirce himself stressed, as long as habits are effective and create no problems, they become undisputed automatisms. Factual validity, however, is not an exhaustive aspect of the life of the rules. If it were so the validity criteria would be destined to become rigid in social practices, imposing forms of social homologation. In effect, Wittgenstein too had paid considerable attention to the problem of plurality and variation in rules, identifying the occasion of new uses and variations that interrupt the regular operation of human practices.

But Habermas, interested in analyzing the polar space of linguistic intersubjectivity, dwells precisely on the conditions that, against the uniform background of shared practices, are responsible for the emergence of individual demands and positions. In order to define the normative space of the social game in all its complexity, it is indeed necessary to reconstruct the system of general rules; but it is also necessary to observe the dynamics of the movements of the single social actors within this game space. And, if one wishes to focus on the subjective dimension of the dynamics of social action, one absolutely cannot ignore the

²³ The theme is abundantly present and amply discussed in Wittgenstein (1953).

linguistic conditions, and precisely the differentiation of the *I-you* linguistic roles and the logical operator *yes/no*.

The differentiation of the roles codified through the *I-you* indicators is the specifically linguistic condition of the polarization of the social space. *I* and *you* indicate the specific positions in relation to the common social horizon. The space of linguistic intersubjectivity is such as to guarantee both identification between speakers and differentiation. Saying *I*, the speaker affirms his or her own non-identity with respect to *you*; but, from another point of view, crediting *you* with the ability to say *I*, also recognizes the identity that unites them. Linguistic life is a condition of identity and difference, a space of common belonging and sharing, due to the ability of the *I*'s to recognize the identity that unites them. But linguistic life is also a space of affirmation of difference because of the impossibility of superimposing *you* and *I* without residues. Referring to the tenacious individualizing constraint exerted by the linguistic medium, a true "*principium individuationis*", Habermas in this connection quotes G. H. Mead, a social psychologist of pragmatist inspiration from whom he received several promptings; the fact is that Mead insists on the complementarity between individualization (*Individuierung, Vereinzellung*) and socialization (*Vergesellschaftung, Sozialisation*).²⁴

Here I will avoid addressing the psychological point of view on the formation of *I*, but I would like to dwell on the effects of this differentiation from the merely logical point of view. The possibility of the speaker to take up a specific position, recognizable through the *I-you* linguistic coordinates, creates a spatial distance between a social practice, governed by stabilized and shared custom, and the activity of the subject that is able to take an individual initiative. And here the other linguistic operator, the *yes/no* operator, comes into play. Taking up a specific position in relation to the common social fabric, and possibly distancing themselves from it, the speaking and acting subjects may accept, but they may also reject the normative constraints and thus indicate new modalities of acting, saying and knowing. In the space of linguistic intersubjectivity it is not at all to be taken for granted that the *I* will reproduce the rule as it is; he or she can also interpret social behaviours in an unexpected way, resulting in performances with different degrees of conformity, ranging from full identification to innovation and divergence.

The combination of the linguistic operators *yes/no* and *I-you* creates an incalculable potential for differentiation of roles, schemes, rules and norms that multiply the possibilities of action and the orientations of meaning. But the fact that plurality and flexibility of rules enter into human experience with language does not mean that language is only a factor of contingency and laceration of the otherwise homogeneous social space. At the same time in the space of linguistic intersubjectivity conditions operate to check this implacable mechanism generating difference. That is to say, we are talking about the dialogic dimension making it possible to compare the countless modalities of doing, saying and knowing, and of establishing more or less formalized procedures of validating selection.

²⁴ See Habermas (1988, Chap. 8). See also Mead (1934).

The dialogic dimension and the relative operation of validating selection brings out in the social space a different property of rules and constraints, expressible once more in terms of validity, but with a different meaning. We have already explained what is meant by factual validity, rooted in shared practices. But it is a very different matter if we find ourselves dealing with a plurality of competing rules, as can happen in the variegated social space where speaking and acting subjects act in such a way that they can distance themselves from the rules in force through new applications. What rule is to be applied in circumstances in which different orientations of action are to be found? If the only meaning of validity were that of the rule that is in fact in force, there would be no doubt: every form of distancing from the rule that is valid (in actual fact) should be considered “not valid” without any possibility of appeal. But in the social space this possibility of appeal exists and is inscribed in the linguistic constitution of intersubjectivity.

And, in fact, where competing rules emerge in praxis, we observe a dialogic work of comparison between rules that tends to foreground precisely the value and the quality of certain operational modalities in preference to others. Showing that a rule is more valid than others means saying that it is better, preferable from some point of view; in short, it involves claims to justify the good qualities to legitimate its use. Claiming to show in the dialogic dimension that a rule is valid amounts to saying that that rule is good, effective, correct, adequate, etc. Hence we are talking about a validity judgment that brings into play the quality of the rule and has to do not so much with the rules already established in social practices, but with the institution, justification and legitimization of the rules. This implies the mutual commitment of speakers to questioning themselves on the validity, goodness, effectiveness, soundness and justness of the rules. In this case the validity claim made by a speaker is instantiated in the dialogic dynamics of the public space where other speakers can say *yes* or say *no*, that is to say can recognize them or not recognize them as having validity. The operation of evaluation underlying opinion of validity that says *yes/no* penetrates into the very constitution of the social rule. Whether the result is affirmative or negative, whether the rule is accepted or rejected, depends on dialogic-discursive dynamics. Discourse is a form of interplay of different rules: providing arguments, giving reasons and justifications, criticizing, are the operations with which speaking and acting subjects co-operationally commit themselves to defining the reasons of social living. As I have affirmed elsewhere, «Precisely in this phase, because of the linguistic specificity of the procedures of legitimization that involve comparison and justification, it is possible for the rules to be charged with reasons. The validity of the norms therefore implies a component of reasonableness. The moment of institution saves the schema from occasional presentation, turning it into a stable and repeatable criterion, on account of the commitment taken on implicitly or explicitly by the agents, who are reciprocally bound to follow the same rule deemed reasonably valid».²⁵

²⁵ Pandolfo (2010, 94).

Through the description of the space of linguistic intersubjectivity and its polarization between general and individual, we can appreciate all the richness of the notion of validity. And, in order to consider all aspects of sociality, it is necessary to take into account precisely the double claim to validity that exists in the normative constraints: (a) the validity of a constraint implies an aspect of factuality and, in this sense, coincides with a rule effectively being in force; (b) but the validity of a constraint also involves public and dialogic recognition of the good quality and the good reasons for which a rule should be valid for a generality of individuals.

In this connection, according to Habermas we should not «alle kontrafaktischen Momente aus dem Geltungsbegriff zu tilgen, und Geltung mit Akzeptanz, Gültigkeit mit sozialer Geltung gleichzusetzen».²⁶ The fact that the norm exists does not in the least tell us whether that norm is worthy of being recognized.²⁷

The double dimension of the notion of validity makes us aware of

- (a) the dynamics of stabilization of the rules, which implies an inevitable mandatory component; and, indeed, if there were not a mandatory moment when the rule is imposed and exerts power over a group of individuals, the constitution of the social space would not be possible;
- (b) the dialogic dynamics, involving the commitment of the social actors to measuring, criticizing and negotiating the quality of the bonds on the basis of reasons; if there were not this connection between validity and reasonableness, social living would be reduced to an obtuse practice limited to reiterating the constraints already existing, without being able to appreciate its reasons. Instead, in the space of linguistic intersubjectivity, it is the speaking and acting subjects themselves that take upon themselves the reasons of social living.

A social theory is exhaustive if it keeps in mind these two inseparable aspects of normative life. Otherwise one risks constructing a social theory that entirely cancels out the reasonableness of social living; and in a sense, this is precisely the limit of the systemic model, according to which rules are in force and no one knows why; or at least the social actors do not know it, blindly caged as they are in the logic of self-regulation followed by that sort of macro-subject²⁸ that is the social system. Against this, it would not even be possible to entirely wipe out the dimension of imperativeness from social constraints, since otherwise one could never create the public space on the basis of which the reasons of the speaking and acting subjects can begin to take shape.

Highlighting the linguistic roots of the polarization of the social space gives rise to a much richer and more complex vision of sociality, in which the tension between the index of factual validity and the validity index measuring the quality

²⁶ TKH (1981) (1988 edition), Band 1, 56.

²⁷ See Habermas (1983).

²⁸ This too is a recurrent expression in Habermas' works. See e.g. Habermas (1976, 247) and all Chap. 12 of Habermas (1985).

and reasonableness of rules cannot be eliminated. Within this space produced by the tension between factual validity and validity claims, between imperativeness and reasonableness, societies define, fix, establish, deny and renegotiate their constraints, bonds and normative identity.

4 Language and Universal Pragmatics: The System of Validity Claims

We have thus far reconstructed the polarization of the space of linguistic intersubjectivity in which conditions exist for understanding/agreement: a fabric of common rules—the linguistic differentiation of the *I/you* roles—the indication of the proximity/distance with respect to the common constraints through the *yes/no* operators—the distinction between factual validity and the validity claim based on reasons—and the dialogic dynamics of validation.

A clarification on the notion ‘understanding/agreement’ that has real semantic density: this notion can refer both to mere linguistic understanding and to the process involved in reaching understanding. In my opinion, this semantic density is due to the fact that understanding is not a space of static equivalence between imperativeness and validity claims. Understanding is produced precisely by the tension between validity (factuality, imperativeness) and validity claim (a claim that the rule is good), between stabilized validity and problematized validity. Understanding is a process involving a tension between homogeneity and dissimilarity of constraints. It is not by chance that Habermas distinguishes between the following aspects of the dynamics of understanding/agreement:²⁹

- understanding as *Verständigung/Verständnis*, which indicates the activity of people committing themselves to coordinating the forms of their acting, knowing and speaking;
- agreement as *Einverständnis*, which indicates the aim of the accord fully achieved;
- the *grey zone* of incomprehension (*Unverständnis*) and misunderstanding (*Missverständnis*) that dots social life with experience of failure, precariousness and uncertainty.

With universal pragmatics³⁰ Habermas, using Wittgenstein’s theory of language games and Searle’s theory of speech acts, intends to reconstruct the linguistic infrastructure producing the complex validation dynamics, fundamental for the constitution of the social fabric. A clarification is in order. In my opinion, to

²⁹ See VE (1984, 355).

³⁰ The main texts are *Was heisst Universalpragmatik?* (1976), in VE (1984, 353–440), and the 4th Vorlesung *Universalpragmatik—Überlegungen zu einer Theorie der kommunikativen Kompetenz*, in VE (1984, 83–104).

fully understand the value of universal pragmatics it is important to grasp the point of view of the analysis. First of all we need to avoid seeing it as an operation conducted with empirical-descriptive intents; universal pragmatics does not describe what in fact happens. Habermas repeatedly underlines the reconstructive intent.³¹ What does this mean? What is the result of this reconstruction? It is a matter of identifying the factors through which speaker come to a shared understanding of something in their concrete activity. These factors taken together define the logical space of linguistic understanding, from which the all possible forms of understanding derive. And only in this way is it possible to explain how the optimal forms of understanding are possible, but also what factors hinder and threaten it. The reconstruction of the logical space of understanding is a tool for understanding the nature of sociality, but it is also a diagnostic tool for detecting the quality of social constructions.

For Habermas, the starting point of the analysis is the speech act, the true “cell” of understanding, constrained by the factors on which the validity of speaking (*Rede*) is founded. To reconstruct the communicative competence of the speaker, it is necessary to make explicit the set of rules a speaker masters when he or she produces utterances (*Äusserungen*); that is to say, it is a question of understanding what a speaker has to be able to do so that his or her saying will adequately interweave with the weft of experience. Obviously, to analyze this universal core of communicative competence, the methodology has to be differentiated from that of the linguist. Although he attributes full legitimacy to the work of linguists who abstract language (*Sprache*) from utterance activity (*Rede*),³² Habermas distinguishes the different aims of linguistics and pragmatics.³³ Linguistic analysis reconstructs the system of grammatical rules for the formation of correct sentences, formalizing the syntactic properties of propositions (*Sätze*) and the truth conditions that allow a proposition to refer correctly to the world. But this level of grammatical and semantic analysis is not enough if one wants to understand the much broader set of rules the speakers have to master for coming to a shared understanding of something. This need leads Habermas to blend together semantics and pragmatics in such a way as to take analysis of the genesis of meaning back to effective utterance activity. From this point of view, the pragmatic dimension does not extrinsically add anything to the proposition; the way in which a speaker uses language is an integral part of the constitution of the meaning of utterances. Language is indissolubly interwoven with three spheres of experience:

- external reality, an aspect of the world of the speaker;
- social reality, the horizon within which habits, traditions, beliefs and institutions are shared;

³¹ On the reconstructive character of science see VE (1984, 363).

³² On the distinction between *Sprache* and *Rede*, see *Was heisst Universalpragmatik?* in VE (1984, 358 and 393–395).

³³ See VE (1984, 386).

- the reality of the speaker, his or her communicative needs, his or her intentions and experiences, obviously always expressed in the first person.

The speech act as an unit of analysis requires this interplay of language, world, society and speakers' roles. It has meaning if there is a recognizable reference to the world, if there is a common social context, if it is informed by the grammar of a historical-natural language, and if there is a speaker expressing a specific attitude.³⁴ The activities with which the speech act is interwoven are regulated activities: reference to the world implies ways of organizing experience proper to a culture; reference to other social actors implies customs, traditions, norms and common beliefs; the construction of comprehensible sentences implies knowledge of the grammar of a language; the expression of one's own needs involves mastery of the linguistic roles of *I* and *you*. Therefore, the speaking and acting subjects operate within a dense fabric of rules, norms and constraints; the speech act is the move of a speaker only endowed with meaning against the background of a specific language game.

The meaning of the speech act is partly based on propositional content. And, in fact, taking up Searle's schema³⁵ F(p), Habermas distinguishes the propositional content (p) from illocutionary force (F) of the speech act. The latter indicates the speaker's specific initiative in the presence of interlocutors. In speaking (*Rede*) we observe a *double structure* (*Doppelstruktur*) that makes two levels recognizable «auf denen sich Sprecher und Hörer gleichzeitig verständigen müssen, wenn sie ihre Intentionen einander mitteilen wollen [...] (a) die Ebene der Intersubjektivität, auf der Sprecher und Hörer durch illokutive Akte die Beziehungen herstellen, die ihnen erlauben, sich miteinander zu verständigen, und (b) die Ebene der Erfahrungen und Sachverhalte, über die sie sich in der durch (a) festgelegten kommunikativen Funktion verständigen möchten».³⁶ The capacity to keep these two levels united, using the reflexivity (*Selbstbezüglichkeit*)³⁷ of language, is a crucial aspect of communicative competence.

The reflexive structure of language³⁸ allows a double operation: it allows people to experience reality and to act in it; but, at the same time, it allows them to observe, appraise, share, refuse or criticize the initiatives of speaking and acting subjects. The double structure of the speech act indicates that the speaker is never alone in experiencing reality. In this sense Habermas departs from referentialist semantic theories, according to which anchorage to the world is the only source of meaning of a given utterance. The speech act certainly possesses a representative function; in this sense Habermas does not entirely favour Wittgenstein, who would have radically segregated language games from all reference to the world

³⁴ See VE (1984, 437–439).

³⁵ See Searle (1969).

³⁶ VE (1984, 406).

³⁷ See VE (1984, 407).

³⁸ Habermas' revisiting in a linguistic key of the modern notion of reflection is amply dealt with in the fine volume by Swindal (1999).

(*Weltbezug*).³⁹ But in the speech act, the level of experience of states of things is not monologic, since the illocutionary dimension is based on the commitment that speakers take on in sharing their experience. In this connection, every stance implies that the speaker explicitly or implicitly commits himself or herself to saying what he or she says sensibly, and therefore to having reasons for maintaining its validity. The speech act possesses a dialogic structure. Habermas wants to furnish a formalization of dialogic dynamics; indeed, he maintains that like language (*die Sprache*) discourse too (*die Rede*) is accessible to formal analysis.⁴⁰ If the illocutionary level indicates the space of the stance of the speaking individual, it appears legitimate to clarify the configuration of this illocutionary space. Let us therefore ask ourselves: in relation to what do speaker and listener commit themselves? What delimits the space of the validity of discourses? The speaker certainly takes up a stance in relation to his or her own experience. But, as we have seen, human experience results from a set of regulated activities that organize and give shape to different spheres: the world, society, the identity of the subject. These three spheres delimit the space of action of language, and are co-implicated in the human activity of coming to a shared understanding of something. With respect to this multiple horizon the speech act makes a claim, in a way that is explicit or implicit (but always capable of being made explicit),⁴¹ to be a valid and sensible move and for this reason acceptable by interlocutors. On the basis of these premises, Habermas believes that in the illocutionary component of the speech act, with which the speaker commits himself or herself to saying what he or she says, a fourfold⁴² validity claim that imitates the articulations of human experience is expressed.

According to this fourfold articulation it is possible to distinguish in the illocutionary dimension of the speech act the four following validity claims:⁴³ claim to truth, claim to normative rightness, claim to truthfulness and claim to intelligibility. Habermas states: «Dieses [Einverständnis] beruht aber nicht nur auf der intersubjektiver Anerkennung eines einzigen, thematisch hervorgehobenen Geltungsanspruchs. Vielmehr wird ein solches Einverständnis gleichzeitig auf drei Ebenen erzielt. Diese lassen sich intuitiv leicht identifizieren, wenn man bedenkt, dass ein Sprecher in kommunikativen Handeln einen verständlichen sprachlichen Ausdruck nur wählt, um sich *mit* einem Hörer *über* etwas zu verständigen und dabei *sich selbst* verständlich zu machen. Es liegt in den kommunikativen Absicht des Sprechers, (a) eine im Hinblick auf den gegebenen normativen Kontext *richtige* Sprechhandlung zu vollziehen, damit eine als legitim anerkannte interpersonale Beziehung zwischen ihm und dem Hörer zustande kommt; (b) eine

³⁹ See Habermas (1988, Chap. 5).

⁴⁰ See VE (1984, 339).

⁴¹ On the massive presence of the implicit in communication see Sbisà (2007).

⁴² Sometimes Habermas oscillates between three and four validity claims; I find most convincing the list of the four claims contained in VE (1984), for the reasons that I will explain.

⁴³ On the extendibility of the validity claims see Leist (1977, 106), Cooke (1997, 80–81).

wahre Aussage [...] zu machen, damit der Hörer das Wissen des Sprechers übernimmt und teilt; und (c) Meinungen, Absichten, Gefühle, Wünsche usw. wahrhaftig zu äussern, damit der Hörer dem Gesagten Glaube schenkt».⁴⁴ Before a detailed analysis is undertaken, a clarification is needed. The validity claims identified by Habermas in “Was heisst Universalpragmatik?” are four, though in the quotation from *Theorie des kommunikativen Handelns* he speaks of three planes of understanding: world (claim to truth), social relationship (claim to normative rightness) and intention of the speaker (claim to truthfulness). There is no contradiction, in that the fourth claim, the claim to intelligibility, as will be seen below, concerns precisely language, which constitutes the ground of the utterance activity, delimited precisely by world, society and itself. Let us analyze in detail the four validity claims:

- (a) The claim to truth (*Wahrheitsanspruch*) indicates the commitment of the speaker to justifying the rules of access to the world, present in all speech acts, but particularly in statements, descriptions, classifications, evaluations and previsions. Habermas rather insists on the pervasive role of the claim to truth, which does not constitute a prerogative of the assertive propositions but affects all utterance activity: «Wahrheit ist ein universaler Geltungsanspruch: seine Universalität spiegelt sich in der Doppelstruktur der Rede».⁴⁵ Every speech act claims to furnish a valid access to shareable experiences. Even an order, a threat, a witticism, a story or a metaphor always implies a coherent reference to reality and the notion of truth,⁴⁶ essential for interaction among speakers. Any opacity and uncertainty could interrupt the flow of the communicative relationship. This notion of truth has a double merit: on one side, it takes language back to the pragmatic and social dimension, and avoids the theoretical narrowness of referentialism, which identifies linguistic meaning with representation of the state of things; on the other hand, the pervasive character of the claim to truth prevents idealistic or relativistic drifts, since reference to reality remains an essential point of anchorage for speakers,⁴⁷ whatever they do with words. The dialogic character of the claim to truth implies the idea that knowledge of the conditions of truth is intersubjectively controllable. Truth does not concern the direct encounter between language and world, but is inscribed in the language-world-community triangulation.
- (b) The normative claim to rightness (*Richtigkeits-, Angemessenheits- Anspruch*) indicates the commitment of the speaker to acting within a context of shared rules. Habermas affirms: «Die Geltung eines normativen Hintergrundes von Institutionen, Rolle, soziokulturell eingewöhnten Lebensformen, d.h. von

⁴⁴ TKH (1981) (1988 edition), Band 1, 412–413.

⁴⁵ VE (1984, 420).

⁴⁶ Rather enlightening on the pervasive character of the rule of truth are the observations by Lo Piparo (1998), formulated in a context of Wittgensteinian analysis.

⁴⁷ The theme of the blend of gaze on the world and gaze to other speakers is amply dealt with in Habermas (1999).

Konventionen, wird immer schon vorausgesetzt».⁴⁸ The speech act establishes a relationship among participants in the communicative interaction, and in proposing it the speaker believes that this is a correct modality of relationship. Like the claim to truth, the claim to normative rightness too imposes on the speaker an obligation of justification and on the interlocutor the possibility of attributing or not attributing validity to the rule followed by the speaker in the specific utterance circumstance. Obviously, the speaker can act in compliance with common constraints, just as he or she can have reasons for proposing new and original moves; whether they are good or bad reasons depends on the dialogic dynamics activated by that utterance. The claim to normative rightness too has a pervasive character, but takes on a preponderant role in prayers, orders, exhortations, promises, agreements and apologies. But beyond these speech acts, every utterance activity establishes a relationship among speakers presupposing a normative context; and hence in the speech act there always exists an interplay between validity claims put forward by the speaker and the validity of the normative background presupposed.

- (c) With the claim to truthfulness (*Wahrhaftigkeitsanspruch*) the speaker attributes to himself or herself the responsibility of the linguistic initiative belonging to him or her with all the consequences involved. More simply, the claim to truthfulness indicates that the speaker intends to say what he or she says. The “proof” of truthfulness can be derived by the interlocutor, observing the continuity and consistency between the acting and saying of the speaker. This validity claim presupposes the possibility of the change of perspective connected to the *I-you* relationship; in this connection, taking on a role implies performing it in the first person in the presence of a *you*, with the awareness that the interlocutor is able to grasp the commitments connected to the role of the utterer.
- (d) The claim to intelligibility (*Verständlichkeitsanspruch*) indicates the commitment of the speaker to the clarity of his or her saying and is founded on competence in the historical-natural language. In this case the speaker assumes as a guarantee of the intelligibility of the speech act grammar and the good working of the language. Intelligibility too can be challenged by the interlocutor, if opacity requires an interpretative dialogue that thematizes the grammar of the language.

The fourfold structure of validity claims doesn't mean that inside every speech act there is an uniform distribution, as if every validity claim were entitled to a quota equal to 25 %. The four validity claims indicate, rather, a logical space occupied by a potential of reasons linkable to one another in various ways. In this logical space the speaker is offered an incalculable possibility of action and invention that establishes between the shores of experience a fertile combinational

⁴⁸ VE (1984, 422).

game, which variously selects the validity bases of the discourse. The incalculable plurality of ways of orienting linguistic action is enacted within a vast space but one delimited by constraints and bonds able to give sensibleness and validity to utterances.

The rooting of validity claims in the speech act disregards the good or bad faith of the speaker and his or her psychological constitution. The speaker may be more or less available for dialogue, may be more or less aware of the reasons for saying something, may have honest intentions, or may have intentions of disguise and manipulation. And in actual fact even the speaker intending to disguise, cheat, deceive or manipulate has to have a very clear awareness of the system of validity claims, precisely in order to pass off as true, correct and sincere what is possibly not. Moreover, such a person, if he or she were called on to express the reasons for saying what he or she says, would have to take care in indicating the sources that could accredit what is said. But even if a claim, for instance the truth claim, were eluded, this would not mean it would cease to play a role in the definition of meaning. For instance, if a speaker said “today it is raining” on a splendid sunny day, this would not mean it would immediately have to be set outside an effective communicative space. The logic of that move might be to express internal uneasiness; in this case the truth claim is not eluded or cancelled out. Indeed, the speaker knows full well that the sun is shining and accesses an experience which is objectively shareable; but he or she highlights his or her own point of view in expressive terms, precisely stressing the contrast between his or her emotional situation and that objective evidence. And he or she probably speaks in these terms precisely because he or she is appealing to the solidarity of a friend.

The system of validity claims shows the radical willingness of the speakers to enter into a dialogue. Even in the incalculable variety of speakers’ moves, what speech acts have in common is the typology of the potential of reasons expressible in the dialogue space. It is precisely the linguistic infrastructure that creates the dynamics of validation through which the social fabric is laboriously woven. Learning to speak means knowing where to track down and how to use the reasons that make what is said valid and sensible. The pragmatic rules that the speaker has to master, in order for what he or she says to be effective, simultaneously operate at the level of cognitive, social and linguistic skills. In order to speak sensibly the speaker has to know how to anchor what he or she says to the world (truth claim), he or she has to know how to insert what he or she says in the context of the social constraints and bonds (claim to normative rightness), he or she has to have grammatical competence in the historical-natural language (claim to intelligibility), and has to know how to play in the intersubjective space the role of *I* in relation to *you* (claim to truthfulness). The guarantees that make the speech act valid are furnished each time by common experience of the world, by a common normative background, by common competence in the language, and by the common possibility for the speakers of playing the role of utterer/listener. The speaker draws on these sources when he or she formulates an utterance and when he or she is called on by the interlocutor to account for its meaning. These are the

banks that delimit the space of the linguistic understanding/agreement, which precisely requires an incessant, sometimes pleasant and intriguing, sometimes fatiguing and frustrating, game of tuning into the other's wavelength.

And this game goes on, as long as the interlocutor recognizes the validity of the speaker's move. If the interlocutor accepts the move, he or she can set the ball rolling again with another utterance. But the game may stop, if the interlocutor does not perceive the sensibleness or the legitimacy of what the speaker says. The speech act may appear unacceptable from one of the following points of view:

- it may prove not to adhere very closely to experience of reality, and be incapable of satisfying the claim to truth. And in that case the listener can ask the reason for the violation of common experience of facts;
- it may violate the constraints and the bonds of the community, proving incapable of satisfying the claim to normative rightness. In this case too the interlocutor can set going a dialogic solution to the discrepancy between what we would expect and the speaker's actual stance;
- it may prove opaque and confused and therefore incapable of satisfying the claim to intelligibility. Therefore, an explanation is required to disambiguate the grammar of the text;
- finally, it may violate the practice of speaking in the first person. In this case the interlocutor can ask how the speaker intends to take on himself or herself the responsibility for what he or she says, inviting him or her to play the role of *I* correctly.

In all cases, discussion of the validity of the speech is intended to render explicit the reasons presupposed by the speaker to attribute meaning to what he or she says.

With this presentation of universal pragmatics we have focused on the rooting of the dynamics of validation on which language is based. The linguistic activity of speakers is not limited to blindly and mechanically following the rules of signification already given. If it were so, the validity of linguistic meaning would merely be factual validity; in this sense, the rules and constraints of signification, in shared use, would end up imposing themselves imperatively on speakers. Instead, the space of linguistic intersubjectivity, polarized both on the side of generality, and on the side of singularity, does not deny but indeed increases the potential of signification of which individuals are bearers. And the boundless richness of the activity of signification, always able to trace out new itineraries, is accompanied by the game of mutual validation with which speakers appraise the reasons for their speaking, questioning themselves on the truth, correctness, intelligibility and truthfulness of their speech acts.

That is the reason why, according to Habermas, there is an inseparable bond between meaning and acceptability, when he affirms: «Ich möchte deshalb die These vertreten, dass die illokutive Kraft, mit der der Sprecher in Ausführung seines Sprechakts auf den Hörer einwirkt, nur zu verstehen ist, wenn man über einzelne Sprechakte hinaus die Ja- und Nein-Stellungnahmen des Hörers zu den vom Sprecher mindestens implizit erhobenen Geltungsansprüchen in die Analyse

einbezieht».⁴⁹ And he takes it up again in TKH: «*Wir verstehen einen Sprechakt, wenn wir wissen, was ihn akzeptabel macht. [...]. Ein Sprechakt soll dann ‘akzeptabel’ heißen dürfen, wenn er die Bedingungen erfüllt, die notwendig sind, damit ein Hörer zu dem vom Sprecher erhobenen Anspruch mit> Ja <Stellung nehmen kann».*⁵⁰ In the light of these peremptory affirmations we assume that the listener really understands the meaning of the speech act if he or she identifies what can make that utterance valid. Only at this point is the listener able to accept the meaning of that move; and this is an indispensable condition for the game to be able to continue. But understanding the meaning of a speech act is also to identify what can make an utterance invalid; in this case, the interlocutor dissenting from the validity of an utterance is inclined to ask the reasons for that move, so that the game can continue. A constitutive trait of communicative competence is the awareness that saying elicits *yes* and *no* replies.

The connection between understanding of meaning and acceptability shows a connection between semantics and pragmatics which is much closer than Wittgenstein showed. The identification between meaning and use does not exhaust all aspects of linguistic life. As I have maintained elsewhere: «Meaning, already constituted in grammar or predetermined by use, would make utterance activity like a mere exchange of *fiches*, whose value is already established once and for all and for whose value/meaning the speaker has no responsibility. In the Habermasian perspective, by contrast, the pragmatic dimension of the utterance commitment⁵¹ is not extrinsically added to the semantic dimension. Indeed, meaning takes form and shape in the initiative of the speaker who proposes a certain configuration of the world to the interlocutors, committing himself or herself to maintaining its validity. In the dialogic game of understanding, the activity of signification is not to be compared to the transmission of contents already structured, but to the cooperative game, whether peaceful or conflictual, in which people test out the reasons for the constraints and the bonds with which they try to give shape, coherence and meaning to the world».⁵² And then if the activity of signification systematically implies the selection of reasons with which we place ourselves in the linguistic, praxis and cognitive space, the taking up of a *yes/no* stance is part of the life of meaning, a crucial aspect of its effectiveness.

It seems to me that this is the important outcome of a pathway that overcomes the limits of referentialist semantics and avoids reducing meaning to the conditions of truth, and to repeating a naïve model of knowledge enmeshed to the “myth of the given”.⁵³ Habermas, instead, following Peirce,⁵⁴ highlights the role of the

⁴⁹ VE (1984, 432).

⁵⁰ TKH (1981) (1988 edition), Band 1, 400–401.

⁵¹ On the pragmatic component of commitment in the theory of meaning see Brandom (2000, Chap. 6). For a comparison between Habermas and Brandom see Giovagnoli (2003).

⁵² Pandolfo (2010, 134).

⁵³ For an efficacious criticism of the myth of the given see Sellars (1997).

⁵⁴ In order to know more about Peirce’s heritage in Habermas’ work see Habermas (1968).

community of speakers that dictates the rules through which men can share their experiences of the world. Objectivity is constituted in the dialogue between speakers.⁵⁵ Nevertheless, if the source of validity does not lie exclusively in the world, it is not even entirely identified with the specific game context, since relativist and contextualist outcomes would be risked. Habermas succeeds in breaking away from the ill-formulated alternative that continually causes semantic theories to oscillate between an objective and subjective pole.

If the criterion of the validity of meaning and linguistic meanings were to depend on the factual consensus of the community, every speaker would be trapped in his or her own context and only what the community in fact considers such could be judged valid. But this would significantly reduce the space for the critical performances of speakers, wiping out the possibility of challenging the rules of a certain context. In this connection, let us remember that a validity judgment also implies an evaluation of the quality of the rule. And, in fact, in the dialogic dynamics it is brought into play not only in what a speaker says but also in the way and in the rule followed to say what he or she says. The possibility of formulating this judgment is the condition for accessing the space of criticism where it is possible to establish forms of understanding/agreement founded on reasons. The space of linguistic intersubjectivity creates the conditions for factual validity stabilized in the social context to be discussed anew in the name of a different validity criterion.

In order to avoid validity criteria being stiffened in social practices, Habermas distinguishes between *de facto* communities and the so-called *ideal speech situation*.⁵⁶ The formulation of this notion is one of the most opaque aspects of Habermas' theory, which is understandably very controversial. Though it is not possible here to review the critical debate, I would like to confirm what results from it, and that is to say that it is not a well-defined notion or one whose theoretical status is very clear⁵⁷ (it is not an empirical notion and it is not transcendental). I believe, however, that we should save and re-launch the theoretical drive underlying this ill-formulated notion. And, then, rather than hypostasizing the ideal of a rational community that "doesn't exist", here Habermas could have maintained the reconstructive attitude tested out so far. In this sense it appears legitimate to consider linguistic intersubjectivity responsible for the critical potential allowing speakers to avoid being immersed and trapped in their cultural horizons.⁵⁸ If anything, it is necessary to render explicit the conditions that make critical exercise effectively possible, if it is attenuated. And, then, rather than statically describing the ideal speech situation, it is necessary to reconstruct in procedural terms the practices allowing speakers to rationally transform factual

⁵⁵ See VE (1984, 382).

⁵⁶ See *Wahrheitstheorien*, VE (1984, 127–183).

⁵⁷ Ambiguity is explicitly referred to by Dal Canton (2002, 20–21).

⁵⁸ Cunico (2006, 33) clearly explains Habermas' need to save the perspectivist point of view without falling into relativism.

understanding into motivated understanding. Speakers must be able to get back the possibility of participating in the social practices of constitution of meaning. This possibility is given in the space of linguistic intersubjectivity, as long as there is a residual tension between factual validity and validity claims. This is the matrix of human discursive activity, which frees the critical potential able to disapprove the stabilized criteria in order to create new forms of understanding/agreement differently motivated and justified with reasons that claim to appear good. Only on these conditions can speakers exert control over social constraints and bonds, possibly overcoming inadequate linguistic and conceptual systems currently in force within a community. It is indispensable for the speaker to be able to *reopen a discourse*⁵⁹ and to make validity claims.

5 Conclusions: From Theory of Language to Social Theory

We will conclude by examining the cognitive and ethical–political effects of interplaying theory of language and social theory that sets up the matrix of normative constraints and bonds in the space of linguistic intersubjectivity.

While atomism and functionalism respectively connect social action with the intentions and strategies of individuals and the system, the *Theorie des kommunikativen Handelns* is able to account not only for the power, competition and conflict, but also for the dialogic dynamics of sociality.

But now it is necessary to clarify the notion of *communicative action*. This notion, as is well known, has set up an intense debate that has mainly challenged the primacy of communicative action over strategic action. Certainly if Habermas merely had descriptive intents, it would be easy to attribute him an abstract and idealizing way of looking at reality. However, if the reconstructive intent of the logical space of sociality is considered, more interesting and fertile interpretative prospects open up.

In the space of linguistic intersubjectivity, speaking and acting subjects can move along a continuum of possibilities of action at whose extreme we find: (a) the typology of communicative action, characterized by full use of the resources of linguistic understanding (*action oriented toward understanding*); (b) the typology of strategic action, characterized by deactivation of communicative resources, with a view to attainment of objectives of intervention in the world (*action oriented toward success*). The dialogic extremity and the monologic one delimit the space of linguistic intersubjectivity that covers all the possibilities of action, from the most collaborative to the most competitive formes. In the same space the subject is offered the possibility of relating his or her action to other agents thanks to the linguistic resources of understanding or else to forego every form of coordination. Communicative action and strategic actions are not two parallel circuits of action

⁵⁹ See VE (1984, 177–178).

that reciprocally exclude one another. If it were so we would have falsifying categories, merely able to interpret societies in dichotomic terms, painting reality in black and white so as to entrust every society to its irremovable “angelic” or “diabolic” destiny. On the contrary, human action is not entirely describable either as action oriented toward understanding or as action oriented toward success. Two modalities interconnect, and different forms of society depend on the type of interconnection.

In this sense, from my point of view, communicative action must not be seen as the original mode of social action; it is one of the polarities of the space of linguistic intersubjectivity, where linguistic energies can absorb and re-determine human action, orienting it toward forms of understanding/agreement that bring into play the validity claims of saying, doing and knowing. Communicative action does not eliminate the teleological components, but can submit egocentric orientations of action to the structural limitations of a language shared intersubjectively.⁶⁰

Habermas therefore affirms: «Der Begriff des *kommunikativen* Handelns [...] bezieht sich auf die Interaktion von mindestens zwei sprach—und handlungsfähigen Subjekten, die (sei es mit verbalen oder extraverbalen Mitteln) eine interpersonale Beziehung eingehen. Die Akteure suchen eine Verständigung über die Handlungssituation, um ihre Handlungspläne und damit ihre Handlungen einvernehmlich zu koordinieren. [...] In diesem Handlungsmodell erhält die Sprache [...] einen prominente Stellenwert».⁶¹

It seems to me that seeing communicative action as one of the polarities of the logical space of linguistic intersubjectivity is a more realistic and more fertile approach. It is realistic because it does not naïvely affirm that this is the primary and most widespread modality; but it is also realistic because, on the other side, it does not cynically affirm selfish brutality as the only source of human action. In this connection, orientation toward understanding is a possibility intrinsic to linguistic nature. And then one cannot exhaustively describe societies without considering the weight that the linguistic dynamics of validation have on the constitution of their normative identity. The latter determine the dimensions of sociality that have to do with the qualities of and reasons for social living, without which societies could exist as organisms only integrated in virtue of coercive mechanisms. And, instead, communicative action is the resource in virtue of which people choose, appraise and reason on the validity of the constraints and the bonds around which they converge.

It is clear, at this point, in what sense the notion of communicative action is fraught with ethical and political implications. The fact is that it is an indispensable diagnostic and critical tool for appraising how much the ethical and political spaces of a society valorise the dialogic dynamics of understanding. It is not by chance that in his production after TKH Habermas works out a theory of

⁶⁰ See Habermas (1988, Chap. 4).

⁶¹ Habermas (1981) (1988 edition), Band 1, 128.

morality⁶² and a theory of law⁶³ that valorise the dimension of communicative action, as a reservoir on which to draw in order to reactivate in associated life the circuit of reasons, maintaining the correct tension between factual validity and validity claims.

References

- Apel, Karl Otto. 1989. Normative Begründung der “Kritischen Theorie” durch Rekurs auf lebensweltliche Sittlichkeit? Ein transzendentalpragmatisch orientierter Versuch, mit Habermas gegen Habermas zu denken. In *Zwischenbetrachtungen: Im Prozess der Aufklärung. J. Habermas zum 60. Geburtstag*, ed. Honneth, Mc Carty, Offe, Wellmer. Frankfurt am Main: Suhrkamp.
- Bar-Hillel, Yehoshua. 1971. Out of the pragmatic waste-basket. *Linguistic Inquiry* 2: 401–407.
- Bianchi, Claudia. 2003. *Pragmatica del linguaggio*. Laterza: Roma-Bari.
- Brandom, B. Robert. 2000. *Articulating reasons. An introduction to inferentialism*. Cambridge: Harvard University Press.
- Cooke, Maeve. 1997. *Language and reason. A study of Habermas’s pragmatics*. Cambridge: The MIT Press.
- Cunco, Gerardo. 2006. *Oltre Saturno. Dialoghi per il tempo plurale*. Reggio Emilia: Diabasis.
- Dal Canton, Ilaria. 2002. *Conflitto e comunicazione. Per una critica di Jürgen Habermas*. Milano: Mimesis.
- De Simone, Antonio. 2010. *Passaggio a Francoforte. Attraverso Habermas*. Perugia: Morlacchi.
- Dews, Peter (ed.). 1999. *Habermas. A critical reader*. Oxford: Blackwell.
- Ferrara, Alessandro. 1987. A critique of Habermas’s consensus theory of truth. *Philosophy and Social Criticism* XIII 1: 39–67.
- Giovagnoli, Raffaella. 2003. On normative pragmatics: A comparison between Brandom e Habermas. *Teorema* XXIII: 51–68.
- Habermas, Jürgen. 1968. *Erkenntnis und Interesse*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1976. *Zur Rekonstruktion des Historischen Materialismus*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1983. *Moralbewusstsein und kommunikatives Handeln*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1984. *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1985. *Der philosophische Diskurs der Moderne. Zwölf Vorlesungen*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1988. *Nachmetaphysisches Denken. Philosophische Aufsätze*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1991. *Erläuterung zur Diskursethik*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1992. *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1999. *Wahrheit und Rechtfertigung. Philosophische Aufsätze*. Frankfurt am Main: Suhrkamp.
- Habermas, Jürgen. 1981 (ed. 1988). *Theorie des kommunikativen Handelns*. Frankfurt am Main: Suhrkamp.

⁶² See Habermas (1983, 1991).

⁶³ See Habermas 1992.

- Hentig, von Hartmut. 1973. Konsenstheorie. Über die Schwierigkeit, gemeinsam nützlichen Wahrheiten näherzukommen. *Neue Sammlung* 13: 265–283.
- Honneth, Axel, Thomas Mc Carty, Claus Offe, Albrecht Wellmer (herausgegeben von). 1989. *Zwischenbetrachtungen: Im Prozess der Aufklärung. J. Habermas zum 60. Geburtstag*. Frankfurt am Main: Suhrkamp.
- Ilting, Karl Heinz. 1976. Geltung als Konsens? *Neue Hefte für Philosophie* X: 20–50.
- Kuhlmann, Wolfgang. 1985. *Reflexive Letztbegründung. Untersuchungen zur Transzendentalpragmatik*. Freiburg/München: Alber Verlag.
- Leist, Anton. 1977. Was, was heisst Universalpragmatik? *Germanistische Linguistik* 5–6: 79–112.
- Lo Piparo, Franco. 1998. Cosa accade quando capiamo una frase. La verità come regola generatrice di senso. *Sicilorum Gymnasium, Rassegna semestrale della Facoltà di Lettere e filosofia dell'Università di Catania*, nn. 1–2, gennaio-dicembre, Catania.
- Mead, Georg Herbert. 1934. *Mind, self and society*. Chicago: Chicago University Press.
- Outhwaite, William. 1994. *Habermas. A critical Introduction*. Cambridge: Polity Press.
- Pandolfo, Alessandra. 2010. *Le regole dell'intesa. Attraverso Habermas uno studio sulla normatività umana*. Pisa: ETS.
- Pedroni, Virginio. 1999. *Ragione e comunicazione. Pensiero e linguaggio nella filosofia di Karl-Otto Apel e Jürgen Habermas*. Milano: Guerini e Associati.
- Petruciani, Stefano. 2000. *Introduzione a Habermas*. Laterza: Roma-Bari.
- Rasmussen, David. 1990. *Reading Habermas*. Oxford: Basil Blackwell.
- Rorty, Richard. 1998. *Truth and progress. Philosophical papers*. Cambridge: Cambridge University Press.
- Sbisà, Marina. 2007. *Detto non detto. Le forme della comunicazione implicita*. Laterza: Roma-Bari.
- Schlieben Lange, Brigitte. 1975. *Linguistische Pragmatik*. Stuttgart: Kohlhammer.
- Searle, John R. 1969. *Speech acts. An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Sellars, Wilfrid. 1997. *Empiricism and the philosophy of mind*. Cambridge: President and Fellow of Harvard College.
- Swindal, James. 1999. *Reflection revisited: Jürgen Habermas's discursive theory of truth*. New York: Fordham University Press.
- Wellmer, Albrecht. 1989. Was ist eine pragmatische Bedeutungstheorie. Variationen über den Satz Wir verstehen einen Sprechakt, wenn wir wissen, was ihn akzeptabel macht. In *Zwischenbetrachtungen: Im Prozess der Aufklärung. J. Habermas zum 60. Geburtstag*, ed. Honneth, Mc Carty, Offe, Wellmer. Frankfurt am Main: Suhrkamp.
- Wittgenstein, Ludwig. 1953. *Philosophische Untersuchungen*. Oxford: Basil Blackwell.

For in Psychology there are Experimental Methods and *Conceptual Confusion*: From Embodied Cognition to Wittgenstein on Language and Mind

Felice Cimatti

Abstract Embodied Cognition (EC) is a new psychological version of an old philosophical idea: human cognition is grounded in sensorimotor experience. According to EC there is not such an entity as abstract and disembodied knowledge, that is, the root of every form of human knowledge is an acting body in the world. In this chapter I will try to show that existing extensions of EC to language partly miss the point because do not fully account for the social and performative nature of language. Therefore a thorough embodied theory of language requires to consider the Wittgenstein legacy, which stresses at least two main points: (a) a coherent theory of language is not possible if not embedded in a more comprehensive description of human way of living; (b) the meaning of a word is not an internal and psychological entity but its social use, it is the *action* we do using that word/tool. In this chapter I will analyze EC literature showing that it needs to be complemented with Wittgenstein ideas on language and mind.

Keywords Ludwig Wittgenstein · Embodied cognition · Language · Words as tools · Sociality · Linguistic meaning · *Homo sapiens*

1 Wittgenstein is Back

It is curious how contemporary analytic philosophy forgot Wittgenstein's ideas on language and mind (Tripodi 2009). That could have been quite reasonable when cognitive sciences seemed to be ready to solve some ancient philosophical

Wittgenstein, 1953, Eng. Transl. 2001, II, p. 197.

F. Cimatti (✉)
Università della Calabria, Cosenza, Italy
e-mail: felice.cimatti@gmail.com

problems, as the mind–body one: cognitive sciences are based on the radically anti-Wittgensteinian idea of “mentalese”, that is, an internal private language. From this premise no dialogue was possible between cognitive sciences and the philosopher who holds that no language exists that is not social. But now a ‘new’ psychology is raising,¹ the so called Embodied Cognition, which holds views somewhat similar to Wittgenstein’s ones. In particular the idea that cognition does not exist which is not context-bound, and that language is a form of action (Gallese 2009). But there are still a lot of theoretical points that Embodied Cognition shares with cognitive sciences: in particular, methodological individualism, and dualism of mind and body (or semantics and pragmatics). In this chapter I should want to show that in order to overcome these and other difficulties it is necessary to come back to Wittgenstein.

2 Embodied Cognition

The basic idea of EC (Garbarini and Adenzato 2004) is (a), that «cognition is inherently perceptual, sharing systems with perception at both the cognitive and the neural levels», and (b) that there are not «amodal» forms of cognitive representations, that is representations that are «inherently nonperceptual» (Barsalou 1999, p. 577). EC theory is based on a particular kind of cognitive entity, «perceptual symbol»: «subsets of perceptual states in sensory-motor systems are extracted and stored in long-term memory to function as symbols. As a result, the internal structure of these symbols is modal, and they are analogically related to the perceptual states that produced them» (*ibid.*, p. 578). «Perceptual symbols» are the building blocks of the whole human cognitive system.

According to EC theory the cognitive process begins with a «perceptual state» which contains «two components: an unconscious neural representation of physical input, and an optional conscious experience». However the notion of «neural representation» does not solve the traditional problem that every EC theory encounters, because there are many «neural representations» how many different brains. The usual answer to this problem is that «a perceptual symbol contains only a schematic aspect» of what represents (*ibid.*, p. 583); this solution does not solve the *logical* problem posed by the individuality of each «perceptual symbol» which is a «record of the neural states that underlie perception» (*ibid.*, p. 482). But this is just an assumption, because it does not explain *how* such a schematic perception can occur.

Let’s consider the linguistic version of this problem: when someone listens to a word, for example the word “apple”, surely a peculiar «neural representation» will correspond to this «physical input». The problem is that if there exist n brains there will exist n different meanings corresponding to the very same «physical input». In

¹ In fact Embodied Cognition is a form of neo-Piagetianism which ignores its own origin.

this case, how can language comprehension be possible? There are not n different meanings of the English word “apple”. Language comprehension is only possible if all speakers of a language use words according to the same rule.

EC theory simply does not consider the *logical* problem posed by the contrast between the individuality of «perceptual symbols» and the universality of language use; «like a perceptual symbol, a linguistic symbol is a schematic memory of a perceived event, where the perceived event is a spoken or a written word. A linguistic symbol is *not* an amodal symbol, nor does an amodal symbol *ever* develop in conjunction with it. Instead, a linguistic symbol develops just like a perceptual symbol» (*ibid.*, p. 592). What EC theory wants to rule out is the existence of amodal symbol, that is the very existence of a nonperceptual based form of cognition. But a cognitive entity only which works as an amodal entity can solve the problems we just raised. From this point of view the main problem that EC theory has to face is the nature of language: EC theory considers language as a means for expressing internal concepts: «language comprehension can be viewed as the construction of a perceptual simulation to represent the meaning of an utterance or text» (*ibid.*, p. 605). Linguistic meaning is an internal «perceptual simulation». This is a very traditional view, language is an expressive tool which convey «deep conceptual information» (Barsalou et al. 2008, p. 251). Therefore linguistic entities are mainly mere vehicles of conceptual information; in this perspective «symbolic operations» are only possible if internal «simulations» are re-activated: «attempting to perform symbolic operations on linguistic forms alone would be like manipulating symbols in an unfamiliar language, with no true comprehension» (*ibid.*).

It is quite strange that EC does not consider language as a peculiar form of bodily action *by itself*, a view that should be very sympathetic with the general assumptions of EC theory. In the following parts of this chapter I will try (a) to show how Wittgenstein legacy could deeply improve EC theory on language and mind, (b) to integrate Wittgenstein views with current EC research on language.

3 EC and Classical Cognitive Sciences

According to classical cognitive sciences a specific linguistic module exists which only processes linguistic input and output. On the contrary, a very important EC evidence on language is that language processing recruits cerebral *motor* system (Glenberg and Kaschak 2003; Scorilli and Borghi 2007; Sato et al. 2008). For example, when I hear a sentence as *Mary kicks the ball* the motor system of my brain, in particular that which controls the foot, is mainly activated. Action verbs used by the utterer are literally understood by the listener through the mediation of his/her foot. According to EC this means that there is no a strong separation between a cognitive module (syntax and semantics) and an executive module (pragmatics); language processing is a form of (more or less mediated) *bodily action* (Rizzolatti and Arbib 1998; Rizzolatti and Sinigaglia 2006).

This is a very important innovation compared to disembodied Cartesian view of mind typical of classical cognitive sciences. On the contrary EC stresses the connection between mind and body; there is not such a thing as a mind operating *in vacuum*. Mind is always situated, in a specific body and in a spatial–temporal context. Nevertheless EC still belongs to cognitive sciences, therefore it shares with them their basic characteristics: cognitive individualism and a persistent and even if concealed form of dualism. Let's take the first point, *cognitive individualism*. We just observed that when I hear a sentence like *Mary kicks the ball* the motor system of my own brain, in particular that which controls my foot, is mainly activated. It is not too incorrect to maintain that according to EC I understand this sentence with my foot. The point that EC theorists want to stress is clear and fair but a *very* difficult problem comes up with such a solution: my own brain is different from yours, then how can we understand each other? And what about if I am born legless? Linguistic *meaning* is the same for all speakers of a language, while individual representations are always different from those of other individuals. EC model doesn't seem to be able to cope such a problem, which is a logical problem not a psychological one.

As for dualism, EC holds that putting the notion of *action* at the very center of its own theory eliminates it, because there could not be any more a mind separated by the body. Actually dualism survives in concealed forms: for example it survives in the distinction between semantics (mind) and pragmatics (body), or between abstract and concrete concepts. The very notion of *concept* as a separate mental entity should be quite suspect for a coherent and consequent EC theorist. Where are concepts? The usual answer is: they are in the mind. It is not such a big difference if concepts are innate or acquired, the point is the very existence of a special class of *mental* entities. EC neither solves the two main problems of cognitive sciences, individualism and dualism.

4 The «Natural History» of Rule and Meaning

Wittgenstein, in order to solve the *logical* problem of cognitive individualism, moves away from minds (together with its own private representations) to «language-games». This is not a simple terminological change. Wittgenstein wants us to think of language as a natural behavior rather than a set of explicit rules we have to learn. We have not to teach a child to play: children play, playing is part of our *biological* nature:

It is sometimes said that animals do not talk because they lack the mental capacity. And this means: “they do not think, and that is why they do not talk”. But—they simply do not talk. Or to put it better: they do not use language—if we except the most primitive forms of language.—Commanding, questioning, storytelling, chatting, are as much a part of our natural history as walking, eating, drinking, playing (Wittgenstein 1953, Eng. Transl. 2001, I, Sect. 25).

A «language game» is not the application of an internal mental rule. The basic idea of this notion is that language is not an external instrument we can or cannot use. We should consider «language games» as peculiar human behaviors, much like walking or breathing. We do not think to how to breathe; the same holds for «language-games». We ‘use’ language as we ‘use’ our own foot for walking: there is the same intimacy relation between ‘us’ and our body as between ‘our’ minds and the language we speak. We do not breathe because it is the best way to oxygenate our lungs: we breathe because the body we are breathe. Something similar holds for language: we do not speak because this is the best way for conveying informations, we speak because... we speak, because speaking it is our nature. In this sense as playing is self-rewarding, so language use is self rewarding.

«Language-games» notion forces us to completely change the way we used to think of language: much more as a physiological behavior than an acquired set of explicit conventions. The constitutive individualism of cognitive sciences is the first victim of this approach: children can immediately begin playing because a mutual relationship *already* exists between them. There is no logical need of deciding that they are playing: *they play*, that’s all. First is the playing, the «language game», then the individual players, the individual speakers. Language is not a way for communicating human’s private thoughts, language is the common field between them:

What we are supplying are really remarks on the natural history of human beings; we are not contributing curiosities, however, but observations which no one has doubted, but which have escaped remark only because they are always before our eyes (*ibid.*, I, Sect. 415).

It is not surprising that the discovery of mirror neurons confirmed Wittgenstein’s philosophical analysis: a mirror neuron is a neuron which discharges both when the body whose it is part executes a certain action and the same body sees another body executing the very same action (Rizzolatti et al. 1996). Mirror neurons are the physiological bedrock of language (Fogassi and Ferrari 2007; Schilhab 2007; Corballis 2009), they ensure the basic «intercorporeity» (Gallese 2009, p. 493) which allows the very possibility of mutual comprehension. In «language games» conventions between speakers are of no use, because the departure point, the presumed internal states, are not shared:

how do words *refer* to sensations? - There doesn’t seem to be any problem here; don’t we talk about sensations every day, and give them names? But how is the connexion between the name and the thing named set up? This question is the same as: how does a human being learn the meaning of the names of sensations? - of the word “pain” for example. Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour. “So you are saying that the world ‘pain’ really means crying?” - On the contrary: the verbal expression of pain replaces crying and does not describe it (Wittgenstein, 1953, Eng. Transl. 2001, I, Sect. 244).

The fundamental «language-game» of the expression of internal states *pre-supposes*, in order to get started, that a natural way exists in which human beings

live the experience of pain. An adult does not need to teach a child what to do when he bangs his head against the edge of a table. Without this spontaneous behavior, the entire “language-game” of the expression of pain simply could not begin. It is also necessary that this behavior be, somehow or other, similar in different humans, in the sense that, faced with a person who is crying in pain, for example, human beings *react* in a manner which is reciprocally recognizable:

look at a stone and imagine it having sensations.—One says to oneself: How could one so much as get the idea of ascribing a *sensation* to a *thing*? One might as well ascribe it to a number!—And now look at a wriggling fly and at once these difficulties vanish and pain seems able to get a foothold here, where before everything was, so to speak, too smooth for it. And so, too, a corpse seems to us quite inaccessible to pain.—Our attitude to what is alive and to what is dead, is not the same. All our reactions are different. [...] (*ibid.*, I, Sect. 284).

Wittgenstein writes «reactions» with regard to the pain of others and this is a highly important observation (which, besides, anticipates by many decades the discovery of mirror neurons) A reaction is not learned. It is a form of behaviour specific to the human species: “think of the recognition of *facial expressions*. Or of the description of facial expressions—which does not consist in giving the measurements of the face! Think, too, how one can imitate a man’s face without seeing one’s own in a mirror” (*ibid.*, I, Sect. 285). So, without this capacity, which is naturally shared among (normal) members of the *Homo sapiens* species, the «language-game» of the expression of internal states would not be possible:

..now, what about the language which describes my inner experience and which only I myself can understand? *How* does I use words to stand for my sensations?—As we ordinarily do? Then are my words for sensations tied up with my natural expressions of sensations? In that case my language is not a ‘private’ one. Someone else might understand it as well as I.—But suppose I didn’t have any natural expression for the sensation, but only had the sensation? And now I simply *associate* names with sensations and use these names in descriptions.—(*ibid.*, I, Sect. 256).

Let us imagine this case. A child trips and falls to the ground, and feels a certain internal sensation, without, however, this being accompanied by any natural expression. How is it possible for this child to learn to use the linguistic expression, by means of which, in his community, reference is made to that internal state? How is it literally possible for the «language-game» of the expression of internal states to begin? An adult witnesses the episode, and asks him *Does it hurt?* How will the child understand *what* the adult is talking about? It would be like wanting to teach a cat who is licking a paw hurt in a fight with another cat that what it is feeling in that paw is called, in Germany for example *Schmerz*. How can the child associate the word *Schmerz* with something that he does not even now how to express (notwithstanding it is possible to feel something that one is in no manner capable of expressing). Yet, and this is even more important, the child would not even understand the *why* of this operation. If it does not come naturally to him to express pain, why should he do it in an artificial manner? What would be the sense of this game? The relation with the other, therefore, does not come about

by means of the «language-game», rather, this presupposes a natural relation, not learned and not explicit:

“what would it be like if human beings shewed no outward signs of pain (did not groan, grimace, etc.)? Then it would be impossible to teach a child the use of the word ‘tooth-ache’.”—Well, let’s assume the child is a genius and himself invents a name for the sensation!—But then, of course, he couldn’t make himself understood when he used the word.—So does he understand the name, without being able to explain its meaning to anyone?—But what does it mean to say that he has ‘named his pain’?—How has he done this naming of pain?! And whatever he did, what was its purpose?—When one says “He gave a name to a sensation” one forgets that a great deal of stage-setting in the language is presupposed if the mere act of naming is to make sense. And when we speak of someone’s having given a name to pain, what is presupposed is the existence of the grammar of the word “pain”; it shewes the post were the new word is stationed (*ibid.*, 1, Sect. 257).

The «language-game» is not based upon an explicit convention established among its participants because a convention requires the presupposition of something that is not in dispute, and which all *preliminarily accept*. Otherwise, the discussion could not even begin. Let us suppose that we wish to establish the rule that when pain is felt the English expression *pain* is used. At the same time, it is necessary that everyone knows how to recognize the spontaneous expression of pain on the part of others. Without this natural capacity, it is impossible to be certain that when someone uses the expression *pain*, he is using it in the same manner in which others could use it; for this reason, «the expression of doubt has no place in the language-game; but if we cut out human behaviour, which is the expression of sensation, it looks as if I *might* legitimately begin to doubt afresh. My temptation to say that one might take a sensation for something other than what it is arises from this: if I assume the abrogation of the normal language-game with the expression of sensation, I need a criterion of identity for the sensation; and then the possibility of error also exists» (*ibid.*, 1, Sect. 288).

Wittgenstein naturalistic stance helps us to find a way out from the two problems EC inherits from cognitive sciences, individualism and dualism. The basic notion for understanding human language is «language-game», a behavior which is part of human «natural history» and which is naturally social. If we want to understand human language we have to look for in ethology more than in psychology.

5 What is a Word?

The same logical relation holds between an hammer and a nail I want to stick into a wall, and the utterance *I love you* said to the woman I want to marry. In both cases what I have to do requires the necessary mediation of a tool, a physical one in the first case, a linguistic one in the second case. A tool that I have to use according to the social norms that regulate it: as I have to grab the hammer by its wooden handle, so I have to use the linguistic tool *I love you* respecting its use

rules and appropriate contexts of use. A tool is not an entity whose use could be arbitrarily established by anyone. A tool is a *normative* entity: « think of the tools in a tool-box: there is a hammer, pliers, a saw, a screw-driver, a rule, a glue-pot, nails and screws.—The functions of words are as diverse as the functions of these objects. (And in both cases there are similarities.)» (*ibid.*, I, Sect. 11).

This not so much a new definition of *linguistic meaning* as a completely different way to conceive language as a way of living: according to Wittgenstein language is neither a cognitive instrument (a way of thinking) nor a means for communicating (a way to express thoughts): it is the peculiar way of *living* of human beings. As birds fly and fishes swim we talk each other. It is still a very unusual way of defining language, but a way that should be very liked by EC theorists. Wittgenstein definition is an *anthropological* one rather than a linguistic or cognitive one: «I shall [...] call the whole, consisting of language and the actions into which it is woven, a “language-game”» (*ibid.*, I, Sect. 7). There is no way, in Wittgenstein analysis, of dividing language from what humans do with it. In this sense language is the peculiarly human way of acting in the world: «here the term «language-game» is meant to bring into prominence the fact that the *speaking* of language is part of an activity, or of a life-form» (*ibid.*, I, Sect. 23). The traditional view separates language as set of expressions from the set of things or thoughts these expressions denote. After Wittgenstein such a separation does not hold any more and the implicit dualism of semantics (mind) and pragmatics (body) fails. For this reason defining language as a means of communicating is so misleading, because it is an unaware reproduction of classical dualism.

The traditional view holds that for each word a corresponding thing or thought has to exist, otherwise the word would be meaningless: as one of the main EC theorists asserts «linguistic system» does not «contain [...] its own semantics» (Barsalou et al. 2008, p. 250). This is, as we have previously seen, the explicit EC model of language. But such a model simply does not apply to actual uses of language:

let us first discuss *this* point of the argument: that a word has no meaning if nothing corresponds to it.— It is important to note that the word “meaning” is being used illicitly if it is used to signify the thing that ‘corresponds’ to the word. That is to confound the meaning of a name with the *bearer* of the name. When Mr. N.N. dies one says that the bearer of the name dies, not that the meaning dies. And it would be nonsensical to say that, for if the name ceased to have meaning it would make no sense to say “Mr. N.N. is dead” (Wittgenstein, 1953, Eng. Transl. 2001, I, Sect. 40).

A word simply it is not significant because a thing it is attached to it; the dualistic relation of reference does not explain the meaning of a word. The whole semantic value of a linguistic entity coincides with its *anthropological* value, that is, its use into a community: «every sign *by itself* seems dead. *What* gives it life?— In use it is *alive*. Is life breathed into it there?—Or is the *use* its life?» (*ibid.* I, Sect. 432). Wittgenstein proposes a radical de-psychologization of the notion of meaning, that is, he proposes to consider meaning as not separable by what we do when we use language in everyday life: «for there isn’t anything hidden—don’t we see the whole sentence? The function must come out in operating with the word.

[(Meaning-body.)]» (*ibid.* I, Sect. 559). There is no more meaning on one side (the mental or psychological one), and use on the other, and other similar couples as signal and content, inner and outer, semantics and pragmatics: there is *only* «meaning-body», that is, *language in action*. As we have previously seen this is a view of language that is more similar to an ethological than a to psychological one. From this premise Wittgenstein arrives at a strong conclusion: «for a large class of cases—though not for all—in which we employ the word “meaning” it can be defined thus: the meaning of a word is its use in the language. And the *meaning* of a name is sometimes explained by pointing to its *bearer*» (*ibid.* I, Sect. 43). This is not a new conception of semantics rather the elimination of the very notion of meaning as an autonomous entity: from now on there is no more place for the independent existence of semantics as a distinct component of human cognitive architecture. A view that should be very attractive to EC theorists, because stresses the importance of perception and mainly *action* in explaining language. But EC theorists define themselves as part of the large family of cognitive sciences, and they do not find them at easy in accepting a theory that somewhat eliminates the very existence of their professional field.

6 From EC to Music

Wittgenstein proposes a radical changing in our way of conceiving language. According to him language is neither a means of communicating nor a cognitive tool (Mirolli and Parisi 2009). It is obviously true that language is both of them but its very nature is neither the first nor the latter one. Wittgenstein tries to see to language with a fresh look, getting rid of the traditional categories through which we used to conceive it. In particular Wittgenstein doesn't want to see language through the glasses of communication. If we stop thinking to language as an expressive tool we are forced to find completely new categories to conceive it. EC is on the road of such a change, but its cognitivist inheritance prevents it to fully realize it.

Mirror neurons discovery pave the way for this completely different approach to the comprehension of language. In many non human animals mirror neurons system allows the establishing of mutual relations between them. The idea is that language is the human transformation of such a system which establishes a species-specific sharable space between us: language is based on a «intercorporeity» system, a «mandatory, pre-rational, non-introspectionist functional mechanism» which permits to each of us of establishing a relation with others, a common space that «is therefore not necessarily the result of a willed and conscious cognitive effort, aimed at interpreting the intentions hidden in the overt—and supposedly intentionally opaque—behavior of others, but rather a basic functional mechanism of our brain» (Gallese 2009, p. 493).

If language is not a means of communicating, the dualism of mind and body, semantics and pragmatics, content and vehicle vanishes. Words are no more

intended as mere signals rather as *gestures*, that is, a *unitary* entity (like the Saussure's sign). From this point of view the classical distinction between the so called concrete and abstract concepts also could be abandoned. According to the received view a concept is abstract whether it is not connected to some previous embodied experience; it is concrete in the other case. This is a distinction that is only acceptable if words are conceived as mere meaningless vehicles of some extrinsic semantic content. This is the usual EC's idea on language: «attempting to perform symbolic operations on linguistic forms alone would be like manipulating symbols in an unfamiliar language, with no true comprehension» (Barsalou et al. 2008, p. 251). This is the major theoretical limit of EC: by itself a word is just a *empty* envelope for conceptual meaning. But if *all* words are intended as bodily tools, as particular way of acting in the world, then there is no more theoretical reason to distinguish between abstract and concrete concepts. Each linguistic act is a meaningful gesture:

it is like looking into the cabin of a locomotive. We see handles all looking more or less alike. (Naturally, since they are all supposed to be handled.) But one is the handle of a crank which can be moved continuously (it regulates the opening of a valve); another is the handle of a switch, which has only two effective positions, it is either off or on; a third is the handle of a brake-lever, the harder one pulls on it, the harder it brakes; a fourth, the handle of a pump: it has an effect only so long as it is moved to and fro (Wittgenstein, 1953, Eng. Transl. 2001, I, Sect. 12).

As there is no intrinsic reason to distinguish between abstract and concrete tools, the same holds for gesture-words: there are tools that directly affects the world, and there are tools that require a longer path to reach their goal. In this perspective language becomes a form of social action: «to imagine a language means to imagine a life-form» (*ibid.*, I, Sect. 19). Understanding a language does not require referring to a particular psychology or semantics, rather *to live* in a certain way. This is not equivalent to sustain that we should prefer pragmatics over semantics, but that we should give up in separating human mind from its behavior. The very distinction between semantics and pragmatics it is nothing more that a semiotics variant of the classical metaphysical distinction between mind (soul) and body. Wittgenstein pushes us to come out from this tradition, when he compares the process of language comprehension to that of music comprehension:

understanding a sentence is much more akin to understanding a theme in music than one may think. What I mean is that understanding a sentence lies nearer than one thinks to what is ordinarily called understanding a musical theme. Why is just *this* the pattern of variation in loudness and tempo? One would like to say “Because I know what it's all about.” But what is it all about? I should not be able to say. In order to ‘explain’ I could only compare it with something else which has the same rhythm (I mean the same pattern). (One says “Don't you see, this is as if a conclusion were being drawn” or “This is as it were a parenthesis”, etc. How does one justify such comparisons?—There are very different kinds of justifications here.) (*ibid.*, I, Sect. 527).

References

- Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660.
- Lawrence W. Barsalou, Ava Santos, W. Kyle Simmons, Christine Wilson. 2008. Language and simulation in conceptual processing. In *Symbols, embodiment, and meaning*, eds. M. De Vega, A.M. Glenberg, A.C. Graesser, A., 245–283. Oxford: Oxford University Press.
- Corballis, Michael. 2009. Language as gesture. *Human Movement Science* 28: 556–565.
- Fogassi, Leonardo, and Francesco Ferrari. 2007. Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science* 16(3): 136–141.
- Gallese, Vittorio. 2009. Motor abstraction: A neuroscientific account of how action goals and intentions are mapped and understood. *Psychological Research* 73: 486–498.
- Garbarini, Francesca, and Mauro Adenzato. 2004. At the root of embodied cognition: cognitive science meets neuropsychology. *Brain and Cognition* 56: 100–106.
- Glenberg, Arthur, and Michael Kaschak. 2003. The body's contribution to language. *The Psychology of Learning and Motivation* 43: 93–126.
- Mirolli, Marco, and Domenico Parisi. 2009. Language as a cognitive tool. *Minds and Machines* 19: 517–528.
- Rizzolatti, Giacomo, L. Fadiga, Vittorio Gallese, and Leonardo Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131–141.
- Rizzolatti, Giacomo, and Michael Arbib. 1998. Language within our grasp. *Trends in Neuroscience* 21: 188–194.
- Rizzolatti, Giacomo, and Corrado Sinigaglia. 2006. *So quel che fai. Il cervello che agisce e i neuroni specchio*. Milano: Raffaello Cortina Editore.
- Sato, Marc, Marisa Mengarelli, Lucia Riggio, Vittorio Gallese, and Giovanni Buccino. 2008. Task related modulation of the motor system during language processing. *Brain and Language* 105: 83–90.
- Schilhab, Theresa. 2007. Interactional expertise through the looking glass: A peek at mirror neurons. *Studies in History and Philosophy of Science* 38: 741–747.
- Scorolli, Claudia, and Anna Borghi. 2007. Sentence comprehension and action: Effector specific modulation of the motor system. *Brain Research* 1130: 119–124.
- Tripodi, Paolo. 2009. *Dimenticare Wittgenstein. Una vicenda della filosofia analitica*. Bologna: Il Mulino.
- Ludwig, Wittgenstein. 1953. *Philosophical investigations*. The German text. With a revised English translation: 2001. (trans: Anscombe, G.E.M.). Oxford: Blackwell Publishing.