Alessandro Capone
Franco Lo Piparo
Marco Carapezza   *Editors*

Perspectives on Linguistic Pragmatics

Springer

Perspectives in Pragmatics, Philosophy & Psychology

Volume 2

Alessandro Capone · Franco Lo Piparo
Marco Carapezza
Editors

Perspectives on Linguistic Pragmatics

Springer

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

Franco Lo Piparo
Marco Carapezza
University of Palermo
Palermo
Italy

Printed on acid-free paper

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

Acknowledgments

Contents

Introduction

We are pleased to introduce the second volume of the series Pragmatics, Philosophy and Psychology, which explicitly deals with linguistic pragmatics. Unlike the first volume of the series, it mainly collects papers by professional linguists, who have been influenced by philosophy (in particular by pragmatic ideas) but work within linguistic departments and thus have integrated philosophical ideas based on pragmatics within well-consolidated paradigms of linguistic research. In this book we have some (critical) theoretical considerations on meaning (see Burton-Roberts' contribution) and its relation to implicatures and explicatures. Within linguistics, Grice's ideas were absorbed and critically developed by Sperber and Wilson, whose 1986 book has spurred much research in GB and across the world. I do not present this paradigm here because it is a popular theory but because it has the potential for explaining away many problematic features of language use—to mention one case where relevance-based ideas were very fruitful, consider the attributive/referential distinction by Donnellan, on which a number of authors have written extensively from the point of view of Relevance Theory. Relevance Theory is clearly interested in the speaker's point of view and in the Hearer's point of view (despite influential recent criticism). In fact, both the speaker and the Hearer cooperate in constructing meaning on the assumption that both the hearer's interpretation work and the speaker's codification/construction work are constrained by the Principle of Relevance, mainly the idea that Relevance is a function of cognitive rewards and an inverse function of cognitive efforts. An interpretation that guarantees rich cognitive effects and is accompanied by few cognitive efforts is certainly to be preferred. The chapter on Relevance Theory by Nicholas Allott magisterially introduces this important framework and allows readers to familiarize with technical details and linguistic terminology. The chapter by Alison Hall, in addition to introducing the readers to Relevance Theory, also fulfills the function of presenting one aspect of the theory, namely the reflections on the semantics/ pragmatics debate and the concept of (conversational) explicatures. One of the main ideas of (at least some) pragmalinguists today is that semantics is mainly underdetermined and that pragmatics serves to bridge the gap between underdetermined logical forms and full propositional content. Now, of course, there is a debate on how wide the gap is or should be and not all pragmaticians propend for the view that semantic logical forms are necessarily underdetermined (see the

important work by Predelli 2005). However, if there are cases of pragmatic
intrusion (and I myself, as well as many others have amply shown that there are
important cases where semantics underdetermines content and pragmatics needs to
intervene to provide full propositional forms), we need the notion of explicature
and we need to accept that occasionally or rather quite often pragmatics intervenes
to construct full propositional forms. Now what form should a pragmatic theory of
pragmatic intrusion take? Here as well there are numerous options—one of these is
the one advocated by Hall: pragmatic free enrichment plays a key role in
pragmatic intrusion. And I myself agree with this idea, provided that we also
accept that at least in some cases contextual considerations saturate elements of
meanings which are by their nature open slots which need to be filled (take the
case of pronominals or demonstratives).

How can we characterize explicatures? Are they cancellable or uncancellable
components of meaning? In my own paper, I argue that explicatures are NOT
cancellable. Of course, it could well be argued that that there are two types of
explicatures and that only one of them is cancellable. However, in this paper I
favor the view that the most clear cases of explicatures are those where the
inference cannot be cancelled.

The considerations by Relevance Theorists can and have to be integrated into a
chapter to be called 'Theory of Mind'; of course, on this issue, the community of
scholars is divided, as some assert while others deny that the human mind has a
module devoted to 'theory of mind'. While I myself have found this idea plausible,
I agree that the topic needs to be explored further. Anyway, Louise Cumming's
judicious considerations offered in this book will allow us to understand more
about his intriguing and hot topic.

Pragmatics is mainly a theory of (linguistic) use; thus it should not surprise us
that some authors approach it as such (see Gregoromichelaki and Kempson), for
example. The considerations by Yan Huang and Corazza on point of view are also
important because they show that the linguistic resources of human languages
encode point of view and such encodings are of importance to a theory of language
use and interpretation. In the remaining chapters the authors deal with reference
(Keith Allan), common ground (Allan), Presupposition (Kecskes and Zhang,
Mandy Simons, Geurts and Maier), definiteness (von Heusinger), pragmatic
inference (Mazzone), pragmemes, language games (Carapezza and Biancini), as
well as with other issues relating to language use. Of particular relevance to
linguistic theorizing is the idea by Mandy Simons that, at least in some cases,
presuppositions are of a pragmatic nature—not only in the sense that the
presupposing expression requires some contextual assumptions to be satisfied
(imposes constraints on context), but also in the sense that conversational
implicature is responsible for the triggering of the presupposition. While we are
still expecting more systematic work on this idea, I think that the importance of
this idea has not been noted enough in the literature and I hope that this volume
will at least serve to propagate it. The last paper in the collection by Carapezza and
Biancini discusses the relationship between language games and Jacob Mey's
notion of 'pragmeme.' The paper contains important considerations on language

use including a discussion of a recent word introduced into language use by the compulsive efforts of the media: 'Bunga Bunga.' While the paper is actually a discussion of some of Wittgenstein's ideas voiced in 'Philosophical Investigations,' it can anchor readers to modern theories such as the view that pragmemes are an essential part of language use—(some) linguistic expressions work on the basis of rich contextual assumptions without which they would be inefficacious.

Needless to say, pragmatics deals with the speaker's intentions, which need to be grasped by a hearer. The speaker must take the hearer into account when he codifies a message and must be able to predict which features of the context will be picked by the hearer in completing or enriching the message. The hearer must take into account the speaker's intentions as evinced by the linguistic materials of the utterance and by the arrangement of clues and cues available for inspection. Semantics and Pragmatics work in tandem (as Levinson 1983 said), however, there are places, which I myself called 'loci of pragmatic intrusion' where contextual information is needed to construct explicature. Very often indeed free enrichment is responsible for communicating the full propositional form of an utterance.

Alessandro Capone
Franco Lo Piparo
Marco Carapezza

References

Capone A. (2001). *Modal adverbs and discourse*. Pisa: ETS.
Levinson S. (1983). *Pragmatics*. Cambridge: CUP.
Predelli S. (2005). *Contexts*. Oxford: OUP.

Linguistic Pragmatics

Noel Burton-Roberts

The paper questions the assumption (widespread in semantic—and indeed pragmatic—theory) that linguistic expressions have meaning in virtue of possessing semantic properties/content. Problems created by this assumption are discussed and an alternative account of meaning is developed. Although Burton-Roberts talks informally of words 'having meaning', the argument is that meaning is not a property—and has in fact nothing to do with the intrinsic properties of the things we say 'have meaning'. Rather, meaning is a relation. More specifically, it is a semiotic relation. In the case of linguistic expressions it is, yet more specifically, a symbolic—not indexical or iconic—relation (in the sense of Peirce). The overarching intention of the paper is set linguistic meaning in the context of meaning in general.

Louise Cummings

Theory of mind (ToM) describes the cognitive ability to attribute mental states both to one's own mind and to the minds of others. In recent years, ToM has been credited with playing a significant role in developmental and acquired pragmatic disorders. In this way, ToM deficits have been linked to pragmatic deficits in individuals with autism spectrum disorders (e.g. Martin and McDonald 2004), emotional and behavioural disorders (e.g. Buitelaar et al. 1999), intellectual disability (e.g. Cornish et al. 2005), right-hemisphere damage (e.g. Winner et al. 1998), schizophrenia (e.g. Brüne and Bodenstein 2005), traumatic brain injury (e.g. McDonald and Flanagan 2004) and neurodegenerative disorders such as Alzheimer's disease (e.g. Cuerva et al. 2001). In 'Clinical Pragmatics and Theory of Mind', Louise Cummings examines the central role of ToM reasoning in utterance interpretation. The chapter addresses what is known about ToM development during childhood and adolescence as well as changes in ToM skills as part of the aging process. The role of ToM in developmental and acquired pragmatic disorders is discussed. The contribution of ToM research into pragmatic

disorders is critically evaluated. Finally, several ToM theories are examined. The question is addressed of which, if any, of these theories is able to capture the pragmatic features of utterance interpretation.

Nicholas Allott

Relevance theory is a wide-ranging framework for the study of cognition, proposed (Sperber and Wilson 1986; 1987) primarily in order to provide an account of communication that is psychologically realistic and empirically plausible. This paper (i) presents relevance theory's central commitments in detail and explains the theoretical motivations behind them and (ii) shows some of the ways in which these core principles are brought to bear on empirical problems. The core of relevance theory can be divided into two sets of assumptions. Assumptions relating to cognition in general include the definition of relevance as a trade-off between effort and effects; the cognitive principle of relevance, i.e., the claim that cognition tends to maximise relevance and the view that human beings possess a 'deductive device' playing a central role in spontaneous inference. Core assumptions related specifically to communication include the Gricean claim that understanding an utterance is a matter of inferring what the speaker intended to convey from what she utters; the claim that there are exactly two speaker's intentions that are central to communication, namely the informative intention and the communicative intention and finally the communicative principle of relevance and the presumption of optimal relevance, which mandate the relevance-theoretic comprehension procedure, a heuristic that guides the search for the correct interpretation of utterances. Relevance theorists try to give explanations for communication in terms of the working of the relevance-theoretic comprehension procedure. There are, in addition, several strategies that guide the explanation of phenomena in relevance theory, including: (i) Grice's Modified Occam's Razor, in a stronger form; (ii) the possibility of dividing what is linguistically encoded between conceptual and procedural information; (iii) the interpretive/descriptive distinction and (iv) the use of ad hoc concepts.

Alison Hall

The distinction between pragmatics and semantics is widely agreed to be between, respectively, meaning that is recovered by inference, and meaning that is determined largely by linguistic mechanisms. However, the fact that there is often much interaction between linguistic and inferred meaning, particularly at the level of explicitly communicated content, has given rise to a variety of different positions on how and where to draw the distinction. Hall takes as her starting-point relevance theory's view (detailed in Carston 2002, 2008a) that semantics

corresponds to linguistically encoded meaning, and pragmatics to any context-dependent meaning, even that which is mandated linguistically and contributes to explicit content. The reason for drawing the distinction this way is that, due to extensive contribution from context-sensitive processes at the level of explicit content, there is no useful level of representation in utterance comprehension that both differs from encoded linguistic meaning, and can be considered 'semantic'. Hall examines a number of theories that do appear to make sense of the idea of semantic content: the hidden indexical theory defended by Stanley (2000) and others, and the semantic minimalism defended in quite different forms by Cappelen and Lepore (2005), Bach (2001), Borg (2004) and Korta and Perry (2008). Hall concludes that, while certain of these do isolate a notion of content that is genuinely semantic, by virtue of being extremely minimal, there is no case for revising the distinctions drawn by relevance theory.

Alessandro Capone (Explicatures are NOT Cancellable)

In this paper, Capone argues that explicatures are not cancellable on theoretical grounds. He takes that explicatures are loci of pragmatic intrusion, where pragmatics mimics semantics. He attempts to differentiate explicatures from conversational implicatures on logical grounds. He answers some objections to Capone (2009) by Seymour (2010) and he also responds to Carston (2010). The crucial problem addressed in this paper is whether by cancellability of explicatures we should intend the evaporation of an explicature from an act of saying when a different context is considered. He discusses the logical problems which this view gives rise to. In this paper, he explores the consequences of considering cancellability of an explicature a language game. He concludes that the cancellability test proposed by Carston can never be unified with the other side of cancellability (explicit cancellability cannot be unified with cancellability due to an aspect of the context that cancels the inference). Furthermore, he considers that cancellability à la Carston is neither a definitional, nor a constructive nor an eliminative language game. The paper makes use of important considerations by Burton-Roberts (Forthcoming) on intentionality and also discusses some of his examples.

Capone Alessandro (The Pragmatics of Indirect Reports and Slurring)

According to Volosinov (1971) there is a tension between two indirect discourse practices; one in which the reported message's integrity is preserved and the boundaries between the main message and the embedded reported message are formally marked and one in which such boundaries are dissolved as the reporting context allows the reporting speaker to intrude to a greater extent and transform

the message by stylistic interpolations. This tension is clearly resolved, in the context of my paper on indirect reports, through the recognition of pragmatic principles which assign default interpretations (according to which the boundaries between the reporting message and the reported message are clearly visible and the reported speaker's voice prevails at least within the embedded message), while allowing context to create priorities which override the default interpretations and make the otherwise costly violations of the pragmatic principles worthwhile thanks to the facilitation and subordination of the information flow to the exigencies of the embedding context.

Eleni Gregoromichelaki and Ruth Kempson

In the last 50 years, there has been general agreement in the domain of Theoretical Linguistics that theories of language *competence* must be grounded in the description of sentence-strings and their literal semantic content without any reflection of the dynamics of language *performance*. However, recent research in the formal modeling of dialogue has led to the conclusion that such bifurcations—language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation—are all based on an arbitrary and ultimately mistaken dichotomy, one that obscures the unitary nature of the phenomena because it insists on a view of *grammar* that ignores essential features of natural language (NL) processing. The subsequent radical shift towards a conception of NL grammars as procedures for enabling interaction in context (Kempson et al 2010) now raises a host of psychological and philosophical issues: The ability of dialogue participants to take on or hand over utterances mid-sentence raises doubts as to the constitutive status of Gricean intention-recognition as a fundamental psychological mechanism (Gregoromichelaki et al 2011). Instead, the view that emerges, rather than relying on mind-reading and cognitive state meta-representational mechanisms, entails a reconsideration of "signalling" (or "ostensive communication") in a naturalistic direction and a non-individualistic view on meaning (see, e.g., Millikan 1993, 2005). Coordination/alignment/intersubjectivity among dialogue participants is now seen as relying on low level mechanisms (see, e.g., Pickering and Garrod 2004; Mills and Gregoromichelaki 2010) like the grammar (appropriately conceived).

Yan Huang

Logophoricity refers to the phenomenon whereby the 'perspective' or 'point of view' of an internal protagonist of a sentence or discourse, as opposed to that of the current, external speaker, is being reported by using some morphological and/or syntactic means. The term 'perspective' or 'point of view' is used here in a

technical sense and is intended to encompass words, thoughts, knowledge, emotion, perception and space-location (e.g. Huang 2000a: 173, 2001, 2002, 2006/2009, 2010a). The aim of this article is threefold. In the first place, Huang will provide a cross-linguistic, descriptive analysis of the phenomenology of logophoricity. Second, he will present a pragmatic account of logophoricity and the related use of regular expressions/pronouns in terms of conversational implicature, utilizing the revised neo-Gricean pragmatic theory of anaphora developed by Huang (1991, 1994/2007, 2000a, b, 2004, 2007, 2010a, c) (see also e.g. Levinson 2000). Finally, he will argue that (i) the neo-Gricean pragmatic analysis of logophoricity and the related use of regular expressions/pronouns in terms of pragmatic intrusion made here provides further evidence in support of the thesis that contrary to the classical Gricean position, pragmatics does 'intrude' or enter into the conventional, truth-conditional content of a sentence uttered, (ii) pragmatic intrusion into logophoricity is a conversational implicature rather than an explicature/impliciture and (iii) it involves 'pre'-semantic neo-Gricean pragmatics.

Eros Corazza

Corazza will argue that the notion of viewpoint plays central stage in our understanding and interpretation of many utterances. He will claim that such a notion is best characterized on the background of indexical reference; yet it cannot be reduced to it. He will thus show how points of view can be unarticulated (roughly, unmentioned) and yet play an important role in our linguistic practice inasmuch as the understanding of some utterances rests on the grasping of the point of view associated with them. Finally, he will mention how the notion of viewpoint (as an unarticulated linguistic phenomenon) plays an essential role in the understanding and interpretation of utterances containing anaphoric reflexive pronouns.

Keith Allan (Referring)

As defined here, a speaker's act of referring is the speaker's use of a language expression in the course of talking about its denotatum. This pragmatic definition of reference is defended against more traditional usage that contrasts "referring", "denoting", "describing", "alluding", "attributing", etc. It is proposed that the various differences in meaning supposedly captured by the different applications of these terms are better dealt with in other ways that can make shaper distinctions. What the hearer recognizes as the speaker's referent necessarily only 'counts as the referent' because it is on many occasions not identical to what the speaker identifies, indeed the speaker and hearer might even have entirely contradictory conceptions of the referent and yet the language expression used by the speaker can be said to successfully refer. For instance, if the Archbishop of Canterbury says to

Richard Dawkins *I will offer proof of the existence of God* and Dawkins replies *But God does not exist*, the deity that they are both referring to only counts as the same referent, because for the Archbishop God exists and for the author of *The God Delusion* God does not; in fact they have almost contradictory conceptions of the referent. This essay argues that an expression *e* frequently cannot identify exactly the same referent r for speaker and hearer, and that it is in fact unnecessary for it to do so; all that is required is that the referent counts as the same referent for the purpose of the communication. This is why mistaken reference like *Who's the teetotaller with the glass of water?* spoken of a man quaffing a glass of vodka can often successfully communicate who it is that is being spoken of.

Bart Geurts and Emar Maier (Layered Discourse Representation Theory)

Layered Discourse Representation Theory (LDRT) is a general framework for representing linguistic content. Different types of content (e.g. asserted, presupposed, or implicated information) are separated by putting them on different layers, all of which have a model-theoretic interpretation, although not all layers are interpreted uniformly. It is shown how LDRT solves so-called 'binding problems', which tend to arise whenever different kinds of content are separated too strictly. The power of the framework is further illustrated by showing how various kinds of contextual information may be accommodated.

Keith Allan (Common Ground)

Language is primarily a form of social interactive behaviour in which a speaker, writer or signer (henceforth S) addresses utterances (U) to an audience (H). It requires S to make certain assumptions about H's ability to understand U. This includes choice of topic, language, language variety, style of presentation and level of presentation. These assumptions constitute what can conveniently be called "common ground". They have been subsumed to context (e.g. Allan 1986; Duranti 1997); and at least a part of the common ground constitutes what Lewis 1969 referred to as "common knowledge", a term adopted by Stalnaker 1973. Schiffer 1972 called it "mutual knowledge*". Prince 1981 rejected "shared knowledge", preferring "assumed familiarity". Following Grice 1981, Stalnaker 2002 named it "common ground", which he described as "presumed background information shared by participants in a conversation … "what speakers [take] for granted— what they [presuppose] when they [use] certain sentences". A fatal flaw was carried over from Schiffer's definition of mutual knowledge* into Stalnaker's definition of common ground: "It is common ground that φ in a group if all

members accept (for the purpose of the conversation) that φ, and all *believe* that all accept that φ, and all *believe* that all *believe* that all accept that φ, etc.". The recursion within this definition would necessitate infinite processing on the part of each of S and H. This flaw has been accepted and repeated by many since. Clark 1996 attempted to circumvent it but his definition includes a clause that calls itself, thus creating an endless loop. In this essay, Allan suggests a way, inspired by Lee 2001, to characterize common ground from the points of view of both S and H which does not admit runaway recursion. In line with Stalnaker's mingling of presupposition and common ground, it refers to the preconditions on illocutions.

Mandy Simons

This paper, originally published in 2001, deals with the question of the source of presuppositions, focussing on the question of whether presuppositions are conventional properties of linguistic expressions, or arises as inferences derivable from ordinary content in combination with some general conversational principles. Simons argues that at least some presuppositions should be analysed as conversational inferences, on the grounds that they show two of the hallmarks of such inferences: contextual defeasibility and non detachability. She makes this case for the presuppositions associated with change of state predicates and with factives. She argues further for the need for a general principle for deriving presuppositions as inferences by illustrating a variety of cases of presupposition-like inferences not clearly involving a lexical presupposition trigger. In the second half of the paper, she moves towards the development of a general conversational account of the relevant presuppositions. Building on a brief comment in Stalnaker 1974, she develops the following pair of ideas: first, that an utterance embedding a proposition P may be seen as raising the question whether P; and second, that P may be related to a further proposition Q in such a way that it would make sense to raise the question whether P only if one already believed Q to be true. It is these required prior beliefs that constitute conversationally derived presuppositions. Although the account developed here is only a preliminary attempt, the relevance of contextually salient questions, or sets of alternatives, to an account of presupposition has been taken up in subsequent work, notably Abusch 2010 and Simons et al. 2010.

Klaus von Heusinger

The salience theory of definiteness has three historical sources: Lewis 1979 criticizes Russell's Theory of Descriptions and sketches an alternative theory using salience. Sgall et al. 1973 describe the information structure of a sentence with a hierarchy of "activated" referents. Grosz et al. 1995 argue on the basis of their analysis of discourse model in artificial intelligence that we need a salience

structure. Egli & von Heusinger 1995 and von Heusinger 1997 give a formal account of salience in terms of choice functions, and Peregrin and von Heusinger 1997 embed this into a dynamic semantics. Schlenker 2004 uses this semantics for definite noun phrases and conditionals.

Kecskes and Zhang

The goal of this paper is to redefine the relationship between common ground and presupposition within the confines of the socio-cognitive approach (SCA). SCA (Kecskes 2008; Kecskes and Zhang 2009; Kecskes 2010) adopted in this paper offers an alternative view on communication, which claims that communication is not an ideal transfer of information, and cooperation and egocentrism are both present in the process of communication to a varying extent. The SCA emphasizes the dynamics of common ground creation and updating in the actual process of interaction, in which interlocutors are considered as "complete" individuals with different possible cognitive status being less or more cooperative at different stages of the communicative process. Presupposition is a proposal of common ground, and there is a vibrant interaction between the two. They enjoy a cross relation in terms of content and manners in which they are formed, and their dynamism is inherently related and explanatory to each other. This claim has important implications to the solution to presupposition accommodation. After the introduction, Chap. 2 describes the socio-cognitive approach. Chapter 3 reviews the assumed common ground, and Chap. 4 introduces the speaker-assigned presupposition. Chapter 5 discusses the dynamism of presuppositions and common ground, and claims that their dynamic observations are coherent and explanatory to each other. Chapter 6 readdresses the accommodation problem with redefinition of the relations.

Alan Libert

This paper argues that there can be a pragmatics of artificial languages, even though most such languages have seen little or no use. Several areas of pragmatics are examined in relation to artificial languages: politeness (including pronouns, terms of address, honorifics and imperatives), formal language and other types of language, conversational implicature, non-descriptive meaning (including conjunctions, interjections and illocutionary force), and metaphor/non-literal language. Texts in several artificial languages (aUI, Sotos Ochando's Lengua Universal, Hom-idyomo, and Esata) are presented and briefly discussed. It is concluded that in most respects artificial languages are not very different in their pragmatics from natural languages, in spite of the fact that on the surface some artificial languages appear quite exotic; this is perhaps to be expected, but nevertheless is a interesting finding.

Sorin Stati

After presenting the contexts in which researchers speak about implicit elements—
a list that highlights the conceptual diversity of meanings attributed to the term
'implicit'—Stati will focus on the actual topic of this article: the property
'implicit' as it functions in argumentative texts. Or, to put it another way, how do
implicit propositions manifest themselves on the argumentative discourse level.
Stati dwells on interesting inferential phenomena involving the argumentative
roles of portions of text. He differentiates between a casual overhearer and the
intended addressee, speculating on the differences in interpretative behavior. The
inferential behavior triggered by argumentative relations within a text very often
involves the recovery of implicit materials.

Marco Mazzone

In utterance understanding, both personal and sub-personal aspects appear to be
involved. Relevance theory (starting from Sperber and Wilson 1986/1995) and
Recanati (2004) have respectively explored two alternative ways to conceive of
those aspects and their interaction. Here a third account is proposed, in the light of
the automatic-controlled distinction in psychology, and of recent views concerning
the cooperation between these two modes of processing. Compared to Recanati
(2004), the account proposed here assigns a larger role to automatic, associative
processes; at the same time, it rejects the view that consciousness applies only to
what Recanati calls secondary pragmatic processes. Consciousness is rather held to
cooperate with associative processes in any aspect of pragmatic processing,
irrespective of the pragmatic distinction between explicatures and implicatures. On
the other hand, a close consideration of how associative and conscious processes
plausibly interact makes it appear unnecessary the hypothesis of a specialized
process for utterance understanding—such as the automatic, inferential mechanism
put forth by Relevance theory.

Dorota Zielinska

In this the author searches for the mechanism correlating linguistic form with
content in order to explain (in the sense of the word 'explain' used in empirical and
modern social sciences) how sentence meaning contributes to the utterance
meaning. She does that against the background of two currently dominating
positions on that issue: *minimalism* and *contextualism*. Minimalists regard
language as a self-standing abstract system and claim that only weak pragmatic
effects are involved in interpreting sentences. Contextualists believe that language

can be described adequately only within a theory of language understanding and that strong pragmatic effects are also involved in interpreting sentences. The resultant controversy, presented in Chap. 1, has been pronounced by Michel Seymour the most important one in the 20th century.

Zielinska proposes a specific model of the form-meaning correlation process, based on a novel mechanism of a linguistic categorization, which is compatible with a bio-social developmental perspective, advocated in Chap. 2. On this view, the utterance meaning is dependent both on the conventional meaning of the construction components conveying it, and on the specific social function of the whole construction (a relevant *pragmeme*). She finishes the paper by preliminarily testing the mechanism of the form-content correlation process both qualitatively and quantitatively. The latter tests meet the standards of empirical sciences.

Marco Carapezza and Pierluigi Biancini

The authors have tried to make the potentiality inherent in the concept of the linguistic game evident by taking it back to its original context in the work of Wittgenstein. This paper aims to re-examine some features of Wittgenstein's thought, considering in particular the notion of 'language-game'. The authors believe that the language-game might play a role in overcoming once and for all the classic distinction between semantics and pragmatics. We deal with the exegetical discussion of the notion 'language-game' as it was interpreted in two different senses: as a synonym of calculus or as a minimal unit of linguistic activity that is directed to obtaining certain pragmatic effects in a societal context. The latter, broader interpretation, is characterized by three different features: topicality, broader normativity and multimodality. Starting from an interpretation of language game as a pragmatic act, the authors work out a possible parallel between language games and the notion of pragmeme as presented by Mey. Both language game and pragmeme refer to an extended notion of the linguistic symbol seen as a non-linear, multimodal concept that overlaps the mere verbal unit of expression and is now considered as a set of diverse expressive resources (such as gesture, tone of voice and so on). This comparison will also work for a problem common to both language-game and pragmeme, that is the need to set a boundary to these units of analysis, thanks to which they could be identified. The authors advance a possible solution to this problem, which is rooted in a rethinking of Wittgenstein's notions. The proposal consists in focusing on the topic for which the language game is played. The topic is taken to be the organizing aspect of understanding of the game. The societal rules, the worldly knowledge, often taken to be the ground of understanding in our discourse are considered as merged together in a holistic unit called language game.

Meaning, Semantics and Semiotics

Noel Burton-Roberts

Abstract This paper questions the assumption, widespread in linguistic theory and pragmatics, that linguistic expressions have meaning in virtue of possessing semantic properties/content. Problems created by this assumption are discussed and an alternative, semiotic, account of meaning is developed that places 'linguistic meaning' in the context of meaning in general.

1 Introduction

In a discussion of relevance theory, Dan Wedgwood comments:

> Relevance theorists have tended to assume that RT can be used… as an adjunct to fairly conventional approaches to other parts of linguistic theory…. But in making the move away from the moderate contextualism of Gricean approaches, RT has more radical consequences, whether we like it or not. In effect, this constitutes a break from conventional perspectives on semantics… the nature of encoded meaning cannot be understood without active consideration of inferential contributions to meaning. This is nothing less than a reversal of conventional methodology, which tends to abstract away from inferential pragmatic processes as much as possible…. Once we reject [conventional semantic analysis], as RT does in principle, the nature of encoded meaning becomes an entirely open question…. A great deal of contemporary syntactic theorising is motivated by a wish to account for the perceived semantic character of a given sentence (hence the regular use of levels of representation of syntactic representation like LF, LOGICAL FORM) (2007, 679).

N. Burton-Roberts (✉)
School of English Literature, Language and Linguistics, Newcastle University,
School of English, Newcastle upon Tyne, NE1 7RU, United Kingdom
e-mail: noel.burton-roberts@newcastle.ac.uk

These are the thoughts I develop in this paper. I will argue that pragmatic theory—and relevance theory in particular—should be seen as offering a challenge to 'conventional' linguistic wisdom.

What is at issue is the nature of meaning, no less. I will make a case for—or aim to reinstate—an account of meaning at odds with 'conventional' linguistic theorising, by which I will assume Wedgwood means Chomskyan generative grammar (CGG). The relevant CGG assumption is evinced by Brody (1995, 1) when he writes 'it is a truism that grammar relates sound and meaning'. CGG seeks to achieve this by positing two interface levels of linguistic representation, Phonetic Form (PF) and Logical Form (LF), attributing to expressions two sorts of property, phonological and semantic. This double-interface assumption is assumed to be necessary to the modelling of 'language as sound with a meaning' (Chomsky 1995, 2).

In reconstructing the idea that expressions 'have meaning' by assigning them semantic properties, CGG effectively equates semantics and linguistic meaning. I will argue that this insulates linguistic meaning from meaning in general and is unexplanatory. Furthermore, it is the basis of Grice's 'moderate contextualism'. As Wedgwood observes, relevance theory (RT) is radically contextualist in principle but shares Grice's and CGG's assumption that there is such a thing as linguistic semantics and Logical Form (LF) thought of as a level of linguistic encoding. To that extent RT's radical contextualism is qualified.

CGG's double-interface assumption is a legacy of Saussure's concept of sign (Burton-Roberts and Poole 2006). I approach the topic of this paper through a discussion of the Saussurean sign (in Sect. 2) because I want to bring semiotics to bear on the subject of meaning and to highlight some problems for CGG's (Saussurean) reconstruction of 'sound with a meaning'. I reject that CGG account in favour of a (roughly) 'Peircean' concept of sign. This, I argue, is more general in its application, more explanatory and more consistent with the cognitive naturalism of CGG and RT.

In the light of that, Sect. 3 offers a general account of what meaning is, developing ideas touched on in Burton-Roberts (2005, 2007). This account denies that linguistic expressions have semantic properties. It distinguishes between meaning and semantics. Assuming, with RT, that only thoughts have ('real') semantic content, it argues that meaning is a relation (of something, potentially anything) to the semantic content of a thought, a cognitive relation effected by inference. The crucial idea here is that meaning is not a (semantic) *property* but a semiotic *relation* (to semantic properties). This goes for meaning generally; my aim is to situate linguistic meaning within a semiotic account of meaning in general. My own view is that the picture of meaning that emerges is consistent with RT. One might even go as far as to say that relevance, as defined in RT, *is* meaning.[1]

---

[1] I won't push this thought further but the reader may want to bear it in mind in Sect. 3. Incidentally, Sperber and Wilson (1986/1995: Chap. 1) assume that any *semiotic* approach to meaning aims to reduce all meaning to a code model of communication. That may have been true of 20th century extensions of Saussurean semiotics (they cite Levi-Strauss and Barthes)—in

2 The Saussurean Sign

There are two well-known features of the Saussurean linguistic sign.

The first is that the linguistic sign is constituted by the conjunction of a concept and a sound image. Sound image and concept combine to make a further entity, the sign itself, which is a 'double entity' (Saussure, 65), 'a two-sided psychological entity' (66).[2] The relation between concept and sign and that between sound image and sign are part ∼ whole relations. In modern parlance, the concept is the sign's semantics; the sound image is its phonology. The relation between semantics and phonology is thus a part ∼ part relation. Saussure (e.g. 67) explicitly refers to the sign as 'the whole' and to the sound image and concept as its 'parts'. The same idea is evident in CGG's double-interface assumption: 'there are sensorimotor systems that access one aspect of an expression and there are conceptual-intentional systems that access another aspect of an expression, which means that an expression has to have two kinds of symbolic objects as its parts' (Chomsky 2000b, 9).

The formal study of part ∼ part and part ∼ whole relations is called 'mereology'. I won't be invoking formal mereology, but I'll borrow the term 'mereological' in discussing the Saussurean sign. In an intuitively equivalent formulation, concept and sound image are respectively the semantic and phonological *properties of* (or 'aspects' of) the linguistic sign.

Saussure has little to say about where syntax figures in this. This is addressed in CGG, where phonological and semantic features are treated as properties (parts, aspects) of *syntactic* entities (Burton-Roberts 2011). This adds substance to the 'double entity' idea: if you believe in the existence of a further entity jointly constituted by phonological and semantic features, there should indeed be something substantive to say about it—e.g. that it is syntactic—over and above what can be said about the phonological and the semantic.

The second feature of the Saussurean sign is that it involves a *semiotic* relation, the signifier ∼ signified relation. The sound image signifies the concept. Significantly, nothing in CGG reflects this feature of Saussure's thinking, a matter I address below.

(Footnote 1 continued)

connection with which they write (with justification in my view) 'The recent history of semiotics has been one of simultaneous institutional success and intellectual bankruptcy' (p. 7)—but nothing could be further removed from my aim here. Whereas, in their terms, 'The semiotic approach to communication… is a generalisation of the code model of verbal communication' (p. 6), my aim is precisely the opposite: to situate linguistic meaning within a more general semiotic—i.e. inferential—account of meaning. My claim is that all meaning (though not all communication, see note 9 below)—and all 'decoding' involved in the construction of meaning—is inferential in character.

[2] All references to 'Saussure' are to Wade Baskin's (1959) translation of the *Cours de Linguistique Generale*.

Fig. 1 The Saussurean sign

Mereological: co-part, or co-property ◄──► co-part, or co-property

Semiotic: signifier ──► signified
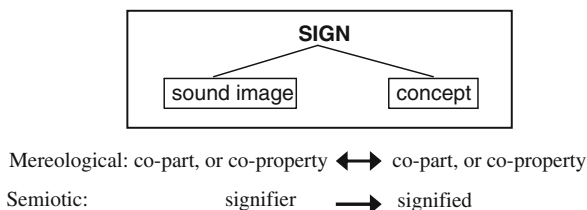
Now, the mereological (part ~ part) relation and the semiotic (signi-fier ~ signified) relation are at least different. Since things generally are related in more ways than one, it might seem that sound image and concept can be related in both these ways, as Saussure assumed (see Fig. 1).

Against this, I will argue that the two relations are so different as to be incompatible; we must choose between them.

Notice first that the part ~ part relation is symmetric, in the sense that they are *co*-parts, *co*-constitutive of the Saussurean sign. Saussure (113) draws a general analogy with a sheet of paper: in a language ('a system of signs'), sound image and concept each relate to the sign as the two sides of a sheet of paper relate to the sheet. As we shall see, this symmetry was crucial for Saussure. In sharp contrast, the semiotic relation is antisymmetric. One term is the signifier, not the signified; the other is the signified, not the signifier. It's a relation *from* the sound image (signifier) *to* the concept (signified). The sheet-of-paper analogy is completely inappropriate in this connection.

Notice furthermore that on the mereological conception there is no *direct* relation between sound image and concept. Their (part ~ part) relation to each other is entirely derivative: they are in that relation only because each is, primarily, in the (antisymmetric) part ~ whole relation to the sign. In that respect, the relation between sound image and concept is not self-explanatory. The semiotic relation, by contrast, is direct and primary. It does not follow from any other more direct relation. In that sense, by comparison with the mereological relation, the semiotic relation between sound image and concept is self-explanatory.

A further difference is that the mereological account treats as an *object* ('the sign') what the semiotic idea, by itself, treats purely as a *relation*. The mere-ological idea, by definition, entails that there is an object distinct from both sound image and concept constituted by their combination. Not so the semiotic idea. Unless mereological, a relation between two objects implies no further object.[3] Notice that 'property-of' talk in connection with the mereological conception is questionable on the semiotic conception. If *x* stands in a semiotic relation to *y* (i.e.

---

[3] Other than an abstract set-theoretical entity. See Burton-Roberts (2011) for discussion of the issue in CGG. Saussure explicitly denies the sign is an abstract entity (15, 102), insisting on its being a 'concrete entity'. His reference to 'parts' and 'wholes' (e.g. 67) would be inappropriate were he thinking in merely set-theoretical terms. Sets don't have parts, they have members.

Fig. 2 The 'purely semiotic'
(Peircean) sign
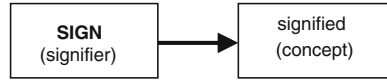
signifies *y*), there is no conceptual necessity to posit an object *Z* (the 'sign') for *x* and *y* to be properties of.

This doesn't mean that, on a purely semiotic conception, we have to abandon any idea of 'sign'. There is an alternative, non-mereological, conception of sign compatible with the signifier ~ signified relation. On this conception it is the *signifier* itself that is the sign (Fig. 2).

I call it 'Peircean' because it is consistent with the thought of Peirce (1933). But I won't be invoking the whole panoply of Peircean semiotics. Furthermore, I will assume—with Saussure, not Peirce—that the signified is always a concept (Sect. 3).

Saussure (67) explicitly rejects this (Fig. 2) concept of sign, insisting on the mereologically constituted sign (Fig. 1). Against this I will suggest that the mereological account is conceptually (a) insufficient and (b) unnecessary, even (c) assuming it is possible. And there are other objections to it, notwithstanding its influence on CGG's double-interface assumption and moderate contextualism's 'encoded meaning/semantics'.

The mereological account is conceptually insufficient, I hold, because there is nothing actually *semiotic* about it, in and of itself. Mereological (part ~ part and part ~ whole) relations have nothing to do with meaning. What I mean is that they fail to differentiate the relation we are concerned with from e.g. that between the barrel and the nib of a pen, the two sleeves of a shirt, the seat and back of a chair, which *are* mereological. This mereological idea offers no explanation of why a sound image (phonology) and a concept (semantics) should actually be related. It might be objected that this ignores the fact that, for Saussure, the relation between sound image and concept isn't only mereological, it's also semiotic. But that just goes to show the conceptual insufficiency of the mereological idea in this context. It *needs* to be supplemented by the semiotic idea.

As regards conceptual necessity: having supplemented it with the semiotic relation, what conceptual/theoretical work remains for the mereological account? It adds nothing. The semiotic idea, we have just seen, is necessary—and it is in itself sufficient. Furthermore, it yields a concept of sign that is at least more obvious: what is a sign if not a signifier?[4] It is also more parsimonious. It calls for just one relation between just two entities. The mereological account multiplies beyond necessity. It posits three relations (two part ~ whole relations and a part ~ part relation) and three entities- and needs the semiotic relation.

---

[4] It is reasonable to ask what signs are signs *of/for*. On the Fig. 2 (purely semiotic) notion of sign, the question receives an answer (though a general one): they are signs of/for concepts. But on the Fig. 1 (Saussurean) notion of sign, the question is simply incoherent: the mereologically constituted sign isn't a sign *of/for* anything. See below on a necessary precondition for signification (and bootstrapping).

We might even question whether in this context the mereological idea is possible, on the grounds that phonological and semantic properties are *sortally* distinct. Sortally distinct properties are such that nothing can have both sorts of property. For example: sore throats, earthquakes, epidemics, years, prime numbers, marriages, mortgages—none of these is of a sort that can be sky blue, right-angled, bisyllabic or constitute a proof of Fermat's last theorem. Against this it might be argued that, since they are both mentally constituted, phonological and semantic properties cannot be sortally incompatible. However, given how phonology and semantics are grounded, their respective contents are sortally incompatible: articulatory/acoustic versus conceptual-intentional. Arguably, the sortal incompatibility of sound image and concept is the basis of Saussurean arbitrariness: whatever species of relation holds between sortally incompatible properties, it can only be arbitrary (non-natural). Within CGG this sortal incompatibility is effectively acknowledged in its assumption that phonological and semantic properties are mutually un-interpretable (Burton-Roberts 2011).

Given this sortal consideration, can we really allow that something could be mereologically constituted as both bisyllabic and prime, e.g. the putative 'double entity' *seven*? That question doesn't even arise on the Peircean conception. The Peircean sign is bisyllabic, not prime. Sortally distinct, and separate, from the sign itself is the numerical concept it signifies, which is prime not bisyllabic. I suggest the relation can only be semiotic, not mereological.

Furthermore, since the signified is mind-internal (a concept), the mereological account must insist the signifier is also mind-internal. Otherwise we would be committed to something (the Fig. 1 'sign') constituted mind-internally in one part and mind-externally in its other part. I take that to be incoherent. Of course, Saussure does insist on the mind-internal nature of the sound image—reasonably, assuming it's phonological. However, the purely semiotic account (Fig. 2), while compatible with a sign being mind-internal, is compatible with a sign being a mind-external phenomenon. There's no reason why a semiotic relation can't hold between a mind-external phenomenon and a (mind-internal) concept/thought. A wide range of mind-external phenomena do function as signs for us.[5] Potentially at least, all phenomena do. I take this to be the substance of RT's (very general) First Cognitive Principle of Relevance (Carston 2002, 379).

In that respect, the purely semiotic (Peircean) conception has wider application, applying not just to linguistic signs but to signs in general. Notwithstanding Saussure's suggestion (e.g. 16) that linguistics might form part of a more general 'semiology', his account actually applies only to linguistic signs. The symmetry of his mereological conception is crucial here.

Saussure insists on this symmetry on the grounds that sound image (qua signifier) and concept (qua signified) are related by mutual implication (e.g. 103).

---

[5] The nearest Saussure comes to allowing this is in the idea that (mentally constituted) signs are 'realized' in phonetic phenomena. But even here, it is surely just the signifier (the sound image) that's 'realized'. Since the Saussurean sign is constituted in part by a concept, it's hard to accept it could be 'realized' in phonetic phenomena.

Recall the sheet-of-paper analogy. However, while the *terms* 'signifier' and 'signified' are mutually implying, it doesn't follow that the things those terms refer to (sound image, concept) are mutually implying. What I mean is that a concept is still a concept whether or not it happens to be signified by a sound image. In fact, surely, a conceptually necessary precondition for signification is that the signified exist independently of the fact that it is signified. (There is again an asymmetry here, for nothing similar can be said of the phonological sound image, the whole rationale of which lies in its being a signifier.) But all this is precisely what Saussure seeks to deny. The symmetry of his mereological conception is motivated by his view that concepts only exist as constituents of (arbitrary) linguistic signs. He explicitly denies that 'ready made ideas exist before words' (65, see also 112). Assuming concepts necessarily figure in thought/ideas, the Saussurean contention is that thought is couched only in the signs—the 'code' (14)—of some particular language. The motivation, in short, is an extreme version of the so-called Sapir-Whorf hypothesis.

This must be rejected, I believe. As Sperber and Wilson (1986/1986, 192) and Carston (2002, 30–42) argue, what is thought/thinkable extends well beyond what is linguistically encoded/encodable. This is not a merely philosophical or theoretical matter; it's the stuff of common experience. Chomsky (2000b, 76) puts it well: '…very often, I seem to be thinking and finding it hard to articulate what I am thinking. It is a very common experience… to try to express something, to say it and to realize that is not what I meant… it is pretty hard to make sense of that experience without assuming that you think without language. You think and then you try to find a way to articulate what you think and sometimes you can't do it at all;… if you are thinking, then presumably there's some kind of conceptual structure there.' The universality of such experience, the necessary precondition for signification mentioned earlier and the related problem of how the Saussurean sign could actually evolve/arise other than by the merest (most circular) bootstrapping (Fodor 1975), all lead me to agree there must be 'some kind of conceptual structure there', a language of thought 'used internally' (Chomsky 2006, 9), if subconsciously, logically prior to a speaker's particular language, as argued by Fodor and assumed in RT.[6]

How is the set of concepts delimited? Since the relation between sound image and concept is arbitrary, if concepts only exist as constituents of (arbitrary) linguistic signs then the set of concepts must be arbitrary (and, notice, semantics = linguistic

---

[6] I would argue it is phylogenetically and ontogenetically prior. The passage from Chomsky, incidentally, is consistent with a more nuanced version of the Sapir-Whorf hypothesis. I am not the first to suggest there is more than one kind of thought. There is (a) the kind of thought explicitly referred in that quote and (b) another, also in evidence there, which consists in gaining more conscious *access* those thoughts by trying to articulate them. See Burton-Roberts (2011).

Grice takes the Saussurean line: 'A plausible position is that…language is indispensable for thought' (1989: 353). But (355) he is upfront—and unusually clear—on the dilemma this poses re the precondition for signification (his 'intelligibility'). Less clear (to me) is his proposed solution (355–356).

semantics). Another response, more consistent with the naturalism of CGG and RT, is that the set of concepts is the set of humanly entertainable concepts, a set delimited by nature—human nature. This is consistent with a Peircean conception of sign but not with what motivates the Saussurean sign.

2.1 CGG and Pragmatics

I've suggested we must choose between the Saussurean (mereological) sign and the Peircean (purely semiotic) sign and have sought to show we must choose the latter. All this would be of merely historical interest were it not that CGG makes the diametrically opposite choice. It is the mereological idea that is embodied in CGG's idea of a (syntactic) object with phonological and semantic properties ('aspects…parts'). And *only* the mereological idea. CGG doesn't appeal to 'sign'. It eschews all reference to semiotic concerns. In CGG we are asked to make do *just* with a mereological (part $\sim$ part) account of the relation between phonology and semantics.

I have argued this is at least conceptually insufficient—uninformative and un-explanatory. Mereological relations have nothing to do with meaning. CGG's response might be that, in attributing semantics to expressions, it assigns them their meanings. This equates 'having meaning' with 'having semantic properties' in the linguistic context. The next section presents an account of meaning that undermines that equation. In the meantime, I here offer some observations on the mereological account in the context of CGG and RT.

In addition to the above objections to it, the mereological account poses two problems internal to CGG (Burton-Roberts 2011). The first is this. As noted, phonological and semantic properties are acknowledged in CGG to be mutually un-interpretable. If lexical items (and expressions composed of them) are [phon + sem] 'double entities', their phonological properties are not interpretable at LF and their semantic properties are not interpretable at PF. As a consequence, neither of the interfaces that the linguistic computation serves is actually capable of interpreting the double entities—words and thus anything composed of words—it is generally thought to manipulate.

The second (related) problem is this. I assume, with CGG, that the linguistic computation is universal (invariant) and natural (innate). But how could such a computation possibly operate with objects mereologically constituted by phono-logical and semantic properties, given that the relation between those properties is arbitrary (non-natural)? Being arbitrary, such relations are cross-linguistically variable, not innate but learned in the course of acquiring a language. This indeed is what Chomsky assumes: 'there is something like an array of innate concepts and… these are to a large degree merely "labeled" in language acquisition' (2000a, 65). Notice that the label metaphor (see also pp. 61, 66) anyway seems less consistent with CGG's mereological picture than with the semiotic. It is super-fluous, and surely wrong, to posit something constituted both by the label itself and

what-it-is-a-label-for. If the label metaphor is appropriate, why not allow it is (morpho-) phonologically constituted *signifier* for the concept?

To my knowledge Chomsky has only ever mentioned semiotics in a brief and scathing dismissal (2000b, 47–8). This dismissal may relate to a methodological objection to semiotics within 'conventional' linguistics. I suggested earlier that the semiotic account is recommended over the mereological on the grounds of its wider application, to meaning in general. Within autonomous linguistics (incl. 'semantics'), this—effectively, the diffuseness of semiotics—would count as a positive dis-recommendation. 'Semantic theory' is thought of as concerned exclusively with *linguistic* meaning. This restriction of the scope of 'semantics' might be thought an advantage if linguistics is autonomous, concerned with a module of mind, to be insulated from dealing with interpretation in general. In connection with the project of constructing an 'interpreter' that 'accepts non-linguistic as well as linguistic inputs', Chomsky writes 'the study of the interpreter… is not a topic for empirical enquiry…: there is no such topic as the study of everything' (2000a, 69). See also Fodor's First Law of the Non-Existence of Cognitive Science: 'the more global…a cognitive process is, the less anybody understands it' (1983, 107).

In the light of that, however, one might wonder whether, in assigning 'semantic' properties (and Logical Form) to expressions in a module of mind, CGG does in fact claim to deal with meaning/interpretation as generally understood. In fact, Chomsky is actually quite sceptical about such an enterprise (2000a, 21)—it is not the job of science to elucidate folk-scientific notions like 'meaning' (notwithstanding his 'language as sound with a meaning')—and indeed sceptical about 'semantics' in generative grammar. 'It is possible that natural language has only syntax and pragmatics; it has "semantics" only in the sense of "the study of how this instrument… is actually put to use"' (2000a, 132).

I believe pragmatic theory does underwrite scepticism regarding 'the semantics of natural language' and 'knowledge of meaning' (Larson and Segal 1995) within an autonomous sub-personal generative model of language, independent of how it is put to use (i.e. independent of pragmatics) and, indeed, independent of meaning in general. As for more recent CGG itself, strong minimalism anyway seeks to go 'beyond explanatory adequacy' (Chomsky 2004), attributing as little as possible to (modular) unexplained features of UG, seeking explanation in other aspects of human cognition. In the light of that enterprise, situating linguistic meaning within an approach to meaning in general might well be explanatory. 'Meaning' may be 'folk-scientific' but it can't be dismissed in the absence of a theory that explains it away.

Carston (2002, Introduction) is a vigorous defence of RT against the charge that, as a theory of interpretation, it is scientifically impracticable/vacuous. Commenting on her defence in my review (2005, 389), I wrote 'RT is a theory of something quite specific, however general in its application, namely all that is implied by "optimal relevance"…'. One needs to go further in its defence of course, but how much further? As I understand it, RT's account of linguistic communication is rooted in a more general theory of how humans interpret the world. Carston's defence seems to retreat from that, portraying RT as about 'a

mental module… domain specific in that it is activated exclusively by ostensive stimuli' (7). I think this is unfortunate—and even questionable. Stimuli, even humanly produced, don't come ready-labelled '± ostensive'. The assumption that a given stimulus is ostensive is just that—an assumption, inferentially derived. If the claim is that *linguistic* stimuli do come thus ready-labelled, their interpretation still requires a host of occasion/context-specific inferences, as RT itself has shown in response to the problems of what has come to be known as 'Gricean pragmatics'.

Notwithstanding Grice's preoccupation with meaning in general, he sought (particularly in his 'Retrospective Epilogue') to erect a corral around one 'central' species of meaning/signification—formal, 'dictive' (in the main), timeless, linguistic-type meaning, un-relativised to speakers, occasions or contexts, which 'will authorize the assignment of truth conditions to… expressions' (1989, 364), as assumed in 'semantic theory'. Gricean pragmatics is in fact a defence of this and yields a moderate contextualism.

Carston by contrast (2002, 4) suggests that, as 'cognitive pragmatics', RT is 'no longer to be seen as an adjunct to natural language semantics….' As Wedgwood observes, this is exactly the right conclusion in principle but not obviously true in practice. Carston immediately qualifies it with '…though it clearly continues to have essential interaction with semantics'. RT assumes and depends on 'natural language semantics', endorsing the attribution of semantic properties to expressions as non-contextual types and their deterministic decoding by a dedicated module of mind. Logical Form as a level of encoded linguistic representation figures crucially in RT's explicature/implicature distinction. Although Fodorian in spirit, RT has not pursued Fodor's suggestion that 'English has no semantics' (1998, 6) or his contention 'that LF is a level of description not of English, but of Mentalese…' (2008, 78). At least, while agreeing with Fodor in locating semantics in thought, RT also posits a 'linguistic semantics'. Equally, Carston stops short of Recanati's 'radical claim that there is no lexical meaning in the sense of stable encoding' (Carston 2002, 375) and his suggestion (1998, 630) that 'the only meaning which words have is that which emerges in context', taking a less 'extreme', more 'conservative' line 'on which words do encode something, albeit something very schematic…'. I have discussed this and its problems elsewhere (2005, 2007).

It may be true (I think it is) that there can be no naturalistic, causal, mechanistic, sub-personal, modular account of interpretation/meaning. And it may follow that no account of interpretation/meaning is natural science. I leave it to others to decide whether that matters at this stage of the game. As Chomsky has often observed, what counts as natural science has changed in the past. It may in the future, perhaps in the light of questions (and answers) now counting as non-scientific. It is not as if contemporary pragmatics—or cognitive science or linguistics (of the more thoughtful kind)—can claim to have jumped entirely free of 'philosophising'. Carston's defence of RT *as natural science* (on the above terms) is ambitious yet constraining in that it unduly moderates RT's contextualism and, I believe, fails to reflect what is potentially the true scope of the theory.

3 Meaning

Meaning is not exclusively linguistic, nor is 'sound with a meaning'. If we seek an explanatory account of 'linguistic meaning', it should be laid in the nest of these (cognitive, semiotic) truisms.

I don't claim that all of what follows is new. What it aims to show is that it is unnecessary and unexplanatory (not to say plain wrong) to attribute *any* kind of semantic property/content to linguistic expressions. To that end, Sect. 3.1 argues for a *distinction* between meaning and semantics. In the light of that, Sect. 3.2 discusses their *relation*.

3.1 Distinguishing Meaning and Semantics

A phenomenon *P* is a sign iff it is significant. It signifies something (*X*). Thereby *P means X*. Consider first some well-worn examples of 'natural signs'. (i) 'Those spots mean measles' (Grice 1989, 213). (ii) Pattering on the window means rain. (iii) Red litmus paper means acid. (iv) Smoke means fire. These are examples of Peirce's category of 'indexical' sign, involving Grice's 'natural meaning'. (ii) is a case of 'sound with a meaning'. What makes *P* a natural sign of *X* here is a natural (agent-less) causal relation between *X* and *P*. *Pace* Peirce, whose terminology suggests that all semiotic relations are representational, these are *not* representational. As a (natural, indexical) sign of fire, smoke is not a representation of fire. By the same token, they are not ostensive.[7] They are signs-*of*, not signs-*for*.

It's not nearly that simple of course. None of (i)–(iv) is objectively true, i.e. true in the absence of a subject *S* noticing *P* on an occasion. A particular phenomenon *P* is a sign only if it signifies something (*X*) *to S*. And meaning-*X*-to-*S* depends on *S* making inferences based on her beliefs. Red litmus paper in a liquid will mean to *S* that the liquid is acid only if *S* assumes it is litmus paper and has the belief *B* that acid causes litmus paper to turn red. 'Red litmus paper means acid' abstracts away from such crucial details, on the assumption that *B* is (general) knowledge. Those spots may mean measles to a doctor but not to Peter. 'General knowledge' and even what passes for 'the evidence of one's eyes (ears, etc.)' don't obviate the need for inference, however subliminal (Searle 1983).

Nor is it true that *P* means *X* to *S*—not if '*X*' stands for the causing phenomenon. Rather, *P* means *that-X* to *S*. Meaning is by its nature *communicative*.

---

[7] However, they can be exploited ostensively in a way that either depends on their indexicality or overrides it. A Glaswegian accent is an indexical sign of Glaswegian provenance but not if deliberately adopted (in which case, if not intended to deceive, it will be representational). Susan putting on her coat is an indexical sign of her intention to go out (perhaps) but she may exploit that with the semiotic intention of (ostensively) representing her desire to do so. White smoke rising from the Vatican ostensively represents a Papal election and, if so recognised, that will override (iv).

Subject to cognitive conditions just described, smoke is meaningful to *S* because it communicates to her that there's a fire—where 'that there's a fire' identifies a thought of *S*'s. I take 'communicates', when it involves meaning, to be equivalent to 'leads *S* to entertain a thought'.[8]

What follows from all this, even if obvious, is important. Although we say that *P*, as a sign, 'is meaningful' or 'has' meaning, and although we talk of 'sound with a meaning', such talk does not identify any *property* of *P*. More to the point—and this *is* my point—no one would want to say that their 'having meaning' identifies a *semantic* property of those phenomena. On an occasion of pattering-on-a-window, it is true (false) that it is raining but no one would attribute any truth condition to the pattering. If truth conditions are what you're after, they pertain to *S*'s thought that-*X*. What might be said to be 'the meaning of *P*' is entirely extrinsic to *P*. Despite the predicational similarity, 'means *X*' is utterly different from 'rises when heated' or 'is mercury/heavy/edible'. In at least the cases discussed so far—indexical signs in general—meaning and semantics are clearly distinct.

With indexical signs we are far from language but, in the respects that matter here, the above holds quite generally. Take Peirce's category of *symbolic* signs—Grice's non-natural meaning (meaning$_{nn}$). I illustrate this first by non-linguistic examples but we are close enough to language here, I suggest, for it to have a direct bearing on 'linguistic meaning'. My argument is that, as with indexical signs so with symbolic signs (whether linguistic or not): meaning and semantics need to be distinguished. Illustrative phenomena are a red flag (meaning artillery practice is in progress or swimming is forbidden), a gunshot (to start a race) and a ringing bell (to stop the bus, raise the alarm, etc.), the last two being 'sound with a meaning'. What makes these 'symbolic' is the fact that there's no natural (i.e. agent-less) causality involved here. These are non-natural in depending on other things: **(a)** on a certain convention, **(b)** on semiotic intention, and **(c)** on inferentially derived recognition by *S* of (a) and (b). Unlike indexicals, in virtue of (a)–(c), symbolic signs *are* representational; they are ostensive. The difference is that, symbolic signs being representational, that-*X* is *meant*. (We don't say that-there's-a-fire *is meant* by smoke.) Here 'meant' = 'intended'; and 'intended' here is short for 'intended to mean (to *S*)'. In these symbolic cases, an agent *A* communicates that-*X* to *S* by means of *P*. In other words, *A* arranges for the occurrence of some phenomenon *P* intending that its occurrence will communicate to *S* (lead *S* to have the thought) that-*X*. Since we are not telepathic, some perceptible *P* is necessary.

It was Grice who instated intention and its recognition at the centre of symbolic (nn) meaning ('m-intention', indeed). A convention, though important when involved, only makes for the *possibility* of meaning.[9] But possible meaning is not

---

[8] I am concerned with communication but only insofar as it can held to involve *meaning* (and thus thought and inference). There are of course other kinds of communication—e.g. among honey bees and (chemical) among cells—as Deirdre Wilson reminds me. I am not concerned with these.

[9] As the first two examples of note 8 above show, while all non-natural (representational, ostensive) signs depend on intention, not all depend on convention.

meaning (any more than a 'possible queue' on a motorway is a queue). A red flag actually means nothing to anyone in its box. Even when flying and intended/meant thereby to mean/communicate e.g. that artillery practice is in progress, it won't *actually* mean that if no one (*S*) notices it other than *F*, the person who flew it. I'm not sure it *actually* means that to *F* himself, crucial though its *potential* meaning was to his intention in hoisting it. (And, to *F*'s commanding officer, it might mean only that his order (to *F*, to hoist the warning flag) had been carried out). It is the implementation of a convention on an occasion that makes for actual meaning and then only if some (other) *S* assumes that it was meant/intended to mean/communicate—to her—what it potentially means (rather than, in the case of the ringing bell, as a demonstration in fire practice). As regards 'artillery practice is in progress' *or* 'swimming is forbidden', a flying flag might be said to be 'ambiguous' in virtue of there being distinct conventions governing what it potentially means. But, for that reason, such 'ambiguity' has the same nugatory status as 'possible meaning'. It's not ambiguity—i.e. *actual* ambiguity. Actual ambiguity is an occasion-specific subjective mental state of (inferential, interpretative) indecision of *S* (Burton-Roberts 1994).

There are of course big differences between indexical and symbolic signs. Nevertheless, I see no reason not to say that, when *P* functions as a symbolic sign, *P* 'has meaning' in the same sense as when *P* is an indexical sign. It is the cognitive mechanics that subserve the meaning that differ, (a)–(c) above in the symbolic but not the indexical case. Indexical signs involve just one person (*S*) while symbolic signs involve more than one (*A* and *S*) but that doesn't affect the general fact about meaning. It concerns only whether the meaning involves intention (was meant) and convention. Both communicate an idea to *S*, lead her to have a certain thought. It is true that, with symbolic signs, *someone* (by means of *P*) is communicating rather than merely *something* (*P*) communicating—but it's inferentially dependent communication in both cases.[10]

Symbolic signs also differ from indexical signs in that, when symbolic, the whole rationale of *P* lies in its *being* (intended as) *a sign*. Since *A* arranges for it, the very occurrence of *P* is motivated by that. But that doesn't (or shouldn't) tempt us into thinking that *P*'s status as sign (for *S*) involves any intrinsic property of *P*. Nor would we attribute any truth condition to a flying flag. Truth conditions lie elsewhere. My point is this: impressive though the differences are, *P* as a symbolic sign 'has meaning' but it no more has semantic properties than any indexical sign.

---

[10] Sperber and Wilson (1986/1995, 22–23) deny this. Their discussion involves an example of the ostensive use of an indexical sign: 'Suppose that Mary intends to inform Peter of the fact that she has a sore throat. All she has to do is let him hear her hoarse voice…' (22). Of this they write (23) 'This should not be regarded as a form of communication'—a denial that must apply, *a fortiori*, to (non-ostensive) indexical signs as such. For myself, I see no reason to deny, if Peter is led to think that Mary has sore throat from hearing her hoarse voice (whether or not by speaking in that voice she intended to inform him of it), that her hoarse voice (*P*) does *communicate* to him that she has a sore throat.

Indeed, what makes for meaning with symbolic signs is, if anything, doubly extrinsic to P.

Subject to sorting out issues with 'word' (below), I see no reason not to say that words are symbolic signs—and thus that all the above applies to them. In some quarters, that's uncontroversial. But Grice wrote '…the distinction between natural and non-natural meaning is, I think, what people are getting at when they display an interest in a distinction between "natural" and "conventional" signs. But I think my formulation is better. For some things which can mean$_{nn}$ something *are not signs* (e.g. *words are not*)' ([1989](), 215, my emphasis). In the absence of any explanation or argument, I can only speculate why Grice (or anyone) should deny that words are signs. He was interested in Peirce's semiotics (Chapman [2005](), 71) and opened 'Meaning' ([1989](), 213–223) by discussing indexical and symbolic signs (though 'Peirce', 'semiotics', 'indexical' and 'symbolic' make no appearance). So, to speculate, could Grice's denial have been in aid of driving a wedge between (semiotic) meaning in general and word meaning? This would be consistent with the semantic 'corral' mentioned earlier. For, if the denial was motivated by the assumption that *words have meaning in virtue of having semantic properties* ('having meaning' in a literal, objective, property sense), that certainly would distinguish their meaning from that of signs generally, as I have sought to show. But, widespread (well-nigh universal) though it is, that assumption (words are unique in 'having meaning' in virtue of having semantic properties) has actually never been defended explicitly. In the absence of explicit supporting argument, a distinction between 'meaning-as-semantics' and semiotic meaning (the former 'linguistic' but not explained beyond that) is both conceptually profligate and uninformative.

A more persuasive reason for denying that words are signs is this. At least as treated thus far, signs are perceptible phenomena ('P' above), but words are not perceptible phenomena. Hence words are not signs. I'm prepared to accept that— on one interpretation of 'word'. However, the relevant perceptible phenomena here are the phonetic phenomena we think of as 'uttered words'. The question is: When do such phenomena count as uttered words? And, since 'uttered word' suggests the existence of words independent of and prior to utterance, what *is* a 'word' anyway?

The CGG (and RT) answer is that a word is a syntactic object constituted by phonological and semantic properties. But I doubt whether proponents of this answer really believe that what we physically utter literally has syntactic and semantic properties. It's agreed, surely, that what speakers utter are sounds—brute sounds. In which case, what speakers utter can no more have semantic properties than the pattering of rain or a starter's gunshots can.

Phonetic phenomena functioning as linguistic signs involve a phonology (more strictly, a morpho-phonology). But I don't want to say (and I don't think phonologists do) that uttered sounds themselves have phonological properties. Actual sounds counting as 'linguistic' are *implementations* of a phonology. What a phonology does is license the production of phonetic phenomena (actual sounds). That much is uncontroversial. But in what sense does a phonology 'license'

sounds? I suggest it licenses them for use as symbolic (representational) signs.[11] It is in virtue of this, not the possession of semantic properties, that we might talk of linguistic 'sound with a meaning'.

On this account, then, a word is a phonologically constituted license for the use of sounds as symbolic signs. A word is not an object (with semantic properties) but a phonologically constituted semiotic license. It's a rule, if you like. As such, words are indeed not signs and don't themselves 'have meaning'. They make for the *possibility* of signs and thus the possibility of meaning. *P* counts as an 'uttered word' and thus an actual sign when assumed to be the physical implementation of a phonologically-constituted semiotic license.

This licensing is subject to the cognitive conditions (a)–(c) above and all the provisos discussed there. Condition (a) made reference to 'convention'. The above amounts to the suggestion that a word is a phonologically constituted convention. By 'convention' I simply mean a relational locus of (non-natural, symbolic) Saussurean arbitrariness. I don't mean it is conventional in the sense of existing in a supposedly objective 'public language' external to individuals [c.f. Sperber and Wilson's (1998) 'public words']. As Chomsky argues (e.g. 1986), there are no such public languages ('E-languages'). Conventionality in this 'public' sense, insofar as it exists as a linguistic phenomenon, is rooted in the personal assumptions of individuals (Pateman 1987). These are *I-linguistic* assumptions, collectively amounting to an I-assumption about others' words—effectively, the I-assumption that others implement the same conventions as 'I'. We notice this in inferential interpretation consciously only when faced with its fallibility. A famous fictional example (from Sheridan's *The Rivals*) is Mrs Malaprop's "Sir, you are the pineapple of politeness!" Here the need for inference regarding her I-language (her I-conventions/licenses) and thus her I-assumptions about 'the public language' is more apparent (intentionally so, it being fictional) and it is generally successful (Davidson 1986).

What does *pineapple* mean? The 'semantic' answer, of supposedly objective necessity, runs: '*pineapple* means PINEAPPLE'. But this is either vacuous ('*pineapple* means what it means') or—if it implies that *pineapple* always and for everyone means the same thing—simply false. There is no literally 'objective/public' fact here, i.e., no fact independent of particular I-assumptions and inferences about others' I-assumptions and intentions on occasions, however well-evidenced those assumptions might seem to be. Nor is there any call here for talk of 'a coming apart of speaker meaning and linguistic meaning' (Carston 2002, 18). What is the use of words—including Mrs Malaprop's—if not linguistic? Words mean what speakers mean by them on occasions of use (Recanati 1998)—and that in the case of Mrs Malaprop's *pineapple* is what I, and I believe others, mean by *pinnacle*.

---

[11] This is the 'representational' (as against 'realizational') conception of phonology proposed by the Representational Hypothesis (e.g. Burton-Roberts 2000, 2011; Burton-Roberts and Poole 2006).

More relevant here, because empirically (non-fictionally) more widely attested, are certain uses of *disinterested*, *infer*, *refute, antisocial*—uses in which they mean what I, and I believe some others, generally mean by *uninterested, imply, reject, unsociable*. But who am I (or anyone) to damn these as 'misuse'—let alone imply they are not 'linguistic'—especially when, on the evidence of my own under-standing of them (given my I-appreciation of their prevalence), they are part of my own I-language?

As work in RT has vividly shown, even when other speakers' uttered words appear to be consistent with our own I-assumptions about 'standard (normal/public) meaning', they seldom if ever actually mean that. See e.g. Sperber and Wilson (1998) on *tired*. Furthermore, as Carston's (2002, Ch. 5) discussion of *open* and *happy* vividly demonstrated, it is not even clear we could actually grasp such 'standard' (objectively encoded) meanings. Meaning (actual meaning) just *is* personal, neither sub-personal nor supra-personal (public). A case in point is "He was upset but not upset", offered by a witness in O.J. Simpson's trial, cited by Carston (2002, 324 and 2004). But Carston's (2004, 839) discussion of it illustrates what I have been questioning. She writes 'As far as linguistically supplied information goes, this is a contradiction, a fact that presumably must be captured somewhere within a semantics for natural language'. I have been urging, in effect, that there is no fact here to be captured within a semantics for 'English', let alone for 'natural language'. What there undoubtedly is here includes the articulatory/acoustic fact of the double occurrence of *upset*. But only a firm commitment to the double-interface [phon + sem] assumption would suggest that this phonetic fact goes, of objective necessity, hand-in-hand with some unique semantic fact. Car-ston herself points out that what's intended and recognised—i.e. the actual meaning—is not contradictory, representing this by her distinction 'UPSET*' versus 'UPSET**'. It is this, as RT itself argues, that's ('really') semantic.

'Convention' needs more discussion than space allows. I'll mention just one issue in this connection. With convention comes 'encoding'. As I have discussed elsewhere (2005, 2007), RT operates with what I've called a 'constitutive' notion of encoding, according to which the encoded meaning of a word is a constitutive semantic property of it, deterministically decoded not inferred (Carston 2002, 322–323), consistent with CGG's Saussurean double-interface assumption. By 'encoding' I mean something different, briefly contemplated by Carston (2002, 363) but not pursued. In illustration of what I mean, and to keep things very simple, take Morse code, where (the convention is that) 'dot-dash' encodes THE LETTER 'A'. The whole point of Morse code is that 'dot-dash' *is not* the letter 'A'. It has none of its properties. It is, quite distinctly, a sign for—a pointer to, a representation of, indeed 'means'—THE LETTER 'A'. I accept there is 'encoded meaning' only in this latter (relational, semiotic, non-constitutive) sense, distin-guishing what's-encoded (X) from what-encodes-it (E(X))—subject always to the above on conventions and in no sense that admits of deterministically decoded, as distinct from inferred, meaning.

I have in effect been arguing with Fodor (1998, 9) that 'English has no semantics'. But Fodor immediately follows that with 'Learning English isn't

learning a theory about what its sentences mean, it's learning how to associate its sentences with the corresponding thoughts'. There is much I would question in this but what concerns me here is the implied equation between (not)-having-semantics and '(not)-having meaning'. If the claim that 'English has no semantics' is to be sustained with any plausibility, we need to avoid any suggestion that what doesn't have semantics doesn't 'have meaning'. I hope I've shown that this does not follow, given a distinction between meaning and semantics. On the contrary, as I will argue in the next section, 'learning how to associate [α]s with…thoughts' *precisely amounts to* 'learning…what [α]s mean'. Jerry Fodor (p.c.) has dismissed this as terminological. Nevertheless, it seems to me important if 'English has no semantics' is to command assent.

3.2 Relating Meaning and Semantics

I have sought to show that meaning is not a *property*. More particularly, it is not a semantic property. Meaning and semantics must be distinguished. I think it follows clearly from Sect. 3.1 that meaning is a *relation*. Talk of 'sound-meaning relations' (or 'relating sound and meaning') has always made me uneasy. What exactly was sound supposed to be related *to*? And what kind/species of relation was it supposed to be? Unease is dispelled in recognising that such talk treats meaning as one term (one of the *relata*) of a relation when it is in fact *the relation itself*.

Meaning is a (the) cognitive, antisymmetric semiotic relation *from* something (α) *to* something else (X). As I show below, α can be anything, not just a mind-external phenomenon P. But for a relation to be *semiotic* (i.e. for it to be meaning) X must be a thought T. (It can be just a concept, as in the Morse code example, but it's thoughts I'm mainly concerned with.) For α to actually mean X to S is for α to communicate that-X to S—i.e. α leads S to entertain T, the thought that-X.

As noted, we're tempted to say that α 'has' meaning and to talk of 'α's meaning'. But notice we're equally tempted to say α's meaning 'lies in' the thought, that the content of T *is* the meaning of α. Well, meaning can't both be a property of α and lie in T. The quandary is explained and resolved if meaning lies neither in α nor in T but in their relation.

Thoughts are where semantics enters the picture. Thoughts I assume are conceptually constituted. Thought involves concepts syntactically complex enough to be *entertained as* representations i.e. those that can be objects of propositional attitudes. An actual thought is one that *is* so entertained. I hold that it is concepts (including those that can be entertained as thoughts) and only concepts that have semantic content.

The relation between meaning and semantics, then, is this. Meaning is a *relation to* semantics—an antisymmetric semiotic relation from α (anything) to conceptual/semantic content. Since semantic content is necessarily one of the terms of the semiotic/meaning relation, it follows that you can't have meaning without semantics. But it doesn't follow that meaning *is* semantics; this relational

account of meaning distinguishes meaning and semantics. As I argue below, even—indeed especially—when we want to say something *both* 'has meaning' *and* has semantic content, we still need to distinguish. Nor does it follow, on the assumption that linguistic expressions 'have meaning', that semantic properties are linguistic, if by 'linguistic' we mean pertaining-to-languages rather than pertaining-to-the-language-of-thought. No-meaning-without-semantics holds across the board: it goes as much for (symbolic) linguistic signs as for non-linguistic signs, be they symbolic or indexical. The semantics of a given concept/thought is constitutive of and only of that concept/thought. It is not the semantics of anything else.

P is an actual linguistic sign (and as such 'has' actual meaning) when it counts as an uttered word or (temporal, linear) concatenation of words—that is, when *P* is produced with the intention of implementing a morpho/phonologically constituted semiotic I-license in aid of representing the syntactically (hierarchically) structured semantic content constitutive of a particular thought. *P* does not, in virtue of *representing* such content/properties, have such content/properties itself. That's why *P*s that count as sequences of uttered words require 'parsing'. I take parsing to be a matter of putting what *lacks* syntactico-semantic properties into (semiotic, representational) 'correspondence' with what *has* such properties.

A brief (all too brief) Fodorian digression is needed here. I have appealed to some notion of a 'language of thought' but how much of the above is Fodorian I hesitate to say. I had better put on record that I depart from Fodor's account of semantic/conceptual 'content'. This, again, involves property vs. relation. His account of conceptual 'content' is externalist and relational—'content is constituted, exhaustively, by symbol-world relations' (Fodor 1998, 14). This is not what I mean. By conceptual 'content' I mean a (indeed the individuating) property of a concept.[12] This is an internalist constitutive notion of conceptual content—and nativist, implying that what you acquire is not a concept but a certain kind of *access* to a concept (from worldly and/or linguistic experience). However, this may be 'terminological', since Fodor himself allows that concept-world relations are determined by something 'mind-dependent', which I assume has to do with the concept itself. 'Being a doorknob is just: striking our kinds of minds in the way that doorknobs do' (Fodor 1998, 162). This seems to me to allow, if not demand, that the concept DOORKNOB has some kind of internal constitutive property. This for me is its 'content', determining non-arbitrarily what external phenomena it 'locks onto' and thereby makes sense of. Concept-world relations (semantics in an externalist, relational sense), it seems to me, arise when a concept sufficiently complex to be entertained as a representation actually *is* so entertained. I think this amounts to saying that the distinction-and-relation between internalist and externalist semantic content is the distinction-and-relation between the language of thought and actual thoughts. End of digression.

---

[12] Fodor (ibid.) comments that his atomistic theory of concepts 'allows' him 'not to hold that one's inferential dispositions determine the content of one's concepts'. It seems to me that it forces him not to hold that and forces the relational account of 'content'. Incidentally, on atomism in RT, see Burton-Roberts (2005, 339).

I earlier undertook to show that, even when something can be said to have both semantic content *and* 'have' meaning, the two must be distinguished. At issue here is whether thoughts (concepts), in and of themselves, 'have meaning'. I've assumed thoughts have semantic/conceptual content; but do they, thereby (in virtue of just that fact), 'have meaning'? My answer is 'No'. I don't recall this question ever having been asked but, intuitively, that seems to me the right answer. Independently of that, my answer must be 'No' given what I'm claiming meaning is. It is signs that 'have meaning' but thoughts are not, in themselves, signs. It is not necessary, not *definitional* of a thought, that it leads you to have a thought. The only sense in which that could be definitional would be if a thought led you have that very thought. Self-causing thoughts are out, I assume. Your having thought *Tj* does not communicate *Tj* to you. In short, semantic content is not, in itself, semiotic (not meaning).

However, although I am claiming that thoughts in and of themselves don't 'have' meaning (are not in and of themselves signs), I am not denying that, nevertheless, thoughts *can* 'have meaning', i.e. can function as signs in the mental life of an individual. They generally do. (It is this that motivated my earlier move from 'P', for external phenomena, to 'α', for anything.) The point is, though, that this 'meaning of' a given thought T, for S, is distinct from its semantic content. The idea is simply this: having one thought T_1 can lead you to have/entertain *another* thought T_2 whose semantic content may be entirely distinct from that of T_1. Having T_1 can communicate to you (the distinct content of) T_2. This strengthens the distinction between semantics and meaning. (The scare quotes around 'have (meaning)' and 'meaning of' are a reminder that the meaning in fact lies in the relation $T_1 \rightarrow T_2$.)

I illustrate this below but I must first address an issue I've not attended to. Let's say Susan (S) sees something α in the kitchen which she takes to be a pile of clean washing. This needs unpacking: α is a visual *stimulus* and S's first thought (T_1) on seeing α is THAT'S (or LO!) A PILE OF CLEAN WASHING. What I'm calling 'first thoughts' are, in psychological terminology, *percepts*. My previous use of 'P' (for phenomenon) obscured the distinction between stimulus and percept (the result of mentally processing the stimulus). Thereby it slid over the issue of whether mere stimuli are signs. Is [stimulus $\rightarrow$ percept] a semiotic relation? In other words, is it meaning? Since I've described the percept as a 'first thought' (T_1), you will have guessed I want to say it *is* meaning. Given how S is internally constituted, the stimulus leads S to have a percept (P for 'percept' now), an inferentially-derived mental representation. P is a conceptual *interpretation* of that stimulus. Without percepts there's no meaning, just meaningless stimuli. The inferential move from α to SMOKE(α) is different from, but not different *in kind* from, the move from SMOKE(α) to CAUSED-BY-FIRE(α). If [stimulus $\rightarrow$ percept] is excluded from (on grounds of being in some sense prior to) any semiotic move, where do we stop? Is CAUSED-BY-FIRE(α) a percept? I'm bothering with this because it might be felt that treating the [stimulus $\rightarrow$ percept] relation as meaning stretches 'meaning' too far, thereby undermining my account. I want to

say it *is* meaning and, given the generality of my project for 'meaning', don't regard this as undermining.

So, Susan's first thought is THAT'S (or LO!) A PILE OF CLEAN WASHING (T_1). Let's say that, given Susan's cognitive context, C_S—her current mental state, her projects and preoccupations—T_1 leads her to have (communicates to her) another thought T_2: JOHN HAS DONE THE WASHING THAT NEEDED DOING. (If T_1 is true then, given C_S, T_2 is true; T_1 by assumption is true, hence T_2 is true.) Now let's say, given C_S, T_2 leads to T_3: I DON'T NEED TO DO ANY WASHING RIGHT NOW and T_3 in turn to T_4: I CAN FINISH MY LECTURE RIGHT NOW. Susan will almost certainly go further—subject to constraints expounded by relevance theory—but I'll stop there. The stimulus and each of T_1, T_2 and T_3 are *signs* for S, given C_S. The 'meaning of' $T1$ for S (and, I claim, 'of' the stimulus itself) ultimately and indirectly 'lies in' T_4. The relations stimulus $\rightarrow T_{1...} \rightarrow ...T_4$ are semiotic. They are meaning relations. But they are clearly not semantic. The stimulus has no semantics and T_1, T_2, T_3 and T_4, are semantically unrelated. The point of this illustration has been to show that even—indeed especially—in the case of what does have semantic content (namely thoughts and only thoughts) *and* 'has' meaning, meaning and semantics are distinct.

As an aside, it's reasonable to ask what *kind* of semiotic relation holds between those thoughts of Susan's. Since it involves neither convention nor semiotic intention, it is clearly not symbolic (not *nn*), not representational. Although we're dealing with an agent, S, the relevant causations of thought are not agentive. So it seems it must be indexical (natural). I reconcile myself to this conclusion as follows. We saw, with our original examples of natural (indexical) signs, that their status as signs depends on a given subjective background. The same holds in the above illustration. The difference is that our original examples were assessed against subjective backgrounds assumed to be shared as (supra-personal) 'general knowledge' and taken for granted as such, whereas in the illustration the background (C_S) is personal, with no claim to be 'general knowledge'. Here what counts as 'general knowledge' is irrelevant. What's relevant is Susan's pre-existing personal projects/preoccupations. Although we, as 'observers' of Susan, won't take C_S for granted, Susan does—necessarily, it being her own current state of mind.

This is not to say that relevant aspects of C_S can't be shared. John may have developed some appreciation of those aspects of C_S. They may be mutually manifest to Susan and John given previous conversation between them. In which case, John may have left the pile of washing in a prominent place in the kitchen ostensively, i.e. with the (semiotic, communicative) intention of leading Susan to have T_2 and intending her further to derive T_3 and T_4. Given C_S, Susan may derive those thoughts whether or not she recognises John's intentions. However, following an account by Susan of her preoccupations, John might instead utter "I've done the washing". The meaning 'of' that utterance/stimulus 'lies in' the semantic content of S's thought T_2 (actually it lies in $U \rightarrow T_2$.) In that case, if Susan recognises his intention, John might be said to have conversationally implicated T_3

and T_4. As a relation, conversational implicature is clearly a meaning relation (i.e. a semiotic relation) but equally clearly (indeed by definition) not a semantic relation.

Incidentally, John might instead march proudly in with the clean washing and simply utter "Done!" with the same effect. Consider also "Kitchen?" in the context of a burning smell or "I've lost my keys!" You don't need 'sentences' (Stainton 2006). The relevance of non-sentential utterances to my general theme is that 'Logical Form' as a semantic level of *linguistic* representation cannot play a (bottom-up) part in their interpretation, since it seems to me that you could only know which LF to assign to them in the light of the thought you take to have communicated by them (i.e. top-down). This makes LF as a level of linguistic representation redundant in utterance processing.

4 Conclusion

I have made a case for thinking of meaning as a relation, entirely extrinsic to the terms related in it. It is a cognitive, semiotic, antisymmetric relation *from* something (potentially anything a subject is aware of, be it an external phenomenon or a thought) *to* semantic content. Semantic content is of-and-only-of thoughts, couched in an internal 'language of thought'. As a relation having semantics as one of it terms, meaning is to be distinguished from (though necessarily related to) semantics. Linguistic expressions, on these terms, can be said to 'have meaning' without being attributed semantic properties (including 'LF' properties). But even that is not quite right, I argued, because linguistic expressions only make for the possibility of meaning and do so only in a given I-language (not in any supposedly objective public language). Actual meaning is personal (neither sub-personal nor supra-personal), specific to the occasion/context of utterance. The scope of pragmatics (contextual inference in interpretation) thus extends into what is often taken to be 'linguistic semantics' and decoding.

References

Brody, M. (1995). *Lexico-logical form*. Cambridge: MIT Press.

Burton-Roberts, N. (1994). Ambiguity, sentence and utterance. *Transactions of the Philological Society, 92*, 179–212.

Burton-Roberts, N. (2000). Where and what is phonology? A representational approach. In N. Burton-Roberts et al (Eds.) *Phonological knowledge* (pp. 39–66). Oxford: Oxford University Press.

Burton-Roberts, N. (2005). Robyn Carston on semantics, pragmatics and 'encoding'. *Journal of Linguistics, 41*, 389–407. (Review of Carston 2002).

Burton-Roberts, N. (2007). Varieties of semantics and encoding: Negation, narrowing and numerals. In N. Burton-Roberts (Ed.) *Pragmatics* (pp. 90–114). Houndmills: Palgrave Macmillan.

Burton-Roberts, N. (2011). On the grounding of syntax and the role of phonology in human cognition. *Lingua, 121*(14), 2089–2102.

Burton-Roberts, N., & Poole, G. (2006). 'Virtual conceptual necessity', feature dissociation and the Saussurean legacy in generative grammar. *Journal of Linguistics, 42*, 575–628.

Carston, R. (2002). *Thoughts and utterances*. Oxford: Blackwell.

Carston, R. (2004). Explicature and semantics. In S. Davis & B. Gillon (Eds.), *Semantics: A reader* (pp. 817–845). Oxford: Oxford University Press.

Chapman, S. (2005). *Paul Grice: Philosopher and linguist*. Houndmills: Palgrave Macmillan.

Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.

Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.

Chomsky, N. (2000a). Language and interpretation. In N. V. Smith (Ed.), *New horizons in the study of language and mind* (pp. 46–74). Cambridge: Cambridge University Press.

Chomsky, N. (2000b). *The architecture of language*. Oxford/New Delhi: Oxford University Press.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.) *Structures and beyond: The cartography of syntactic structures* (Vol. 3, pp. 104–131). Oxford: Oxford University Press.

Chomsky, N. (2006). Approaching UG from below. Ms, MIT. In U. Sauerland, & H.-M. Gartner (Eds.) *Chomsky's minimalism and the view from syntax-semantics* (pp. 1–29). Berlin/New York: Mouton de Gruyter (Appeared (2007).

Davidson, D. (1986). A fine derangement of epitaphs. In E. Lepore (Ed.), *Truth and interpretation: Perspectives on the philosophy of Donald Davidson* (pp. 433–446). Oxford: Blackwell.

De Saussure, F. (1974) *Course in general linguistics*. Translated from the French by Wade Baskin. Revised edition. Fontana books.

Fodor, J. (1975). *The language of thought*. New York: Crowell.

Fodor, J. (1983). *Modularity of mind*. Cambridge: MIT Press.

Fodor, J. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Clarendon Press.

Fodor, J. (2008). *LOT2*. Oxford: Oxford University Press.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press.

Larsen, R., & Segal, G. (1995). *Knowledge of meaning*. Cambridge: MIT Press.

Pateman, T. (1987) *Language in mind and language in society*. Oxford: Oxford University Press.

Peirce, C. S. (1933). *Collected papers*. Cambridge: Harvard University Press.

Recanati, F. (1998). Pragmatics. In *Routledge Encyclopedia of Philosophy* (Vol. 7, pp. 620–633). London: Routledge.

Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. Oxford: Basil Blackwell.

Sperber, D., & Wilson, D. (1998). The mapping between the mental and the public lexicon. In P. Carruthers and J. Boucher (Eds.) *Language and thought: Interdisciplinary themes* (pp. 184–200). Cambridge: Cambridge University Press.

Stainton, R. (2006). *Words and thoughts: Subsentences, ellipsis, and the philosophy of language*. Oxford: Oxford University Press.

Wedgwood, D. (2007). Shared assumptions: Semantic minimalism and relevance theory. *Journal of Linguistics, 43*, 647–681.

Clinical Pragmatics and Theory of Mind

Louise Cummings

Abstract Theory of mind (ToM) describes the cognitive ability to attribute mental states both to one's own mind and to the minds of others. In recent years, ToM has been credited with playing a significant role in developmental and acquired pragmatic disorders. In this way, ToM deficits have been linked to pragmatic deficits in individuals with autism spectrum disorders (e.g. Martin and McDonald, *Journal of Autism and Developmental Disorders, 34*, 311–328, 2004), emotional and behavioural disorders (e.g. Buitelaar et al., *Development and Psychopathology, 11*, 39–58, 1999), intellectual disability (e.g. Cornish et al., *Journal of Intellectual Disability Research, 49*, 372–378, 2005), right-hemisphere damage (e.g. Winner et al., *Brain and Language, 62*, 89–106, 1998), schizophrenia (e.g. Brüne and Bodenstein, *Schizophrenia Research, 75*, 233–239, 2005), traumatic brain injury (e.g. McDonald and Flanagan, *Neuropsychology, 18*, 572–579, 2004) and neurodegenerative disorders such as Alzheimer's disease (e.g. Cuerva et al., *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 14*, 153–158, 2001). In this chapter, I examine the central role of ToM reasoning in utterance interpretation. The chapter addresses what is known about ToM development during childhood and adolescence as well as changes in ToM skills as part of the aging process. The role of ToM in developmental and acquired pragmatic disorders is discussed. The contribution of ToM research into pragmatic disorders is critically evaluated. Finally, several ToM theories are examined. The question is addressed of which, if any, of these theories is able to capture the pragmatic features of utterance interpretation.

L. Cummings (✉)
Nottingham Trent University, Nottingham, UK
e-mail: louise.cummings@ntu.ac.uk

1 Introduction

The study of pragmatic disorders is an important area of work in the clinical communication sciences and an increasingly recognised sub-discipline within linguistic pragmatics (Cummings 2010, 2014a). This study, known as clinical pragmatics, has progressed to the point where certain trends or patterns are discernable. One of the most evident trends has been the pursuit of clinical studies of pragmatics simply because they can be done rather than because they are addressing theoretically significant questions. The lack of a theoretical rationale for many clinical studies has resulted in an abundance of research findings, a substantial number of which throw little light upon the disorders they purport to examine (Cummings 2005, 2007a, b, 2008, 2009, 2011, 2012a, b, 2014a). A recent counter-trend has been the attempt to explain pragmatic disorders in terms of underlying cognitive factors. This trend has largely developed in response to the observation that many populations in which there are marked impairments of pragmatics—adults who sustain a traumatic brain injury, for example—also exhibit significant cognitive deficits. Amongst the cognitive factors examined with respect to pragmatic disorders are deficits of executive functions such as planning and problem solving, visuospatial processing, the generation of inferences and a type of cognitive processing characterised by weak central coherence (see Chaps. 2 and 3 in Cummings 2009). Although studies of these cognitive factors are becoming increasingly common, one factor has received more investigation than all others. This is the ability to attribute mental states to the minds of others, and to use one's knowledge of these states to predict a person's behaviour, a cognitive skill which has become known by the term *theory of mind*. This chapter will examine this key cognitive skill and assess its role within pragmatic disorders in children and adults.

In the preface of their book *Understanding Other Minds*, Baron-Cohen et al. (2000) capture the main tenets of theory of mind (ToM). The developing child is involved in the construction of a theory about the contents of his own mind and the minds of others, a theory which is used to understand and predict the actions of those around him. These contents are none other than mental states such as beliefs, knowledge and desires:

> During the first few years of life, children acquire an understanding of the relations between their own mental states, the world (particularly, the social world), and action. They use this to understand themselves and others. Without obvious effort or formal instruction they learn that other people, just like themselves, have minds and that the behaviour of others, just like their own, reflects their knowledge, thoughts, beliefs, and desires. In the scientific literature, this has been called the child's acquisition of a 'theory of mind'. This term underlines the intellectual achievement of a theory upon which a child can rely (Baron-Cohen et al. 2000: v).

One mental state not mentioned by Baron-Cohen et al., but which is integral to pragmatic interpretation, is intentions. Since Grice set out in his William James Lectures at Harvard University to characterise the exchange of utterances between interlocutors in terms of the communicative intentions that motivate those

utterances, theorists have appreciated that pragmatic interpretation involves cognitive skills that are quite distinct from those employed during the linguistic decoding of utterances. Specifically, pragmatic interpretation requires an ability to establish the mental states, and particularly the communicative intentions, of other participants in communication. Even the briefest acquaintance with utterance interpretation reveals that this is the case. Imagine a scenario in which two friends, Jane and Mary, are having an evening meal together. Jane has spent all afternoon preparing a lamb stew. Mary does not like lamb but nevertheless forces herself to eat the serving of stew that Jane has put on her plate. When their meal is complete, Jane asks Mary 'Would you like more stew?', to which Mary replies 'I would like to leave room for dessert'. Of course, Mary is implicating that she does not want more stew. This implicature can only be recovered by Jane to the extent that she is able to make certain inferences about Mary's mental states. Jane must be able to infer, for example, that Mary believes, and believes that Jane believes, that in order to eat dessert, one must have some remaining appetite. As well as accounting for Jane's ability to recover the implicature of Mary's utterance, ToM also explains why Mary decides to use an implicature to refuse Jane's offer of more stew. It is because Mary conceives of Jane as an agent that can entertain mental states, specifically a range of feelings or emotions, that she decides to decline Jane's offer indirectly by means of an implicature, rather than by issuing the more direct, but less polite response, of 'No!'.

It emerges that pragmatic interpretation draws extensively on ToM skills. Yet, notwithstanding its significance to utterance interpretation, ToM remains a poorly characterised notion in relation to pragmatics in general and pragmatic disorders in particular. In reviewing what is known about the role of ToM in pragmatic disorders in this chapter, several related areas of ToM research must be addressed. To make sense of findings of ToM deficits in developmental pragmatic disorders in Sect. 4, something must be said about the trajectory taken by normally developing children on their way to acquiring a full theory of mind. Developmental studies of ToM have tended to focus on the transition that occurs between 3 and 4 years of age in normally developing children. This is a developmental period during which children move from experiencing failure on tests of false belief (a standard test of ToM) to achieving success on these tests for the first time at around 4 years of age. More recently, ToM research has been extended to examine maturation of ToM beyond the early years. In this way, there is now a small but growing number of studies which are charting ToM skills into adolescence and adulthood, and decline of ToM as part of the aging process. These various ToM developments will be examined in Sect. 3. ToM skills in adulthood and later life provide an important context in which to view findings of ToM deficits in acquired pragmatic disorders in Sect. 5. One of the impediments to an improved understanding of ToM in pragmatic disorders is the poor characterisations of ToM and pragmatic interpretation in clinical studies. We examine some of these characterisations in Sect. 6. Finally, theory of mind is itself the subject of intense theoretical deliberation. We examine theoretical proposals concerning the nature of ToM in Sect. 7. However,

we begin this chapter by saying something more about the role of ToM in utterance interpretation.

2 Theory of Mind in Utterance Interpretation

We have seen that in order to recover the implicature of an utterance, a hearer must be able to attribute mental states, specifically communicative intentions, to the speaker of an utterance. The hearer, who can establish these intentions, is able to determine what a speaker is attempting to communicate which, as we saw in Sect. 1, can be something quite different from what the speaker's utterance literally means. Implicature is not a unique pragmatic concept, however, with respect to the role played by ToM in its recovery. Consider the utterances in (1) to (5) below:

(1) Can you fetch my walking stick?
(2) I plan to leave for Paris tomorrow.
(3) John regretted leaving the team.
(4) Fran helped the old woman across the road. She was very grateful.
(5) What an architectural triumph! (uttered by a speaker upon seeing his son's tree house).

The speaker of the utterance in (1) is using an indirect speech act to request that the hearer fetch the walking stick. The hearer has little difficulty in establishing the illocutionary force of the speaker's utterance because he knows that the speaker knows that the hearer is indeed capable of fetching the stick—let's assume that the speaker and hearer are an elderly resident in a care home and a nurse, respectively. In such a case, the speaker is unlikely to be asking a question about the hearer's ability to fetch the stick. Rather, his utterance is more likely to be motivated by a mental state—a communicative intention—that takes the form of a request to the hearer to fetch the stick. The deictic expressions 'I' and 'tomorrow' in (2) also demonstrate the role of mental state attribution in utterance interpretation. To assign a referent to the pronoun 'I' (person deixis), the hearer must know who is the speaker of this utterance and must be able to attribute that knowledge both to his own mind and to the mind of the speaker. Similarly, the 24-hour period that is the referent of 'tomorrow' (temporal deixis) requires that the hearer attribute to the speaker knowledge of the 24 h henceforth and also possibly a calendar date. The speaker of the utterance in (3), which presupposes that John did leave the team, must also engage in mental state attribution. Specifically, the speaker must attribute to the hearer knowledge that John left the team, an action which the speaker then goes on to inform the hearer was a source of regret for John. Presupposition is a pragmatic mechanism whereby knowledge that is shared by speaker and hearer can be assumed in the asking of a question or the statement of an utterance. Yet, this is only possible in the current case to the extent that the speaker of (3) is aware of the hearer's knowledge of John and specifically the fact that John left the team.

ToM skills also play an important role in assigning a referent to the pronoun 'she' in the utterance in (4). Either of the preceding noun phrases 'Fran' and 'the old woman' is a potential referent of 'she'. What makes 'the old woman' the actual referent of this pronoun, the referent intended by the speaker, is a process of mental state attribution that proceeds as follows. The speaker of these utterances believes that the hearer believes that when a person receives the assistance of others, they usually express gratitude or appreciation. For his part, the hearer attributes similar beliefs and knowledge to the mind of the speaker. Based on their mutual knowledge of each other's mental states, the speaker can proceed to use the non-specific pronoun 'she' in the certain expectation that the hearer will assign to this term the referent intended by the speaker. The irony expressed by the utterance in (5) is only perceptible to a hearer who is both able to see the tree house in question and to make certain inferences about the speaker's mental states. Specifically, the hearer must infer that the speaker of this utterance believes that the tree house is a somewhat shambolic construction and, furthermore, that the hearer also believes this to be the case. The speaker must also be able to make similar inferences about the hearer's beliefs in order to be sure that his utterance will be understood as the ironic remark it is intended to be. The lesson to emerge from these examples is clear. Whether it is indirect speech acts, deictic expressions, presuppositions, pronoun reference assignment or irony, no account can be given of these key pragmatic notions in the absence of the involvement of ToM. This central idea will be returned to time and again in the following sections.

These examples can be used to demonstrate an important distinction in the mental representations that are the focus of ToM research. Some mental states that were attributed to speakers and hearers in the above examples concerned beliefs about events and states of affairs in the world, e.g. the belief that John had left the team in the utterance in (3). Other mental states that were attributed to the minds of the interlocutors in these examples were not beliefs about the world, but beliefs about an interlocutor's beliefs. These two types of mental state attribution highlight a key distinction between first-order and second-order mental representations, respectively. In utterance interpretation, a hearer must not only make inferences about a speaker's beliefs about the world (first-order ToM reasoning), but he must also make inferences about what the speaker believes about his (the hearer's) beliefs (second-order ToM reasoning). It was second-order ToM reasoning that Jane used to establish the implicature of Mary's utterance in Sect. 1, for example. The central role of second-order ToM reasoning in utterance interpretation presents the following challenge for ToM research. Most developmental studies of ToM have tended to focus on the transition that makes it possible for children to pass tests of false belief for the first time at around 4 years of age. However, these tests and the developmental maturation that they assess concern first-order ToM reasoning, typically an actor's false belief about the world. Given that utterance interpretation demands skills in second-order ToM reasoning, it would seem that developmental studies of ToM are failing to target the mentalising skills that are of most significance to the interpretation of utterances. This point is developed further in Sect. 30.2 of Cummings (2014b).

3 Normal Development and Theory of Mind

It has already been stated that normally developing four-year-olds pass tests of false belief, a developmental achievement which indicates that they are in possession of first-order ToM skills. False-belief tests have been at the centre of developmental and clinical studies of ToM. For this reason, the standard format of these tests will now be described. In an early study by Baron-Cohen et al. (1985), the type of scenario typically presented in these tests is characterised as follows:

> There were two doll protagonists, Sally and Anne. First, we checked that the children knew which doll was which (Naming Question). Sally first placed a marble into her basket. Then she left the scene, and the marble was transferred by Anne and hidden in her box. Then, when Sally returned, the experimenter asked the critical Belief Question: "Where will Sally look for her marble?". If the children point to the previous location of the marble, then they pass the Belief Question by appreciating the doll's now false belief. If however, they point to the marble's current location, then they fail the question by not taking into account the doll's belief. These conclusions are warranted if two control questions are answered correctly: "Where is the marble really?" (Reality Question); "Where was the marble in the beginning?" (Memory Question) (1985: 41).

Baron-Cohen et al. presented this scenario, which is an adaptation of Wimmer and Perner's (1983) puppet play paradigm, to three groups of children: 27 normally developing children (mean CA 4;5 years), 20 autistic children (mean CA 11;11 years), and 14 children with Down's syndrome (mean CA 10;11 years). Naming, Reality and Memory Questions were passed by all three groups of children. However, while 85 % of normal children and 86 % of children with Down's syndrome passed the Belief Question, 80 % of autistic children failed this question, a finding that was highly significant. The four autistic children who passed the Belief Question had chronological ages from 10;11 to 15;10 years. Clearly, the autistic children had a severe deficit in first-order mental state attribution which could not be accounted for by their reduced verbal and nonverbal mental ages (the verbal and nonverbal MAs of autistic subjects were higher than those of the Down's syndrome subjects in the study).

The ToM achievement of the normally developing children in Baron-Cohen et al.'s study is certainly one of the most significant findings in ToM research. However, it is preceded and followed by other ToM developments, about which much is now known. The ability to attribute beliefs and knowledge to the minds of others emerges after children have already acquired skills in manipulating a range of other mental states including desires, emotions and pretence. Ruffman et al. (2002) found that children's desire talk preceded their talk about beliefs. In a study of the reactions of toddlers aged 18–24 months to negative emotion displays, Jenkins et al. (1995) reported spontaneous comments by some children which indicated that they appreciated the emotional states of the actresses who participated in the displays. For example, during the sadness display, one child said 'lady not happy'. There is evidence that infants can appreciate pretence from around 15- or 16-months of age (Bosco et al. 2006; Onishi et al. 2007). It is now known that behaviours which first become evident in infancy serve as developmental

precursors to these early skills in recognising and attributing mental states. One such behaviour is joint attention. Charman et al. (2000) found that joint attention behaviours at 20 months were longitudinally associated with theory of mind ability at 44 months in a group of 13 infants. Similar findings are reported by Nelson et al. (2008) in a study of the joint engagement experiences of toddlers. Other factors which have been found to be facilitative of early ToM development include the presence of older siblings, maternal use of mental state language, ability to engage in fantasy and pretend play and early language skills (Ruffman et al. 1998, 2002; Adrián et al. 2007; Taylor and Carlson 1997; Watson et al. 2001).

   Although normally developing children pass first-order ToM tasks for the first time around 4 years of age, younger children can also pass these tasks under certain conditions. Some of the conditions under which 3-year-olds can pass false belief tests include downplaying the salience of the real state of affairs or making salient the prior mental state of the actor in the scenario (both of which encourage the child to identify the actor's false belief), the child being actively engaged in deceiving the target person, overlearning the key features of the false belief narrative or phrasing the false belief question in certain ways (Wellman and Lagattuta 2000: 25). False belief performance can also deteriorate under certain conditions. Symons et al. (1997) found that the integration of caregiver figures into false belief location tasks did not result in the same age-related improvements in false belief performance that occurred during object identity and object location tasks. These investigators argued that the developing awareness of the minds of others in five-year-olds and the emotional content of the task may have interfered with false belief performance in the caregiver condition. Even amongst normally developing children, investigators have observed considerable individual differences in ToM development and performance on false belief tasks. Some of the factors that have been examined with a view to explaining these differences include a child's language ability and executive functions such as inhibitory control, planning ability and working memory capacity (Carlson and Moses 2001; Carlson et al. 2004; Hughes et al. 2005; Milligan et al. 2007; Mutter et al. 2006). With some exceptions, studies have revealed a correlation between ToM performance on the one hand and language ability and inhibitory control on the other hand.

   ToM development in normally developing children beyond 4 years of age has been the focus of a growing number of studies. Liddle and Nettle (2006) examined higher-order recursive ToM skills in a group of 60 children who were 10 and 11 years of age. As expected, the children in this study mastered first- and second-order ToM problems. However, they performed slightly above chance on third-order problems and at chance on fourth-order problems. There is also evidence that further ToM developments take place in adolescence and even adulthood. Dumontheil et al. (2010) examined the development of ToM into adulthood by administering a computerised task to 177 female subjects in each of five age groups: Child I (7.3–9.7 years), Child II (9.8–11.4 years), Adolescent I (11.5–13.9 years), Adolescent II (14.0–17.7 years) and Adults (19.1–27.5 years). The task required participants to use the perspective of a 'director' and move only

those objects that the director could see. There was an improvement in the performance of this task between the Child I and Adolescent II age groups. Also, the Adolescent II group made more errors than the Adult group, suggesting that ToM use improves between late adolescence and adulthood. Even in adulthood, ToM skills do not remain static. There is now growing evidence that ToM skills undergo decline with increasing age (Pardini and Nichelli 2009; Sullivan and Ruffman 2004). Maylor et al. (2002) examined understanding of ToM stories in young, young-old and old-old age groups (mean ages 19, 67 and 81 years, respectively). The performance of the old-old age group on these stories was significantly worse than the other age groups across all conditions in the study (e.g. memory load present/absent). This age deficit remained significant even after measures of vocabulary and executive functioning, and processing speed were taken into account.

4 Developmental Pragmatic Disorders and Theory of Mind

Pragmatic disorders which have their onset in the developmental period have been the focus of considerable clinical investigation. Notwithstanding this intensive study, few general statements about the nature and extent of these disorders are possible. This is related in large part to the fact that the children who exhibit these disorders form a clinically heterogeneous group. It includes most notably children with an autism spectrum disorder, a neurodevelopmental disorder which has profound implications for the development of language in general, and pragmatic language skills in particular. Pragmatic impairments are also a feature of developmental language disorders such as specific language impairment. In some cases, these impairments appear to be secondary to deficits in structural language skills. But in at least one developmental language disorder—pragmatic language impairment—there is evidence that pragmatic impairment is primary in nature. Children who exhibit intellectual disability have difficulty acquiring the pragmatics of language. In some cases, this difficulty is commensurate with delays in acquiring phonological, syntactic and semantic aspects of language. In other cases, pragmatics is more or less impaired than aspects of structural language. Recently, investigators have begun to document the language and communication skills of children with emotional and behavioural disorders including attention deficit hyperactivity disorder, conduct disorder and selective mutism. Within this small, but growing area of clinical study, researchers have started to investigate the pragmatic disorders of these children. An examination of this as yet undeveloped area of clinical study will complete our survey of developmental pragmatic disorders.

Of course, a pragmatic disorder can result from a number of linguistic and cognitive factors which may act separately or in combination. A child with specific language impairment may be unable to produce indirect speech acts such as the utterance in (1) above because he or she lacks the requisite syntactic skills to

achieve inversion of the subject pronoun and auxiliary verb in this utterance. Alternatively, the child with poor impulse control may inappropriately initiate conversations or contribute utterances that overlap with the turns of others. While in the former case, a linguistic factor is central within the aetiology of the pragmatic disorder, in the latter case a cognitive factor predominates. The cognitive factor that will be examined alongside the pragmatic disorders in this section is theory of mind. As the discussion of Sect. 2 demonstrates, there is reason to believe that ToM will play a significant role within these disorders. However, we will see that the exact nature of that role is still somewhat difficult to determine on the basis of research that has been conducted to date. While studies have revealed significant associations or correlations between ToM functioning and various pragmatic skills, there is still considerable debate about what these correlations reveal. For example, rather than indicating a direct causal role for ToM in pragmatic disorders, these correlations may simply reflect the influence of a third variable (e.g. an executive function) on both ToM and pragmatics. This issue is raised, not with a view to addressing it in this chapter, but in the expectation that researchers in the area will increasingly regard it as a question of true theoretical significance. Certainly, discussions of the causal or other role of ToM in the impairments found autism are already well developed (see Chap. 4 in Cummings 2009).

The ToM and pragmatic impairments of children and adults with autism spectrum disorders (ASDs) have been extensively documented. In relation to ToM, studies have revealed that subjects with ASD display impaired understanding of the perception-knowledge relationship (Lind and Bowler 2010), have diminished awareness of their own and others' intentions (Williams and Happé 2010), and have impaired visual perspective taking (i.e. knowledge that different people may see the same thing differently at the same time) (Hamilton et al. 2009). Individuals with ASD also have difficulty inferring complex emotions and mental states in social contexts and from nonverbal social cues (Golan et al. 2008; David et al. 2010) as well as from faces and voices (Golan et al. 2006; Kleinman et al. 2001; Rutherford et al. 2002). An equally wide-ranging set of findings regarding the pragmatic skills of individuals with ASD is also available. Subjects with ASD have been found to have difficulty comprehending irony and metaphor (Gold et al. 2010; Martin and McDonald 2004), detecting violations of Grice's maxims (Surian et al. 1996), using features of context in utterance interpretation (Loukusa et al. 2007) and synchronising gestures with speech (de Marchena and Eigsti 2010). Conversational and discourse problems are also commonplace. Jones and Schwartz (2009) found that children with autism initiated fewer bids for interactions, commented less often, used fewer conversational turns to continue ongoing interactions and responded less often to communication bids than typically developing children during dinner conversations. Colle et al. (2008) found that adults with high-functioning autism or Asperger syndrome used fewer personal pronouns, temporal expressions and referential expressions than control subjects during narrative production.

The co-occurrence of ToM deficits and pragmatic impairments in subjects with ASD has led investigators to enquire if the former deficits might not be causally implicated in the pragmatic disorders of children and adults with ASD. This question assumes a particular direction in the relationship between ToM and pragmatics which, as Tager-Flusberg (2000: 128) has argued, is by no means universally accepted by investigators. Moreover, any prospect of meaningfully addressing this question demands an altogether more detailed developmental model of these two domains than is currently available, some 10 years on from when Tager-Flusberg identified the lack of such a model as an impediment to progress on this issue:

> …although all researchers agree that pragmatics are closely tied to theory of mind, the direction of this relationship has not been clearly delineated. Some argue that some understanding of mind is a prerequisite for acquiring language…or communication…Others suggest that through verbal interactions with others children come to understand that people have minds with contents different from their own…These positions may not be incompatible; what is needed is a more detailed *developmental model* of how different components of a theory of mind might be causally related at different points in time to specific aspects of pragmatics, communication, and discourse skills… (italics in original).

Until such times as a model of this type is forthcoming, it is difficult to assess the true import of the results of empirical studies that have examined the relationship between ToM and pragmatics in autism. Certainly, some studies have revealed significant associations or correlations between ToM and aspects of pragmatics in subjects with ASD. Martin and McDonald (2004) found that second-order ToM reasoning was significantly associated with the ability of subjects with Asperger's syndrome in their study to interpret non-literal utterances (ironic jokes). Hale and Tager-Flusberg (2005) examined discourse skills—specifically, the use of topic-related contingent utterances—and ToM in 57 autistic children. Over 1 year, autistic children made significant gains in the ability to maintain a topic of discourse. ToM contributed unique variance in the contingent discourse skills of these children beyond the significant contribution made by language skills. However, these correlations have not been replicated in other investigations. Losh and Capps (2003) examined the narrative discourse abilities of 28 high-functioning children with autism or Asperger's syndrome. These investigators found that the narrative abilities of these subjects were associated with performance on measures of emotional understanding, but not with ToM or verbal IQ. If the lack of a developmental model is impeding progress in addressing the question of the relationship between ToM and pragmatics in ASD, then investigators can surely receive little consolation from the somewhat inconsistent findings of the few empirical studies in the area. Clearly, more empirical and theoretical research must be undertaken if researchers are to succeed in representing the true nature of the relationship between ToM and pragmatics in ASD.

Children with developmental language disorder of unknown aetiology have received a number of diagnostic labels over the years, the most recent and widely accepted of which is specific language impairment (SLI). While the severe structural language deficits of these children may account for at least some of their

pragmatic deficits, the existence of a clinical subtype called pragmatic language impairment (PLI) attests to the fact that certain other pragmatic deficits are unrelated to these children's difficulties with linguistic structure.[1] Children with SLI and PLI have problems using context to understand implied meanings (Rinaldi 2000), produce inappropriate conversational responses (Bishop et al. 2000) and have difficulty manipulating inferences which play a role in pragmatic interpretation (Adams et al. 2009; Botting and Adams 2005; Ryder et al. 2008; Spanoudis et al. 2007). Alongside these pragmatic impairments, children with SLI have been found to have ToM-related deficits. Studies have shown, for example, that these children make less frequent use of cognitive state predicates than their mental age peers and exhibit delayed development of visual perspective taking (Johnston et al. 2001; Farrant et al. 2006). Children with SLI perform similarly to same-age peers on false belief tests when the linguistic complexity of these tests is low (Miller 2001). One linguistic feature in particular, the syntax of complement structures, has been found to predict false belief performance in children with SLI (Miller 2004). It is not difficult to see why this is the case. In order to attribute the false belief *Anne believes that the ball of wool is in the cupboard* to the mind of an actor in a false belief test, one must have some appreciation of the embedded clause *the ball of wool is in the cupboard*. Complement structures are an integral part of the mental representations through which we conceive of the mental states of others.

Children with intellectual disability form a large and heterogeneous clinical population, making any general characterisation of the pragmatic and ToM skills of this group all but impossible. Investigators have therefore tended to examine the pragmatic and other features of this population on a syndrome-by-syndrome basis. In this way, it has been reported that subjects with Williams syndrome have difficulty with the comprehension of irony and metaphor, as well as with referential communication, the latter in the context of communicating to a speaker that a message is inadequate (Annaz et al. 2009; John et al. 2009; Sullivan et al. 2003). Individuals with fragile X syndrome produce tangential language during conversation, engage in topic repetition, and have difficulty signalling non-comprehension of language to a speaker (Abbeduto et al. 2008; Murphy and Abbeduto 2007; Sudhalter and Belser 2001). Although individuals with Down's syndrome have less impaired pragmatic skills than subjects with these other genetic syndromes (Laws and Bishop 2004), specific pragmatic impairments including problems with referential communication, and metaphor and idiom comprehension have been reported in subjects with Down's syndrome (Abbeduto et al. 2006; Papagno and Vallar 2001). Alongside pragmatic impairments, a range of ToM deficits have been reported in Williams syndrome, fragile X syndrome and Down's syndrome (Abbeduto et al. 2001; Cornish et al. 2005; Grant et al. 2007; Sullivan and Tager-

---

[1] In a study of the conversational responsiveness of children with SLI, Bishop et al. (2000: 177) make this same point as follows: 'this study lends support to the notion that there is a subset of the language-impaired population who have broader communicative impairments, extending beyond basic difficulties in mastering language form, reflecting difficulty in responding to and expressing communicative intents'.

Flusberg 1999; Yirmiya et al. 1996). Few investigators, however, have attempted to examine ToM in the context of pragmatics in individuals with intellectual disability. One study which does is Abbeduto et al. (2004) who reported that limited narrative language skills in the subjects in their study contributed substantially to the failure of these subjects on a false belief task.[2]

Emotional and behavioural disorders (EBDs), which include attention deficit hyperactivity disorder (ADHD) and conduct disorder, have only recently been the focus of clinical pragmatic studies. These studies indicate that pragmatics is impaired in children with EBDs and, in some cases at least, is more impaired than structural language (Geurts and Embrechts 2008). In an investigation of pragmatic skills in ADHD, Bishop and Baird (2001) reported that 73 % of their child subjects attained a score below the 132 cut-off point indicative of pragmatic impairment on the Children's Communication Checklist (Bishop 1998). The scale measuring inappropriate initiation of conversation revealed particularly poor scores. Other pragmatic impairments in ADHD include difficulty drawing inferences when listening to spoken texts and problems with the comprehension of figurative language (Berthiaume et al. 2010; Bignell and Cain 2007; McInnes et al. 2003). Two-thirds of the children with conduct disorder studied by Gilmour et al. (2004) displayed pragmatic impairments and behavioural features similar to those found in autism. ToM and mentalising deficits have been identified in children with EBDs. Donno et al. (2010) reported poorer mentalising abilities in 26 persistently disruptive children than in comparison subjects. These children also possessed poorer pragmatic language skills than comparisons. Children with ADHD display poorer recognition of emotional facial expressions, lower levels of social perspective taking and worse performance on second-order ToM tasks than normally developing children (Buitelaar et al. 1999; Marton et al. 2009; Pelc et al. 2006). One study which has attempted to examine ToM in the context of pragmatics is Adachi et al. (2004). The children with ADHD in this study displayed problems with the comprehension of metaphor. However, there was no correlation between metaphor comprehension and performance in a ToM task.

5 Acquired Pragmatic Disorders and Theory of Mind

Many pragmatic disorders found in children and adults have their onset outside of the developmental period. These so-called acquired pragmatic disorders may be caused by a range of diseases and injuries. An adult may sustain a cerebrovascular accident or stroke that causes a focal lesion in either the left or right hemisphere of the brain. If there is left-hemisphere damage, pragmatics may be impaired in the presence of a wider aphasia. Even greater pragmatic impairment can result from a

---

[2] Abbeduto et al.'s finding lends support to the second of the views delineated by Tager-Flusberg (2000) in the main text. This is the view that language and verbal interaction play a vital role in the development of ToM.

lesion in the right hemisphere of the brain, a hemisphere that for many years was deemed to be of lesser significance in terms of language and communication. An adult who is involved in a road traffic accident or similar traumatic incident sustains a quite different pattern of brain damage from that incurred in a stroke. The multi-focal brain pathology that occurs in a traumatic brain injury can not only cause pragmatic and discourse impairments, but can also lead to significant cognitive deficits which compromise communication. A number of mental illnesses have their onset in adulthood. These disorders, which include schizophrenia, bipolar disorder and depression, can have serious implications for a person's use of language or pragmatics. The population of adults with neurodegenerative disorders is large and growing. These disorders include the dementias, but also a range of other conditions such as Parkinson's disease, Huntington's disease and multiple sclerosis. Although the pragmatic impairments of individuals with dementia related to Alzheimer's disease are by now well characterised, researchers are increasingly turning their attention to investigating pragmatic impairments in a number of other neurodegenerative disorders. This section will examine what is known pragmatic disorders in each of the above clinical populations and will consider the role of ToM deficits within them.

Traditionally, it was believed that pragmatic deficits in adults with left-hemisphere damage (LHD) were a direct consequence of impairments in structural language skills. So it was argued, for example, that the adult with LHD who cannot engage in inversion of a subject pronoun and auxiliary verb (a syntactic deficit) will struggle to produce an indirect speech act such as 'Can you tell me the time?' (a pragmatic deficit). However, clinical studies of adults with LHD are increasingly revealing that pragmatic deficits in this clinical population are not adequately characterised in terms of linguistic deficits. There is evidence that while structural language skills can improve over time in adults with LHD, pragmatic skills may fail to improve appreciably (e.g. Coelho and Flewellyn 2003). Also, nonverbal pragmatic behaviours can also be impaired in adults with LHD (e.g. Cutica et al. 2006). Both scenarios indicate that a more nuanced explanation of pragmatic deficits in LHD needs to be found. In the meantime, studies of pragmatics in LHD have revealed problems in verbal pragmatic aspects of discourse production (Borod et al. 2000), difficulty with the comprehension of proverbs and implicatures (Chapman et al. 1997; Kasher et al. 1999) and, in patients with left prefrontal lesions, problems with pragmatic inferences (Ferstl et al. 2002). Some studies have also reported preserved areas of pragmatic functioning. In this way, Beeke (2003) describes the case of a man with agrammatic aphasia who exhibited recurrent use of 'I suppose' when his production of subject-verb constructions was generally poor, because he was motivated by an interactional need to produce unproblematic turns at talk. ToM skills in adults with LHD appear to be largely intact (Varley and Siegal 2000; Varley et al. 2001). This has led investigators to claim that ToM is functionally independent of grammar (Siegal and Varley 2006).

Pragmatic disorders in right-hemisphere damage (RHD) have been the focus of numerous clinical studies since Myers (1979) undertook the first formal investigation of impaired communication skills in stroke patients with RHD (see Sect. 3.3

in Cummings (2009) for discussion). Among the pragmatic deficits in this population, investigators have reported reduced sensitivity to violations of Gricean maxims and difficulty varying the production of requests in accordance with the interpersonal and situational features of an interaction (Brownell and Stringfellow 1999; Surian and Siegal 2001). The comprehension of non-literal language in idioms, proverbs and humour is compromised (Brundage 1996; Cheang and Pell 2006; Papagno et al. 2006). Discourse skills are often markedly disrupted in individuals with RHD. These subjects have been reported to produce narratives that have poor information content and lack cohesion and coherence (Marini et al. 2005). Blake (2006) found that tangentiality, egocentrism and extremes of quantity (verbosity or paucity of speech) were features of the discourse produced by the adults with RHD in their study. Although some studies have revealed ToM and mentalising deficits in the RHD population (Griffin et al. 2006; Happé et al. 1999; Weed et al. 2010), a recent review of work in this area by Weed (2008) judged that evidence for a specific ToM deficit in RHD is still inconclusive. One of the few studies to examine the relationship between ToM and pragmatics in RHD is Winner et al. (1998). These investigators found that the subjects with RHD in their study performed significantly worse than controls on one of two measures of second-order belief attribution. Moreover, the ability to distinguish lies from ironic jokes correlated strongly with two second-order belief measures. Winner et al. (1998: 90) concluded that 'the fragility of RHD patients' understanding of second-order mental states underlies a portion of their difficulties in discourse comprehension'.

Subjects who sustain a traumatic brain injury (TBI) can have numerous neuropsychological sequelae related to the multi-focal nature of their brain pathology. Among the language and cognitive problems found in this clinical population, subjects with TBI often present with significant pragmatic and discourse impairments as well as ToM deficits. MacLennan et al. (2002) found pragmatic impairments in 86 % of 144 patients with TBI in their study. On a pragmatic rating scale, components that examined cohesion, repair, elaboration, initiation and relevance displayed the highest frequency of impairment. Other pragmatic impairments reported in subjects with TBI include violations of Gricean maxims (quantity, relation, manner), problems with topic management and the use of politeness markers in conversation, and difficulties in inferencing and intentionality, the latter related to the mental states and intentions involved in pragmatic skills such as the production of speech acts and the understanding of irony (Coelho et al. 2002; Dennis and Barnes 2001; Douglas 2010; Togher and Hand 1998). Although ToM deficits have been extensively reported in the TBI population (Henry et al. 2006; Milders et al. 2006; Muller et al. 2010; Turkstra et al. 2004),[3] there is some uncertainty about the role of these deficits in the pragmatic impairments of this clinical population. Some investigators have argued that ToM deficits are probably

---

[3] Some studies have failed to find evidence of ToM impairments in patients with TBI. Bach et al. (2006) reported that the patients with TBI with and without behavioural disturbance in their study were unimpaired on ToM tasks.

unrelated to the pragmatic skills of clients with TBI (e.g. Muller et al. 2010). For other investigators, a relationship between ToM and pragmatics can be demonstrated in these clients. In this way, McDonald and Flanagan (2004) reported that the second-order ToM judgements of the adults with TBI in their study were related to the ability to understand social or conversational inference.

A number of mental illnesses which have their onset in adulthood can cause marked pragmatic disturbances. The most prominent and extensively investigated of these conditions is schizophrenia. Investigators have reported pragmatic difficulties in the decoding of Gricean conversational maxims, the use of linguistic context during language processing and the interpretation of metaphor and idiom in patients with schizophrenia (Kuperberg et al. 2000; Mazza et al. 2008; Tavano et al. 2008; Tényi et al. 2002). These patients also contribute irrelevant information and engage in derailments during narrative production (Marini et al. 2008). Some of these pragmatic impairments have been examined in relation to the ToM skills of subjects with schizophrenia. Brüne and Bodenstein (2005) investigated the relation of proverb understanding to ToM in 31 patients with schizophrenia. These investigators found that approximately 39 % of the variance of proverb comprehension in these patients was predicted by ToM performance. Mo et al. (2008) studied metaphor and irony comprehension and conducted first- and second-order ToM tasks in 29 patients with schizophrenia who were in remission. These patients had a ToM deficit and were impaired in their comprehension of metaphor and irony. However, only metaphor comprehension was significantly correlated with second-order false belief understanding. Langdon et al. (2002) found impairments in false-belief picture sequencing (a test of ToM) and the understanding of irony and metaphor in the subjects with schizophrenia in their study. Poor ToM performance was associated with poor understanding of irony, but not with metaphor comprehension. The different findings to emerge from these studies may reflect the clinical status of the patients used in these investigations—patients with remitted schizophrenia (Mo et al.) versus patients with formal thought disorder (Langdon et al.).[4]

A large number of neurodegenerative disorders can cause pragmatic impairments in adults. The most extensively investigated pragmatic impairments in this group of disorders are those found in dementia related to Alzheimer's disease (AD). Among the pragmatic and discourse impairments in subjects with AD, investigators have reported problems with referential communication (Carlomagno et al. 2005; Feyereisen et al. 2007), the use of cohesion devices (Ripich et al. 2000), and the comprehension of figurative language (Papagno 2001). The pragmatic features of non-Alzheimer's dementias are increasingly the focus of clinical studies (e.g. Kertesz et al. 2010). Rousseaux et al. (2010) studied verbal and non-verbal communication in patients with Alzheimer's disease, behavioural variant frontotemporal dementia (FTD) and dementia with Lewy bodies (DLB). Pragmatic skills

---

[4] There is some basis for this claim in the findings of several studies that different symptom profiles in schizophrenia appear to be associated with certain patterns of ToM and pragmatic deficits (e.g. Corcoran and Frith 1996; Montag et al. 2011).

were well preserved in patients with DLB. Patients with AD exhibited some pragmatic impairment in greeting behaviour, understanding deictics and using gestures. The most severe pragmatic impairments were found in patients with FTD, particularly in responding to open questions, presenting new information, logically organising discourse, adapting to interlocutor knowledge and emitting feedback. Beyond the dementias, pragmatic disorders have also been documented in patients with Parkinson's disease, including problems in the comprehension of metaphor, irony and speech acts, and difficulty with conversational appropriateness and turn taking (Holtgraves and McNamara 2010; McNamara and Durso 2003; Monetta and Pell 2007; Monetta et al. 2009). A recent study of patients with Huntington's disease has reported impaired performance on complex comprehension tasks which draw upon pragmatic and discourse skills (Saldert et al. 2010).

There is now an extensive body of research findings suggesting the presence of ToM deficits in adults with neurodegenerative disorders. One of the most consistent findings of this research is that patients with frontal or behavioural variant FTD experience significant ToM deficits (Fernandez-Duque et al. 2009; Gregory et al. 2002; Lough et al. 2006; Torralva et al. 2009). However, a range of ToM deficits have also been identified in patients with Alzheimer's disease (Fernandez-Duque et al. 2009; Gregory et al. 2002), Parkinson's disease (Bodden et al. 2010; Saltzman et al. 2000) and motor neurone disease (Gibbons et al. 2007; Girardi et al. 2011). These studies are revealing interesting features of ToM deterioration in subjects with neurodegenerative disorders. In this way, Castelli et al. (2011) found that the ToM skills of their patients with Alzheimer's disease were lost in an order that is opposite to the developmental sequence followed during the acquisition of ToM. Although there is a well-developed literature on ToM deficits in patients with neurodegenerative disorders, as yet few studies have undertaken to examine the relation between these deficits and pragmatic skills. Two studies that have done so are Cuerva et al. (2001) and Monetta et al. (2009). In a study of 34 patients with probable Alzheimer's disease, Cuerva et al. reported a significant association between performance on a test of second-order false belief and pragmatic deficits in the interpretation of conversational implications and indirect requests. Monetta et al. examined ToM performance and irony comprehension in 11 non-demented patients with Parkinson's disease. These investigators found a significant correlation between these patients' ability to interpret an utterance as a lie or an ironic remark and performance on second-order belief questions.

6 Theory of Mind Research in Pragmatic Disorders: Critical Commentary

There can be little doubt that a voluminous literature on theory of mind now exists. Much is known about the emergence of ToM skills in typically developing children, the further maturation of these skills in adolescents and adults and their

deterioration in normally aging individuals. Today, studies of ToM in clinical populations extend well beyond the autism spectrum disorders to include individuals with neurocognitive, neurodegenerative and neuropsychiatric disorders. Even within these populations, investigators are increasingly moving towards characterisations of ToM based on the diagnostic sub-types of a disorder. For example, it is now not uncommon for researchers to examine the ToM skills of patients with schizophrenia in relation to the (positive and negative) symptom profiles of these patients (see footnote 4). This abundance of empirical research into ToM has been more than an academic exercise for theorists and clinicians. The results of this research have been used to devise ToM-targeted interventions, which are increasingly employed in the treatment of children and adults with conditions such as autism spectrum disorders and schizophrenia (Roncone et al. 2004; Swettenham 2000). In short, ToM research has delivered many theoretical insights and clinical gains for both researchers and patients. However, the same cannot be said of one particular aspect of ToM research. That aspect concerns the relationship between ToM and pragmatics. The difficulty here lies not in the lack of empirical research which has directly examined this relationship—although it is certainly the case that few studies have undertaken such an examination—but in the way in which investigators have come to construe the ToM skills that are integral to pragmatic interpretation. This issue will be addressed throughout this section and is developed further in Cummings (2009).

In parallel with the vast expansion that has occurred in ToM research in recent years, there has been an equally dramatic proliferation in the number and type of tests that purport to examine theory of mind. The so-called Sally-Anne experiments[5] that have become synonymous with early ToM research into autism undertook to examine a subject's ability to attribute first-order false beliefs to the mind of an actor in a scenario. First-order false-belief tests now sit alongside second-order false-belief tests, deception tests, tests of imagination, 'seeing leads to knowing' tests and tests of understanding the causes of emotion, to name just a few of the ToM tests currently in use (Baron-Cohen 2000). These tests are not without their difficulties.[6] The particular difficulties that will be examined in this section relate to a subset of ToM tests which, it is claimed, assess ToM skills in pragmatic interpretation. Baron-Cohen (2000) classifies these tests as follows: (1) tests of understanding metaphor, sarcasm, and irony; (2) tests of pragmatics in speech; (3) tests of recognition of violations of pragmatic rules. Categories (2) and (3) include tests that examine faux pas detection and the recognition of violation of Gricean conversational maxims, respectively. In demonstration of the tests subsumed by (1), consider the following investigation by Mo et al. (2008) of metaphor

---

[5] Sally-Anne experiments derive their name from the two dolls—Sally and Anne—that are used in these false-belief tests. See an account of these tests by Baron-Cohen et al. (1985) in Sect. 3.

[6] There is some evidence that these different ToM tests may not examine the same mentalising skills. For example, Spek et al. (2010) found low or absent correlations between the Eyes test and three other ToM tests used in their study of adults with high-functioning autism and Asperger syndrome.

and irony comprehension in patients with schizophrenia. These patients, who were all in remission at the time of the study, also completed first- and second-order ToM tasks. Metaphor and irony comprehension was assessed by means of narrated stories. These stories were followed by questions which were intended to test patients' understanding of the metaphorical and ironical content of the narrated passages. One of the passages used in this study is shown below, along with the questions which were intended to probe the patients' understanding of metaphor and irony:

> Xiao Zhang could never make up his mind about anything. One day when Li Qi and Wang Li asked him if he would like to go to the cinema, Xiao Zhang could not decide. It took him so long to make up his mind that by the time he did, they had already missed the first half of the film…
>
> On the metaphorical presentation, the story continued… Li Qi said: 'Xiao Zhang, you are a ship without a captain!'
>
> **Metaphor** question: What does Li Qi mean? Does he mean Xiao Zhang is good or not good at making decisions?
>
> On the ironical presentation, the story continued…Wang Li said to Xiao Zhang: 'You really are so good at making decisions!'
>
> **Irony** question: What does Wang Li mean? Does he mean Xiao Zhang is good or not good at making up his mind?

This passage and its accompanying probe questions are typical of those used by investigators who examine the relation of ToM to pragmatics in clinical subjects. Yet, neither the metaphor question nor the irony question is successfully tapping the particular pragmatic phenomenon it purports to examine. In the metaphorical presentation of the passage, regardless of how the respondent answers the metaphor question, he or she becomes committed to a presupposition of that question, a presupposition to the effect that the utterance 'You are a ship without a captain!' is stating something about Xiao Zhang's ability to make decisions. In not allowing the respondent the opportunity to deny a metaphorical interpretation of the passage, the probe question is itself not felicitous. This somewhat ironical pragmatic error on the part of these investigators is accompanied by an even greater problem in the irony version of the same passage. The irony question 'What does Wang Li mean?' invites the response 'Wang Li thinks that Xiao Zhang is good/not good at making up his mind'. This question and its response involve a first-order belief, Wang Li's belief about Xiao Zhang's decision-making ability. Yet, this is quite different from the second-order ToM reasoning which is the basis of the interpretation of irony in language. In order for *this* interpretation to be adequately tested, Mo et al. needed to pose a quite different question, a question to the effect 'What does Wang Li think that Xiao Zhang believes the utterance *You really are so good at making decisions* means?'. The response to this question—Wang Li thinks that Xiao Zhang believes the utterance means *x*—is the belief that the respondent must attribute to Wang Li's mind in order to be said to have appreciated the ironic intent of Wang Li's utterance. It is only by asking this latter question that Mo et al. can expect to assess the second-order ToM reasoning and irony comprehension of the subjects with schizophrenia in their study.

Tests of faux pas detection are used extensively by ToM researchers. To assess the extent to which these tests succeed in examining the ToM skills used in 'pragmatics in speech', we turn to a study by Baron-Cohen et al. (1999). These investigators examined the recognition of faux pas by 12 children with Asperger's syndrome (AS) or high-functioning autism (HFA). Sixteen normal control subjects also participated in the study. Ten short stories were presented to the children on an audiotape. In one of these stories, a young girl called Sally is staying at her Aunt Carol's house. Sally has short blonde hair. One day, the doorbell rang. When Sally and her aunt opened the door, they saw a neighbour called Mary. After saying "Hello" to Sally and Carol, Mary went on to say "Oh, I don't think I've met this little boy. What's your name?". Aunt Carol responding by saying "Who'd like a cup of tea?". Subjects were then asked a series of questions. One of these questions examined if subjects had detected the faux pas in the story ('In the story did someone say something that they should not have said?'). A second question required subjects to identify the faux pas ('What did they say that they should not have said?'). A third question tested subjects' understanding of the language used in the story ('Whose house was Sally at?'). A fourth question aimed to assess if subjects were aware that the faux pas was a consequence of a false belief on the part of the speaker in the story ('Did Mary know that Sally was a little girl?').

Tests of faux pas detection of the type just described pose the following problem for ToM researchers. To the extent that these tests are supposed to reveal something of the ToM skills at work in pragmatic interpretation, they need to examine second-order ToM reasoning, i.e. one person's beliefs about another person's beliefs. Yet, in asking the children with AS or HFA in this study to detect if a faux pas has been committed, Baron-Cohen et al. are only examining first-order ToM reasoning. To see this, consider the faux pas committed by Mary in the above scenario. Mary calls Sally a little boy because she entertains a mistaken or false belief to the effect *Mary believes that Sally is a boy*. However, this belief is a first-order false belief about the world, not a second-order belief of the type *Mary believes that Sally believes that x* which is integral to all pragmatic interpretation. The child who recognises that Mary entertains the false belief that Sally is a boy is clearly capable of first-order ToM reasoning. And certainly no pragmatic interpretation would be possible if hearers could not also assume certain beliefs on the part of their interlocutors about the world. However, these first-order beliefs fall well short of the second-order ToM reasoning that is the basis of utterance interpretation. To the extent that the recognition of faux pas is none other than the recognition of first-order false belief, it is unsurprising that while 75 % of normal children passed this faux pas detection test, only 18 % of children with AS or HFA managed to do so—false belief performance is significantly depressed in subjects with AS or HFA, after all. Quite apart from revealing anything about the ToM skills at work in pragmatics, faux pas detection tests serve only to reinforce the well-established finding that children with ASD fail first-order false-belief tests on account of a specific ToM deficit in mental state attribution.

A further category of ToM test examines the recognition or detection of the violation of pragmatic rules. Typically, the rule violations in question take the

form of Gricean conversational maxims of quality, quantity, relation and manner with a number of studies also examining politeness maxims. A study by Surian et al. (1996) examined both types of pragmatic rule violation in high-functioning children with autism. The rationale for this study is presented by Surian et al. (1996: 58) as follows: 'if children with autism have deficits in ascribing mental states, and particularly ascribing intentions, then they should fail to recognise when such Gricean maxims are being violated'. In demonstration of the conversational scenarios presented on audiotape to subjects, consider the following violation of the maxim of relation:

A: What is your favourite programme on telly?
B: My favourite is sandwiches.

In the above exchange, B's response to A's question is clearly violating the maxim of relation—an utterance about sandwiches fails every expectation of relevance that A could have when he poses a question about favourite TV programmes. However, the child with autism who recognises this violation has really not engaged in the type of mental state attribution that is the basis of pragmatic interpretation. To see this, one need only consider the role of Gricean maxims within the recovery of implicatures. Even as the speaker in the above exchange recognises that the hearer has produced an irrelevant utterance, A uses an assumption of mutual adherence to the principle of cooperation to derive an implicature of B's utterance. One such implicature may be that B does not have a single favourite programme, as is presupposed by A's question, but actually likes a number of programmes on the theme of food and cooking. Alternatively, B may be seeking to implicate that he finds A's question trivial and does not wish to provide a serious response. Whichever implicature B is attempting to generate, it is clear that the recognition that the relation maxim has been violated is nothing more than a first step on the road to recovering this implicature. It is only when this maxim violation is used by A to recover the intended implicature of B's utterance, a process that requires A to attribute a particular communicative intention to B's mind, that the interlocutors in this exchange can be said to be engaging in the type of mental state attribution that is the basis of pragmatic interpretation. Surian et al. found that while most children with autism performed at chance on this maxim task, all children with specific language impairment and all normal controls performed above chance. However, given that this task is not even assessing the ToM skills used in pragmatic interpretation, it emerges that this finding lacks any real implications for our knowledge of ToM.

The discussion in this section has revealed a number of shortcomings in ToM research into pragmatics in general and pragmatic disorders in particular. Firstly, tests of the comprehension of pragmatic phenomena such as irony were shown not to assess the second-order ToM skills that are integral to pragmatic interpretation. These tests, it was argued, targeted instead an actor's beliefs about the world, i.e. first-order ToM reasoning. Secondly, in examining a subject's ability to detect an actor's mistaken or false belief about the world, faux pas detection tests were found to be little more than first-order false-belief tests. As such, they failed to

assess any part of the ToM skills used in pragmatic interpretation and simply reinforced the long-established finding that performance on false-belief tests is impaired in subjects with ASD. Thirdly, studies examining the recognition of maxim violations were shown to hold no lessons for the ToM skills used in pragmatic interpretation. The mere recognition that a maxim had been violated, it was argued, was simply the first step on the road to recovering the implicature of an utterance and could in no way come to represent the attribution of communicative intentions that is the essence of pragmatic interpretation. Clearly, these various failings of ToM studies in pragmatics need to be addressed if theorists and clinicians are to have confidence in the results to emanate from these studies. I have argued that the place to begin this revision of ToM research is in something quite fundamental, the notion of what constitutes pragmatic interpretation itself (Cummings 2007b, 2009). It is only when we are clear on the nature and extent of this notion that we can expect to devise a ToM framework that is capable of representing the mentalising skills involved in the interpretation of utterances.

7 Theory of Mind Theories

Empirical studies of ToM are important because their findings play a key role in ToM theory construction. The set of mentalising skills that investigators have identified as ToM are consistent with different theoretical accounts of these skills. This section will examine the three main contenders to a theoretical account of ToM: ToM as a cognitive module (modular theory of ToM), ToM as theory construction (theory theory account), and ToM as imaginative projection (simulation theory). Although these theoretical accounts make quite different claims about the development and nature of ToM, each account receives some degree of validation from empirical studies. This discussion is therefore less concerned with attempting to decide among these alternative theories than it is with teasing out the features of each which may hold particular significance for an account of ToM that has relevance to pragmatics. It does so in the knowledge that not every aspect can be addressed in detail. The reader is referred to Cummings (2009) for a more extensive discussion of this area.

Simon Baron-Cohen and Alan Leslie are leading proponents of the view that a cognitive module underlies the human ability to attribute mental states to the minds of others. This module has certain features which will be recognized by any reader familiar with Jerry Fodor's notion of modularity (Fodor 1983). Such a module is informationally encapsulated, meaning that there are restrictions on the availability of information. In this way, information within the module may not be available outside the module, even though the module's output is available to the mind's central system, for example. Similarly, information outside the module may not be accessible to the workings of the module itself. A ToM module is domain specific, that is, it contains specialized representations and computations which relate only to

the mental states of intentional agents.[7] A further feature of a ToM module is its innateness, that is, it forms part of a human being's genetic endowment. A ToM module is also fast (it engages in rapid processing of intentional information) and mandatory (it cannot choose not to process intentional stimuli). Furthermore, a ToM module may be selectively impaired, as we have seen in the case of autism, for example (Scholl and Leslie 1999). These features of modular ToM, theory theorists have argued, are antithetical to the developmental changes that young children go through on their way to acquiring a theory of mind.[8] Modular ToM theorists counter this challenge by arguing that there is no requirement within modularity itself which precludes the possibility that modules can develop from within (even though the requirement for informational encapsulation restricts the information that will be available to such an internal developmental process) (see Scholl and Leslie (1999) for discussion of this developmental process in relation to modular ToM).

According to theory theorists, the young child who is acquiring ToM skills is effectively constructing a theory of the actions and mental states of those around him. In much the same way that the scientist constructs theories to explain and predict events in the natural world, the child is constructing a theory which he uses to explain and predict the behaviour of others in the social world. This scientific analogy, and the theoretical position it seeks to capture, is characterised by Gopnik et al. (2000: 51) as follows:

> [W]e propose that our ordinary understanding of the mind proceeds by the formation, revision and replacement of successive theories of the mind. […] Like scientists, children understand the world by constructing coherent views of it and changing those views in the light of new evidence that they obtain. Children play an active role in this process by making predictions, seeking explanations and considering evidence that is relevant to the mind.

Certain other features of the theory theory account are noteworthy. Theories in one domain can influence theories in another domain, a transfer of information and knowledge that is not possible on a modular theory of ToM on account of the informational encapsulation of cognitive modules. The succession of theories on this account is made possible by general inferential mechanisms, which are not available to cognitive modules (such mechanisms, after all, are likely to be part of the mind's central system on a modular approach). The theory theory account of ToM also posits innate structures in the form of theory-formation mechanisms. These mechanisms are evolutionarily determined and enable the developing child

---

[7] It is worth remarking that the original proponent of modularity—Fodor (1983)—would disavow the attempt to locate ToM skills within a domain-specific cognitive module. For Fodor, such skills are located within a non-modularised central system in the mind. Frye (2000: 149) makes this same point as follows: 'An odd aspect of the view that theory of mind is domain specific is that it is one Fodor's (1983) own approach to modularity would explicitly disclaim'.

[8] That one theory theorist, Alison Gopnik, cannot even conceive of modularity as permitting of developmental change is evident in her use of the word 'indefeasible' when characterising a ToM module. A ToM module, she states, is 'a genetically-determined and indefeasible way of understanding the mind' (Gopnik et al. 2000: 51).

to make inferences about the underlying structure, particularly causal structure, of the world based on his observation of events. A further innateness component in the theory theory account is the claim that babies are born with initial, starting-state theories about the mind and other aspects of the world. These theories are 'genuinely theoretical' in that they are 'specific, substantive, coherent, abstract, representations of the world that allow babies to make predictions, and to interpret, and even perhaps explain, what they see around them' (Gopnik et al. 2000: 52). These representations are revised as the child encounters new evidence in much the same way that scientists need to revise the theories they work with if they no longer accord with the data in an area. So the initial theories that the child sets out with do not constitute the architecture of the fully mature representational system. The process of theory revision that takes the child from his initial theories to this final state is the basis of the developmental changes in ToM that were described in Sect. 3.

According to simulation theorists, our mentalising abilities are not explained by cognitive modules or by theory construction. Rather, when we simulate, we are imaginatively projecting from our own mental activity (what we would think, believe or desire in a situation) to what someone else is likely to think, etc. in a similar situation. Carruthers and Smith (1996: 3) capture this key notion within simulation theory as follows:

> According to this view, what lies at the root of our mature mind-reading abilities is not any sort of theory, but rather an ability to project ourselves imaginatively into another person's perspective, *simulating* their mental activity with our own (italics in original).

Proponents of simulation theory differ with respect to the details of how simulation comes about. According to Goldman (1993), simulation requires first-person awareness of one's own mental states, with the inference from these states to the mind of another taking the form of an argument from analogy. Alternatively, simulation theorists like Gordon (1996) argue that recognition of one's own mental states is not a requirement of simulation and that the type of imaginative identification that occurs in simulation can take place without introspective self-awareness. In one version subscribed to by Gordon (1986), our practical reasoning system is taken 'off-line' and is fed pretend inputs such as images and suppositions. Based on these inputs, the system arrives at a decision which is not acted upon (the system is off-line, after all), but becomes the basis of our anticipation of another person's behaviour:

> Our decision-making or practical reasoning system gets partially disengaged from its "natural" inputs and fed instead with suppositions and images (or their "subpersonal" or "sub-doxastic" counterparts). Given these artificial pretend inputs the system then "makes up its mind" what to do. Since the system is being run off-line, as it were, disengaged also from its natural output systems, its "decision" isn't actually executed but rather ends up as an anticipation…of the other's behaviour (Gordon 1986: 170).

Which of these theoretical approaches to ToM—if any—is best able to capture the mentalising skills used in utterance interpretation is a question of interest to researchers in both theoretical and clinical pragmatics. We conclude this chapter

with some comments about the pragmatic plausibility of each of these approaches. On the (unproblematic) assumption that ToM cannot be any less dynamic and resourceful than pragmatic interpretation itself, it is difficult to see how a modular account of ToM can capture the mentalising skills at work in utterance interpretation. Of necessity, hearers must be able to draw upon information and knowledge of any type as they proceed to attribute communicative intentions to the minds of speakers during the interpretation of utterances. Imagine how difficult it would be to constrain in advance of the interpretation of an implicature the information or knowledge that a hearer can have access to in the recovery of that implicature (of course, I would argue that it is not just difficult but completely impossible). Yet, just such restrictions on the type and flow of information are exactly what the modular ToM theorist is offering us through his requirements for informational encapsulation and domain specificity. Of course, a modular account of ToM receives strong support from two prominent pragmatic theorists, Dan Sperber and Deirdre Wilson. According to Sperber and Wilson (2002), their relevance-based procedure for the interpretation of utterances forms a sub-module of the mind-reading module. This view is consistent with the cognitive scientific character of their relevance-theoretic account more generally, an account that I first challenged in Cummings (2005). It is difficult to see how a relevance-based procedure located within a mind-reading module evades any of the criticisms that were made in that context.

A theory theory account of ToM certainly lacks the prominence that modular accounts of ToM have enjoyed within pragmatics. Yet, there are reasons to believe that a theory theory account of ToM may be more in tune with the open texture of pragmatic interpretation than other theoretical approaches to ToM. As the theory theory position has been expounded by Gopnik and others, it is clear that theories in one domain can influence theories in another domain. This flow of information or knowledge between domains forms the very essence of pragmatic interpretation and must be embraced by any theoretical account of ToM that hopes to capture the mentalising skills involved in such interpretation. In my determination of the intended referent of the demonstrative pronoun 'that' in the utterance *That was unexpected*, I must surely draw upon several theories including my theories of visual perception, word meaning and grammar, and physical space and movement to establish that the speaker intended to refer to the man who has just fallen into a hole in the ground. The general inferential mechanisms that theory theorists argue are the basis of theory revision may prove ultimately to be more akin to the type of inferencing at work in pragmatic interpretation than the highly specialised mechanisms of inference that are posited to exist within cognitive modules. Certain findings to emerge from empirical studies lend tentative support to this suggestion. One such finding is that clinical subjects who exhibit significant pragmatic and ToM deficits often also present with more general impairments of inference. For example, individuals with schizophrenia have been shown to have difficulty with deductive reasoning as well as various types of inductive or probabilistic inference (Armstrong et al. 2012; Averbeck et al. 2011; Corcoran 2003; Speechley et al. 2010; Titone et al. 2004).

Intuitively at least, simulation theory captures a plausible idea that resonates with our commonsense understanding of how the interpretation of utterances proceeds. There is a clear sense in which as communicators, we come to attribute certain communicative intentions to the mind of a speaker exactly because we can imagine ourselves holding just those same intentions within the situation in which the speaker finds himself. It is because I can envisage myself experiencing certain negative thoughts about a child who is behaving badly at a party that I can go on to attribute an ironic communicative intention to the speaker who utters *What a delightful child!* Notwithstanding its initial plausibility, simulation theory still experiences a significant difficulty as an account of the ToM skills that are involved in pragmatic interpretation. To appreciate this difficulty, we need to return to Gordon's account of how simulation takes place, an account in which our practical reasoning system is taken 'off-line' and fed certain inputs. Let's imagine that the person at the party to whom the above utterance is directed feeds certain inputs into their 'off-line' practical reasoning system, such as the beliefs that the child is behaving badly, that people dislike badly behaved children, and so on. Based on these input beliefs, the hearer's practical reasoning system might reasonably be expected to arrive at the decision that the badly behaved child at the party is anything but delightful. This decision is then attributed as a communicative intention to the mind of the speaker of the utterance. However, in selecting certain beliefs and feeding these into the hearer's practical reasoning system, we have not explained pragmatic interpretation, but rather transformed the problem of interpretation into the problem of how certain beliefs are selected for the simulation. It is this selection process that a pragmatically acceptable theory of ToM must explain, and about which simulation theory says nothing (see Cummings (2009) for further discussion).

8 Summary

In this chapter, the role of ToM in pragmatic disorders has been examined. It was argued that ToM is an indispensable component of all pragmatic interpretation and, as such, can be expected to contribute to pragmatic disorders. For some years, investigators have examined the emergence of ToM skills in normally developing children and, more recently, the maturation and decline of these skills in adults and aging subjects. This large and growing literature has produced important findings, several of which were discussed. Children and adults with a range of clinical disorders present with concomitant pragmatic impairments and ToM deficits. In reviewing studies of these disorders, it was consistently remarked that few investigators had examined the relationship between ToM deficits and pragmatic impairments. It was also argued that ToM research is not without difficulties which threaten to stall further progress in our understanding of the role of ToM in pragmatic disorders. Several flaws in ToM research were examined. Finally, three

theoretical accounts of ToM were examined and assessed for their relevance to pragmatic interpretation.

References

Abbeduto, L., Pavetto, M., Kesin, E., Weissman, M. D., Karadottir, S., O'Brien, A., et al. (2001). The linguistic and cognitive profile of Down syndrome: Evidence from a comparison with fragile X syndrome. *Down's Syndrome Research and Practice, 7*, 9–15.

Abbeduto, L., Short-Meyerson, K., Benson, G., & Dolish, J. (2004). Relationship between theory of mind and language ability in children and adolescents with intellectual disability. *Journal of Intellectual Disability Research, 48*, 150–159.

Abbeduto, L., Murphy, M. M., Richmond, E. K., Amman, A., Beth, P., Weissman, M. D., et al. (2006). Collaboration in referential communication: Comparison of youth with Down syndrome or fragile X syndrome. *American Journal on Mental Retardation, 111*, 170–183.

Abbeduto, L., Murphy, M. M., Kover, S. T., Giles, N. D., Karadottir, S., Amman, A., et al. (2008). Signaling noncomprehension of language: A comparison of fragile X syndrome and Down syndrome. *American Journal of Mental Retardation, 113*, 214–230.

Adachi, T., Koeda, T., Hirabayashi, S., Maeoka, Y., Shiota, M., Wright, E. C., et al. (2004). The metaphor and sarcasm scenario test: A new instrument to help differentiate high functioning pervasive developmental disorder from attention deficit/hyperactivity disorder. *Brain & Development, 26*, 301–306.

Adams, C., Clarke, E., & Haynes, R. (2009). Inference and sentence comprehension in children with specific or pragmatic language impairments. *International Journal of Language & Communication Disorders, 44*, 301–318.

Adrián, J. E., Clemente, R. A., & Villanueva, L. (2007). Mothers' use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: A longitudinal study. *Child Development, 78*, 1052–1067.

Annaz, D., Van Herwegen, J., Thomas, M., Fishman, R., Karmiloff-Smith, A., & Rundblad, G. (2009). Comprehension of metaphor and metonymy in children with Williams syndrome. *International Journal of Language & Communication Disorders, 44*, 962–978.

Armstrong, K., Kose, S., Williams, L., Woolard, A., & Heckers, S. (2012). Impaired associative inference in patients with schizophrenia. *Schizophrenia Bulletin,38*, 622–629

Averbeck, B. B., Evans, S., Chouhan, V., Bristow, E., & Shergill, S. S. (2011). Probabilistic learning and inference in schizophrenia. *Schizophrenia Research, 127*, 115–122.

Bach, L. J., Happé, F., Fleminger, S., & David, A. S. (2006). Intact theory of mind in TBI with behavioural disturbance. *Brain and Cognition, 60*, 196–198.

Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 3–20). New York: Oxford University Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37–46.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29*, 407–418.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford: Oxford University Press.

Beeke, S. (2003). 'I suppose' as a resource for the construction of turns at talk in agrammatic aphasia. *Clinical Linguistics & Phonetics, 17*, 291–298.

Berthiaume, K. S., Lorch, E. P., & Milich, R. (2010). Getting clued in: Inferential processing and comprehension monitoring in boys with ADHD. *Journal of Attention Disorders, 14*, 31–42.

Bignell, S., & Cain, K. (2007). Pragmatic aspects of communication and language comprehension in groups of children differentiated by teacher ratings of inattention and hyperactivity. *British Journal of Developmental Psychology, 25*, 499–512.

Bishop, D. V. M. (1998). Development of the children's communication checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*, 879–891.

Bishop, D. V. M., & Baird, G. (2001). Parent and teacher report of pragmatic aspects of communication: Use of the Children's communication checklist in a clinical setting. *Developmental Medicine and Child Neurology, 43*, 809–818.

Bishop, D. V. M., Chan, J., Adams, C., Hartley, J., & Weir, F. (2000). Conversational responsiveness in specific language impairment: Evidence of disproportionate pragmatic difficulties in a subset of children. *Development and Psychopathology, 12*, 177–199.

Blake, M. L. (2006). Clinical relevance of discourse characteristics after right hemisphere brain damage. *American Journal of Speech-Language Pathology, 15*, 255–267.

Bodden, M. E., Mollenhauer, B., Trenkwalder, C., Cabanel, N., Eggert, K. M., Unger, M. M., et al. (2010). Affective and cognitive theory of mind in patients with Parkinson's disease. *Parkinsonism & Related Disorders, 16*, 466–470.

Borod, J. C., Rorie, K. D., Pick, L. H., Bloom, R. L., Andelman, F., Campbell, A. L., et al. (2000). Verbal pragmatics following unilateral stroke: Emotional content and valence. *Neuropsychology, 14*, 112–124.

Bosco, F. M., Friedman, O., & Leslie, A. M. (2006). Recognition of pretend and real actions in play by 1- and 2-year-olds: Early success and why they fail. *Cognitive Development, 21*, 3–10.

Botting, N., & Adams, C. (2005). Semantic and inferencing abilities in children with communication disorders. *International Journal of Language & Communication Disorders, 40*, 49–66.

Brownell, H., & Stringfellow, A. (1999). Making requests: Illustrations of how right-hemisphere brain damage can affect discourse production. *Brain and Language, 68*, 442–465.

Brundage, S. (1996). Comparison of proverb interpretations provided by right-hemisphere-damaged adults and adults with probable dementia of the Alzheimer type. *Clinical Aphasiology, 24*, 215–231.

Brüne, M., & Bodenstein, L. (2005). Proverb comprehension reconsidered—'theory of mind' and the pragmatic use of language in schizophrenia. *Schizophrenia Research, 75*, 233–239.

Buitelaar, J. K., van der Wees, M., Swaab-Barneveld, H., & van der Gaag, R. J. (1999). Theory of mind and emotion-recognition functioning in autistic spectrum disorders and in psychiatric control and normal children. *Development and Psychopathology, 11*, 39–58.

Carlomagno, S., Santoro, A., Menditti, A., Pandolfi, M., & Marini, A. (2005). Referential communication in Alzheimer's type dementia. *Cortex, 41*, 520–534.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology, 87*, 299–319.

Carruthers, P., & Smith, P. K. (1996). Introduction. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 1–8). Cambridge: Cambridge University Press.

Castelli, I., Pini, A., Alberoni, M., Liverta-Sempio, O., Baglio, F., Massaro, D., et al. (2011). Mapping levels of theory of mind in Alzheimer's disease: A preliminary study. *Aging & Mental Health, 15*, 157–168.

Chapman, S. B., Ulatowska, H. K., Franklin, L. R., Shobe, A. E., Thompson, J. L., & McIntire, D. D. (1997). Proverb interpretation in fluent aphasia and Alzheimer's disease: Implications beyond abstract thinking. *Aphasiology, 11*, 337–350.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development, 15*, 481–498.

Cheang, H., & Pell, M. (2006). A study of humour and communicative intention following right hemisphere stroke. *Clinical Linguistics & Phonetics, 20*, 447–462.

Coelho, C. A., & Flewellyn, L. (2003). Longitudinal assessment of coherence in an adult with fluent aphasia. *Aphasiology, 17*, 173–182.

Coelho, C. A., Youse, K., & Le, K. (2002). Conversational discourse in closed-head-injured and non-brain-injured adults. *Aphasiology, 16*, 659–672.

Colle, L., Baron-Cohen, S., Wheelwright, S., & van der Lely, H. K. (2008). Narrative discourse in adults with high-functioning autism or Asperger syndrome. *Journal of Autism and Developmental Disorders, 38*, 28–40.

Corcoran, R. (2003). Inductive reasoning and the understanding of intention in schizophrenia. *Cognitive Neuropsychiatry, 8*, 223–235.

Corcoran, R., & Frith, C. D. (1996). Conversational conduct and the symptoms of schizophrenia. *Cognitive Neuropsychiatry, 1*, 305–318.

Cornish, K., Burack, J. A., Rahman, A., Munir, F., Russo, N., & Grant, C. (2005). Theory of mind deficits in children with fragile X syndrome. *Journal of Intellectual Disability Research, 49*, 372–378.

Cuerva, A. G., Sake, L., Kuzis, G., Tiberti, C., Dorrego, F., & Starkstein, S. E. (2001). Theory of mind and pragmatic abilities in dementia. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 14*, 153–158.

Cummings, L. (2005). *Pragmatics: A multidisciplinary perspective*. Edinburgh: Edinburgh University Press.

Cummings, L. (2007a). Pragmatics and adult language disorders: Past achievements and future directions. *Seminars in Speech and Language, 28*, 98–112.

Cummings, L. (2007b). Clinical pragmatics: A field in search of phenomena? *Language & Communication, 27*, 396–432.

Cummings, L. (2008). *Clinical linguistics*. Edinburgh: Edinburgh University Press.

Cummings, L. (2009). *Clinical pragmatics*. Cambridge: Cambridge University Press.

Cummings, L. (2010). Clinical pragmatics. In L. Cummings (Ed.), *The Routledge pragmatics encyclopedia* (pp. 40–43). Abingdon, Oxon: Routledge.

Cummings, L. (2011). Pragmatic disorders and their social impact. *Pragmatics and Society, 2*, 17–36.

Cummings, L. (2012a). Pragmatic disorders. In H. -J. Schmid (Ed.), *Cognitive pragmatics*. [Handbook of pragmatics, vol.4], Berlin and Boston: Walter de Gruyter, 291–315.

Cummings, L. (2012b). Theorising context: The case of clinical pragmatics. In R. Finkbeiner, J. Meibauer & P. Schumacher (Eds.), *What is context? Theoretical and experimental evidence*, Amsterdam: John Benjamins, 55–80.

Cummings, L. (2014a). Clinical pragmatics. In Y. Huang (Ed.), *Oxford handbook of pragmatics*. Oxford: Oxford University Press.

Cummings, L. (2014b). Pragmatic disorders and theory of mind. In L. Cummings (Ed.), *Cambridge Handbook of communication disorders* (pp. 559–577). Cambridge: Cambridge University Press.

Cummings, L. (2014c). *Communication disorders*. Houndmills, Basingstoke: Palgrave Macmillan.

Cutica, I., Bucciarelli, M., & Bara, B. (2006). Neuropragmatics: Extralinguistic pragmatic ability is better preserved in left-hemisphere-damaged patients than in right-hemisphere-damaged patients. *Brain and Language, 98*, 12–25.

David, N., Aumann, C., Bewernick, B. H., Santos, N. S., Lehnhardt, F. G., & Vogeley, K. (2010). Investigation of mentalizing and visuospatial perspective taking for self and other in Asperger syndrome. *Journal of Autism and Developmental Disorders, 40*, 290–299.

de Marchena, A., & Eigsti, I.-M. (2010). Conversational gestures in autism spectrum disorders: Asynchrony but not decreased frequency. *Autism Research, 3*, 311–322.

Dennis, M., & Barnes, M. A. (2001). Comparison of literal, inferential, and intentional text comprehension in children with mild or severe closed head injury. *Journal of Head Trauma Rehabilitation, 16*, 456–468.

Donno, R., Parker, G., Gilmour, J., & Skuse, D. H. (2010). Social communication deficits in disruptive primary-school children. *British Journal of Psychiatry, 196*, 282–289.

Douglas, J. M. (2010). Relation of executive functioning to pragmatic outcome following severe traumatic brain injury. *Journal of Speech, Language, and Hearing Research, 53*, 365–382.

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*, 331–338.

Farrant, B. M., Fletcher, J., & Maybery, M. T. (2006). Specific language impairment, theory of mind, and visual perspective taking: Evidence for simulation theory and the developmental role of language. *Child Development, 77*, 1842–1853.

Fernandez-Duque, D., Baird, J. A., & Black, S. E. (2009). False-belief understanding in frontotemporal dementia and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 31*, 489–497.

Ferstl, E. C., Guthke, T., & von Cramon, D. Y. (2002). Text comprehension after brain injury: Left prefrontal lesions affect inference processes. *Neuropsychology, 16*, 292–308.

Feyereisen, P., Berrewaerts, J., & Hupet, M. (2007). Pragmatic skills in the early stages of Alzheimer's disease: An analysis by means of a referential communication task. *International Journal of Language & Communication Disorders, 42*, 1–17.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: The MIT Press.

Frye, D. (2000). Theory of mind, domain specificity, and reasoning. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 149–167). East Sussex: Psychology Press.

Geurts, H. M., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders, 38*, 1931–1943.

Gibbons, Z. C., Snowden, J. S., Thompson, J. C., Happé, F., Richardson, A., & Neary, D. (2007). Inferring thought and action in motor neurone disease. *Neuropsychologia, 45*, 1196–1207.

Gilmour, J., Hill, B., Place, M., & Skuse, D. H. (2004). Social communication deficits in conduct disorder: A clinical and community survey. *Journal of Child Psychology and Psychiatry, 45*, 967–978.

Girardi, A., Macpherson, S. E., & Abrahams, S. (2011). Deficits in emotional and social cognition in amyotrophic lateral sclerosis. *Neuropsychology, 25*, 53–65.

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 36*, 169–183.

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'Reading the Mind in Films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 38*, 1534–1541.

Gold, R., Faust, M., & Goldstein, A. (2010). Semantic integration during metaphor comprehension in Asperger syndrome. *Brain and Language, 113*, 124–134.

Goldman, A. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences, 16*, 15–28.

Gopnik, A., Capps, L., & Meltzoff, A. N. (2000). Early theories of mind: What the theory theory can tell us about autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 50–72). New York: Oxford University Press.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language, 1*, 158–171.

Gordon, R. M. (1996). 'Radical' simulationism. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 11–21). Cambridge: Cambridge University Press.

Grant, C. M., Apperly, I., & Oliver, C. (2007). Is theory of mind understanding impaired in males with fragile X syndrome? *Journal of Abnormal Child Psychology, 35*, 17–28.

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., et al. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: Theoretical and practical implications. *Brain, 125*, 752–764.

Griffin, R., Friedman, O., Ween, J., Winner, E., Happe, F., & Brownell, H. (2006). Theory of mind and the right cerebral hemisphere: Refining the scope of impairment. *Laterality, 11*, 195–225.

Hale, C. M., & Tager-Flusberg, H. (2005). Social communication in children with autism: The relationship between theory of mind and discourse development. *Autism, 9*, 157–178.

Hamilton, A. F., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition, 113*, 37–44.

Happé, F., Brownell, H., & Winner, E. (1999). Acquired 'theory of mind' impairments following stroke. *Cognition, 70*, 211–240.

Henry, J. D., Phillips, L. H., Crawford, J. R., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: The role of emotion recognition and executive dysfunction. *Neuropsychologia, 44*, 1623–1628.

Holtgraves, T., & McNamara, P. (2010). Pragmatic comprehension deficit in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 32*, 388–397.

Hughes, C., Jaffee, S. R., Happé, F., Taylor, A., Caspi, A., & Moffitt, T. (2005). Origins of individual differences in theory of mind: From nature to nurture? *Child Development, 76*, 356–370.

Jenkins, J. M., Franco, F., Dolins, F., & Sewell, A. (1995). Toddlers' reactions to negative emotion displays: Forming models of relationships. *Infant Behavior and Development, 18*, 273–281.

John, A. E., Rowe, M. L., & Mervis, C. B. (2009). Referential communication skills of children with Williams syndrome: Understanding when messages are not adequate. *American Journal on Intellectual and Developmental Disabilities, 114*, 85–99.

Johnston, J. R., Miller, J., & Tallal, P. (2001). Use of cognitive state predicates by language-impaired children. *International Journal of Language & Communication Disorders, 36*, 349–370.

Jones, C. D., & Schwartz, I. S. (2009). When asking questions is not enough: An observational study of social communication differences in high functioning children with autism. *Journal of Autism and Developmental Disorders, 39*, 432–443.

Kasher, A., Batori, G., Soroker, N., Graves, D., & Zaidel, E. (1999). Effects of right- and left-hemisphere damage on understanding conversational implicatures. *Brain and Language, 68*, 566–590.

Kertesz, A., Jesso, S., Harciarek, M., Blair, M., & McMonagle, P. (2010). What is semantic dementia?: A cohort study of diagnostic features and clinical boundaries. *Archives of Neurology, 67*, 483–489.

Kleinman, J., Marciano, P. L., & Ault, R. L. (2001). Advanced theory of mind in high-functioning adults with autism. *Journal of Autism and Developmental Disorders, 31*, 29–36.

Kuperberg, G. R., McGuire, P. K., & David, A. S. (2000). Sensitivity to linguistic anomalies in spoken sentences: A case study approach to understanding thought disorder in schizophrenia. *Psychological Medicine, 30*, 345–357.

Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2002). Disturbed communication in schizophrenia: The role of poor pragmatics and poor mind-reading. *Psychological Medicine, 32*, 1273–1284.

Laws, G., & Bishop, D. V. M. (2004). Pragmatic language impairment and social deficits in Williams syndrome: A comparison with Down's syndrome and specific language impairment. *International Journal of Language & Communication Disorders, 39*, 45–64.

Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology, 4*, 231–246.

Lind, S. E., & Bowler, D. M. (2010). Impaired performance on see-know tasks amongst children with autism: Evidence of specific difficulties with theory of mind or domain-general task factors? *Journal of Autism and Developmental Disorders, 40*, 479–484.

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders, 33*, 239–251.

Lough, S., Kipps, C. M., Treise, C., Watson, P., Blair, J. R., & Hodges, J. R. (2006). Social reasoning, emotion and empathy in frontotemporal dementia. *Neuropsychologia, 44*, 950–958.

Loukusa, S., Leinonen, E., Kuusikko, S., Jussila, K., Mattila, M.-L., Ryder, N., et al. (2007). Use of context in pragmatic language comprehension by children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 37*, 1049–1059.

MacLennan, D. L., Cornis-Pop, M., Picon-Nieto, L., & Sigford, B. (2002). The prevalence of pragmatic communication impairments in traumatic brain injury. *Premier Outlook, 3*, 38–45.

Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U. (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language, 93*, 46–54.

Marini, A., Spoletini, I., Rubino, I. A., Ciuffa, M., Bria, P., Martinotti, G., et al. (2008). The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research, 105*, 144–155.

Martin, I., & McDonald, S. (2004). An exploration of causes of non-literal language problems in individuals with Asperger syndrome. *Journal of Autism and Developmental Disorders, 34*, 311–328.

Marton, I., Wiener, J., Rogers, M., Moore, C., & Tannock, R. (2009). Empathy and social perspective taking in children with attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology, 37*, 107–118.

Maylor, E. A., Moulson, J. M., Muncer, A.-M., & Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology, 93*, 465–485.

Mazza, M., Di Michele, V., Pollice, R., Casacchia, M., & Roncone, R. (2008). Pragmatic language and theory of mind deficits in people with schizophrenia and their relatives. *Psychopathology, 41*, 254–263.

McDonald, S., & Flanagan, S. (2004). Social perception deficits after traumatic brain injury: Interaction between emotion recognition, mentalizing ability, and social communication. *Neuropsychology, 18*, 572–579.

McInnes, A., Humphries, T., Hogg-Johnson, S., & Tannock, R. (2003). Listening comprehension and working memory are impaired in attention-deficit hyperactivity disorder irrespective of language impairment. *Journal of Abnormal Child Psychology, 31*, 427–443.

McNamara, P., & Durso, R. (2003). Pragmatic communication skills in patients with Parkinson's disease. *Brain and Language, 84*, 414–423.

Milders, M., Ietswaart, M., & Crawford, J. R. (2006). Impairments in theory of mind shortly after traumatic brain injury and at 1-year follow-up. *Neuropsychology, 20*, 400–408.

Miller, C. A. (2001). False belief understanding in children with specific language impairment. *Journal of Communication Disorders, 34*, 73–86.

Miller, C. A. (2004). False belief and sentence complement performance in children with specific language impairment. *International Journal of Language & Communication Disorders, 39*, 191–213.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*, 622–646.

Mo, S., Su, Y., Chan, R. C. K., & Liu, J. (2008). Comprehension of metaphor and irony in schizophrenia during remission: The role of theory of mind and IQ. *Psychiatry Research, 157*, 21–29.

Monetta, L., & Pell, M. D. (2007). Effects of verbal working memory deficits on metaphor comprehension in patients with Parkinson's disease. *Brain and Language, 101*, 80–89.

Monetta, L., Grindrod, C. M., & Pell, M. D. (2009). Irony comprehension and theory of mind deficits in patients with Parkinson's disease. *Cortex, 45*, 972–981.

Montag, C., Dziobek, I., Richter, I. S., Neuhaus, K., Lehmann, A., Sylla, R., et al. (2011). Different aspects of theory of mind in paranoid schizophrenia: Evidence from a video-based assessment. *Psychiatry Research, 186*, 203–209.

Muller, F., Simion, A., Reviriego, E., Galera, C., Mazaux, J.-M., Barat, M., et al. (2010). Exploring theory of mind after severe traumatic brain injury. *Cortex, 46*, 1088–1099.

Murphy, M. M., & Abbeduto, L. (2007). Gender differences in repetitive language in fragile X syndrome. *Journal of Intellectual Disability Research, 51*, 387–400.

Mutter, B., Alcorn, M. B., & Welsh, M. (2006). Theory of mind and executive function: Working-memory capacity and inhibitory control as predictors of false-belief task performance. *Perceptual and Motor Skills, 102*, 819–835.

Myers, P. S. (1979). Profiles of communication deficits in patients with right cerebral hemisphere damage: Implications for diagnosis and treatment. In *Clinical aphasiology conference* (pp. 38–46). Phoenix, AZ: BRK Publishers.

Nelson, P. B., Adamson, L. B., & Bakeman, R. (2008). Toddlers' joint engagement experience facilitates preschoolers' acquisition of theory of mind. *Developmental Science, 11*, 847–852.

Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica, 124*, 106–128.

Papagno, C. (2001). Comprehension of metaphors and idioms in patients with Alzheimer's disease: A longitudinal study. *Brain, 124*, 1450–1460.

Papagno, C., & Vallar, G. (2001). Understanding metaphors and idioms: A single-case neuropsychological study in a person with Down syndrome. *Journal of the International Neuropsychological Society, 7*, 516–528.

Papagno, C., Curti, R., Rizzo, S., Crippa, F., & Colombo, M. R. (2006). Is the right hemisphere involved in idiom comprehension? A neuropsychological study. *Neuropsychology, 20*, 598–606.

Pardini, M., & Nichelli, P. F. (2009). Age-related decline in mentalizing skills across adult life span. *Experimental Aging Research, 35*, 98–106.

Pelc, K., Kornreich, C., Foisy, M.-L., & Dan, B. (2006). Recognition of emotional facial expressions in attention-deficit hyperactivity disorder. *Pediatric Neurology, 35*, 93–97.

Rinaldi, W. (2000). Pragmatic comprehension in secondary school-aged students with specific developmental language disorder. *International Journal of Language & Communication Disorders, 35*, 1–29.

Ripich, D. N., Carpenter, B. D., & Ziol, E. W. (2000). Conversational cohesion patterns in men and women with Alzheimer's disease: A longitudinal study. *International Journal of Language & Communication Disorders, 35*, 49–64.

Roncone, R., Mazza, M., Frangou, I., De Risio, A., Ussorio, D., Tozzini, C., et al. (2004). Rehabilitation of theory of mind deficit in schizophrenia: A pilot study of metacognitive strategies in group treatment. *Neuropsychological Rehabilitation: An International Journal, 14*, 421–435.

Rousseaux, M., Sève, A., Vallet, M., Pasquier, F., & Mackowiak-Cordoliani, M. A. (2010). An analysis of communication in conversation in patients with dementia. *Neuropsychologia, 48*, 3884–3890.

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. A. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology, 34*, 161–174.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mother's mental state language and theory-of-mind understanding. *Child Development, 73*, 734–751.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders, 32*, 189–194.

Ryder, N., Leinonen, E., & Schulz, J. (2008). Cognitive approach to assessing pragmatic language comprehension in children with specific language impairment. *International Journal of Language & Communication Disorders, 43*, 427–447.

Saldert, C., Fors, A., Ströberg, S., & Hartelius, L. (2010). Comprehension of complex discourse in different stages of Huntington's disease. *International Journal of Language & Communication Disorders, 45*, 656–669.

Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of mind and executive functions in normal human aging and Parkinson's disease. *Journal of the International Neuropsychological Society, 6*, 781–788.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and 'theory of mind'. *Mind & Language, 14*, 131–153.

Siegal, M., & Varley, R. (2006). Aphasia, language, and theory of mind. *Social Neuroscience, 1*, 167–174.

Spanoudis, G., Natsopoulos, D., & Panayiotou, G. (2007). Mental verbs and pragmatic language difficulties. *International Journal of Language & Communication Disorders, 42*, 487–504.

Speechley, W. J., Murray, C. B., McKay, R. M., Munz, M. T., & Ngan, E. T. C. (2010). A failure of conflict to modulate dual-stream processing may underlie the formation and maintenance of delusions. *European Psychiatry, 25*, 80–86.

Spek, A. A., Scholte, E. M., & Van Berckelaer-Onnes, I. A. (2010). Theory of mind in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders, 40*, 280–289.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language, 17*, 3–23.

Sudhalter, V., & Belser, R. C. (2001). Conversational characteristics of children with fragile X syndrome: Tangential language. *American Journal on Mental Retardation, 106*, 389–400.

Sullivan, S., & Ruffman, T. (2004). Social understanding: How does it fare with advancing years? *British Journal of Psychology, 95*, 1–18.

Sullivan, K., & Tager-Flusberg, H. (1999). Second-order belief attribution in Williams syndrome: Intact or impaired? *American Journal of Mental Retardation, 104*, 523–532.

Sullivan, K., Winner, E., & Tager-Flusberg, H. (2003). Can adolescents with Williams syndrome tell the difference between lies and jokes? *Developmental Neuropsychology, 23*, 85–103.

Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are children with autism deaf to Gricean maxims? *Cognitive Neuropsychiatry, 1*, 55–72.

Surian, L., & Siegal, M. (2001). Sources of performance on theory of mind tasks in right hemisphere-damaged patients. *Brain and Language, 78*, 224–232.

Swettenham, J. (2000). Teaching theory of mind to individuals with autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 442–456). New York: Oxford University Press.

Symons, D., McLaughlin, E., Moore, C., & Morine, S. (1997). Integrating relationship constructs and emotional experience into false belief tasks in preschool children. *Journal of Experimental Child Psychology, 67*, 423–447.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 124–149). New York: Oxford University Press.

Tavano, A., Sponda, S., Fabbro, F., Perlini, C., Rambaldelli, G., Ferro, A., et al. (2008). Specific linguistic and pragmatic deficits in Italian patients with schizophrenia. *Schizophrenia Research, 102*, 53–62.

Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*, 436–455.

Tényi, T., Herold, R., Szili, I. M., & Trixler, M. (2002). Schizophrenics show a failure in the decoding of violations of conversational implicatures. *Psychopathology, 35*, 25–27.

Titone, D., Ditman, T., Holzman, P. S., Eichenbaum, H., & Levy, D. L. (2004). Transitive inference in schizophrenia: Impairments in relational memory organization. *Schizophrenia Research, 68*, 235–247.

Togher, L., & Hand, L. (1998). Use of politeness markers with different communication partners: An investigation of five subjects with traumatic brain injury. *Aphasiology, 12*, 755–770.

Torralva, T., Roca, M., Gleichgerrcht, E., Bekinschtein, T., & Manes, F. (2009). A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. *Brain, 132*, 1299–1309.

Turkstra, L. S., Dixon, T. M., & Baker, K. K. (2004). Theory of mind and social beliefs in adolescents with traumatic brain injury. *NeuroRehabilitation, 19*, 245–256.

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Current Biology, 10*, 723–726.

Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase, 7*, 489–493.

Watson, A. C., Painter, K. M., & Bornstein, M. H. (2001). Longitudinal relations between 2-year-olds' language and 4-year-olds' theory of mind. *Journal of Cognition and Development, 2*, 449–457.

Weed, E. (2008). Theory of mind impairment in right hemisphere damage: A review of the evidence. *International Journal of Speech-Language Pathology, 10*, 414–424.

Weed, E., McGregor, W., Nielsen, J. F., Roepstorff, A., & Frith, U. (2010). Theory of mind in adults with right hemisphere damage: What's the story? *Brain and Language, 113*, 65–72.

Wellman, H. M., & Lagattuta, K. H. (2000). Developing understandings of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 21–49). New York: Oxford University Press.

Williams, D., & Happé, F. (2010). Representing intentions in self and other: Studies of autism and typical development. *Developmental Science, 13*, 307–319.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language, 62*, 89–106.

Yirmiya, N., Solomonica-Levi, D., Shulman, C., & Pilowsky, T. (1996). Theory of mind abilities in individuals with autism, Down syndrome, and mental retardation of unknown etiology: The role of age and intelligence. *Journal of Child Psychology and Psychiatry, 37*, 1003–1014.

Relevance Theory

Nicholas Allott

Abstract Relevance theory is a framework for the study of cognition, proposed primarily in order to provide a psychologically realistic account of communication. This paper (1) presents relevance theory's central commitments in detail and explains the theoretical motivations behind them; and (2) shows some of the ways in which these core principles are brought to bear on empirical problems. The core of relevance theory can be divided into two sets of assumptions. Assumptions relating to cognition in general include the definition of relevance as a trade-off between effort and effects, and the claim that cognition tends to maximise relevance. Assumptions about communication include the claims that understanding an utterance is a matter of inferring the speaker's communicative and informative intentions; and that the communicative principle of relevance and the presumption of optimal relevance mandate the relevance-theoretic comprehension procedure, a heuristic that guides the search for the intended interpretation of utterances. Relevance theorists model communication in terms of the working of this comprehension procedure. There are, in addition, several strategies that guide the explanation of phenomena in relevance theory, including: (1) a stronger form of Grice's Modified Occam's Razor, (2) the possibility of dividing what is linguistically encoded between conceptual and procedural information; (3) the interpretive/descriptive distinction; (4) the use of *ad hoc* concepts.

N. Allott (✉)
CSMN, University of Oslo, Oslo, Norway
e-mail: nicholas.allott@gmail.com

1 Introduction: The Relevance-Theoretic Research Programme

Relevance theory is a rather wide-ranging framework (or 'research programme'— see below) for the study of cognition, devised primarily in order to provide an account of communication that is psychologically realistic and empirically plausible. It was originally proposed by Sperber and Wilson (1986b; 1987). Other key publications include Blakemore 1987; Sperber and Wilson 1995; Carston 2002 and Wilson and Sperber 2012.

For some time relevance theory has been one of the leading programmes of research in pragmatics. There has been work within the relevance-theoretic framework[1] on such central topics as scalar implicatures (Carston 1998; Breheny et al. 2006; Noveck and Sperber 2007), bridging (Wilson and Matsui 1998; Matsui 2000), speech acts and mood (Sperber and Wilson 1986b, pp. 243–254; Wilson and Sperber 1988; Jary 2007; Jary 2010), disambiguation (Sperber and Wilson 1986b, pp. 183–193), discourse particles (Blakemore 1987, 2000, 2002, 2004; Iten 2005), evidentials (Ifantidou 2001), loose talk (Sperber and Wilson 1986a; Carston 1997a; Wilson and Sperber 2002), literary language (Sperber and Wilson 1986b, Chap. 4; Clark 1996; Pilkington 2000; Sperber and Wilson 2008), genre (Unger 2006), translation (Gutt 1991), non-verbal communication (Wharton 2009), the referential/attributive distinction (Rouchota 1992; Bezuidenhout 1997; Powell 2001, 2010) and rhetorical tropes such as metaphor (Sperber and Wilson 1986b, pp. 231–237, 2008; Carston 1997a, 2010b; Vega Moreno 2007) and irony (Sperber and Wilson 1981, 1986b, pp. 237–243, 1998a; Wilson and Sperber 1992; Wilson 2006). The theory has had considerable influence in the disputed borderlands between semantics, pragmatics and philosophy of language, including ongoing debates about the distinction between what is explicitly and what implicitly communicated, and the extent to which pragmatic inference affects the proposition expressed by an utterance (Wilson and Sperber 1981; Sperber and Wilson 1986b, Chap. 4; Carston 1988, 2002, Chap. 2–5, 2010a). Relevance theory has also inspired considerable work on the application of experimental and developmental evidence to pragmatics and related questions in the psychology of reasoning, helping to shape the emerging field of experimental pragmatics (Jorgensen et al. 1984; Happé 1993; Sperber et al. 1995; Bezuidenhout and Sroda 1998; Nicolle and Clark 1999; van der Henst et al. 2002a, b; Happé and Loth 2002; Noveck and Sperber 2004, 2007; Breheny et al. 2006; Chevallier et al. 2010, 2011).

Despite its reach and popularity, however, relevance theory is poorly understood beyond its practitioners. There is confusion among both linguists and

---

[1] The references given here are far from exhaustive. For many more references, sorted by author and by subject matter, see Francisco Yus' online relevance theory bibliography at http://www.ua.es/personal/francisco.yus/rt.html.

philosophers about what relevance theorists are committed to and what kinds of explanations they attempt to give.[2]

This paper attempts to clarify these issues by i) presenting relevance theory's central commitments in detail and explaining the theoretical motivations behind them; and ii) showing some of the ways in which these core principles are brought to bear on empirical problems.

As Wilson and Sperber say:

> Like other psychological theories, [relevance theory] has testable consequences: it can suggest experimental research, and is open to confirmation, disconfirmation, or fine-tuning in the light of experimental evidence. As with other theories of comparable scope, its most general claims can only be tested indirectly. For example, the Cognitive Principle of Relevance suggests testable predictions only when combined with descriptions of particular cognitive mechanisms (e.g. for perception, categorization, memory, or inference). (Wilson and Sperber 2004, pp. 625–626)

There are echoes here of the model of scientific research proposed by Imre Lakatos. According to Lakatos, scientists work within competing research programmes, and each research programme has two components: (1) a 'hard core' of fundamental theoretical commitments; and (2) auxiliary hypotheses (Lakatos 1968, p. 168 ff.).[3] Most of the predictions relevance theory makes do not follow from the hard core on its own: they only follow once the auxiliary hypotheses are also taken into account.[4] Each research programme also has a positive heuristic, specifying strategies for forming theories outside of the hard core: i.e. suggesting what 'paths of research' to pursue (Lakatos 1968, p. 168).

This paper describes the central assumptions of relevance theory in detail and then sketches some of the strategies that relevance theorists use in developing theories beyond that core.

1.1 The Central Assumptions and Positive Heuristic of Relevance Theory

The core of relevance theory can be divided into two sets of assumptions. Assumptions in the first set relate to cognition in general, assumptions in the second to communication more specifically, particularly to utterance interpretation.

---

[2] See Sperber and Wilson's replies to comments on their precis of 'Relevance' in Behavioral and Brain Science (Sperber and Wilson 1987), and Wedgewood (2007) and Kjøll (2010) who have argued that certain recent criticisms of relevance theory in the philosophy of language literature are based on fundamental misunderstandings about relevance theory's commitments.

[3] See also Lakatos 1970. Lakatos' papers on the methodology of science are collected in volume 1 of Lakatos et al. 1978. For critical commentary see Hacking 1979.

[4] Lakatos also claims that the core commitments are to be kept, while auxiliary hypotheses should be modified or disposed of in response to empirical challenges. (He calls this the 'negative heuristic': Lakatos 1968, p. 169.) I return to this point briefly in the conclusion of this paper, where I discuss some changes that have occurred in the core of relevance theory.

The central assumptions that relevance theory makes about human cognition include the definition of relevance as a trade-off between effort and effects; the cognitive principle of relevance, which is the claim that cognition tends to max-imise relevance; and the views, shared with other work in cognitive science, that cognition is a matter of (or at least can be well modelled as) computations over mental representations, and that human beings possess a 'deductive device' which plays a central role in spontaneous inference. I set out these core assumptions relating to cognition in Sect. 2 of this paper.

The core of relevance theory as it relates specifically to communication includes the Gricean claim that understanding an utterance is a matter of inferring what the speaker intended to convey from what she utters (in what way, in what circumstances). Another fundamental of relevance theory, departing somewhat from Grice, is that there are exactly two speaker's intentions that are central to communication, namely the informative intention and the communicative inten-tion. The last main part of the hard core relating specifically to communication is entirely original to relevance theory: the communicative principle of relevance and the presumption of optimal relevance, which mandate the relevance-theoretic comprehension procedure, a heuristic that guides the search for the correct (i.e. intended) interpretation of utterances. I examine the core assumptions that are specific to communication in Sect. 3.

The characteristic approach of relevance theory to the explanation of com-municative phenomena is a corollary of its central commitments. Relevance the-orists try to give psychologically realistic explanations and to understand communicated meaning in terms of the working of the relevance-theoretic com-prehension procedure. This way of working is at the heart of relevance theory's 'positive heuristic', but in Sect. 4 I show that there are several additional strategies that guide the explanation of phenomena in relevance theory including: (1) Grice's Modified Occam's Razor, in a stronger form; (2) the possibility of dividing what is linguistically encoded between conceptual and procedural information; (3) the interpretive/descriptive distinction; (4) the use of *ad hoc* concepts.

2 Relevance Theory and Cognition

The central claim of relevance theory is that, as a result of constant selection pressures, the human cognitive system has developed a variety of dedicated (innate or acquired) mental mechanisms or biases which tend to allocate attention to inputs with the greatest expected relevance, and process them in the most relevance-enhancing way. (Wilson 2009, p. 394)

2.1 The Cognitive Principle of Relevance

At the centre of the hard core of relevance theory are the cognitive principle of relevance and the definition of relevance as a trade-off of cognitive benefit against processing cost. The cognitive principle is the hypothesis that cognitive systems tend to maximise relevance.

2.1.1 Cognitive Principle of Relevance

Human cognition tends to be geared to the maximisation of relevance. (Sperber and Wilson 1986b, p. 260)

'Relevance' here is a technical term. It is defined as a property of inputs to cognitive systems: an input is more relevant the more cognitive effects it yields, and less relevant the more mental effort it takes to process.

2.1.2 Relevance of an Input to an Individual

(a) Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.

(b) Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time. (Wilson and Sperber 2004, p. 609; c.f. the original formulation, at Sperber and Wilson 1986b, p. 153)

On this definition of relevance, the cognitive principle is the claim that human cognitive systems tend to work with their input in such a way as to yield the maximum cognitive benefit for the least mental effort. The reach of this principle is rather broad. For its purposes, cognitive systems include (at least) those that are centrally involved in perception, memory and reasoning as well as those that underpin the production and interpretation of utterances.

The definition of relevance obviously raises two questions: (1) what constitutes cognitive effects; and (2) what causes mental effort? Relevance theory gives definite, although not necessarily exhaustive answers to these two questions, and I set them out below.

A less obvious question concerns the cognitive principle: *How* do cognitive systems maximise relevance? Is it, for example, by systematically minimising effort or by systematically maximising benefit? It is compatible with the cognitive principle that different cognitive systems implement different approaches to maximisation. However, we will see below (a) that relevance theory has a general account of how the mind as a whole directs effort to tasks that yield cognitive effects, and (b) that much more specific claims are made about how the system for

interpreting utterances seeks relevance (for discussion see Sperber 2005). But before I go into these answers, I want to sketch out the intuitive reasons for the core assumptions set out above.

Relevance theory starts from the idea that there is normally much more going on in the environment of any human being than it could pay attention to, and certainly much more than it could mentally process fully. (For discussion, see Sperber and Wilson 1996; Allott 2008, Chap. 3.) If this were not the case, there would be no need to consider a trade-off between the effort put in and the benefit extracted from doing so. We could process each input fully to extract all the cognitive benefit it might yield, and theories of cognition could ignore processing effort. However it is highly plausible that the environment is too full, and processing too costly, for this abstraction to be justified, particularly considering that by 'environment', here, one must understand not just physical objects, but also sources of information such as utterances made by other human beings, books, the internet, advertisements etc.[5] (Sperber and Wilson 1996, p. 530; Todd and Gigerenzer 2000, pp. 729–730). This crucial assumption which underlies relevance theory—that we cannot maximise by considering all options and processing each of them as deeply as possible—is shared with work on 'bounded rationality', pioneered by Herbert Simon, and including research on 'simple heuristics' by Gerd Gigerenzer and colleagues (Simon 1957; Cherniak 1981; Gigerenzer and Goldstein 1996; Gigerenzer and Todd 1999). I return to these parallels in the discussion of the relevance theoretic comprehension procedure in Sect. 3.7 below.

A further assumption is required to justify the conclusion that our cognitive systems tend to get a good return on effort expended. That assumption is, roughly, that our cognitive systems are well-adapted to their normal environments. In lectures, Sperber quotes the biologist Dobzhansky: "nothing makes sense in biology except in the light of evolution" (Dobzhansky 1964, p. 449). Human beings are evolved creatures; and complex subsystems including physical organs like the heart, brain and skin and cognitive systems such as memory, face-recognition, ability to communicate etc. must therefore be seen as having been subject to selection pressure. In addition, children's abilities and knowledge develop from infancy, assuming that the child is in an appropriate environment. Thus we should expect 'normal' adults, on average, to be well adapted to normal environments.

There is an analogy with an animal that forages for food, such as a monkey living in the canopy of a rainforest. It will look for things that have a high nutritional payoff: ripe fruit probably contain more energy than leaves, for example. But the monkey cannot just be built to pursue high-energy food at any cost. There must be some balancing of the nutritional payoff against the costs required to obtain and process the food. Fruit that are far away and hard to reach are not as good as fruit that are to hand. A well-adapted creature should tend to eat nearby fruit first, before

---

[5] The 'environment' of each cognitive system is still richer, since it includes outputs from other cognitive systems. For example, our general reasoning is fed by memory, not just by our perceptions of the external environment.

investigating food that is up at the end of narrow branches and difficult to reach. Equally, we would expect it to go for food that can be eaten straightaway if it can find it, rather than fruit or nuts with hard shells that require a great deal of effort to open. That is not to say that monkeys never bother with fruit that are difficult to process: in fact, some of them use stones to smash open tough fruit, seeds and tubers (de Moura and Lee 2004, p. 1909), but presumably they only do this if the tough food is much more nutritious than the other available food sources.[6]

According to relevance theory, something very similar applies to human cognition. The cognitive system should (if it is well adapted) be so constructed that it seeks and processes inputs that are cognitively valuable, all other things being equal; and, on the other hand, that it looks for things that are easy to process, all else being equal. If something is difficult to process, then it will only be worth attending to if the payoff is big enough (where how big that is depends on the other possible sources of cognitive nutrition in the environment, and on the organism's general state of alertness and stores of energy). Conversely, if an input has a low payoff then it will only be worth processing if that is easy to do (where, again, how easy that needs to be depends on the other potential sources of relevance, and alertness and energy).

As well as these parallels with foraging theory, relevance theory's fundamental dependence on notions of cost-benefit trade-off and maximisation make it an intellectual cousin of game theory and rational decision theory, areas which study decision making on the assumption that agents are rational maximisers.[7] The parallel is closer with fields such as foraging theory and evolutionary game theory than with standard game theory (Allott 2006, p. 147). The basis of the models in these fields (as of the cognitive principle of relevance) is not that agents or their cognitive systems are aware of all the potentially relevant details of the structure of the environment, nor that they use this information to maximise rationally—the 'Common Knowledge and Rationality' assumptions of standard game theory. Insofar as the cognitive principle of relevance is a principle of rational maximisation, the kind of rationality involved is of the evolutionary, adaptive sort: that is, it is assumed that evolution and development have selected for systems which produce behaviour that tends to maximise return in normal environments by working with limited information and taking shortcuts.

It should also be clear that no higher-level rationality is necessarily involved, that is, the kind of rationality that requires awareness of and openness to reasons, the ability to reflect on actions and their consequences and so on (Evans and Over 1996; Sloman 1996). Of course, human beings are (sometimes) capable of such reflection, but it is not our reflective abilities that are supposed to underwrite the adaptive rationality summarized in the cognitive principle. Rather, the cognitive principle is supposed to apply to all aspects of human cognition, including such largely automatic,

---

[6] De Moura and Lee say that the capuchins they studied, "living in a harsh dry habitat, survive food limitation and foraging time constraints through their extensive tool use." (p. 1909). On animal foraging more generally, see Emlen 1966; Stephens and Krebs 1986; Stephens et al. 2007.

[7] Optimal foraging theory is also in this intellectual territory, since it can be seen as an application of rational decision theory.

non-reflective systems as the face-recognition module and our innate tendency to attend to loud noises, as well as to reflective, conscious, 'person-level' reasoning.

2.2 The Payoff: Cognitive Effects

> A positive cognitive effect is a worthwhile difference to the individual's representation of the world (Wilson and Sperber 2004, p. 608).

> … the addition of new information which merely duplicates old information does not count as an improvement; nor does the addition of new information which is entirely unrelated to old information. The sort of effect we are interested in is a interaction between old and new information. (Sperber and Wilson 1986b, p. 106)

In relevance theory, benefit to cognition is seen as a matter of the positive cognitive effects—the worthwhile changes in the individual's cognitive system, including improvements in her representation of the world—that are produced in an individual by processing an input in a context.[8] Changes in the representation that make it less good for "the fulfilment of cognitive functions or goals" (Sperber and Wilson 1995, p. 265) (such as changes that take it further away from accurately representing the world) are cognitive effects, but not positive ones, and they contribute not to actual relevance but (in some cases) to how relevant an input *seems* (Sperber and Wilson 1995, p. 263ff).[9]

What counts as a improvement in an individual's representation of the world? Sperber (2005, p. 65) lists several ways that our knowledge can be fruitfully revised as a result of processing new inputs:

> adding new pieces of knowledge, updating or revising old ones, updating degrees of subjective probability in a way sensitive to new evidence, or merely reorganizing existing knowledge so as to facilitate future use.

Simplifying a bit, the three types of cognitive effect normally discussed in relevance theory are as follows:

2.2.1 Cognitive Effects

1. to support and strengthen an existing assumption;
2. to contradict and rule out an existing assumption;
3. to interact inferentially with existing assumptions to produce a new conclusion.

---

[8] Cognitive effects are sometimes called contextual effects, particularly in Sperber and Wilson 1986b.

[9] This is a change from the definition of cognitive effects in Sperber and Wilson 1986b. This is a considerable change in principle, but it may not imply much difference in processing: Sperber (2005, p. 65) suggests that in practice "the brain would be roughly right in treating any and every cognitive effect as a positive effect, in other words, as a cognitive benefit."

In this definition, assumptions are mental representations of aspects of the world: propositions that are believed by the individual, or at least given some degree of credence.

The first kind of cognitive effect is to raise the degree of credence that an individual accords to a particular assumption: e.g. from *probable* to *almost certain.* For example, Mary, who is about to enter King's Cross station, believes that it is probable that there will be a train to Newcastle within the hour (since she believes that there are several each hour during the daytime, and that it is daytime, and has no good reason to think that there is a rail strike, etc.). Entering the station she sees that it is 9.20 and there is a 9.46 train for Newcastle listed on the departure board. Her original belief is strongly reinforced.

The second type of cognitive effect is to reduce to nil the credence that the individual attaches to an assumption. Suppose that when Mary looks at the departure board the first Newcastle train listed is at 11.20, or that the board is displaying a notice saying 'All trains cancelled'. Either of these bits of input would contradict her original belief and—in normal circumstances and absent contrary evidence—either would be credible enough to rule it out.[10]

As an illustration of the third type of cognitive effect, suppose now that Mary knows that there is a newspaper shop in the station, and has normal beliefs about how long it takes to buy a newspaper, and, once she is in the station, can see how far it is from the shop to the train. When she enters the station at 9.20 and sees that the next train is at 9.46 she may infer that she has time to buy a newspaper before boarding the train. This is a cognitive effect of the third type. The new input—the time of the next train—interacts inferentially with assumptions that were already available to Mary—about the availability of newspapers, and the time taken to get one and to get to the train—to yield a new conclusion.

Note that it is part of the criterion for this to be a cognitive effect that the interaction between the beliefs is inferential. From *It is 9.20*; *The next train is at 9.46*; and *It takes no more than 10 min to buy a newspaper here*, it follows that there is time to buy a newspaper, so this is a *bona fide* cognitive effect. In contrast, an input that causes a new assumption in a purely associative way does not count as a cognitive effect: e.g. the thought that the train is at 9.46 reminds you of granny since she lives at number 46, which in turn reminds you that you should visit her soon. I return below to the assumptions that relevance theory makes about the role of inference in cognition.

Note also that it is a deliberate feature of Sperber and Wilson's characterisation of cognitive effects that learning new information that has no relation to any previously held assumption does not count as a cognitive effect, even if the new

---

[10] One might wonder why the lowering of credence in an assumption only counts as a cognitive effect if it lowers it to zero i.e. eliminates it as an assumption. Briefly, it is because Sperber and Wilson assume that "[mere] weakening is always a by-product of a more basic contextual effect" (Sperber and Wilson 1986b, pp. 294, fn d), for example the elimination of another assumption which provided support for the one that is weakened—and so mere weakening does not need to be counted separately.

information is true. They say that new information that "is entirely unconnected with anything in the individual's representation of the world … can only be added to this representation as isolated bits and pieces, and this usually means too much processing cost for too little benefit." (Sperber and Wilson 1986b, p. 48)

2.3 The Cost: Processing Effort

What is meant in relevance theory by 'processing effort' is the effort required to process an input *to the point that its cognitive effects are derived*. More specifically, this is the effort taken to "to represent the input, access contextual information and derive any cognitive effects" (Wilson 2009, p. 394). This effort is therefore a sum of the effort involved in perception, memory and inference (Wilson 2009, p. 394).

Beyond this general characterisation, relevance theory does not try to define sources of processing effort a priori. Instead it works with the results of the fields of psychology which study perception, memory and inference. Relevant research includes work on attention in perception (e.g. Lavie 1995, 2001; Pashler 1998), in psycholinguistics on retrieval of word senses and disambiguation, which has tended to focus on effort factors, (e.g. Meyer and Schvaneveldt 1971; Neely 1991; Forster and Chambers 1973), and in the psychology of reasoning on the varying costs of different types of inference (e.g. Braine 1978; Braine and O'Brien 1998; Rips 1983; Johnson-Laird 1983).

Different stimuli will in general require different amounts of processing effort. For example, a longer sentence will (other things being equal) require more effort to process than a shorter one. An uncommon word, or an uncommon sense of an ambiguous word, requires more effort to process than a common one (Forster and Chambers 1973).

A more subtle point is that the same stimulus in different contexts will generally require different amounts of processing effort. This is because in different contexts the stimulus may be more or less salient (i.e. more or less easy to perceive); the contextual assumptions required to process it may be more or less accessible (i.e. more or less easy to retrieve from memory or derive); the inferences required to draw out its implications may be more or less involved and demanding, and, indeed, what implications it supports will also depend on the context (Wilson and Sperber 2004, p. 609).

2.4 How Do We Maximise Relevance?

Within relevance theory, the problem is not so much to assess contextual effects and processing effort from the outside, but to describe how the mind assesses its own achievements and efforts from the inside, and decides as a result to pursue its efforts or reallocate them in different directions. (Sperber and Wilson 1986b, p. 130)

The picture of cognition that relevance theory assumes is of a number of possible inputs dealt with by a number of mental processes running in parallel. Processes and inputs that are cognitively productive—e.g. returning a lot of effects for reasonable effort, or returning reasonable amounts of effects for low effort—will be preferentially given resources:

> cognitive resources tend to be allocated to the processing of the most relevant inputs available .... human cognition tends to be geared to the maximisation of the cumulative relevance of the inputs in processes. It does this not by pursuing a long-term policy based on computation of the cumulative relevance achieved over time, but by local arbitrations, aimed at incremental gains, between simultaneously available inputs competing for immediately available resources (Sperber and Wilson 1995, p. 261).[11]

In a system like this, there is no need for the cognitive systems to calculate ahead of time what the relevance of an input is going to be. That is just as well, since that would probably be self-defeating, requiring huge processing effort (Sperber 2005, p. 64). The reason is that it is very costly to calculate an optimal stopping point for a search. Simple heuristics that process until some target is achieved, or threshold reached, are much less computationally expensive (Sperber and Wilson 1986b, pp. 130–131; Todd and Gigerenzer 2000, pp. 729–730; Gigerenzer 2004, p. 391; Allott 2008, pp. 170–172).

In fact, relevance theory takes an even stronger line here. It claims that generally we do not mentally represent processing effort or cognitive effects, so they could not enter into calculations of whether to proceed in processing, and that when represented at all they are represented as comparative (not absolute or quantitative) judgments. Our awareness of mental effort and effects, Sperber and Wilson speculate, may depend on our awareness of "symptomatic physico-chemical changes" that they cause (1986b, p. 130) in much the way that we have a sense of how much physical effort is being taken up in lifting a certain object, or how filling a meal is (see also Sperber 2005, pp. 64–66).

In support of the assumption that we do not, in general, mentally represent mental effort or effects, Sperber and Wilson argue that we are not in fact able to "compare the contextual effects and processing effort involved in any [arbitrary] pair of mental performances" (1986b, p. 131) and that it is "implausible that human beings might have a system for computing and representing the strength of assumptions which is both wholly unconscious and radically more sophisticated than anything that is reflected in their conscious intuitions" (1986b, p. 79).

On the assumption that effects and effort are not mentally represented, it follows that relevance, which is defined in terms of them, is also a non-representational notion. For Sperber and Wilson, "relevance is a property which need not be represented, let alone computed, in order to be achieved" (1986b, p. 132). As with effort and effects, our sense of relevance is intuitive and comparative, rather than absolute.

---

[11] See also Sperber 1994a, pp. 46–50 and Sperber 2005, p. 63ff. Similar models include the pandemonium model (Selfridge and Neisser 1960), and 'enzymatic computation' (Barrett 2005).

To summarize: there are two reasons why the search for relevance cannot be driven by calculations of how profitable it will be to process an input: (1) the processing required would be too costly; (2) the quantities required for the calculation are not mentally represented, and therefore are not available to be computed over.

Instead, then, our search for relevance is fed by our internal sense of how cognitively profitable and demanding a certain input or task has been proving. Of course, the allocation of resources should also be guided by expectations of future cost and benefit. Monkeys may anticipate good returns from foraging in a guava tree, and we surely expect more cognitive nutrition from a book by Chomsky than from one by Dan Brown. As Sperber and Wilson put it, as well as *retrospective* intuitions, we have *prospective* intuitions about the effort a task will take and the effects that will be achieved (1986b, p. 130). As we will see, it is central to relevance theory's account of utterance interpretation that in communicative interactions there is a very specific expectation about the degree of relevance that each utterance should attain.

2.5 Cognition, Mental Representation and Inference

The assumptions that relevance theory makes about the way that human cognition performs inference are fundamental to the explanations that it gives of utterance interpretation, and more generally to understanding its definitions of cognitive effects and therefore relevance. They are set out in some detail in Chap. 2 of Sperber and Wilson (1986b), but they have not been much discussed in subsequent work. For the purposes of the brief summary possible here, it is convenient to divide them into two postulates, one of which—the Computational/Representational Theory of Mind (Fodor 1975)—is adopted in some form across much work in cognitive science, and one of which—the deductive device—is more specific to relevance theory.

The first assumption, then, is that cognition can be modelled in terms of computations performed on mental representations. This assumption is what I am calling the Computational/Representational Theory of Mind (C/RTM). This theory has two central commitments. The first is that the form of a mental representation determines the way that it is processed, since the computational rules that operate on mental representations are sensitive only to their formal (i.e. syntactic) properties. For example, the following representation has the form 'P and Q':

1. John studies linguistics and Mary studies philosophy.

Starting from any representation with the form 'P and Q' as a premise, one can deduce P as a conclusion: in this case *John studies linguistics*. (Of course, one can also deduce Q as a conclusion.)

The second central commitment of C/RTM is that the mind's syntactic operations generally preserve semantic value. For systems that take propositional input and produce propositional output the value preserved will be *truth*. We can see that this is the case for the example given, because any situation in which the premise (John studies linguistics and Mary studies philosophy) is true is one in which the conclusion (John studies linguistics) is also true.

The point of C/RTM is that it provides some insight into the way that one thought leads to another in reasoning and inference: the syntax of a thought "determine[s] the causes and effects of its tokenings in much the way that the geometry of a key determines what locks it will open" (Fodor 1985, p. 93). Representations of the form 'P and Q' both lead to (i.e. cause) and logically entail representations of the form 'Q'. Thus, Fodor writes, "the syntactic theory of mental operations provides a reductive account of the *intelligence* of thought." (1985, p. 98. His emphasis.) I do not go into more detail here about C/RTM. It has been discussed at length in the philosophy of mind and, as noted, something along these lines is commonly assumed in cognitive science.[12]

Sperber and Wilson's second postulate about the cognitive realisation of inference is that human beings possess a *deductive device.* On their view, human beings are not only equipped with the ability to make logical deductions based on the form of mental representations. In addition, (1) these rules are "spontaneously brought to bear in the deductive processing of information" (Sperber and Wilson 1986b) and (2) this also plays a central role in spontaneous non-demonstrative reasoning. I explain the first of these points here, and return to the second point in Sect. 3.8 below, after discussing the role of non-demonstrative reasoning in utterance interpretation.

The deductive device starts with some input (the premises) and performs all deductive inferences that are possible from each premise and from the premises taken in conjunction, recursively (i.e. also operating on the output of the rules), where what is possible is determined by the set of rules possessed by the deductive device, and by the context in which the input is processed. According to Sperber and Wilson, the deductive device has elimination rules like the one mentioned above that takes input of the form 'P and Q' and returns 'P' as output. So if you start with 'John studies linguistics and Mary studies philosophy' you immediately have available to you 'John studies linguistics'. Suppose now that you already believe, falsely, that *if anyone studies linguistics then he/she is a polyglot*. In that case you will put that together with what you have just deduced and infer, perhaps wrongly, that John is a polyglot. In the terminology of relevance theory, in this example 'John studies linguistics' is an *analytic consequence* of the input (i.e. one that can be reached purely through the use of elimination rules); 'If someone studies linguistics then he/she is a polyglot' is a *contextual assumption*; and 'John is a polyglot' is a *contextual*

---

implication of the input (i.e. an implication of the input taken together with one or more contextual assumptions). Contextual implications are identical to the third type of cognitive effect in the definition of cognitive effects given in Sect. 2.2 above.

There are a lot of details to spell out about how the deductive device works, and not enough room here to go into them in depth. There are similarities to the system for deduction postulated in 'mental logic' theories of reasoning (e.g. Braine 1978; Rips 1983; Braine and O'Brien 1998). One important detail is original to relevance theory. In order to stop overgeneration, Sperber and Wilson postulate that the deductive device does not have introduction rules (1986b, p. 96). For example, given a representation of the form 'P', it does not generate a representation of the form 'P or Q', although this would be a logically impeccable deductive inference. Another crucial detail is that the output of the deductive rules is monitored for redundancy (generating something that is already present) and for contradiction (generating a mental representation that is the logical negation of one that is already present), and in each case, suitable action is taken (Sperber and Wilson 1986b, p. 95).

The deductive device is important for relevance theory in several ways. One important point is that the three types of cognitive effect discussed in Sect. 2.2 above are a corollary of the way that the deductive device is defined. That is, given the specifications of the deductive device, one can show that there must be at least those three types of cognitive effect (1986b, pp. 108–109). So the assumptions made about the deductive device feed into the definition of cognitive effects and therefore into the definition of relevance.

There is another reason why the deductive device is important to relevance theory's account of utterance interpretation. Sperber and Wilson postulate that it plays a central role in non-demonstrative inference (1986b, p. 108). In deduction, if the premises are true and the deductive rule is sound and correctly applied then the conclusion must be true, as with the inference discussed above, from *P and Q* to *P*. Non-demonstrative inference is different, in that it is inference that is uncertain. One type of non-demonstrative inference is inference to the best explanation, in which there is some event or state of affairs and we want to know why that event happened or how the state of affairs came to be. On the Gricean view of communication adopted by relevance theory, interpreting an utterance is just this sort of problem. In the next section I set out the central assumptions that relevance theory makes about communication, and I return to explaining the role of the deductive device in inference to the best explanation once I have shown why utterance interpretation is seen this way.

3 Communication and Relevance

Relevance theory may be seen as an attempt to work out in detail one of Grice's central claims: that an essential feature of most human communication, both verbal and non-verbal, is the expression and recognition of intentions. (Wilson and Sperber 2004, p. 607, referring to Grice 1989: Essays 1–7, 14, 18; and Retrospective Epilogue).

3.1 The Problem: Inference About Intentions

Turning to the second part of the core of relevance theory, we come to the problem that relevance theory was devised to solve. How do human beings communicate? More specifically: How is it possible that in saying a phrase and/or making gestures, a human being can convey certain propositions to a conspecific? Conversely, how can the conspecific who has perceived the utterance work out what are the propositions that the producer of the utterance had in mind?

The way I have stated these questions already implicitly narrows down the field of phenomena to be explained. We are concerned here with deliberate communication—utterances made on purpose—rather than the sort of accidental information transfer that results from non-deliberate signs or signals: one's accent, posture, pheromones etc. Relevance theory adopts this more precise and narrowly focussed version of the problem from the work of the philosopher Paul Grice (and refines it somewhat for the study of communication, a shift from Grice's interest in 'speaker meaning', as discussed below). In relevance theory, this kind of deliberate, open communication is called 'ostensive-inferential' communication.

Crucially, relevance theory also accepts Grice's characterisation of an utterance as the utterer's expression of certain intentions. There are two strands of Grice's work that are relevant here: his theory of conversation and his theory of meaning. In the latter work, Grice tried to give a definition of meaning and the verb 'mean' for cases of communicative meaning, or as he called it, *speaker meaning* (excluding another use of the word 'mean' which is typified by such examples as 'Smoke means fire' and 'Black clouds mean rain').

According to Grice, when a speaker means something by an utterance the speaker has a set of nested intentions. The first of these is the intention to produce a certain response in the hearer. In the terms of a cognitive theory, we can think of this as an intention to modify the hearer's mental representation of the world by providing the hearer with information about the speaker's representation of the world. To take a simple case, when a speaker says 'It is sunny', she may intend her addressee to come to think it is sunny. The reason that the hearer comes to think this (if he does) is that the utterance provides *prima facie* evidence that the speaker thinks that it is.

The second intention is that the first intention be recognised. This criterion rules out cases in which an agent wants to bring about a change in the hearer in some other way than by openly producing an utterance. Famously, Grice discusses a case in which Mr X is anonymously informed of his wife's affair by means of a photograph that has been left lying where he will see it. In this case, the person who places the photograph intends to affect Mr X's beliefs (so has the first intention), but does not want Mr X to know that she intended any such change in his beliefs (nor indeed that she had anything at all to do with the photograph), so she lacks the second intention.

This basic structure is adopted by relevance theory as characteristic of ostensive-inferential communication, in the following form:

3.1.1 Ostensive-Inferential Communication

(a) *The informative intention*:
   The intention to inform an audience of something.

(b) *The communicative intention*:
   The intention to inform the audience of one's informative intention. (Wilson and Sperber 2004, p. 611. See also Sperber and Wilson 1986b, pp. 46–64.)

As demonstrated by Grice's photograph example, the presence of the communicative intention is a criterion for whether the speaker intends to communicate in the deliberate, purposive sense that we are discussing. Moreover, the success of this intention is sufficient for successful communication.[13] That is because if this intention succeeds, then by definition the hearer has recognised the informative intention: i.e. he realises what it is that the speaker intended him to come to think. The success of the informative intention, by contrast, is not required for successful communication. Believing what a speaker has communicated is a different matter from understanding. For example, an utterance of 'It's sunny' is understood when the hearer grasps that the speaker intended him to think that the weather is sunny. Whether he trusts the speaker enough to believe her is a separate matter. (Sperber et al. 2010 discuss this last point thoroughly).

For Grice, the decision to include this second intention in his theory of speaker meaning rests on intuitions about whether it helps to capture the intended sense of 'mean': e.g. we might be reluctant to say that the anonymous photograph-placer *meant* that Mr X was having an affair (or anything at all) by the photograph (or her leaving of it where Mr X would see it).[14] For Sperber and Wilson, whose concern is not conceptual analysis but the foundation of a scientific account of communication, the criterion is different. The hope is that the presence of the informative and communicative intentions marks out a natural class of phenomena ('ostensive stimuli' as they are called in relevance theory) which fall under interesting generalisations and laws and can be productively studied.

Scientific study of any area works towards lawlike generalisations relating to that area. Phenomena that fall under a particular natural law are described as a

---

[13] But not necessary, according to Sperber and Wilson. They suggest that unintentional ostensive communication is possible in cases where an utterer has the informative intention but not the communicative intention (as in the photograph example) but acts so ineptly that the intended audience infers that the informative intention is present (1986b, pp. 63–64).

[14] Grice's definition of speaker meaning includes a third intention, expressly to rule out from counting as meaning such cases as openly showing a photograph as evidence of an affair, showing a plaster cast as evidence of a broken leg, or showing a severed head as evidence that the person whose head it was, is dead. While it might be that these are not happily called cases of *meaning* (which as discussed, was Grice's concern), there is no doubt that they are cases of *communication*, in the deliberate, open sense that we have been discussing, so this third intention is not needed in relevance theory (Sperber and Wilson 1986b, pp. 53–54).

natural kind.[15] The existence of a general term (such as 'communication') in itself provides no guarantee that there is any such natural kind. Sperber and Wilson give *locomotion* as a counter-example (Sperber and Wilson 1986b, pp. 2, 3). There is, they say, no general theory of locomotion. There are specific theories of certain modes of locomotion—of aerodynamics, of the biophysics of walking, of flight and of swimming—and more general theories that are relevant, including laws of motion and of the conservation of energy. But there are no interesting, lawlike generalisations that hold at the level of locomotion rather than at a more specific or more general level. In that sense, locomotion is not a natural kind.

What is the evidence that ostensive stimuli form a natural kind? As in all scientific research, the ultimate criterion is the success of the theory which is founded on the assumption, relative to competing research programmes. If sufficient progress is not made, then the conclusion should eventually be drawn that foundational assumptions are wrong (or at best, unproductive).[16]

There are some pre-theoretic intimations that the choice of ostensive stimuli as an area of study will be productive, and considering them leads directly to the next core assumption of relevance theory, the inferential model of communication. Communication seems to require separate study from linguistic syntax and semantics for two reasons, both implicit in Grice's work. As Levinson writes,

> Grice's theory gives us an account both of how we can communicate without conventional signals at all… and of how we can communicate something distinct from what the conventional signals actually mean. (Levinson 2006, p. 50)

The first point here is that it is intuitively clear that both gesturing and uttering linguistic material are (or rather, can be) means of communicating. Indeed most spoken utterances involve both simultaneously. Crucially, speaker intentions are normally taken as criterial for communicative gestures as well as for linguistic utterances. When we see someone pointing, and want to know what she meant by it, then what we are interested in is finding out what she *intended* to point to. There may be many objects and parts of objects in the direction she pointed in; but what matters is which one she had in mind and wanted her audience to come to have in mind.[17] This is parallel to the case of assigning reference to indexical linguistic items such as pronouns. If a speaker says (for example) "It'll be here later", then questions about what 'it' means in that utterance are really questions about what

---

[15] The term 'natural kind' comes from Quine, 1969. The criteria for natural-kind-hood are debated. Bird and Tobin 2010 discuss various criteria.

[16] Sperber and Wilson give semiotics as an example of a field that has failed to progress partly because its fundamental assumptions do not pick out a natural class: there are no interesting generalisations, they say, to be obtained over the totality of languages, fashion, novels, road signs etc. seen as coded signals (Sperber and Wilson 1986b, p. 6ff).

[17] What the producer of an utterance intends to communicate is constrained (like other intentions) by what she can rationally hope to achieve. For example, I could not normally expect an addressee to work out that I intend to talk about cats using the word 'dog', nor that I intend to refer to my cat by pointing at a passing dog, so I cannot normally intend these interpretations (Grice 1971; Sperber and Wilson 1986b, p. 169; Neale 1992, p. 551).

the speaker intended to refer to. So it seems that we need a theory that covers both non-verbal and verbal communication and relates them both to speaker intentions. Grice's work on speaker meaning provided the basic framework for such a theory.

A further point is that gestures need not have any encoded meaning (in Levinson's terms, there need not be any 'conventional signal'). Sperber and Wilson give the example of raising one's empty glass in a pub, so as to draw a friend's attention to it (2008, p. 89). There is no code or convention that says that raising one's glass means 'Please get me another drink', but in the right circumstances the gesture would be understood as conveying that. Again, concern with the meaning of the gesture on a particular occasion comes down to interest in the utterer's intentions. One might ask the utterer: *What did you mean by raising your empty glass like that?* or *What were you trying to convey?*

Note also that if the speaker had no intention to convey information but was (e.g.) holding up the glass to better examine it in the light, then we would say that the gesture was not a communicative act at all. So the intentions of the maker of the utterance seem to be more fundamental to communication than are language or codes more broadly.

More precisely, as Sperber and Wilson put it:

> Grice's greatest originality was not to suggest that human communication involves the recognition of intentions. That much … is common sense. It was to suggest that this characterisation is sufficient: as long as there is some way of recognising the communicator's intentions, then communication is possible. (Sperber and Wilson 1986b, p. 25)

Even in the cases of linguistic utterances (and utterances of gestures that encode meaning[18]), the meaning of the utterance may differ from the encoded meaning of the phrase or gesture uttered. Here it is Grice's theory of conversation that is directly relevant. Grice discussed examples in which intuitively what the speaker means includes something that the speaker intentionally implies by (or in) making her utterance: that is, an implicature. Grice intended this category to unify such apparently diverse phenomena as indirect answers to questions and ironic utterances.

2. Mary: Have you done the hoovering?
   John: I've only just got in from work.
3. What lovely weather! [said in a downpour]

The intuition that Grice trades on here is that in both cases at least part of what the speaker means is something quite different from what she says. He coined the word 'implicature' as a term of art for this sort of thing: an intended implication of an utterance.

As Grice pointed out, utterances of a given sentence may have different implicatures (or none) on different occasions, in different contexts. Also, implicatures

---

[18] Some but not all gestures encode meanings. For example, thumbs-up encodes something like 'Good!' (or, for divers, 'Let's surface').

can be cancelled (e.g. if John's reply in (2) were 'I've only just got in from work, but in that short time, yes, I've already done it', the 'but'-clause cancels the implicature of the previous clause). Implicatures, then, are not something that should be treated by linguists as encoded in the words uttered, but are instead things that the speaker communicates by relying on the hearer's ability to work out that the speaker intended to convey something distinct from what she said.

To summarize, communication is distinct from linguistic encoding in that it can be accomplished by gestures with no conventional meaning, and in that speakers often communicate something different from what is encoded by the words or gestures they utter. In establishing these points, Grice implied that communication cannot be purely a matter of encoding and decoding (or in more Gricean terms, the deploying and retrieving of 'timeless' meanings of words) and that recognition of speaker intentions is sufficient for communication. But then how does this work?

In his theory of conversation, Grice outlines a way in which implicatures could be inferred by hearers, and therefore that speakers can rationally intend to convey them. The details of Grice's theory do not matter here (but see remarks below on contrasts with relevance theory's communicative principle of relevance). What is crucial is that his theory of conversation proposes that hearers must *infer* what it is that speakers intend to convey. Thus, as Wilson and Sperber put it, "Grice laid the foundations for an inferential model of communication, an alternative to the classical code model." (2004, p. 607)

3.2 The Inferential Model and the Code Model

Work on communication in relevance theory is a thorough exploration of the view that the linguistic material in an utterance serves as a clue that the speaker offers the hearer about her communicative and informative intentions and that the hearer uses this evidence to infer an appropriate interpretation of the speaker's utterance. This is a radical departure from pre-Gricean accounts of communication, which effectively treat communication as purely a matter of coding and decoding of a message.

It is worth noting two ways in which the inferential model differs from a code model. First, the code model is a one-stage model of utterance interpretation. The hearer just decodes the signal and retrieves the message. In contrast, in Sperber and Wilson's inferential model there will often be two stages to utterance interpretation, since the clues provided by the speaker will often include a phrase of some language, and language is a code. In such cases the hearer will have to (1) decode the phrase used, and (2) infer what the speaker intended to convey by using it.[19]

---

[19] It does not follow that in interpreting an utterance all the decoding is done first, followed by the pragmatic inference. As is well known from psycholinguistics, processing of utterances proceeds 'online', that is, in real time, as the words are heard or read.

Secondly, according to the inferential model, communication is fallible (and therefore risky but also creative) in ways that purely coded communication is not.[20] The code model claims that where the code is shared by speaker and hearer and the encoded message is received intact, the message will be decoded precisely as it was sent (Sperber 1994b). On this model, the norm is a kind of reproduction of the speaker's thought in the mind of the hearer.

Contrast this with the inferential model. The kind of inference involved in utterance interpretation is inference to the best explanation. The hearer has to infer intentions that the speaker had and which led her to make the utterance. The input to the inference is something like (e.g.): Mary said: "John isn't here yet" (with a certain intonation, perhaps accompanied by certain gestures). The question, then, is *What best explains the production of these words and gestures (at this time, in this way)?* The answer will generally be of the form, *Mary wanted to convey Σ*, where Σ is the hearer's best estimate of the intended interpretation.[21] As discussed above, this sort of inference is unlike logical deduction in that the explanation reached is not guaranteed to be the right one.

As also noted, the inferential model allows room for creativity. The speaker may leave open to some extent just what she wants to convey in making a certain utterance. Then the hearer will have to take on some responsibility for the interpretation that he derives. Relevance theory's treatment of this point makes use of two related notions introduced by Sperber and Wilson: manifestness and strength of communication.

3.3 Manifestness and Strong and Weak Communication

Roughly, an assumption is manifest to an individual in a context if he could represent the assumption mentally (on the basis of memory, perception or inference) and accept it as true or probably true. Some assumptions are not manifest at all, while among some assumptions that are manifest some are more highly manifest than others. That is, manifestness is both a classificatory and a gradable notion. (Compare, e.g. *poisonousness*: substances may be anywhere from mildly to highly poisonous, or not poisonous at all.[22]) According to relevance theory, utterances do not necessarily make the addressee mentally entertain the assumptions communicated. Rather they make it manifest that the speaker intended to make them manifest (Sperber and Wilson 1986b, p. 197), and the degree to which they are made manifest is variable. Some implicatures are so highly manifest that

---

[20] This is also a Gricean observation. Grice lists indeterminacy as a property of implicatures, a conequence of the fact that they must be inferred non-demonstratively (Grice 1975, p. 58).

[21] I use 'Σ' (for 'sum') because in general an interpretation is a bundle of propositions. See Sects. 3.4 and 3.5 below.

[22] Example suggested by Deirdre Wilson (p.c.). See also Sperber and Wilson (1986b, pp. 39, 79–80).

the speaker will almost certainly recover them—and the utterance would not be relevant enough without them. These are strong implicatures. Others may be only weakly manifest: the utterance provides the hearer with some evidence that the speaker intended to convey them but that evidence is not conclusive, and the relevance of the utterance does not depend on any particular one of them. These are called weak implicatures.[23] Consider (4) (Sperber and Wilson 1986b, p. 194):

4. Peter: Would you drive a Mercedes?
   Mary: I wouldn't drive ANY expensive car.

Mary's utterance strongly implicates (5) and (6). If Peter does not grasp them then he has not understood the utterance.

5. A Mercedes is an expensive car.
6. Mary wouldn't drive a Mercedes.

In addition it has a number of weaker implicatures, including, from stronger to weaker, (7) a–c. These contribute to the relevance of the utterance, but Peter need not entertain any particular one of these to get the point Mary is making.

7a Mary wouldn't drive a Rolls-Royce.
 b Mary wouldn't drive a Lexus.
 c Mary wouldn't drive a Saab.

3.4 Implicated Premises and Implicated Conclusions

> We assume that a crucial step in the processing of new information, and in particular of verbally communicated information, is to combine it with an adequately selected set of background assumptions—which then constitutes the context… (Sperber and Wilson 1986b, pp. 137–138)

Mary's utterance in (4) exemplifies another important relevance-theoretic distinction. According to Sperber and Wilson, it is possible to implicate contextual assumptions such as (5), as well as contextual implications of the utterance, such as (6).[24] The latter are *implicated conclusions*; implicated contextual assumptions are called *implicated premises*. According to relevance theory, all implicatures are of one of these two types (Sperber and Wilson 1986b, pp. 194–195).

Why assume that some contextual assumptions are communicated? Consider the example again. Given (5), but not otherwise, it follows from what Mary asserts

---

[23] Relevance theory also distinguishes between i) the strength with which an utterance *implicates* an assumption, and ii) the strength with which an assumption is *implied* (Sperber and Wilson 2008, §7), but there is no space in the current paper to discuss this distinction.

[24] The notions of contextual assumption and contextual implication were introduced in Sect. 2.5 above.

that she won't drive a Mercedes, i.e. (6). The assumption in (5) has to be supplied to make sense of her utterance, so Mary must have intended to make both (5) and (6) manifest.

This illustrates the key point about implicated premises and conclusions. They are tightly related to each other in the following way: given the constructed context, i.e. the implicated premises, the explicit meaning of the utterance logically warrants the implicated conclusions. The complete interpretation of the utterance is thus a logically coherent package. I return to this point in Sect. 3.8 below. Now I turn to relevance theory's treatment of the explicit content of utterances.

3.5 Explicatures: Basic- and Higher-Level

Since Grice's work there has been a gradual understanding that the role of pragmatic inference goes beyond the derivation of implicatures to other aspects of what is conveyed by an utterance, particularly the proposition expressed. How far to take this has been controversial (see Hall, this volume). Relevance theorists have been instrumental in this development (Wilson and Sperber 1981; Sperber and Wilson 1986b, Chap. 4; Carston 1988, 2002).

As discussed above, the relevance theoretic position is that any linguistic material uttered is no more than a clue to the interpretation. It follows that not just implicatures, but what is explicitly conveyed by an utterance is pragmatically inferred. Consider an utterance of the sentence in (8). It may be used to express (and in this case, assert) the proposition in (9). Relevance theory calls the proposition expressed an *explicature* (by analogy to 'implicature'). It is an inferential fleshing out of the encoded logical form of the utterance. This fleshing out in general may include reference assignment for indexical elements (e.g. 'I' → Peter; 'it' → the car), disambiguation of ambiguous words or phrases, and enrichment (e.g. 'ready' → ready for the trip to the seaside).

8. Peter: I'll get it ready in time.
9. Peter will get the car ready for the trip to the seaside in time to set off early enough to get there by noon.

Peter's utterance may be intended as a promise. In that case it also conveys (10):

10. Peter promises that he will get the car ready for the trip to the seaside in time to set off early enough to get there by noon.

This is also an explicature of the utterance, given the definition of explicature:

3.5.1 Explicature

An assumption communicated by an utterance is an explicature if and only if it is a development of a logical form encoded by that utterance (Sperber and Wilson 1986b, p. 182. See also Carston 2002, pp. 116–125.)

The proposition in (9) is the basic-level explicature of the utterance; the one in (10) is a higher-level explicature. Higher-level explicatures are embeddings of the basic-level explicature under speech-act descriptors like 'promise that' and 'ask whether', or attitudinal ones such as 'regret that' or 'be pleased that'.

According to relevance theory, the explicatures of the utterance, like the implicatures, may be strongly or weakly communicated, since here also communication is a matter of making it manifest that the speaker wanted to make an assumption manifest. Putting all of this together, we see that in relevance theory an interpretation of an utterance is in general a bundle of propositions—basic- and higher-level explicatures, plus implicated premises and implicated conclusions—each of which the speaker communicates more or less strongly.

Now I turn to relevance theory's explanation of how the hearer arrives at the interpretation.

3.6 The Communicative Principle of Relevance

According to relevance theory, the search for the correct interpretation of each utterance is guided by "the expectation that utterances should meet certain standards" (Wilson 2009, p. 393). This idea originates with Grice, although the way relevance theory develops it is quite different. Grice proposed that conversation is governed by a Cooperative Principle and a number of conversational maxims: do not say things that are false; provide enough but not too much information; be relevant; etc. From the hearer's point of view, these can be seen as expectations: that the speaker will be cooperative, where that includes trying to be truthful, to provide an appropriate amount of information, to be relevant, and so on.

Relevance theory postulates instead that each utterance raises an expectation that it will be optimally relevant. This is because each utterance is an ostensive stimulus, that is, an open attempt to take up some of the hearer's precious attention. This is stated in the communicative principle:

3.6.1 The Communicative Principle of Relevance

Every act of ostensive communication communicates a presumption of its own optimal relevance. (Sperber and Wilson 1995, p. 260)

The presumption of optimal relevance has two clauses, as follows:

3.6.2 The Presumption of Optimal Relevance

The utterance is presumed to be

1. at least relevant enough to be worth the speaker's effort to process it and
2. the most relevant one that is compatible with the speaker's abilities and preferences. (cf. Sperber and Wilson 1995, p. 270)[25]

According to relevance theory, the cognitive and communicative principles are not mentally represented by speakers or hearers nor communicated. They thus have a different status from Grice's Cooperative Principle and maxims, which are supposed to guide behaviour through the awareness of the speaker and hearer that they are in force and should be obeyed.[26] The cognitive and communicative principles are intended to be purely descriptive generalisations, like the laws of physics or biology. The claim is that speakers and hearers conform to these principles without awareness of them and without intending to: "Communicators and audience need no more know the [communicative] principle of relevance to communicate than they need to know the principles of genetics to reproduce." (Sperber and Wilson 1987, p. 704)

The *general* presumption of optimal relevance has the purely descriptive, non-represented status that the cognitive and communicative principles have. It is the specific presumption that comes with each utterance that, according to relevance theory, is communicated. Relevance theory does not claim that the presumption will always be true, nor that it is always taken as true. According to Sperber and Wilson, "It is enough that the presumption of relevance should be communicated—and it always is—to fulfil its most important role: determining the interpretation of the ostensive stimulus." (Sperber and Wilson 1987, p. 704)

How does the presumption of optimal relevance help the hearer to infer the correct interpretation? The two clauses set a lower bound and a higher point respectively for the relevance that the hearer is entitled to. The first clause sets the lower bound. It might seem that this lower bound is not well-specified, or is uninterestingly low. What level of relevance is "enough to be worth the speaker's effort to process" the utterance? If we recall the discussion of the assumptions around the cognitive principle, we see that the lower bound is largely set by the environment. The point of the first clause is that an utterance must be worth attending to amid the other possible sources of cognitive effects in the hearer's environment. This may be quite a high degree of relevance, given the limits on human attention, and the fact that other potential sources of relevance may be (or seem) highly relevant.

---

[25] The formulation of the presumption of optimal relevance given here is stronger than the one originally presented in Sperber and Wilson 1986b. See Sect. 5 below.

[26] The point of calling the maxims 'maxims' is to suggest that, like Kant's maxims, they motivate agents' actions.

The second clause, by contrast, strikes many people as too strong. Why should the hearer be entitled to expect the *most* relevant utterance that the speaker is willing and able to provide? Sperber and Wilson explain this in terms of two thought experiments (1995, pp. 268–269; see also Wilson and Sperber 2002, p. 604). Suppose that a speaker wants her utterance to produce certain cognitive effects in the hearer. Now suppose that there are several possible utterances that she could make that would produce these cognitive effects. One of them would produce just the desired cognitive effects, while the others would produce these plus other cognitive effects. Which utterance should the speaker produce?

Now suppose instead that the speaker has a choice between utterances which would all produce only the desired cognitive effects, but some of which are easier for the hearer to process. Again, the question is: Which utterance should the speaker produce?

The general answer, according to Sperber and Wilson is that "She should choose the utterance that would be (or seem) the most relevant to the addressee" (1995, p. 269). Why? Well, minimizing the hearer's effort is good policy because it makes it more likely that the hearer will pay attention and fully process the utterance, i.e. more likely that the cognitive effects which the speaker wants to cause will occur. And maximising the hearer's returns is also good policy because giving the hearer more information that is of interest to him will also maximise the chances of his paying attention, fully processing the utterance, and remembering the information that the speaker wanted to convey.

Note that none of this assumes anything like Grice's Cooperative Principle. It just follows from the speaker's desire to be understood, and the assumption that the cognitive principle of relevance applies to the hearer: i.e. that the hearer generally seeks maximal relevance. According to relevance theory, speakers exploit that tendency.

Since the communicative principle and presumption of optimal relevance set bounds on the relevance to be expected from any utterance, they obviously help to make tractable the search for an interpretation of an utterance. In fact, relevance theory claims that they licence a specific interpretation procedure, the relevance theoretic comprehension procedure.

3.7 The Relevance Theoretic Comprehension Procedure

if there is one conclusion to be drawn from work in artificial intelligence, it is that most cognitive processes are so complex that they must be modelled in terms of heuristics rather than failsafe algorithms. We assume, then, that communication is governed by a less than perfect heuristic. (Sperber and Wilson 1986b, p. 45)

The relevance theoretic comprehension procedure is as follows:

1. Following a least effort path, construct a (hypothetical) interpretation of the utterance. This interpretation will generally include explicatures, implicated premises and implicated conclusions.
2. Check to see whether the interpretation as a whole satisfies both clauses of the presumption of optimal relevance. That is, it should be (1) relevant enough (i.e. it should provide enough cognitive effects for the effort expended thus far in processing the utterance) and (2) it should be the most relevant one that is compatible with the speaker's abilities and preferences (in the hearer's estimation of those abilities and preferences).
3. If the interpretation hypothesised in step 1 passes the test outlined in step 2, accept it as the intended interpretation.
4. If not, go back to step 1, and work through the steps again: i.e. construct the next most accessible interpretation and check it for optimal relevance. Repeat until an optimally relevant interpretation is found. Alternatively, or in addition, adjust the expectation of relevance: perhaps the utterance is not as relevant as it might have been because the speaker is not fully competent; or the speaker is not benevolent and the utterance is only intended to seem optimally relevant.[27]
If no interpretation that is optimally relevant (or intended to be optimally relevant, or to seem so) is found within reasonable time/effort, the overall cognitive economy will shut down the search.

Note first that the most accessible interpretation will always be checked first, given that this procedure follows a least effort path. So if the most accessible interpretation is relevant enough, it will be accepted as the intended interpretation (i.e. the speaker's intended interpretation). Thus in situations where other sources of potential relevance are largely absent, and where the hearer has low expectations of the speaker's abilities and preferences, the most accessible interpretation will generally be accepted as the correct one.[28] In other cases, the hearer's expectations will make it clear that the speaker intended a more relevant interpretation and the search will continue.

The comprehension procedure is a heuristic in the sense used in the literature on bounded rationality, namely that it is not guaranteed to arrive at the right answer. Like other heuristics that are worth using, it is supposed to find results quickly and without too much effort: it is 'fast and frugal' in Gigerenzer's terms (e.g. Gigerenzer and Goldstein 1996). In common with other fast and frugal heuristics, it uses both blades of Herbert Simon's scissors: that is, both "the structure of task environments and the computational capabilities of the actor" (Simon 1990, p. 7). Full

---

[27] See Sperber 1994b, who suggests that the ability to make these adjustments develops in early childhood.

[28] As Sperber and Wilson point out (1986b, p. 185), this may be part of the reason why much psycholinguistic work (e.g. on disambiguation) has tended to focus only on accessibility (i.e. effort) factors.

computation of all the possible interpretations of an utterance is not necessary, according to relevance theory, because of the environmental regularity described in the presumption of optimal relevance. Instead, because all ostensive stimuli come with a (fallible) guarantee, the hearer's utterance interpretation system just has to come up with the first interpretation that occurs to it, evaluate it, and then perhaps formulate the next most accessible interpretation (which is likely to be a modified variant of the previous interpretation), evaluate that, and so on. In other words, what makes this procedure frugal is that a) it follows a least-effort path, and b) that the first interpretation reached that satisfies the presumption of relevance stops the search.

The reason that the presumption of optimal relevance makes it reasonable for interpretation to follow a least effort path is that relevance varies inversely with effort, so an utterance whose intended interpretation is off the least effort path is less relevant than another utterance that the speaker could have managed to produce. To satisfy clause b of the presumption, speakers have to make their utterances as easy to understand as possible.[29] The reason why the hearer can stop at the first optimally relevant interpretation is that an utterance that has two significantly different interpretations that both yield the expected degree of cognitive effects would fail to be optimally relevant, since the hearer would have to expend effort in choosing between them.[30]

3.8 Utterance Interpretation as Inference to the Best Explanation

We have seen how the relevance theoretic comprehension procedure decides how to stop and accept an interpretation as the correct, intended one. But how are hypothetical interpretations derived? Part of the answer has already been sketched out above. Verbal material in the utterance is decoded, and then hypothetically fleshed out in ways that include disambiguation of ambiguous words or structure, and the assignment of reference to indexical expressions such as pronouns. This derivation of explicatures occurs in parallel, and in 'mutual adjustment' with the derivation of implicated premises and implicated conclusions. The mechanism for the derivation of implicated conclusions was discussed in Sect. 2.5 above: the deductive device, given an input and contextual assumptions, will churn out contextual implications.

---

[29] For more discussion of why the presumption of optimal relevance mandates a least effort path and stopping at the first optimally relevant interpretation see Sperber and Wilson, 1995, p. 272; Wilson and Sperber 2002, p. 605; Allott 2008, pp. 259–260.

[30] This is not meant to rule out puns, *double lecture* and the like. In such cases, Wilson and Sperber say, "it is the fact that the speaker has produced such an utterance that is seen as a communicative act. It receives a higher-order interpretation, which may involve endorsing both lower-order interpretations (if they are compatible), or rejecting both (if they are not)." (2002, pp. 605, fn 6).

Where, though, do the contextual assumptions come from? A hearer has a lot of information available, some highly accessible, some less so. A linguistic utterance raises the accessibility of information associated with the concepts that are encoded by the words used.

Consider (4) again. Peter expects that Mary's utterance is optimally relevant, and that it will answer his question. Mary's use of the phrase 'expensive car' temporarily makes what Peter knows about expensive cars highly accessible. Putting this together with the fact that he is already thinking about Mercedes cars, the proposition that they are expensive is highly accessible. At the same time, Mary's utterance is decoded, and reference is assigned to 'I', yielding the proposition: *Mary would not drive any expensive car*. The deductive device automatically combines this hypothetical explicature with the highly accessible *Mercedes are expensive cars* to yield (6), a conclusion that answers his question. Thus there is a logically coherent package of proposition expressed, implicated premise and implicated conclusion. Taking into account also the weak implicatures noted in the discussion of this example above, the total hypothetical interpretation is optimally relevant, so it is accepted as the intended one.

In this model of inference to the best explanation, the inference process is separated into hypothesis formation and hypothesis testing. The hypothesis formation is entirely mechanical. It is partly a matter of taking the most accessible assumptions, the most accessible disambiguation, the most accessible reference assignment etc. and partly a matter of feeding what results through the deductive device. Hypothesis testing is then just a matter of checking the putative interpretation against expectations of relevance.[31]

4 Beyond the Core

Given that the focus of most research in relevance theory has been on communication, one might wonder whether the broader commitments about cognition (which themselves rest, as discussed, on assumptions about evolution) are strictly necessary.

One way to think about this is to imagine constructing a different theory, which we can call RT′. RT′ adopts the communicative principle and the other assumptions in Sect. 3, but discards the cognitive principle, and with it the evolutionary backstory. The other assumptions in Sect. 2 are kept, including the definition of relevance and the computational/representational theory of mind. The scope of RT′ would be much more restricted. But one can ask what would be lost, from the more specialised perspective of pragmatics, in moving from relevance theory to RT′. A partial answer is that RT′ would lack several important explanatory features. First,

[31] See also the more detailed worked example at Wilson and Sperber 2002, p. 607ff. and discussion at Allott 2008, pp. 65–66.

and most obviously, it would lack a rationale for the communicative principle of relevance. Why should hearers assume that speakers will be, or try to be, optimally relevant, if there is no general tendency of cognition to maximise returns for effort?

In addition, the cognitive principle implies that humans are somewhat predictable in their cognition, and thus helps to explain how speakers can produce utterances which the hearer will process in the way that was intended.

A related point is that RT′ would have no explanation of why information tends to be stored in long-term memory in ways that are useful for understanding utterances (and for cognition more broadly). The cognitive principle implies that memory should tend to be organised so that information is stored in a useful form and so that it will tend be recalled when relevant, and not otherwise. As discussed in Sect. 3.8 above, the accessibility of information plays a considerable role in relevance theory's explanation of utterance interpretation. It is commonly assumed that long-term memory is organized in chunks sometimes called 'frames' or 'schemas' (Sperber and Wilson 1986b, p. 138). Thus, for example, when a restaurant is mentioned, it raises the accessibility of stereotypical information about restaurants, such as facts about waiters. The necessity of this sort of structured memory for utterance interpretation is brought out by 'bridging' cases such as the utterances in (11) (c.f. Wilson and Matsui 1998).

11(a) We went to a Thai restaurant. The waiter was from Bangkok.
(b) ??We spent the day in London. The waiter was from Bangkok.

The cognitive principle provides some explanation why memory is arranged in chunks, and also, therefore, sheds light on how speakers are able to fine-tune their utterances to rely on and exploit such facts about what the hearer is likely to have stored and to quickly retrieve.

4.1 Auxiliary Assumptions and Positive Heuristic

I turn finally to some of relevance theory's auxiliary hypotheses and to its 'positive heuristic'. As noted in the introduction, Lakatos proposed that research programmes (or series of research programmes) have positive heuristics, strategies for forming theories outside of the hard core which specify what 'paths of research' to pursue (Lakatos 1968, p. 168).

> … the positive heuristic consists of a partially articulated set of suggestions or hints on how to develop the 'refutable variants' of the research-programme, how to modify, sophisticate, the protective belt. (Lakatos 1968, p. 171)

The main thrust of relevance theory's positive heuristic is, of course, something like this: confronted with a phenomenon(/data) in the realm of communicative behaviour, try to understand it in terms of the operation of the relevance theoretic comprehension procedure, i.e. in a way that is compatible with, and, to as great an extent as possible, predicted by the communicative and cognitive principles.

However, the relevance theoretic research programme has several additional resources that amount to suggestions on how to form theories. I discuss four. Three of these are parts of the framework that are somewhat logically independent of the core, but which are general in that each underlies several auxiliary hypotheses.[32] They are (1) the possibility of dividing what is linguistically encoded between conceptual and procedural information; (2) the interpretive/descriptive distinction; (3) the postulation of *ad hoc* concepts. The fourth is not itself a hypothesis, but something more like an attitude to pragmatic theorising: a strong economy principle, akin to Grice's Modified Occam's Razor, but with a wider scope. I discuss this first.

4.2 Modified Occam's Razor

Grice's Modified Occam's Razor is the principle that senses should not be multiplied beyond necessity (Grice 1989, p. 47). It amounts to an economy argument in favour of treating meanings as pragmatically derived—and thus for Grice, as implicatures—rather than linguistically encoded, unless there is compelling evidence of linguistic ambiguity. For example, an utterance of the sentence in (12) will typically convey that John's kicking of the dog came after his being slapped by Mary (and perhaps also that it was a result of his being slapped). But 'and' does not always convey temporal or causal relation, as illustrated by (13).

12. Mary slapped John and he kicked the dog.
13. Mary lives in London and John lives in Oxford.

Grice argued that the word 'and' has just the meaning that it contributes to examples like (13), namely logical conjunction, and that what is additionally conveyed by its use in (12) is pragmatically implicated (Grice 1967).

Relevance theorists also invoke Modified Occam's Razor. They have been concerned with some of the same words and examples. For example, in a series of papers, Robyn Carston defends the Gricean simple, univocal semantics for 'and' against various objections (Carston 1988, 1993, 2002, Chap. 3; Blakemore and Carston 2005). The most notable of these is L. J. Cohen's observation that the extra component of meaning cannot be an implicature because it comes under the scope of logical operators (Cohen 1971). Otherwise the following would seem nonsensical or internally contradictory:

14. You are being unfair to Mary. It's not true that she slapped John and he kicked the dog. He kicked the dog and she slapped him.

---

[32] It is tempting to say that in addition to the core, there is both a *mantle* and a *crust*. Then *ad hoc* concepts, the conceptual/procedural distinction and the distinction between interpretive and descriptive use are in the mantle, while the relevance theoretic accounts of pronouns, utterance modifiers, irony, non-declaratives, loose use, hyperbole and metaphor are parts of the crust.

To defend the view that 'and' encodes logical conjunction, Carston takes the non-Gricean position that the pragmatic enrichment in such cases affects the proposition expressed.

Similar use of Modified Occam's Razor is widespread in the work of relevance theorists. The principle is that if something *can* plausibly be done by the pragmatic mechanism, then it *should* be attributed to it, since the pragmatic mechanism is required independently of the analysis of any particular case: it comes for free, one might say. There are real cases of ambiguity which cannot plausibly be analysed as purely pragmatic differences: the lexical ambiguity of 'bank', for example. But note that the pragmatic mechanism will still be required and involved in any genuinely linguistically ambiguous cases, since the hearer has to infer which of the senses the speaker intended.

Relevance theory's use of Modified Occam's Razor clashes with the central positive heuristic of linguistic formal semantics, which is roughly: when one finds a difference in truth-conditions, one should try to show how that difference can be derived compositionally from the encoded meanings of the words in the sentence, postulating complex encoded meanings as necessary. Following this principle leads in the opposite direction to Modified Occam's Razor, locating the explanatory action in syntax and/or semantics and tending to multiply linguistic representations. These are modern variants of Posner's "two competing strategies for the description of verbal communication" (Posner 1980, p. 170). From the point of view of relevance theory, the issue cannot be settled globally, but only case by case, subject to the Gricean presumption that unless there is good reason to propose two or more linguistic representations, one should prefer a pragmatic explanation.

This way of putting it illustrates that the economy principle of Modified Occam's Razor can be stated in terms of representations. For example, an utterance of the sentence in (12) has representations on at least two cognitively significant levels. There will be a linguistic semantic level (sometimes called LF), and the level of the interpretation of the utterance, after[33] pragmatic processing. That the sentence has (at least) two readings is common ground between relevance theorists and ambiguity-theorists, so there is no clear difference between the theories at the post-pragmatic level. But if we postulate that the sentence is linguistically ambiguous, then the string in (12) will correspond to at least two distinct representations at LF,[34] whereas on the Gricean or relevance theoretic position there will be only one. Simply counting the representations we can see that the pragmatic explanation is more economical. And as Carston shows, 'and' has many more than two readings, so the pragmatic account is very much more

---

[33] It is from the point of view of interpretation that this mental representation of what is conveyed by the utterance is *after* pragmatic processing: i.e. the hearer only gets to it once pragmatic processing has been performed.

[34] One might be able to argue that there is only one LF for the sentence if one postulates (as Cohen does) a very complicated lexical entry for 'and', but this shifts the complexity without reducing it and also creates new problems (Carston 1993, p. 35).

economical than the proliferation of senses and representations which would be required by a syntacto-semantic account (Carston 1993, pp. 27–28, 35).

The impulse to simplify carries over, in relevance theory, to whole categories of mental representation, and beyond. As discussed above, relevance theory has only one communicative principle where Grice had the Cooperative Principle and several maxims, and the communicative principle is not mentally represented (except by theorists, of course) whereas in the Gricean framework, speakers and hearers must be aware of the maxims. Then there is also only one way of deriving pragmatically inferred meaning, again in deliberate contrast to Grice's theory, which claims that implicatures arise in at least four ways.[35]

In addition, relevance theory rejects both conventional implicatures (that is, implicatures encoded by certain words) and generalised conversational implicatures (implicatures that are pragmatically implied by default). For relevance theory, all implicatures are of the type that Grice called particularised conversational implicatures, the kind that hearers have to (a) infer, (b) taking into account the specifics of the situation. That is, there is no separate class of default implicatures, and there are no implicatures encoded by lexical items. Finally, relevance theorists do not employ a separate category of presuppositions.[36]

4.3 The Conceptual/Procedural Distinction

> Linguistic decoding provides input to the inferential phase of comprehension; inferential comprehension involves the construction and manipulation of conceptual representations. An utterance can thus be expected to encode two basic types of information: representational and computational, or conceptual and procedural – that is, information about the representations to be manipulated, and information about how to manipulate them. (Wilson and Sperber 1993, p. 1)

According to relevance theory, words can encode two different types of meaning. The first is conceptual meaning. For example, the word 'cat' encodes the concept CAT, and contributes this concept to the proposition expressed by utterances of sentences containing the word, as in (15).

15. Her cat is antisocial. So no one picks him up and pets him.

---

[35] Namely (1) real and (2) apparent violation of maxims, (3) clashes between maxims and even (4) in cases in which there is no violation and no appearance of it.

[36] Much of what would be explained in terms of presuppositions by other theorists is naturally understood in relevance theory as (and unified with) the communication of implicated premises. Some other alleged presuppositions are treated as entailments: e.g. relevance theorists tend to endorse the Russellian/Gricean account of definite descriptions (Carston 2002, pp. 110, 306–11). The roots of relevance theory's view of presuppositions are in Wilson's early work (Wilson 1975a, b). See also Kempson 1975.

Words can also encode procedural meaning, which is to say that they can encode constraints on the way that an utterance is processed. For example, the discourse connective 'so' in (15) encourages a reading in which the second sentence is taken as a conclusion supported by the first. (Contrast the way that 'after all' promotes a reading in which the first sentence is supported by the second.)

16. Her cat is antisocial. After all no one picks him up and pets him.

The idea of procedural meaning and the treatment of discourse connectives in these terms comes from Diane Blakemore (1987). Subsequently Wilson and Sperber (1993) expanded the role of procedural meaning. They propose that procedural meaning can constrain the derivation of explicatures (whereas previously it had been seen as contributing only to non-truth-conditional aspects of interpretation). For example pronouns are taken in relevance theory to encode constraints on explicatures: e.g. 'him' encodes (roughly) *search for a male individual to fill this slot in the proposition expressed.*[37] So the conceptual/procedural distinction allows a surprising partial unification of the semantics of pronouns and discourse connectives. (For a recent review of work on procedural meaning see the papers in Escandell-Vidal et al. 2011, particularly Wilson 2011).

4.4 The Interpretive/Descriptive Distinction

Relevance theory makes a distinction between different ways that sentences (and words) can be used. Consider examples (17) and (18).

17. John: What did the prime minister say?
    Mary: He knew nothing about the leak until this week. But I don't believe him. (cf. Sperber and Wilson 1986b, p. 228, their example 101.)
18. Mary: He's just trying to get himself out of trouble.

An utterance of a sentence can be intended to convey an explicature which is a statement about the world, on the basis that the logical form of the proposition expressed resembles the logical form of a proposition that describes a certain state of affairs. Relevance theory calls this *descriptive use*. Example (18) is most likely to be interpreted this way, i.e. as Mary voicing her own opinion that the prime minister is trying to get out of trouble.

But this is not the only possible use of a sentence. A sentence can be uttered with the intention that its logical form resembles the logical form of a proposition someone is entertaining, or the logical form of an utterance that someone else has made or might make. In this first case the utterance is a representation of a mental state; in the second it is a representation of another utterance. Relevance theory

---

[37] This account of pronouns, as Sperber and Wilson point out, is something like a cognitive version of Kaplan's character/content distinction.

calls such uses of sentences interpretive, because in typical examples like Mary's utterance in (17) the speaker is not presenting her own view of the way things are, but is acting as an interpreter of someone else's view or utterance[38] (Sperber and Wilson 1986b, p. 224ff.).

This distinction is made use of in relevance theory's account of irony. Sperber and Wilson reject the classical definition of verbal irony—the expression of a meaning by using words that usually mean the opposite—and propose instead that irony is interpretive use with a dissociative attitude (Sperber and Wilson 1986b, p. 237ff; Wilson and Sperber 1992). In true irony, the dissociative attitude and the fact that there is interpretive use are both tacit – i.e. not linguistically signalled, but left for the hearer to work out. To see what this amounts to, consider the sentences in (19), uttered in a rainstorm. What they would express is similar, but the first is ironic because the hearer is left to infer the attitude and the fact that the utterance is meant interpretively.

19(a) What beautiful weather!
(b) It was ridiculous to expect beautiful weather.

This account fits well with the observation often made that irony comes at a significant risk of misunderstanding, since on this account ironic utterances require the hearer to infer two pieces of tacitly conveyed information. It also accounts for irony's affinity with quotation. Furthermore, it correctly predicts intuitions for examples that are problematic for the classical definition of verbal irony. Consider (20) uttered in a context in which it is obvious that the car in question has a broken window:

20. Look, that car has all its windows intact. (Grice 1967, p. 53)

This is not generally ironic, as Grice noted. The relevance theoretic account predicts this, since it is hard to process as a (mocking) echo of something someone might say or think. However, as the relevance theoretic account also predicts, it can be uttered ironically in a rather contrived context in which it is manifest that someone has said or thinks something that entails that the car doesn't have broken windows: e.g. in response to someone smugly saying 'There's practically no crime in *this* neighbourhood and certainly no one here would break into a car to steal the radio.' (Sperber and Wilson 1986b, pp. 240–241; Wilson 2006, p. 1732)

This theory of irony is logically independent from the core of relevance theory (and was originally proposed, in a slightly different form, before the core: Sperber and Wilson 1981). One could, therefore, adopt Sperber and Wilson's theory of irony without relevance theory's core assumptions.[39] Conversely, the relevance theoretic account of irony could be abandoned without giving up any of the core

---

[38] Interpretive use also includes the use of sentences to "represent an assumption, without attributing this assumption to anyone" (Sperber and Wilson 1986b, p. 229).

[39] On the other hand, the classical account of irony is incompatible with the communicative principle. A speaker cannot generally communicate just the opposite of what her words mean

assumptions of relevance theory, as (of course) it might be if it comes into conflict with observation.[40]

Making the descriptive/interpretive distinction do explanatory work has been fruitful in the development of a number of other auxiliary hypotheses in relevance theory. In early relevance theory, loose use, hyperbole and metaphor were understood as a kind of interpretive use in which the speaker's utterance is an approximate interpretation of her own thought[41] (Sperber and Wilson 1986a, b, p. 231ff), although that account has now been abandoned in favour of one in terms of *ad hoc* concepts (discussed below). Papafragou attempts a relevance theoretic account of metonymy as a type of interpretive use (1996). Finally, combined with a distinction between the desirable and the actual, the descriptive/interpretive distinction underlies the relevance theoretic account of mood and non-declarative sentences: in assertion there is a descriptive relation between speaker's thought and world; in imperatives the speaker's thought describes a desirable state of affairs; and in interrogatives the speaker's thought is in an interpretive relation to desirable thoughts (Sperber and Wilson 1986b, pp. 231, 243–254; Wilson and Sperber 1988).

4.5 Ad Hoc *Concepts*

The use of the notion of *ad hoc* concepts is a recent development. Relevance theory postulates that strict and literal utterances, loose use, hyperbole and metaphor are not qualitatively distinct phenomena but belong to a continuum. As noted, this was originally explained in terms of interpretive use (Sperber and Wilson 1986a), but Carston (1997a), and Sperber and Wilson (1998b) now advocate an account in terms of *ad hoc* concepts.

Consider the utterance in (21) made in a context in which it is clear that Mary is talking about her husband, who is not canonised. What she expresses is not that Peter is a strict and literal saint, but something else: perhaps that he is very considerate and self-sacrificing.

(Footnote 39 continued)

because it would cause the hearer gratuitous effort, given that she could just have said what she meant.

[40] In fact, it accounts well for intuitions about the examples in the literature (Sperber and Wilson 1998a; Wilson 2006), and has received some corroboration from developmental evidence (Happé 1993).

[41] All utterances are interpretive in this way, since an utterance is meant to represent a thought of the speaker's. What relevance theory usually calls interpretive use is use in which that thought is itself interpretive: i.e. resembles another thought, or an utterance, rather than a proposition which describes a state of affairs (Sperber and Wilson 1986b, p. 231.).

21. Mary: Peter is a saint.

Embedding under logical operators suggests that this is a matter of the proposition expressed, rather than an implicature. A utterance of (22) as a response to (21) is not a denial that Peter is a strict-and-literal saint, but that he is a very nice, considerate etc. individual.

22. Peter's no saint. He always does what he prefers and makes it look like a huge sacrifice.

Therefore relevance theory postulates that the proposition expressed by (21) is PETER IS A SAINT*, where SAINT* is an *ad hoc* concept: a distinct concept from the lexically encoded concept SAINT, and accessed/constructed by pragmatic inference sensitive to the specific occasion. *Ad hoc* concepts may be broader or narrower than the lexicalised concepts from which they derive. In examples such as (23) and J.L. Austin's (24), the concepts communicated (MILES* and HEXAGONAL*) apply to broader sets than the lexicalised concepts: e.g. France is not HEXAGONAL, but it is HEXAGONAL*, along with many other objects that are not strictly speaking six-sided, but are close enough.

23. It's miles to the canteen!
24. France is hexagonal.

Combined broadening and narrowing is seen in (21). The concept SAINT* denotes a set that is both broader than the lexicalised concept (it includes individuals such as Peter, who are not literally saints) and narrower (it will exclude any literal, canonised saints who were not considerate, self-sacrificing etc.).

Recourse to the notion of *ad hoc* concepts has become the primary strategy in relevance theoretic lexical pragmatics. Relevance theorists now argue that lexical pragmatic adjustment is nearly ubiquitous, and "fine-tunes the interpretation of virtually every word." (Carston and Powell 2006, p. 345)

5 Concluding Remarks

Both the core and auxiliary assumptions of relevance theory have developed during its history.[42] To conclude I briefly set out two important early developments in the core, the first of which was mainly driven by the desire to maximise the simplicity, coherence and symmetry of the theory, the second by the aim of bringing relevance theory into line with a development in a related field.[43]

---

[42] For developments in relevance theory see Sperber and Wilson (1995), Carston and Powell (2006) and Clark (2011).

[43] Lakatos is sometimes interpreted as saying that the core of a research programme should never change (e.g. Godfrey-Smith, 2003, p. 105). However, his 'negative heuristic' (see note 5 above)

The presumption of optimal relevance given in Sect. 3.6 above is stronger than the one originally put forward (Sperber and Wilson 1986b, p. 158). Sperber and Wilson present and argue for the updated formulation in the postface added to the second edition of 'Relevance' (1995, p. 267ff.). The original formulation was not symmetrical in effort and effects: its clause b is the presumption that the speaker will maximise the relevance of the utterance, but it treats the intended interpretation (and therefore cognitive effects) as given, so this amounts to an expectation that effort will be minimised. The revision, then, is largely motivated by considerations of simplicity and generality (Sperber and Wilson 1995, p. 270); although Sperber and Wilson also argue that it increases the predictive power of relevance theory (1995, p. 270).

Another important development in the core of relevance theory is the move to the view that there is a dedicated inferential mechanism for utterance interpretation. In early work, Sperber and Wilson say on the one hand that utterance interpretation appears to be "an ordinary central thought process… relatively unspecialised" (1986b, p. 116) and non-modular (1986b, p. 69), but, on the other hand, suggest that it is carried out by "a less-than-perfect heuristic" (1986b, p. 45) one among "a number of heuristics, some of them innate, others developed through experience, aimed at picking out relevant phenomena" (1987, p. 703) and argue that analogies with the slow, deliberative reasoning involved in scientific theorising are unhelpful (1986b, p. 117). In more recent work, they argue that there is a mental module dedicated to utterance interpretation (see also Carston 1997b, 2002). This change reflects considerable rethinking within psychology: of both the nature of central cognition and of the concept of a mental module, and in particular, Sperber's proposal of the massive modularity thesis (Sperber 1994a). It is also partly prompted by the emergence of the view that human beings have dedicated 'theory of mind' or 'mindreading' abilities (Wimmer and Perner 1983; Baron-Cohen et al. 1985; Wellman 1990): Sperber and Wilson (2002) argue that the comprehension module is related to, but distinct from, the general mindreading module. It should be no surprise that significant developments in cognitive science are reflected in changes in the core of relevance theory, given that the main purpose of relevance theory is to provide a account of communication that is psychologically realistic.

(Footnote 43 continued)

merely forbids changing the core in response to empirical problems. It is compatible with this that there be changes to the core not motivated by a direct clash with empirical evidence. Indeed, if changes to the core turn out to be mainly motivated by other considerations that would provide some corroboration for (a version of) his views.

References

Allott, N. (2006). Game theory and communication. In A. Benz, G. Jäger, & R. van Rooij (Eds.), *Game theory and pragmatics* (pp. 123–151). Basingstoke: Palgrave Macmillan.

Allott, N. (2008). *Pragmatics and rationality*. Unpublished Ph.D thesis, University of London.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37–46.

Barrett, H. C. (2005). Enzymatic computation and cognitive modularity. *Mind and Language, 20*(3), 259–287.

Bezuidenhout, A. (1997). Pragmatics, semantic undetermination and the referential/attributive distinction. *Mind, 106*(423), 375.

Bezuidenhout, A., & Sroda, M. S. (1998). Children's use of contextual cues to resolve referential ambiguity: An application of relevance theory. *Pragmatics and Cognition, 6*(1–2), 265–299.

Bird, A., & Tobin, E. (2010). Natural kinds. In N. E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2010 ed.).

Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Basil Blackwell.

Blakemore, D. (2000). Indicators and procedures: *Nevertheless* and *but*. *Journal of Linguistics, 36*(03), 463–486.

Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

Blakemore, D. (2004). Discourse markers. In L. R. Horn & G. L. Ward (Eds.), *The handbook of pragmatics* (pp. 221–240). Malden: Blackwell.

Blakemore, D., & Carston, R. (2005). The pragmatics of sentential coordination with *and*. *Lingua, 115*(4), 569–589.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85*(1), 1–21.

Braine, M. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic*. Mahwah: L. Erlbaum Associates.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100*(3), 434–463.

Carston, R. (1988). Implicature, explicature and truth-theoretic semantics. In R. M. Kempson (Ed.), *Mental representations: The interface between language and reality* (pp. 155–181). Cambridge: Cambridge University Press.

Carston, R. (1993). Conjunction, explanation and relevance. *Lingua, 90*, 27–48.

Carston, R. (1997a). Enrichment and loosening: Complementary processes in deriving the proposition expressed. *Linguistische Berichte, 8*, 103–127.

Carston, R. (1997b). Relevance-theoretic pragmatics and modularity. *UCL Working Papers in Linguistics, 9*, 29–53.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston & S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 179–236). Amsterdam: J. Benjamins.

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.

Carston, R. (2010a). Explicit communication and 'free' pragmatic enrichment. In B. Soria & E. Romero (Eds.), *Explicit communication: Robyn Carston's pragmatics* (pp. 217–285). Basingstoke: Palgrave Macmillan.

Carston, R. (2010b). XIII-metaphor: Ad hoc concepts, literal meaning and mental images. *Proceedings of the Aristotelian Society (Hardback)*, 110(3pt3): 295–321.

Carston, R., & Powell, G. (2006). Relevance theory: New directions and developments. In E. LePore & B. C. Smith (Eds.), *The Oxford handbook of the philosophy of language* (pp. 341–360). Oxford: Oxford University Press.

Cherniak, C. (1981). Minimal rationality. *Mind, 90*(358), 161–183.

Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the theory of mind account of autism. *Neuropsychologia, 49*(3), 507–517.

Chevallier, C., Wilson, D., Happé, F., & Noveck, I. (2010). Scalar inferences in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40*(9), 1104–1117.

Clark, B. (1996). Stylistic analysis and relevance theory. *Language and Literature, 5*(3), 163–178.

Clark, B. (2011). Recent developments in relevance theory. In D. Archer & P. Grundy (Eds.), *The pragmatics reader* (pp. 129–137). Oxford: Routledge.

Cohen, L. J. (1971). Some remarks on Grice's views about the logical particles of natural language. In Y. Bar-Hillel (Ed.), *Pragmatics of natural languages* (pp. 50–68). New York: Humanities Press.

Dobzhansky, T. (1964). Biology, molecular and organismic. *American Zoologist, 4*(4), 443–452.

Emlen, J. M. (1966). The role of time and energy in food preference. *The American Naturalist, 100*(916), 611–617.

Escandell-Vidal, V., Leonetti, M., & Ahern, A. (Eds.). (2011). *Procedural meaning: Problems and perspectives*. Emerald: Bingley.

Evans, J. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove: Psychology Press.

Fodor, J. A. (1975). *The language of thought*. New York: Crowell.

Fodor, J. A. (1985). Fodor's guide to mental representation: The intelligent auntie's vade-mecum. *Mind, 94*(373), 76–100.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627–635.

Gigerenzer, G. (2004). Striking a blow for sanity in theories of rationality. In M. Augier & J. G. March (Eds.), *Models of a man: Essays in memory of Herbert A. Simon* (pp. 389–409). Cambridge: MIT Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Godfrey-Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science*. Chicago: University of Chicago Press.

Grice, H. P. (1967). Logic and conversation: William James lectures. In H. P. Grice (1989), *Studies in the way of words* (pp. 1–143). Cambridge: Harvard University Press.

Grice, H. P. (1971). Intention and uncertainty. In *Proceedings of the British Academy* (vol. 58, pp. 263–279).

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press.

Gutt, E. A. (1991). *Translation and relevance: Cognition and context*. Oxford: Blackwell.

Hacking, I. (1979). Imre Lakatos's philosophy of science. *The British Journal for the Philosophy of Science, 30*(4), 381–402.

Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition, 48*(2), 101–119.

Happé, F., & Loth, E. (2002). Theory of mind and tracking speakers intentions. *Mind and Language, 17*, 24–36.

Ifantidou, E. (2001). *Evidentials and relevance*. Amsterdam: J. Benjamins.

Iten, C. (2005). *Linguistic meaning, truth conditions and relevance: The case of concessives*. New York: Palgrave Macmillan.

Jary, M. (2007). Are explicit performatives assertions? *Linguistics and Philosophy, 30*(2), 207–234.

Jary, M. (2010). *Assertion*. Basingstoke: Palgrave Macmillan.

Johnson-Laird, P. N. (1983). *Mental models*. New York: Cambridge University Press.

Jorgensen, J., Miller, G. A., & Sperber, D. (1984). Test of the mention theory of irony. *Journal of Experimental Psychology: General, 113*(1), 112–120.

Kempson, R. M. (1975). *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.

Kjøll, G. (2010). Content similarity and communicative success. *International Review of Pragmatics, 2*(1), 21–45.

Lakatos, I. (1968). Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society New Series, 69*, 149–186.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). Cambridge: Cambridge University Press.

Lakatos, I., Worrall, J., & Currie, G. (1978). *Philosophical papers*. Cambridge: Cambridge University Press.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451–468.

Lavie, N. (2001). The role of capacity limits in selective attention: Behavioural evidence and implications for neural activity. In J. Braun, C. Koch, & J. L. Davis (Eds.), *Visual attention and cortical circuits* (pp. 49–68). Cambridge: MIT Press.

Levinson, S. C. (2006). On the human 'interactional engine'. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Oxford: Berg.

Matsui, T. (2000). *Bridging and relevance*. Amsterdam: J. Benjamins.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227–234.

de Moura, A. C. A., & Lee, P. C. (2004). Capuchin stone tool use in Caatinga dry forest. *Science, 306*(5703), 1909.

Neale, S. (1992). Paul Grice and the philosophy of language. *Linguistics and Philosophy, 15*(5), 509–559.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale: L. Erlbaum Associates.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery, 19*(3), 113–126.

Nicolle, S., & Clark, B. (1999). Experimental pragmatics and what is said: A response to Gibbs and Moise. *Cognition, 69*(3), 337–354.

Noveck, I. A., & Sperber, D. (Eds.). (2004). *Experimental pragmatics*. New York: Palgrave Macmillan.

Noveck, I. A., & Sperber, D. (2007). The why and how of experimental pragmatics: The case of 'scalar inferences'. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 184–212). Basingstoke: Palgrave Macmillan.

Papafragou, A. (1996). On metonymy. *Lingua, 99*(4), 169–195.

Pashler, H. E. (1998). *The psychology of attention*. Cambridge: MIT Press.

Pilkington, A. (2000). *Poetic effects: A relevance theory perspective*. Amsterdam: J. Benjamins.

Posner, R. (1980). Semantics and pragmatics of sentence connectives in natural language. In J. R. Searle, F. Kiefer, & M. Bierwisch (Eds.), *Speech act theory and pragmatics* (pp. 168–203). Dordrecht: Riedel.

Powell, G. (2001). The referential-attributive distinction: A cognitive account. *Pragmatics and Cognition, 9*(1), 69–98.

Powell, G. (2010). *Language, thought and reference*. Basingstoke: Palgrave Macmillan.

Quine, W. V. O. (1969). Natural kinds. In *Ontological relativity, and other essays* (pp. 114–138). Princeton: Columbia University Press.

Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90*(1), 38–71.

Rouchota, V. (1992). On the referential/attributive distinction. *Lingua, 87*, 137–167.

Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203*(2), 60–68.

Simon, H. A. (1957). *Models of man: Social and rational; mathematical essays on rational human behavior in a social setting*. New York: Wiley.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*(1), 1–19.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3–22.

Sperber, D. (1994a). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 39–67). Cambridge: CUP.

Sperber, D. (1994b). Understanding verbal understanding. In J. Khalfa (Ed.), *What is intelligence?* (pp. 179–198). Cambridge: Cambridge University Press.

Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Structure and contents* (pp. 53–68). Oxford: Oxford University Press.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*(1), 31–95.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., et al. (2010). Epistemic vigilance. *Mind and Language, 25*(4), 359–393.

Sperber, D., & Wilson, D. (1981). Irony and the use-mention distinction. In P. Cole (Ed.), *Radical pragmatics* (pp. 295–318). New York: Academic Press.

Sperber, D., & Wilson, D. (1986a). Loose talk. *Proceedings of the Aristotelian Society New Series, 86*, 153–171.

Sperber, D., & Wilson, D. (1986b). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell.

Sperber, D., & Wilson, D. (1987). Précis of 'Relevance: Communication and cognition'. *Behavioral and Brain Sciences, 10*, 697–754.

Sperber, D., & Wilson, D. (1995). Postface. In D. Sperber & D. Wilson, *Relevance: Communication and cognition* (2nd ed., pp. 255–279). Oxford: Blackwell.

Sperber, D., & Wilson, D. (1996). Fodor's frame problem and relevance theory (reply to Chiappe and Kukla). *Behavioral and Brain Sciences, 19*(3), 530–532.

Sperber, D., & Wilson, D. (1998a). Irony and relevance: A reply to Seto, Hamamoto and Yamanashi. In R. Carston & S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 283–293). Amsterdam: J. Benjamins.

Sperber, D., & Wilson, D. (1998b). The mapping between the mental and the public lexicon. In P. Carruthers & J. Boucher (Eds.), *Language and thought: Interdisciplinary themes* (pp. 184–200). Cambridge: Cambridge University Press.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind and Language, 17*(1–2), 3–23.

Sperber, D., & Wilson, D. (2008). A deflationary account of metaphors. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 84–108). New York: Cambridge University Press.

Stephens, D. W., Brown, J. S., & Ydenberg, R. C. (Eds.). (2007). *Foraging: Behavior and ecology*. Chicago: University of Chicago Press.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton: Princeton University Press.

Todd, P. M., & Gigerenzer, G. (2000). Précis of 'Simple heuristics that make us smart'. *Behavioral and Brain Sciences, 23*(5), 727–741. Discussion 742–80.

Unger, C. (2006). *Genre, relevance, and global coherence: The pragmatics of discourse type*. New York: Palgrave Macmillan.

van der Henst, J. B., Carles, L., & Sperber, D. (2002a). Truthfulness and relevance in telling the time. *Mind and Language, 17*(5), 457–466.

van der Henst, J. B., Politzer, G., & Sperber, D. (2002b). When is a conclusion worth deriving? A relevance-based analysis of indeterminate relational problems. *Thinking and Reasoning, 8*(1), 1–20.

Vega-Moreno, Rosa, E. (2007). *Creativity and convention: The pragmatics of everyday figurative speech*. Amsterdam: John Benjamins.

Wedgewood, D. (2007). Shared assumptions: Semantic minimalism and relevance theory. *Journal of Linguistics, 43*, 647–681.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge: MIT Press.

Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge: Cambridge University Press.

Wilson, D. (1975a). Presupposition, assertion, and lexical items. *Linguistic Inquiry, 6*(1), 95–114.

Wilson, D. (1975b). *Presuppositions and non-truth-conditional semantics*. London: Academic Press.

Wilson, D. (2006). The pragmatics of verbal irony: Echo or pretence? *Lingua, 116*(10), 1722–1743.

Wilson, D. (2009). Relevance theory. In L. Cummings (Ed.), *The pragmatics encyclopedia* (pp. 393–399). London: Routledge.

Wilson, D. (2011). The conceptual-procedural distinction: Past, present and future. In V. Escandell-Vidal, M. Leonetti, & A. Ahern (Eds.), *Procedural meaning: Problems and perspectives* (pp. 5–31). Emerald: Bingley.

Wilson, D., & Matsui, T. (1998). Recent approaches to bridging: Truth, coherence, relevance. *UCL Working Papers in Linguistics, 10*, 173–200.

Wilson, D., & Sperber, D. (1981). On Grice's theory of conversation. In P. Werth (Ed.), *Conversation and discourse* (pp. 155–178). London: Croom Helm.

Wilson, D., & Sperber, D. (1988). Mood and the analysis of non-declarative sentences. In J. O. Urmson, J. Dancy, J. M. E. Moravcsik, & C. C. W. Taylor (Eds.), *Human agency: Language, duty, and value; philosophical essays in honor of J. O. Urmson* (pp. 77–101). Stanford: Stanford University Press.

Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua, 87*(1/2), 53–76.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua, 90*(1), 1–25.

Wilson, D., & Sperber, D. (2002). Truthfulness and relevance. *Mind, 111*(443), 583–632.

Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn & G. L. Ward (Eds.), *The handbook of pragmatics* (pp. 607–632). Malden: Blackwell.

Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge: Cambridge University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128.

Relevance Theory, Semantic Content and Pragmatic Enrichment

Alison Hall

Abstract Work in the last two decades on semantics and pragmatics has given rise to a multitude of different positions on where to draw the distinction between them, whether such a distinction can be drawn at all, and to what extent, if any, pragmatics contributes to semantic content. I outline the relevance-theoretic view that the semantics-pragmatics distinction corresponds to the distinction between linguistically encoded and pragmatically provided meaning, and the reasons for the rejection of any intermediate level of semantic content such as a minimal proposition. In the second part of the paper, I survey a range of recent frameworks (indexicalism; certain versions of minimalism) that potentially avoid those objections, and consider whether these approaches motivate the need for some variety of semantic content as a psychologically real or theoretically valid level of representation distinct from encoded meaning and explicit utterance content.

1 Introduction: Semantics, Truth Conditions, and Explicature

The distinction between pragmatics and semantics is widely accepted to be a distinction between, respectively, meaning that is recovered by principles or maxims of pragmatic inference, and meaning that is determined largely by linguistic mechanisms. However, the interplay of linguistic and inferential factors, particularly at the level of explicitly communicated content, has given rise to a variety of different positions on how and where to draw the semantics-pragmatics distinction.

A. Hall (✉)
UCL Linguistics, London, UK
e-mail: a.hall@ucl.ac.uk

Traditionally, the guiding assumption of many semantic theories (from Frege 1892, through Russell 1905; Davidson 1967; Kaplan 1977/1989, to Larson and Segal 1995) has been that semantics concerns truth conditions (see Carston 1999: Sect. 2 for discussion of the aims of some of these semantic programmes). Grice's (1967) influential work showed that, in addition to what a speaker says, which is both largely conventional, and that content on the basis of which her utterance would be judged true or false, a speaker may also convey implicatures that do not affect the truth-value of what she says; these conversational implicatures are calculated by assuming that the speaker is being cooperative and adhering to certain expected standards of truthfulness, relevance, informativeness, and manner of expression (Grice's Cooperative Principle and Maxims; see Grice 1975/1989: 26–27).[1] This suggests a natural way of drawing the semantics-pragmatics distinction: semantics would correspond to the truth-conditional content of the utterance (Grice's what is said), and pragmatics to any conveyed meaning that falls outside the truth-conditional content (implicatures).

On this way of drawing the distinction, semantic content can be equated with Grice's what is said, and has the following two features: it is the truth-conditional content of the utterance, and it is determined almost entirely by the conventional, encoded meaning of the linguistic expressions used. As Grice acknowledged, the truth-conditional content is not completely free of contextual input: In some brief comments on an utterance of "He is in the grip of a vice," he says, 'for a full identification of what the speaker has said, one would need to know (a) the identity of x, (b) the time of utterance, and (c) the meaning, on the particular occasion of utterance, of the phrase *in the grip of a vice*' (1975/1989: 25). However, he did not seem to see these processes of reference assignment and disambiguation as requiring appeal to his conversational maxims; instead, the idea seems to be that they are resolved more automatically, the requisite values being something like objective features of the context of utterance (see Carston 2002: 105–106 on Grice's appeal to a criterion of best contextual fit).

This equation of semantics with the truth-conditional content of the utterance quickly runs into difficulties when one tries to apply the semantics-pragmatics distinction to the following kinds of examples:

1. Anna: How's Max doing now? Is he any better than last time I saw him?
   Ben: Well, his wife's left him and he's started drinking again.

2. After he bulldozed his way past Berdych in straight sets, Nadal is everyone's favourite to make it three grand slam titles in a row.

3. A: The White House doesn't visit Tip O'Neill in his Congressional office.
   B: Old grudge.

Consider, first, Ben's reply in (1). It seems clear that he is communicating that Max is not doing well, and this communicated meaning is an implicature. What

---

[1] Page references to Grice are to his 1989 collection.

about the truth-conditional content? Grice allowed for reference assignment (and disambiguation), so context can provide a value for the pronouns "he/his/him." But does this result in the truth-conditional content that the speaker intends to express? According to the intuitions of ordinary speaker-hearers—and, as most theorists now agree—the answer is no. The linguistically-encoded denotation of "drinking" covers events of drinking water, tea, cough syrup, and so on, but what the use of "drinking" in Ben's reply is intended to denote is likely to be a subset of the events denoted by the linguistic meaning: the denotation is narrowed to exclude many of them and cover only those events—that is, drinking of alcohol—relevant in the context. There is also a cause-consequence relation between the conjuncts, which Grice would have treated as a conversational implicature (a generalized one: see Grice 1975/1989: 37–38), but Carston (1988) argued that such non-truth-functional meanings conveyed by "and"-conjunctions are often best treated as contributing to the truth-conditional (i.e. explicit) content. Here, this causal meaning does seem to be part of that content on the basis of which one would agree or disagree (a third speaker in the exchange in (1) could directly contradict Ben by saying "No; what happened was he started drinking and his wife left him"),[2] so the truth-conditional content is something like this, with pragmatically supplied elements underlined[3]:

4. MAX'S_i WIFE HAS LEFT HIM_i & AS A RESULT HE_i HAS STARTED DRINKING EXCESSIVE QUANTITIES OF ALCOHOL.

Turning to (2), apart from reference assignment, there is a metaphorical use of "bulldozed"; as I discuss later in the chapter, many such loose uses are best seen as contributing to truth-conditional content. Other elements of this content that appear to arise here independently of semantic (linguistic) motivation or constraints include, arguably, a cause-consequence relation between the two clauses, much as in (1), and the domain of "everyone" would be restricted to something like EVERYONE WHO IS INTERESTED IN THIS MATCH. In (3B), taken from Barton (1990), the underdetermination of truth-conditional content by conventional meaning is even more extreme: linguistic decoding arguably produces just a determiner phrase, OLD GRUDGE, yet the content expressed is, roughly, THAT IS BECAUSE OF AN OLD GRUDGE.[4]

In all three cases, the linguistic meaning, even when saturated with contextual values for the indexical elements, falls far short of determining truth-conditional content. So the problem is that the two criteria—that semantics be conventional and that it be truth-conditional—pull in opposite directions. What corresponds to the intuitive truth-conditional content of an utterance—the proposition that is explicitly expressed by the speaker—is, as illustrated in (1)–(4), often the result of

---

[2] The use of this embedding procedure can be traced back to Cohen (1971).

[3] I use small caps to indicate mental representations.

[4] See Stainton (2006) for an extended defence of the idea that such fragmentary utterances are genuinely subsentential and used to perform speech acts.

extensive pragmatic development, and departs radically from any content that could be considered semantic (these apparently pragmatically motivated contributions to explicit content, known as 'free pragmatic enrichment', will be discussed further in Sect. 3). Thus many theorists reject the idea that the semantics-pragmatics distinction corresponds to the distinction between what is explicitly expressed (the intuitive truth-conditional content), and what is implicated.[5] Still, there continues to be little consensus in the literature as to where the distinction should be drawn.

This chapter will take as its starting-point the relevance-theoretic view that semantics is standing (encoded) linguistic meaning (hence context-invariant), while pragmatic meaning is any meaning that is recovered by appeal to contextual (pragmatic) factors, including all those pragmatic processes that contribute to determining explicit utterance content, that is, including reference assignment and disambiguation (see in particular Carston 2002 and 2008a). In the first part of the chapter, I recap Carston's arguments for drawing the distinction this way and against the positing of some level of semantic content intermediate between linguistically encoded meaning and explicit content (henceforth 'explicature'), focusing on her discussion of Bach (2001) and Cappelen and Lepore (2005). In the second part (Sects. 3,4, and 5), I consider various recent approaches to semantics that potentially avoid Carston's objections and make sense of the idea of semantic content, motivating the idea that semantics does deal in truth conditions or 'content' after all. The first two, which I argue do not motivate a revision of the relevance-theoretic distinction, are the 'hidden indexical' approach defended by Stanley (2007) and the very minimalist versions of minimalism proposed by Borg (2004) and Korta and Perry (2006, 2008). Finally (Sect. 5), I consider a version of the 'semantic relativism' that has gained popularity in the last few years, and provide a brief preliminary assessment of its implications for the issues addressed in this chapter.

---

[5] A further complication is the fact that certain linguistic items seem not to contribute to explicature at all, but are probably best analyzed as constraining the implicit side of communication, by indicating what sort of inferences are to be drawn from the explicit content. Discourse connectives or particles such as "but" and "although" communicate a contrast or contradiction of some sort between the clauses they conjoin, but this contrastive meaning is generally agreed not to affect the truth or falsity of the utterance. In (1) above, the use of "again" does not affect truth conditions but affects the inferences drawn from the explicature: its use here indicates that Max has previously been a heavy drinker. Arguably, "well" is another such expression, though what it communicates is less determinate. Grice (1989) suggested that such items are conventional implicature triggers; see Blakemore (1987, 2002) and Wilson and Sperber (1993) for a relevance-theoretic treatment.

2 Relevance Theory's Semantics-Pragmatics Distinction

Relevance theory (Sperber and Wilson 1986/1995; Carston 2002), along with many contextualist philosophers of language such as Bezuidenhout (2002), Neale (2007), Recanati (2004) and Stainton (2006), takes the view that a theory of utterance comprehension requires only three psychologically real levels of representation. There is the linguistically encoded meaning, or 'logical form' of the utterance, plus the two kinds of communicated assumption introduced in the last section: explicatures, which are arrived at by pragmatically developing the logical form, and implicatures, inferred from explicatures plus contextual assumptions. Explicatures answer to our intuitions about the truth-conditional content of an utterance but, as illustrated above, may have extensive pragmatic input that goes beyond anything mandated by the linguistic meaning, a fact that argues against equating the semantics-pragmatics distinction with the explicature-implicature one. Having dropped Grice's truth-conditionality requirement on semantics, we are left with the criterion of conventionality. The rest of the argument goes as follows. There is no level of representation that is both distinct from conventional—that is, linguistically encoded—meaning, and more closely tied to it than explicature is, nor is there any theoretical use for isolating any such entity. That leaves the distinction between encoded and inferred meaning as the only useful place to draw the semantics-pragmatics distinction (Carston 1999, 2002, 2008a).

   Underlying this approach is a view of linguistic meaning on which its role is not to determine truth conditions, but to act as one clue, among others, to the speaker's meaning. We can communicate propositional—truth-conditional—contents without using language at all, drawing on various cues (gestures such as pointing, facial expressions, other paralinguistic behaviour, the contextual salience of objects), and linguistic meaning is another of these potential cues, albeit one that allows the communication of far more complex contents than would be communicable nonverbally. There is, then, no reason why linguistic (semantic) decoding should produce something that can be given a truth value, or why it should even articulate at logical form all the elements of the truth-conditional content (leaving variables or slots to be saturated with the kinds of values specified by the linguistic meaning). As Carston (2002: 29–30) puts it, 'underdeterminacy is an essential feature of the relation between linguistic expressions and the propositions (thoughts) they are used to express. [...] public-language systems are intrinsically underdetermining of complete (semantically evaluable) thoughts because they evolved on the back, as it were, of an already well-developed cognitive capacity for forming hypotheses about the thoughts and intentions of others on the basis of their behaviour'. Although Carston is defending an 'essentialist' underdeterminacy position here (one shared by Recanati 1987, 1994, 1996, Searle 1978, 1980, 1992, and Sperber and Wilson 1986/1995), this evolutionary justification is also compatible with a weaker version of underdeterminacy, on which the linguistic expressions used do, generally, underdetermine the thought expressed, but 'this is merely a matter of effort-saving convenience for speakers and another sentence

which fully encodes the proposition/thought could always be supplied' (Carston 2002: 29).[6] The acceptance of pervasive underdeterminacy, whether one thinks of it as essential or merely convenient, is based on the recognition that we have a highly developed theory-of-mind capacity to bridge the gap between linguistic meaning and communicated thoughts, so that encoding of fully propositional forms is often unnecessary and uneconomical. Rather, the role of linguistic meaning is just to encode what is necessary, which, as illustrated particularly well by example (3) above (B's utterance of "Old grudge"), is often only a fragment of a propositional form. As Sperber and Wilson say, 'all that is required is that the properties of the ostensive stimulus (e.g. the utterance) should set the inferential process on the right track; to do this they need not represent or encode the communicator's informative intention in any great detail' (1986/1995: 254).

The idea that semantics should determine truth conditions/propositions or at least some genuine level of content, though, continues to exert a strong pull, even while many authors agree that semantics radically underdetermines the *communicated* content. Given the variety of pragmatic processes that are held to contribute to explicature (disambiguation, reference assignment, various kinds of entirely pragmatically-motivated enrichments and meaning modulations illustrated in (1)–(3) above), there are, clearly, several places where it would be possible to delineate a content intermediate between linguistic meaning and explicature. Two such accounts are the minimal propositions of Cappelen and Lepore (2005) and the (sometimes sub-propositional) 'what is said' of Bach (2001); in the rest of this section, I present these accounts and their motivations, and argue against them, drawing mainly on Carston's (2008a) previous discussion. In the following sections, I look at some more recent approaches and consider whether they require any revision or supplementation of the relevance-theoretic picture.

According to Cappelen and Lepore (2005), semantic context-sensitivity is displayed by only a limited set of expressions consisting mostly of overt indexicals and demonstratives (Kaplan's (1977/1989: 489) Basic Set of Context Sensitive Expressions).[7] The result of decoding, disambiguation, and saturating these overt context-sensitive elements is the semantic content. They agree that the semantic content would not generally be what the speaker intended to communicate, and that recovering what the speaker 'said', or 'asserted', or 'claimed' (that is, explicature) usually requires far more pragmatic processing. However, Cappelen

---

[6] Bach seems to hold this version of the underdeterminacy thesis: he writes, '…what is being communicated could have been made fully explicit by the insertion of additional lexical material' (1994: 134).

[7] Kaplan's list of indexicals is this: The personal pronouns "I", "you." "he", "she", "it" in their various cases and number, the demonstrative pronouns "that" and "this" in their various cases and number, the adverbs "here", "now", "today", "yesterday", "tomorrow", "ago", "hence(forth)", and the adjectives "actual" and "present." Cappelen and Lepore (2005: 1) add words and aspects of words that indicate tense, and suggest that they might also include 'contextuals'—common nouns like "enemy", "outsider", "foreigner", "alien", "immigrant", "friend", and "native" as well as common adjectives like "foreign", "local", "domestic", "national", "imported", and "exported."

and Lepore appear to take the view that there is nothing enlightening to say about how explicature is arrived at, since it 'depends on a potentially indefinite number of features of the context of utterance and the context of those who report on (or think about) what was said by the utterance' (2005: 4). An interesting systematic account of how these myriad pragmatic features interact with linguistic meaning is seen as a hopeless prospect. What is tractable, and what semantic minimalists should limit themselves to, is to concern oneself with pragmatics only when necessary—that is, where its interaction with linguistic meaning is mandated by the linguistic meaning itself.

So Cappelen and Lepore are not in the business of explaining how we grasp speaker's meaning—explicature—but do think that their minimal semantic content has a role to play in an account of communication, because this content—a minimal proposition—is required to explain how it is possible to 'share content' across contexts.

The two aspects of Cappelen and Lepore's proposals that have been most disputed, which I'll address here, are whether this semantic content is a proposition, and whether it is shared. First, Cappelen and Lepore's (2005) claim that the minimal semantic content is a proposition initially seems clearly wrong: The result of decoding, disambiguation, and saturation of the overt indexicals often does not produce something that we are able to evaluate the truth or falsity of. The minimal semantic content of an utterance of "He is ready," for example, would just be, say, JOHN IS READY. Intuitively, we need to know more than that to grasp a truth-evaluable content—that is, for a proposition to be expressed. But Cappelen and Lepore (2005: 87–112) appeal to various tests in support of their claim, one of which, the 'Inter-Contextual Disquotational Indirect Report' test, is illustrated here:

Context C1: In a conversation about exam preparation, someone raises the question of whether John is well prepared. Nina utters "John is ready."

Context C2: Three people are about to leave an apartment; they are getting dressed for heavy rain. Nina utters "John is ready."

Cappelen and Lepore claim that they can truly report that "In both C1 and C2, Nina said that John is ready." The content shared across the two contexts of utterance (and the report) is that JOHN IS READY (*punkt*), and that this minimal content is a proposition is established by the fact that the report is supposedly truth-evaluable as it stands, without any felt need to 'complete' it by specifying what John was ready for in each context.

This test, and the conclusions Cappelen and Lepore draw from it, have already been extensively criticized (Bezuidenhout 2006; Carston 2008a; Szabó 2006, among others). Even if we agree with Cappelen and Lepore that the report is truth-evaluable, this does not show that the reported utterance—what is embedded under "said that"—itself is a proposition. Rather, the correct conclusion to draw is that the 'content' of "John is ready" that is shared across the various utterances is the linguistically encoded meaning, free from any pragmatic contribution. Wedgwood (2007) illustrates this by applying the test to a pair of utterances containing an ambiguous expression. Consider two utterances by Nina, in different contexts, of "John went to the bank." Imagine we know that, in the first context, she was

referring to the financial bank, and in the second, to the river bank. We can truthfully report "In both C1 and C2, Nina said that John went to the bank." Given the use they made of this test with the utterances of "John is ready," Cappelen and Lepore should maintain that Nina's two utterances of "John went to the bank" had a shared content (abstracting away from tense, and treating proper names as constants). But the only possible candidate for shared content would include a disjunction of the two meanings of "bank." The minimal proposition expressed by Nina's first utterance was JOHN WENT TO BANK1; that expressed by her second utterance was JOHN WENT TO BANK2. The only level of content that is shared between the two utterances is a level prior to the recovery of the minimal proposition—that is, the linguistically encoded meaning. And as for why we judge the report ("In both C1 and C2, Nina said that John is ready/that John went to the bank") true, I suggest that that must be because we are not entirely disquoting: we are simply judging it as a true report of Nina's words, rather than of the contents of the two utterances.

This leaves Cappelen and Lepore with no argument for the propositionality of their minimal semantic contents, and, more recently, at least one of them appears to have dropped the insistence that this semantic content is propositional (see, for example, Cappelen 2007). A more important issue, though, is the role that they see their semantic content—propositional or not—playing in an explanation of communication. It is essential for philosophy of language to explain how content can be shared across contexts, and, according to Cappelen and Lepore, only semantic minimalism 'can account for how the same content can be expressed, claimed, asserted, questioned, investigated, etc. in radically different contexts. It is the semantic content that enables audiences who find themselves in radically different contexts to understand each other, to agree or disagree, to question and debate with each other. It can serve this function simply because it is the sort of content that is largely immune to contextual variations' (2005: 152).

The minimal semantic content is, then, supposed to be what we can expect people to grasp, even if they do not have enough information about the context to allow them to recover the speaker's full message. The intuitive truth-conditional content cannot play this role, as this level of communicated content incorporates the results of too much pragmatic inference. So semantic content has an important role as a fall-back content that the speaker can rely on the hearer grasping. However, as has been pointed out (Wedgwood 2007; Carston 2008a), this is a role that the linguistically encoded meaning is far better able to play than Cappelen and Lepore's semantic content. This is because linguistic meaning is recovered by decoding, an automatic, algorithmic process, in contrast with pragmatic inference which, by virtue of being a process of hypothesis formation and confirmation, always takes place at some risk. And, since Cappelen and Lepore's semantic content is partly, and often largely, pragmatic—being the result of decoding plus the pragmatic processes of disambiguation and saturation[8]—speaker and hearer

---

[8] They say nothing about disambiguation, but are clear that Gricean mechanisms do a lot of work in getting to their semantic content—i.e. in saturation.

may diverge substantially. The only content that can be guaranteed to be shared is the linguistically encoded meaning; any content that incorporates the results of pragmatic inference cannot serve as a fall-back content that the speaker can rely on the hearer grasping if the hearer lacks knowledge of the context. So this version of semantic minimalism makes no progress on how we manage to faithfully share content well enough to enable us to debate, investigate, hold people responsible for what they say, and so on.

Cappelen and Lepore's semantic minimalism is one of the recent varieties of what Recanati (2004: 51) calls Syncretism, a view that is a compromise between 'Literalism'—by which Recanati seems to mean Grice's view, on which what is said departs as little as possible from 'literal' (i.e. encoded) meaning—and the contextualist and relevance theoretic approach on which what is said is much more pragmatically developed. Syncretism considers both these notions of what is said to be legitimate, giving a four-level picture, with two levels of 'literal meaning'— sentence meaning and the minimal what is said—and two levels of speaker's meaning—the pragmatic what is said (explicature), and conversational implicature.

Several philosophers have defended this four-level picture: Soames (2002) contrasts what is said in the semantic sense with what is said in the pragmatic sense, which is what the speaker states or asserts, while Salmon (1991: 88) distinguishes 'saying or asserting in the strict and philosophical sense' from 'saying in the loose and popular sense'; the latter 'what is said' is the content of the speaker's speech act, but does not cover what is merely implicated in Grice's sense. Both these authors' conceptions of the semantically expressed proposition include the results of saturation, like Cappelen and Lepore's, so are susceptible to the same objection: as Stanley (2007: 233) says in his review of Recanati (2004), 'The central problem for the Syncretic View is that the notion of semantic content appealed to in the theory threatens to be an idle wheel in an explanation of linguistic practice'.

A different version of Syncretism, however, is defended by Bach (1994, 2001), who, while sharing with Cappelen and Lepore and Soames the view that a semantic content is needed as a shared fall-back content, delineates semantic content differently than other syncretists. He distinguishes a level of what is said, separate from explicature (which he calls impliciture[9]) and implicature, on the basis of the kind of context needed to derive semantic content versus other contents. I consider here just his (2001), which represents his latest view. For discussion of his (1994)— similar to Cappelen and Lepore (2005) except that he does not insist that semantic contents are propositions—see Carston (2002; Sect. 2.5).

According to Bach (2001), 'what is said' results from saturating 'pure' indexicals, which are expressions such as "I," "today," "here," and "now." The result of

---

[9] Impliciture and the relevance-theoretic explicature consist of the same amount of content in most cases, except that Bach is not inclined to count figurative uses of expressions as cases of impliciture, whereas relevance theorists believe that at least some figurative uses are cases of modulating the encoded meaning of an expression, for example "bulldozed" in example (2), and the modulated meaning is part of the explicature. See Bach (2010) on the differences between explicature and impliciture.

saturating other indexicals ("he," "that," "we," etc.), and of 'completion' and 'expansion'[10] is impliciture. What is said may be subpropositional (a 'propositional radical'), and so need not be a communicated assumption. Within communicated assumptions, Bach recognizes the same binary distinction as relevance theorists and other contextualists, between developments of logical form (Bach's impliciture, RT's explicature), and conversational implicatures.

Bach delineates his 'what is said' on the basis of the kind of context necessary for recovering it: It consists of linguistically encoded meaning plus saturation of pure indexicals, whose referents, supposedly, can simply be read off of the context—e.g. "I" is automatically assigned the speaker as referent; "now" is assigned the time of utterance, and no pragmatic inference comes in. Only 'narrow' context is required—that is, the objective features of the context, such as speaker, addressee, location and time of utterance. Pure indexicals contrast with other indexicals and demonstratives, such as "he" and "that," which require consideration of speaker intentions and other contextual features ('broad context').

Bach's reason for drawing this distinction is to single out 'the linguistically determined input to the hearer's inference to what, if anything, the speaker intends to be conveyed in uttering the sentence' (2001: 15). The thought here is that, since the value of a pure indexical is information that simply 'arises from the act of uttering' without consideration of speaker intentions, it can be considered the linguistically determined input, and can be guaranteed to be shared among interlocutors.

This appears, initially, to be a shared content that can serve as the input to pragmatic reasoning and that avoids the problems of Cappelen and Lepore's approach by distinguishing between two types of indexicals (and the two types of context that their saturation requires). But the problem with this is that narrow context is not enough to fix the value of "here" (which the speaker could be intending to use to mean in this room, in this building, in this city, or any number of other values) or "now" (in the 21st century, at 2 pm on August 30 2008?). The idea that narrow context is sufficient is only remotely plausible for "I," but even with "I," there are several kinds of cases for which pragmatic inference, and appeal to wide context, is necessary for working out the referent. Cases familiar from the literature include these:

---

[10] 'Completion' and 'expansion' are Bach's terms for different kinds of 'free' (that is, non-linguistically mandated) pragmatic processes, some of which were illustrated in examples (1)–(3). 'Completion' occurs when the sentence, after saturation, is semantically incomplete, that is, does not express a full proposition. An utterance of "He is ready," for example, after saturation of the indexical "he," is still semantically incomplete, and requires the hearer to work out what the person is ready for (assuming here, with Bach, that "ready" is not, and does not encode, an indexical or variable). Expansion occurs when the sentence, after saturation (and maybe completion) is semantically complete, but that complete proposition is not the intuitive asserted content. An example would be a mother saying to a child screaming about a grazed knee, "You're not going to die." This expresses a proposition (after saturation): You_x ARE NOT GOING TO DIE {*ever*}, but the asserted content, or impliciture, would be You_x ARE NOT GOING TO DIE FROM THAT CUT.

5. Professor X is out of his office. A colleague, seeing students calling at Prof X's office to find him, writes a note saying "I am not here now" and sticks it on Prof X's door (Predelli 1998).
6. George Bush utters "The founding fathers invested me with the power to appoint members of the Supreme Court" (Nunberg 1993).

So, given that even assigning reference to "I" can involve appeal to broad context, the most natural move would be to accept that the linguistically determined input is just the standing—context-invariant—linguistic meaning, and class "I" together with the rest of the indexicals.

After introducing the relevance-theoretic distinction between semantics and pragmatics, in the rest of this section, I've considered the following two alternative ways of drawing the distinction:

(a) Semantics as linguistically encoded meaning plus contextual values for all indexicals (i.e. semantics as 'minimal proposition'); pragmatics as the rest of speaker meaning.
(b) Semantics as linguistically encoded meaning plus contextual values for pure indexicals; pragmatics as the rest of speaker meaning.

The discussion so far clearly supports Carston's conclusion that 'not only are these not good ways to draw the distinction between semantics and pragmatics, they are not worthwhile distinctions of any other sort either; that is, they do no useful work' (2008a: 322). However, several other recent theories avoid the objections discussed above. In the next sections, I look at these in turn, and consider whether they motivate a revision to the relevance-theoretic distinctions, by justifying the positing of a level of genuine content (as opposed to linguistically encoded meaning) that is semantic.

3 Covert Indexicals versus Free Pragmatic Enrichment

The first of these views is in much the same spirit as Cappelen and Lepore's minimalism, in that it sees semantic content as being propositional, and resulting only from decoding and disambiguation plus saturation of all indexicals. Where it differs from minimalism is in its claim that what the semanticist should be trying to account for is the intuitive truth-conditional content. On this, it agrees with relevance theorists and contextualists that our intuitions are evidence that the content on which these truth-value judgments are based is a psychologically real level of representation, while there is no evidence for the existence of the more minimal content posited by Cappelen and Lepore and other syncretists.

Recall that the two Gricean criteria on semantic content (Grice's what is said) were that it be conventional (departing little from linguistic meaning, allowing only for saturation and disambiguation), and that it be the speaker-meant truth-conditional content. These two criteria proved contradictory, and one had to be

dropped, as a host of examples—such as (1)–(3) in Sect. 1—showed that (intuitive) truth-conditional content incorporated the results of more pragmatic processes than just saturation and disambiguation. Another example that is central to the debate addressed in this section is (7):

7. It's raining.

The truth value of weather reports is agreed to depend on the location that the weather in question is supposed to occur at, so (7), when appropriately contextualised, would have the explicature IT IS RAINING IN LONDON. This location constituent appears to be provided entirely pragmatically, on grounds of relevance; if so, then this example would reinforce the message of the above examples, that semantic content cannot be both conventional and truth-conditional. But the approach that I discuss here manages to reconcile these apparently incompatible requirements.

This approach claims that explicature is fully linguistically articulated, and the only pragmatic processes affecting it are linguistically mandated. Stanley (2000: 391) states the view thus: "all truth-conditional effects of extra-linguistic context can be traced to logical form." Given the agreement on what the object of explanation is—the non-minimal, intuitive conception of truth-conditional content—the challenge for the defender of this view is to account for the elements of explicature that do not appear to be the values of anything in the overt linguistic form. His explanation is that there are covert indexicals attached to certain lexical items, which are, therefore, present in logical form whenever the item is tokened. For instance, weather verbs might encode a location variable, and, on Stanley and Szabó's (2000) proposal, every nominal encodes a pair of domain indices, which accounts for domain restriction and the completion of definite descriptions, and so on. (8)–(10) show, in rough form, the kinds of structures envisaged:

8. It is raining <at Loc L>
9. The <candidate, f(i)> was late
10. Every <bottle, f(i)> is empty

If this approach could be extended to account for all effects of context on explicature, then, apart from disambiguation, the only pragmatic process involved in getting from logical form to explicature would be saturation of (overt and covert) indexicals: In accordance with what Stanley suggests in the above quote, there would be no contribution from pragmatic processes that are not linguistically mandated.

The motivation for this approach, which I label here 'indexicalism', is as follows. The traditional, straightforward account of how an utterance's content is grasped is that we work out the semantic values of the expressions used (decode and disambiguate; assign referents to indexicals, demonstratives, tense indicators, etc.) and combine these values according to the rules of semantic composition that are part of our linguistic knowledge. The attractiveness of such a systematic account of semantic content is at least partly what motivates many of the minimalist approaches discussed in this paper. If it could be extended to the intuitive,

rather than minimal, content, then it would enable a systematic and constrained account of the relation between utterances and the propositions they explicitly express. But if free—that is, non-linguistically mandated—pragmatic processes 'intrude' on explicature, then the prospect of a systematic account of how we get from linguistic meaning to explicature diminishes. The indexicalist is sceptical that these alleged free pragmatic processes can be shown to be adequately constrained by purely pragmatic factors, and claims that they would massively overgenerate interpretations of utterances at the level of explicature. For example, Stanley (2005: 225–226) asks, if free pragmatic enrichment can supply the quantifier domain so that an utterance of "Every Frenchman is seated" can have the explicature EVERY FRENCHMAN IN MY PHONOLOGY CLASS IS SEATED, what prevents it from enriching the utterance in a different way, e.g. to EVERY FRENCHMAN OR DUTCHMAN IS SEATED? This is an interpretation which, according to Stanley, is impossible but is predicted to be able to occur by free pragmatic enrichment.

I do not address this systematicity objection in this paper; for more detail, see Stanley (2002, 2005), and for responses to it, see Carston and Hall (in preparation), Hall 2008, 2009, 2014, Recanati (2010: introduction and Chap. 1). My concern with it here is its role in justifying the idea that explicature—intuitive truth-conditional content—should correspond to semantic content and so requires the positing of extensive covert indexicality in the linguistic logical forms of sentences. This approach, in principle, promises to better substantiate the idea of semantic content than do the minimalisms discussed in Sect. 2, as it suggests that what is agreed to be a psychologically real level of representation—that level to which our truth-value judgments about the utterance are responsive—is the result of only linguistic decoding plus processes that are mandated and constrained semantically.

In practice, though, for this approach to work requires all pragmatic contributions to explicature to plausibly be traceable to linguistic form. Otherwise, what can be considered the semantic content (the product of decoding, disambiguation, and saturation) will fall short of explicature—a result that undermines both of the justifications for positing it (that it is a psychologically real level of content, and that free pragmatic effects on explicature must be excluded thanks to their unconstrained nature). Below, I present a range of types of pragmatic contributions to explicature for which, I argue, it is not plausible that they are underpinned by covert structure.

Before that, it should be noted that there does seem to be linguistic evidence that a small number of pragmatic effects on explicature are linguistically mandated by a variable in the linguistic meaning. This evidence concerns relational terms such as 'local', 'home', 'enemy', 'foreign':

11.a. [Every reporter]$_i$ was sponsored by her$_i$ local bar
   b. *Her$_i$ local bar sponsored [every reporter]$_i$

12.a. Every reporter was sponsored by a local bar
   b. *A local bar sponsored every reporter (Carston 2002: 200)

The expression "local" in (12) seems to behave syntactically very like the overt pronoun "her" in (11), giving rise to so-called 'weak crossover effects': neither expression, in the (b) sentences here, can be bound by the quantified phrase "every reporter," and Carston (2002) and Recanati (2004) agree that these are plausible cases of covert variables (though for some doubts about the weak crossover evidence, see Pupa and Troseth 2011). Other likely cases are "ready" and "enough," which seem inherently underdetermined, and require a specific value (ready for x, enough of y) to be provided on every occasion of utterance. A linguistic variable mandating saturation, then, may well be the right analysis of these expressions.

But the expressions for which there is such evidence are very limited in number, leaving the indexicalist with no linguistic evidence for many of the hidden indexicals he is positing.[11] The strongest remaining argument for his overall approach is his pessimism about the prospects of a sufficiently constrained account if one posits free pragmatic enrichment. Providing a positive account of how free enrichment is constrained is beyond the scope of this paper; instead, what I do in the rest of this section is show that, if that concern were valid, then the indexicalist would be just as susceptible to it as the pragmatic enrichment approach, because there are several kinds of contextual contribution to truth conditions that are not plausibly traceable to linguistic logical form. In the rest of this section, I present a selection of the examples used by Hall (2008: Sect. 3) to make this case.

The first phenomenon I'll consider is what Stanley himself acknowledges is the central worry for the indexicalist, which is deferred reference, or metonymy, illustrated by (13a) [with the explicature given in (13b)] and (14):

13.a. The ham sandwich wants his bill.
   b. THE PERSON X WHO ORDERED THE HAM SANDWICH WANTS HIS$_X$ BILL.

14. I'm parked out back.

Following Nunberg (1995), most authors accept that the deferred content [e.g. PERSON WHO ORDERED THE HAM SANDWICH in (13)], rather than the encoded, or 'literal', meaning, contributes to the explicature of the utterance. Supporters of this approach include those on both sides of the present debate: Recanati (1993, 2004), Carston (2002), Sag (1981), Stern (2000, 2006), and Stanley (2005). Stanley also agrees with relevance theorists and contextualists in dismissing the idea that there is a covert metonymy operator or indexical in logical form, on the model of Stern (2006)'s metaphor operator. One reason that a metonymy operator/indexical is unappealing is that it would place no constraints on the sort of deferred

---

[11] Stanley (2000) suggests another kind of evidence for certain hidden indexicals, the argument from binding. An utterance of "Every time John lights a cigarette, it rains" has an interpretation where the location of the raining co-varies with the location of the cigarette lighting, an interpretation which, according to Stanley, could not arise from pragmatic inference alone, meaning that it must be underpinned by a linguistic variable. However, Stanley offers no argument for this claim, which has been rejected by Carston (2002), Neale (2004), and Recanati (2002).

interpretation the hearer arrives at, because the relation between literal and deferred meaning is different from case to case. Consider just the two examples above: (13) requires a move from reference to a culinary item to reference to its orderer; (14), from a property of a car to a property of a person. As Stanley (2005: 229–230) says, beyond the fact that the property encoded by the metonymically used expression must provide a guide of some sort, the content is determined entirely by pragmatics, and the syntax/semantics places no constraints on what the pragmatics can do. A metonymy operator would be, therefore, redundant. A second reason for rejecting the metonymy operator approach is that to account for such figurative effects semantically would require a massive multiplication of linguistic operators for which there is no evidence: Almost every simple or complex expression can be used in a variety of figurative ways, so practically every word and phrase in the language would require not just an operator for metonymy, but also one for metaphor, and perhaps for other figures such as hyperbole and meiosis. In summary, Stanley agrees with the defenders of pragmatic enrichment that metonymy contributes to truth conditions, and that it is not underpinned by a linguistic operator, and this amounts, I think, to a concession that this is a genuine case of a free pragmatic effect on explicature.

One might feel uncomfortable relying on figurative cases to support pragmatic enrichment, feeling that, in 'normal', non-figurative speech, the indexicalist thesis should hold. But, as I discuss below, there are many kinds of non-figurative utterances to make the case, which metonymy, therefore, simply reinforces.

One set of examples where truth conditions and semantics diverge are (15a)–(17a), where the overt meaning alone can determine a proposition (modulo saturation), but this is clearly not a proposition that the speaker wants to communicate, or that the hearer recovers [likely explicatures are given in (b)]:

15.a. It will take time for that cut to heal.
    b. IT WILL TAKE A LONG TIME FOR THAT CUT TO HEAL

16.a. Mary has a brain.
    b. MARY HAS A VERY GOOD BRAIN

17.a. You're not going to die.
    b. YOU ARE NOT GOING TO DIE FROM THAT SCRATCH.

The propositions determined by the linguistic meaning of (15a) and (16a) are trivially true, while (17a), taken literally, is patently false. Utterances of such sentences, though, are not perceived as obviously non-literal and are accepted as non-trivially true in appropriate contexts, so the extra elements of meaning, such as FROM THAT SCRATCH in (17), are unlikely to be mere implicatures. Rather, what forms the intuitive truth-conditional content in each case is (b), on the basis of which the utterance would be evaluated, agreed/disagreed with, etc. (e.g. in reply to (15a), one might say "No it won't, I'm having the stitches out tomorrow"). However, hidden indexicals or other covert elements are highly improbable in these cases. The interpretations of "time" and "brain" in these utterances are not instances of domain restriction, so the domain variables allegedly attached to the

nominals would not account for these effects, and the meaning of "die" seems very unlikely to encode a variable for 'cause of death'. That these are highly occasion-specific effects of context makes it very difficult to accept that they should be provided for in the linguistic meaning by an indexical triggering mandatory saturation (which would presumably have to occur even in the many cases where the provision of a value would be redundant). These examples look like further cases of genuine free pragmatic effects on truth-conditional content, which cannot be traced to an element of the syntax or semantics.[12]

Finally in this section, I look at a construction that both camps in the debate have discussed in detail, and argue that the indexicalist solution proposed so far is inadequate. An utterance of (18) would often be understood as communicating a causal relation between the events referred to by the conjuncts:

18. If Hannah insulted Joe and Joe resigned, then Hannah is in trouble.

Since the causal connection falls within the scope of the operator 'if…then', practically everyone agrees that it contributes to truth conditions. King and Stanley are no exception here: they write: '[(18)] seems to express the proposition that if Hannah insulted Joe and Joe resigned as a result of Hannah's insult, then Hannah is in trouble' (2005: 158). This requires, for them, that the linguistic meaning of the sentence contain some element that can take AS A RESULT as its context-specific value. Here, they appeal to Stalnaker (1999)'s work on the semantics of indicative conditionals. According to Stalnaker, an indicative conditional is true if and only if the consequent is true in every one of the most relevantly similar worlds in which the antecedent is true, and this similarity relation counts only those non-actual worlds compatible with the mutually accepted background assumptions as similar worlds for purposes of semantic evaluation (King and Stanley 2005: 154). If this is going to help King and Stanley account for examples like (18), the idea has to be that the conditional structure would come with a linguistic parameter, requiring saturation, which specifies the similar-worlds constraint, thus requiring the selection of the most relevantly similar worlds in the context set. In a context in which 'the speaker has in mind a causal relationship' between the events described in the conjuncts, the most relevantly similar worlds will be just those worlds in which that causal relationship holds. This predicts the reading of (18) on which a causal relation is part of the truth-conditional content (ibid: 160).

---

[12] Other good candidates for free pragmatic enrichments are loose uses such as "This steak is raw," "Jane has a round face" and "Holland is flat". The same arguments apply as given in the text for examples (15)–(17): they are optional pragmatic effects (not occurring on every use of the expression) and are intuitively part of the truth-conditional content. See Hall (2008: Sect. 3) for more discussion. Another plausible case is provided by referential uses of definite descriptions, whose truth conditions are widely agreed to be distinct from those of attributive uses (see, among many others, Recanati 1993; Larson and Segal 1995; Bezuidenhout 1997; Neale 1999, and King and Stanley 2005), but there is no argument for a hidden indexical or parameter here. Assuming, along with these authors, that the encoded meaning is attributive, then the move from encoded meaning to referential truth conditions is a free pragmatic effect.

One important problem with this solution is this (see Hall 2008: Sect. 3 for others): Since it is the semantics of the conditional that is supposed to account for the incorporation of the causal relation into truth conditions, it follows that, for King and Stanley, the explicature of the unembedded conjunction (19a) does not include the causal relation. Instead, this would be conversationally implicated, and is, according to them, calculable from the explicature using Gricean maxims:

19.a. Hannah insulted Joe and Joe resigned.
   b. Explicature: HANNAH INSULTED JOE & JOE RESIGNED.
   c. Implicature: JOE RESIGNED BECAUSE HANNAH INSULTED HIM.

Is this supported by our intuitions about truth conditions? I don't think so, but intuitions are perhaps not entirely clear with isolated utterances such as (19a). They seem to be sharper when (19a) is presented as one premise of an argument, as in (20), which looks like an obvious case of modus ponens, and which most people would judge to be a valid argument:

20.a. If Hannah insulted Joe and Joe resigned, then Hannah is in trouble.
   b. Hannah insulted Joe and Joe resigned.
   c. Hannah is in trouble.

For King and Stanley, the proposition expressed by (20b) must have a different propositional form than that of the antecedent of the proposition expressed by (20a). That is, the argument is not in the modus ponens form 'If P then Q. P. Therefore Q', but is rather 'If P then Q. R. Therefore Q', which is not a valid inference. So, the indexicalist account makes a wrong prediction.[13]

This outcome does not sit well with King and Stanley's concern to respect intuitions about truth conditions, which are what they take to be the primary object of semantic theorizing (see, especially, Stanley and Szabo 2000: 240 and King and Stanley 2005: 141). They offer no explanation for why our intuitions about the validity of the argument in (20), which depend on the intuitions about the truth conditions of (18) and (19), should not be respected, but are forced into this position by the fact that none of the overt expressions in the unembedded conjunction are plausible candidates for carrying an appropriate hidden indexical or a semantic parameter that could pick up a similarity relation. So (19) is another case where an aspect of intuitive truth-conditional content cannot be accounted for linguistically but seems to rely on a process of free pragmatic enrichment.

In conclusion, then, there are several types of case where context clearly does affect truth conditions, but where this effect cannot be accounted for by any linguistic trigger requiring contextual contribution. In other words, there are elements of meaning that, according to the truth-conditions criterion shared by both camps, are part of explicature, but that cannot be considered semantic.

---

[13] This argument was first used by Carston (2004) against Levinson's (2000) treatment of 'and'-conjunction utterances, which shares the relevant features of King and Stanley's (2005).

The indexicalist approach denies the existence of free pragmatic effects on explicature—a position motivated by the suspicion that the only way that a systematic, explanatory account of our grasp of explicit utterance content will be tractable is to assume that it is semantic content: that is, that the role of the linguistic meaning is to provide the truth-conditional content of the utterance, modulo contributions that are triggered and constrained semantically. The examples surveyed in this section, though, have shown that covert indexicality, while plausible for certain phenomena such as relational expressions and perhaps even domain restriction, cannot be as extensive as is required to support the indexicalist thesis. Ultimately, then, this approach does not rescue the idea of semantic content.

4 Genuinely Minimal Minimalisms and Multipropositionalism

Having discussed in Sect. 2 some views that posit minimal semantic contents, which are not the 'enriched' contents (explicatures) that speakers communicate, but which also differ from standing linguistic meaning in that some or all of the overt indexical and other contextual elements are saturated, the conclusion was that no convincing case has been made for such a level of content: It is an 'idle wheel' in an account of communication, and any semantic (or semantically-mandated) content falls well short of the communicated content. Nonetheless, there is an enduring feeling that semantic theory should deal in truth conditions and propositions—that the function of linguistic meaning is to determine these, even if they are not what are judged the intuitive truth conditions of the utterance. Despite the arguments presented in the last two sections, this view can be reconciled with the relevance-theoretic position on the semantics-pragmatics distinction. What are called for here are genuinely minimal propositions, determined solely by standing linguistic meaning. I discuss two recent accounts that posit such entities—Borg (2004), and Korta and Perry (2006, 2008).

Borg (2004) presents a semantic theory that she sees as meshing with Chomskyan/Fodorian views on mental architecture. These views hold that the perception and language encoding/decoding systems are modules, translating visual, linguistic, and other perceptual signals into representations in a common format—a language of thought or 'Mentalese'—which these modules deliver to the central system (or systems[14]), which is where reasoning processes, including pragmatic inference, take place. (Language, of course, is also an output system, encoding central-systems representations). These peripheral input systems are

---

[14] It seems likely that the central inferential system itself has a far more modular structure than Fodor (1983, 2000, etc) has been prepared to admit (see Sperber 2002, Carruthers 2006). But whatever view on its internal structure is correct does not affect anything I say here, so I remain neutral: the distinction between perceptual and language modules on the one hand, and the central inferential system(s) on the other, is sufficient.

'encapsulated', that is, they do not have access to information in the central system (or in the other input systems), so a given input into one of these modules will always deliver the same output to the central system.

Semantics deals with linguistic decoding, so must, for Borg, be part of the language faculty, encapsulated from central-systems processes. This means that no pragmatic inference can affect semantics: Semantics can only be that portion of communicated meaning that is formally tractable—that can be described without any influence from pragmatics. As discussed above (Sect. 2 on Bach 2001), and as Borg agrees, assignment of reference to indexicals is not formally tractable, because of its inevitable dependence on recognition of speaker intentions (she considers and rejects various ways of making it so in Chap. 3 of her 2004 book). So Borg concludes that semantics is purely the result of linguistic decoding, and does not incorporate anything else often included under the label 'semantics', such as assigning referents (not even to "I").

This much is something that relevance theorists and many contextualist philosophers (in particular Carston, Recanati, Sperber and Wilson) would fully endorse. The difference between Borg and these authors is that she assumes the output of semantics must be propositional. Indexicality is the obvious problem for this view, since, as she argues at length, the output of a modular semantics cannot include the results of reference assignment: In the case of indexicals, semantics would just give the Kaplanian 'character' (Kaplan 1977/1989), which is their context-independent standing meaning (for example, that "I" refers to the person who utters it). Kaplan distinguishes 'character' from 'content': the former is not a constituent of a proposition; instead, it is an indication of how to recover the content (the referent), which is what does form a constituent of a proposition. So it would appear that the result of semantics cannot be a complete proposition: it contains variables with attached instructions or constraints, specifying what kind of value should saturate it (for instance, the result of decoding "he" would be something we can represent as [$X_{male,\ singular,\ animate}$]—a placeholder for the concept that will be pragmatically supplied, where the 'X' is a variable or slot requiring saturation).

Borg's solution is to suggest that the output of linguistic decoding in the case of indexicals and demonstratives is a syntactically-provided singular concept. So the result of decoding an utterance of "That is red" is α IS RED. α is a singular concept, the semantic content of "that," and comes with a constraint on how it is to be integrated with the hearer's 'wider cognitive environment'—an instruction for what sort of concept is to be pragmatically supplied to grasp the content that the speaker intends to assert. In effect, the semantic content of an indexical or demonstrative on this proposal is the character of the expression under a Kaplanian dthat operator. For example, the semantic content of "That is red," $\boldsymbol{\alpha}$ IS RED, can

also be characterized as [dthat (salient object) is red].[15] Although, as a result only of semantic processing, the hearer will not be able to work out the intended referent (and may not be able to assign a referent at all if he lacks sufficient knowledge of the context), the syntactically-tokened concept is the kind of thing that can have content; it is object-dependent. As the utterance has object-dependent truth-conditions, the semantic content that the hearer recovers counts as being truth-evaluable (propositional) even if the hearer is not in a position to know what the truth conditions are.

It is a contentious part of Borg's view that these very 'thin' singular concepts, which result purely from semantic decoding processes, can endow the semantic content with truth-conditions and propositionality: one could just as well say that the subpropositional logical form assumed by relevance theorists has truth conditions and genuine content in this sense.[16] The reason for her insistence on the output of semantics being propositional is not clear; that aside, though, her approach seems compatible with that of relevance theory, drawing the semantics-pragmatics distinction in the same place and recognizing that there may be a wide gap between the output of semantics and the proposition explicitly expressed. And, in fact, her thin, singular concepts might turn out to be a useful way of cashing out the relevance-theory talk of indexicals triggering slots, or placeholders for concepts, as suggested by Carston (2008b).

How Borg's minimalist semantics differs from the RT picture, then, is not that it posits some extra level of semantic content between encoded linguistic meaning and explicature, but that it gives a slightly different, but probably compatible, account of how the linguistic meaning is represented and integrates with the pragmatic inference system. On the RT view, linguistic decoding produces an incomplete propositional template or schema, with a number of slots corresponding to indexicals and demonstratives (plus similar elements such as tense operators). Each slot consists of a variable-like element with constraints on what sort of concepts are to be supplied as their value. Borg's approach to indexicals and demonstratives suggests a different way of thinking about what their decoding results in. This variable-plus-constraint complex is what an indexical or demonstrative is likely to be when considered as part of our linguistic knowledge, independent of any utterance of it—the format in which its meaning would be stored in the lexicon. However, when an indexical is uttered the result is not that

---

[15] Carston (2008b: 364) notes that there is a worry about whether Borg's account can deal with descriptive uses of indexicals and demonstratives, for example, where the speaker points at a massive footprint and utters "He must be a giant," or holds up a book and says "This is my favourite author." Borg (2002) argued that even for descriptive cases, the semantic content is singular, and seemed to see the descriptive content as an implicature. Intuitively, the descriptive content is part of the explicit content, but it is not clear how this could be accommodated by Borg given that dthat is a rigidifier.

[16] Carston (2008b: 365) questions whether the semantic content really does meet the '(Davidsonian) truth-conditional desideratum' that Borg appears to want semantic content to meet, which requires that 'the language user who grasps [the] truth condition [be able] to tell worlds in which it is satisfied from worlds in which it is not' (Borg 2004: 235).

the standing meaning, which is not a concept, simply appears unaltered in the output of semantics, but rather that the act of utterance brings about the tokening of a singular concept, which carries with it the constraint (or 'character') of the indexical. This seems to provide a plausible alternative to the conception of logical form so far assumed by relevance theory.

A different kind of genuinely minimal propositional content, directly determined by the rules of the language, completely independently of the speaker's meaning, is the 'reflexive' proposition (cf. Perry 2001). It is 'reflexive' because it makes reference to the utterance itself in describing the truth conditions: an utterance u of "I am here now" expresses the reflexive proposition that THE UTTERER OF u IS AT THE PLACE u IS UTTERED AT THE TIME u IS UTTERED. As Recanati (2004: 66–67) says, this way of preserving the notion of 'what the sentence says', in a purely semantic sense, does not support the syncretic view (e.g. Cappelen and Lepore, Soames) with its four levels, because the reflexive proposition 'does not incorporate those contextual ingredients whose provision is linguistically mandated; […] it is directly and immediately determined by the linguistic meaning of the sentence.' So we appear to retain the RT picture with its three levels of content: the standing linguistic meaning (and the reflexive proposition it directly determines), explicature, and implicature.

Korta and Perry (2006, 2008), however, criticize the RT and contextualist picture, because, as they illustrate, often what would seem to be the explicature is not what serves as input to the hearer's inference to implicatures. What is required instead may be one of various levels of partly token-reflexive or descriptive content that they propose,[17] from (and including) the reflexive proposition up to (but not including) the 'proposition expressed' by which they mean the pragmatic what is said, or explicature, incorporating the result of saturation, disambiguation, and enrichment. It is this proposition, as opposed to the partly token-reflexive and descriptive propositions, that they see as the 'official content'. Korta and Perry work through several examples to demonstrate how these propositions of varying degrees of token-reflexivity enter into comprehension: In some cases, one of these propositions might have a role to play when there is some difficulty in recovering the proposition expressed, but in other cases, it seems that a partly token-reflexive proposition would be equally or even better suited as input to the hearer's inference to implicatures than the fully saturated (enriched, etc.) proposition expressed would be. This initially looks to motivate the idea of more levels of content in between standing linguistic meaning (plus fully reflexive proposition) and explicature.

Here are some of Korta and Perry's examples. They consider Grice's (1975/1989: 32) example of the motorist (A) who, having run out of petrol, is standing by his car and is approached by B:

---

[17] It appears that this token-reflexive content can include reflexive descriptions of elements that are not part of the linguistic meaning, such as IN WHATEVER DOMAIN THE SPEAKER OF u INTENDS (Korta and Perry 2006).

21. A. I'm out of petrol.
     B. There's a garage round the corner.

They concentrate on how B would interpret A's utterance, of which they write, 'It is natural to take A's opening remark as implicating that he would like some help in finding petrol for his car'. Then they ask us to suppose that A is Harold Wilson,[18] and continue:

According to the theories of names and indexicals that are now widely accepted, A would have then expressed the same proposition in this scenario [i.e. 21′ below], the singular proposition individuated by Harold Wilson and the property of being out of petrol.

[21']. A. Harold Wilson is out of petrol.
        B. There is a garage round the corner. (Korta and Perry 2006: 169)

That is, the same proposition is expressed by A in both scenarios (21) and (21′), but, while B's utterance in (21) is an appropriate thing to say in reply, Korta and Perry comment that 'In scenario [21′] there is no motivation for B's remark. What does the proximity of a garage to the participants in the conversation have to do with Harold Wilson's being out of petrol?' (ibid: 170)

Korta and Perry's point is this: It is the implicature of A's utterance in (21), that A would like help finding petrol, which prompts B's reply. But this implicature is not inferred from the proposition expressed by A's utterance. If it was, then A's utterance in (21′), since it expressed the same proposition, should have led to the recovery of the same implicature, to which B's reply would be an appropriate response. But B's response in (21′) does not seem to be motivated. So the implicature of A's utterance must have been inferred from something other than the proposition expressed. For this, they suggest, what the hearer (B) starts with is the reflexive proposition, THE SPEAKER OF u IS OUT OF PETROL AT THE TIME OF u, from which, together with the contextual information that the speaker is the person B is now looking at, he infers a proposition THE PERSON I AM NOW LOOKING AT IS OUT OF PETROL NOW. This proposition is something less than the 'official' proposition expressed, in that it doesn't have HAROLD WILSON assigned as referent of "I" (or a specific time assigned to the tense indicator): But it's this *description* of the proposition expressed that is the basis for inferring the implicature, and thus for B's helpful reply.

A second example has 'JP' saying to 'KK', while KK is driving,

22. He is going to drive his car into yours.

They say, 'JP is, let us imagine, referring to FR, a famous philosopher, who is careening down the street in the opposite direction. There is a pretty clear implicature that KK would do well to engage in evasive maneuvers, to avoid

---

[18] British Prime Minister in Grice's day.

getting hit. But how is KK supposed to figure this out?' The mode of presentation under which KK needs to be thinking of the person JP is referring to, in order to figure out the implicature, is something like THE PERSON DRIVING THAT CAR THAT IS WEAVING ACROSS THE STREET. As Korta and Perry say, 'To take proper evasive action, it is unnecessary, and perhaps counterproductive, to recognize JP's referent as the famous philosopher FR. KK might become awestruck [etc.]' (ibid: 180–181).

The apparent lesson of these examples (and others given in Korta and Perry 2006, 2008) is that we need to recognize various levels of propositional content somewhere between the reflexive proposition (equivalent to standing linguistic meaning) and the official proposition expressed. These in-between levels are utterance-bound and descriptive to varying degrees: rather than having the referents assigned to the indexicals, demonstratives, and tense indicators, what hearers sometimes need in order to work out the implicatures, and respond appropriately, is a proposition with, e.g. "I" assigned the value THE PERSON STANDING IN FRONT OF ME, or "that" assigned the value THE THING THE UTTERER OF u INTENDED TO REFER TO. Does this require a revision of RT's three-level picture, which has linguistic meaning being developed into explicature (also known as the proposition expressed, what is said, etc.), with no levels of semantic or sentence content in between?

Despite initial appearances, it does not undermine the RT semantics-pragmatics distinction presented earlier, because RT takes the explicature to be, roughly, the representation that the hearer does in fact entertain as the development of the linguistic meaning, and this is different from what Korta and Perry describe as 'the (official) proposition expressed' or 'official content'. Recall their discussion of Grice's example, involving an utterance of "I am out of petrol." They claimed that, if the speaker is Harold Wilson, then the proposition expressed is the singular proposition individuated by Harold Wilson and the property of being out of petrol. That may be the case, but it is not relevant to an account of how what is communicated here is recovered. In the sense in which Korta and Perry are using the notion, the 'proposition expressed' by A's utterance is HAROLD WILSON IS OUT OF PETROL, by virtue of Harold Wilson being the speaker, irrespective of whether the hearer even knows the speaker's identity. Let's assume that the hearer does not know who the speaker is (and that the speaker has no reason to assume that he does). Then the official proposition expressed might be as Korta and Perry claim, but this is not what the hearer recovers, or is expected to recover, as the explicature of the utterance. The explicature is not the (external) content, but the representation that hearers recover of that content (which fulfils the speaker's communicative intention).

In example (21), then, the explicature of A's utterance is just what Korta and Perry say is the proposition that forms the basis for the hearer's inference to the implicature: it can be glossed as something like THE PERSON I AM NOW LOOKING AT IS OUT OF PETROL NOW, though may be better represented as α IS OUT OF PETROL NOW, with 'α' being a singular concept with the relevant person as content, and with a mode of presentation that can be paraphrased as 'the person I am now looking at'. It might be that, in some cases, two or more such propositions, of different degrees

of utterance-bound-ness/descriptiveness, are required as the basis for inference to different implicatures of the same utterance, or that the hearer would probably represent the full referential content but also a less than fully referential content. In such cases, relevance theorists would simply say that the utterance has more than one (basic[19]) explicature, which are all speaker-meant, unlike the semantic contents posited by Cappelen and Lepore, Soames, and Bach as intermediate between linguistic meaning and speaker meaning.

To summarize this section, both of the views discussed here are compatible with relevance theorists' and other contextualists' views on the semantics-pragmatics distinction described in Sect. 2. Borg draws the distinction in the same place as relevance theory, but provides a slightly different take on the nature of the representation of the linguistically encoded meaning than that assumed so far by relevance theory and contextualists; it remains to be seen which alternative is correct. Korta and Perry highlight the need to recognize that explicatures need not be 'fully' enriched, saturated, etc., but that any of a number of different varieties of representation of different degrees of utterance-boundness and descriptiveness may be the explicature(s).

The approaches discussed in this paper so far do not provide any grounds for revising relevance theory's semantics-pragmatics distinction or justify a level of semantic content distinct both from linguistic meaning (and propositions directly and immediately determined by it) and from explicature. However, a position that remains to be considered in relation to these questions is the relativism about utterance content that has sparked intense debate in the last few years. I leave for another time a detailed comparison of relativism with relevance theory and contextualism. In the following, final, section I just present some of the general features of relativism(s), drawing mainly on the 'moderate' version defended by Recanati (2007), and make some brief preliminary remarks to try and assess its implications for the issue discussed in this paper.

5 Moderate Relativism, Semantic Content, and Unarticulated Constituents

The central idea of semantic relativism is that certain elements that appear to contribute to the truth-conditions of utterances, so on a relevance-theoretic or contextualist account would probably be considered part of explicature, are better treated as lying outside of that content. Rather than being implicatures, though, they are parameters in the circumstances of evaluation or context of assessment,

---

[19] Relevance theory makes a distinction between the basic explicature (= asserted content, intuitive truth-conditional content, etc) and higher-level explicatures, which are speech-act or propositional-attitude descriptions. Where the term "explicature" is used without any qualification, it refers to a basic explicature of the utterance.

against which the content is evaluated for truth. Common examples are taste predicates ("Seaweed is delicious"), knowledge claims ("I know that my car is on the drive"), and epistemic modals ("Fred might be in Boston"). Assuming the values of indexicals are fixed, the content of each of these utterances does not vary across contexts; instead, what varies is a standard of taste or of knowledge. When Mary utters sincerely "Semantics is fun," the content is true relative to Mary's taste; when Bob replies "Semantics is not fun," the content is true relative to Bob's taste. This explains why Mary and Bob can give the appearance of disagreement, while one would not want to say that either of them is at fault.[20]

Here, I look at a variety of relativism that is relevant to the issues discussed in this paper because it directly addresses the question of free pragmatic processes contributing to explicature, processes whose existence I have used as an argument against the positing of a level of semantic content distinct from linguistic meaning. Relevance theorists distinguish, broadly, two types of free pragmatic enrichment. In some cases, such as when "It's raining" is given a location-specific interpretation, the explicature contains an extra constituent, an 'unarticulated constituent' (henceforth abbreviated to UC), that is not the value of an element of linguistic form. Similarly when "and" is interpreted as & AS A RESULT, as seen in several of the above examples. However, in other cases, what is going on seems better construed as not the addition of extra concepts, but the adjustment of an encoded concept so that the concept understood as communicated by the use of a word is different than the encoded concept—it may be narrower, looser, or some combination of the two. Examples of this pragmatic adjustment of word meaning or 'lexical modulation' seen above include "drink" being used in example (1) to denote only alcoholic drinks, and "raw" being used of a steak that has received some (inadequate) amount of cooking; a few more examples are given here[21]:

23. To buy a house in London you need *money*.
24. Put the *empty* bottles in the garbage.
25. This water's *boiling*.

Grasping the explicature of each of these examples is very likely to involve an optional process of meaning modulation (of the concept encoded by the italicized word in each case). In most contexts, the proposition that buying something requires (some, or any amount of) money will be trivial and uninformative, so the lexically encoded concept MONEY is likely to be narrowed to a concept, represented as MONEY*, that denotes just those quantities that would count as sufficiently large amounts of money in the context of London house-buying. If we imagine (24) uttered in a context where people are clearing up after a party, then the bottles in question may well contain dregs of beer, cigarette ash, etc., so will not be strictly

empty; similarly, the concept encoded by "boiling" in (25) may be adjusted if the expression is used as an approximation or hyperbole.

The addition of extra components—unarticulated constituents—and lexical pragmatic modulation are the two types of free pragmatic enrichment that relevance theorists distinguish.[22] The latter is a relatively new idea, and it is not yet clear how many of the cases previously treated as unarticulated constituents should be reanalyzed as modulation, though some clearly do not look susceptible to this treatment. Consider again the case of weather predicates, discussed earlier, where "It's raining" has the explicature IT IS RAINING IN LONDON. It is difficult to construe RAIN-IN-LONDON as a narrowed form of the concept RAIN. Similar remarks apply to the enrichment of certain "and"-conjunctions, such as the cause-consequence interpretation in (1) [and the similar causal interpretation in (2)]. So at least some cases are likely to remain best analysed as unarticulated constituents rather than lexical pragmatic modulation. Anyhow, the idea is that both types of enrichment contribute to explicature, either changing or adding a constituent of the representation. However, Recanati (2007) takes a different view: he suggests that, while modulation contributes to the explicit content (what he labels the 'lekton'), unarticulated constituents do not; instead, they are part of the situation of evaluation. Lekton and situation of evaluation together form the entire truth-conditional content, which Recanati calls the 'Austinian proposition'.

To illustrate, consider weather predicates, on which much of the discussion is focused. The truth of an utterance of "It's raining" virtually always depends on the situation at a specific location. However, for Recanati, the location UC is not part of the lekton but a parameter of the circumstances of evaluation. There are two variants of this scenario: first, in much of our weather talk, as when we look out the window and utter "It's raining" to someone else in the same room, we are not thinking about the location we are in as opposed to any other location.[23] So in this case, the location has no cognitive significance, and is not represented at all. Yet the truth of the utterance still depends on the location.

The second variant is where we do cognitively articulate the location, for example, if someone is on the phone to a friend in another city, who reports "It's raining," the hearer, in comprehending the utterance, needs to entertain a concept of the location being talked about, as opposed to his own and any other salient locations. Recanati rightly claims that, from the fact that the location is represented, it does not follow that it is part of the lekton (2007: 224–230)—which is to

---

[22] Probably at least one more type should be distinguished: the transfer involved in referential uses of definite descriptions, and the metonymy illustrated in (13)–(14).

[23] Perry (1986) considers the case of Z-land. Z-landers are a small, isolated group who are unaware of any locations other than the small area they live in. They do not have a concept of Z-land as opposed to other places. When a Z-lander utters or hears "It's raining," he does not have in mind the location at which it is raining: the location is not articulated even in his thought. While the rest of us do have a concept of the place where we are as opposed to other locations, this does not mean that we contrastively represent our location when we utter weather-predicates; it may be that, as Perry says, 'there is a little of the Z-lander in the most well-travelled of us'.

say, it does not follow that its representation is incorporated into explicature. Instead, the location remains part of the situation of evaluation. Other elements of the overall truth-conditional content that would submit to the same treatment are the judges of taste and available information that appear to affect the truth values of utterances of "This is delicious" and "He might be in Boston," mentioned above.

In earlier sections I rejected two candidates for a semantic content distinct from merely the linguistically encoded meaning: Cappelen and Lepore's minimal propositions resulting from linguistic decoding plus saturation of all indexicals, and Bach's what is said resulting from linguistic decoding plus saturation of pure indexicals. But now we have a third candidate, the lekton, which is, according to Recanati, a genuine level of content—the content developed out of logical form that is explicitly represented in thought, or, the semantic interpretation of the sentence relative to context. In the case of utterances, all constituents of the lekton are the value of an element of the uttered sentence; the difference from the minimalist and indexicalist positions is that the lekton includes not only the values of indexicals, but also the results of lexical pragmatic modulation, illustrated in (23)–(25) above. What advantages does this have over the standard contextualist and RT view? Firstly, it could be more plausible for the cases in which we are not contrastively thinking of the location we are at, or of our own taste or knowledge as opposed to others'. Second, Recanati (2010) suggests, the 'no UCs in the lekton' position may also have the advantage of allowing the preservation of a compositional account of our grasp of explicit content, and thus should find favour with (or less opprobrium from) those indexicalists (Stanley, Szabó, Martí, King) and others who are sceptical about the prospects of a systematic account incorporating free pragmatic enrichment.

This is only one among several approaches that are discussed under the label 'relativism'. I leave a detailed assessment to further work; here, I just raise a few questions about the idea that there are no UCs in the lekton/explicit content.

First, consider again the normal utterances of "It's raining" where, Recanati suggests, the location has no cognitive significance: we are not thinking of our location as opposed to any other. I agree with the intuition here, that we do not have a contrastive representation of the location, so the explicature should probably not be characterized as, for example, IT IS RAINING IN LONDON. However, this would mean that there is no difference in the cognitive significance of the location in these two scenarios: (a) when someone looks out the window and utters "It's raining," and (b) when someone utters "Mary's singing" (in a context where the location is irrelevant to the truth of the utterance).[24] This does not seem quite right, as the former case requires the thought to be anchored to a location in a way that the latter doesn't. An alternative to the explicature given above, then, would be one that does incorporate a representation of the location in the weather case, but,

---

[24] The location can, of course, be relevant in certain contexts, e.g. "I'm going to the concert at the school tonight. Mary's singing."

rather than a full-fledged concept, it would be more like a mental indexical or demonstrative, much like singular concepts of individuals that would be expressed with, e.g. "that guy," "him," or "this." As Korta and Perry (2006)'s examples, discussed in the previous section, showed, often what is needed in the representation that the hearer entertains is not a fully identifying conception of the object or person referred to by a referring term; the same should apply to locations, times, etc. If this is correct, then the explicature would be better characterized as IT IS RAINING-HERE (where HERE is not a natural language indexical but a mental one whose content is fixed).

The more relevant issue here, though, is whether the location constituent, even if represented (whether by the full-fledged, contrastive concept of that location, or by a demonstrative-like element) is incorporated into explicature. One reason to be sceptical is that a lot of the motivation for relativism appears to come from the phenomenon of faultless disagreement, for example where Mary says "Snails are delicious" and John says "Snails are not delicious." Something similar happens with the other paradigm cases for relativism: for example, Mary says "Fred might be in Boston" and John says "He can't be: I saw him just twenty minutes ago." Despite John's contradiction (which we assume to be correct) of Mary, there is a sense in which what she said seems correct or even true in that it was compatible with her knowledge at the time of utterance. Yet in the case of weather and locations, there is no similar phenomenon: If Mary, reporting on the situation where she is in Paris, utters "It's raining," and Fred utters "It's not raining," then there is no appearance of faultless disagreement: either one of them is wrong, or else they are not talking about the same location, thus not disagreeing. So in the former cases, it is easy to see how the standard of taste/speaker's grounds can be considered a parameter relative to which the content is evaluated; in the "rain" case, the location does not behave the same way, but as a constituent of the content to be evaluated. The same applies to other likely UCs, such as the causal and temporal interpretations of various kinds of conjunctions, and completion of incomplete descriptions.

This leaves the question of how pragmaticists should treat taste predicates, epistemic modals, and so on: do we need to recognize a content (lekton) versus circumstances of evaluation (or context of assessment) distinction for these? The obvious way for RT to go would seem to be to say that the reference to the judge, rather than being a parameter of evaluation, is incorporated into explicature. Several authors, under the label 'contextualism', have given indexicalist accounts of taste predicates (e.g. Glanzberg 2007) and epistemic modals (e.g. von Fintel and Gillies (2011). The way that RT would implement a contextualist account would likely be to analyse the reference to a judge as an unarticulated constituent. I leave this open, but my tentative suggestion is that, in most of the cases at issue, neither solution applies, because the relevant parameter (the judge according to whose

taste or knowledge the claim is made) is not communicated. Consider this dialogue, already mentioned above:

26. *Mary*: Semantics is fun.
    Tom: Semantics is not fun.

Both utterances in this dialogue seem to be presented as objective statements; their explicatures could be represented as SEMANTICS IS FUN/NOT FUN *punkt* (or *tout court*). The fact that we are unwilling to judge either speaker wrong, despite their disagreement, gives the appearance that the truth-conditional content being evaluated—hence the explicature—includes a standard of taste, but if the disagreement is genuine, then this is likely to be an illusion. There are two possibilities for accounting for it. The first is that what we are evaluating here is not the truth of the utterances, as we lack the means to determine which is true (if either: in some cases it may be that there is no fact of the matter). Instead, when asked to judge truth or falsity, all that we can evaluate in such cases is something closely related, which is the appropriateness of the utterance; the appropriateness depends on the sincerity of the speaker, which depends, in turn, on her personal taste. On this count, both utterances in (26) are faultless, though not both can be true. The second alternative is that my claim that no reference to the judge is communicated is quite compatible with some versions of relativism—in which case both utterances may be true from the relevant point of assessment. After all, the issue of how what is communicated is assessed for truth or falsity is not the concern of pragmatic or linguistic theory. So if it is correct that no reference to the judge is communicated, then this would be a reason to reject the proposals by Lasersohn (2005) and Stephenson (2007) that are contextualist-relativist hybrids, employing a linguistic parameter that can be bound to a judge in the context of assessment.

In conclusion, then, I have given some reasons to be sceptical of the import of relativism for the issues discussed in this paper, and possibly for pragmatic theory more generally. These brief considerations are, of course, far from decisive, and a much more detailed comparison of relevance theory and contextualism with a wider variety of relativist views will be necessary to establish whether the RT position on the semantics-pragmatics distinction and the non-existence of semantic content can ultimately be maintained.

References

Bach, Kent. (1994). Conversational impliciture. *Mind and Language, 9*, 124–162.
Bach, K. (2001). You don't say. *Synthese, 128*, 15–44.

Bach, K. (2010). Impliciture vs explicature: What's the difference. In B. Soria & E. Romero (Eds.), *Explicit communication: Robyn Carston's pragmatics* (pp. 126–137). Basingstoke: Palgrave.

Barton, E. (1990). *Nonsentential constituents*. Amsterdam: John Benjamins.

Bezuidenhout, A. (1997). Pragmatics, semantic underdetermination and the referential/attributive distinction. *Mind, 106*, 375–409.

Bezuidenhout, A. (2002). Truth-conditional pragmatics. *Philosophical Perspectives, 16*, 105–134.

Bezuidenhout, A. (2006). The coherence of contextualism. *Mind and Language, 21*, 1–10.

Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Blackwell.

Blakemore, D. (2002). *Linguistic meaning and relevance: The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

Borg, E. (2002). Pointing at Jack. Talking about Jill: Understanding deferred uses of demonstratives and pronouns. *Mind and Language, 17*, 489–512.

Borg, E. (2004). *Minimal semantics*. Oxford: Oxford University Press.

Cappelen, H. (2007). Semantics and pragmatics: Some central issues. In G. Preyer & G. Peter (Eds.), *Context-sensitivity and semantic minimalism* (pp. 3–22). Oxford: Oxford University Press.

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell.

Carruthers, P. (2006). *The architecture of the mind*. Oxford: OUP Press.

Carston, R. (1988). Implicature, explicature and truth-theoretic semantics. In R. Kempson (Ed.), *Mental representations: The interface between language and reality* (pp. 155–181). Cambridge: Cambridge University Press.

Carston, R. (1999). The semantic-pragmatics distinction: A view from relevance theory. In K. Turner (Ed.), *The semantics/pragmatics interface from different points of view (CRiSPI 1)* (pp. 85–125). Oxford: Elsevier.

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.

Carston, R. (2004). Truth-conditional content and conversational implicature. In C. Bianchi (Ed.), *The semantics/pragmatics distinction* (pp. 65–100). Stanford: CSLI.

Carston, R. (2008a). Linguistic communication and the semantics-pragmatics distinction. *Synthese, 165*, 321–345.

Carston, R. (2008b). Review of Minimal Semantics by Emma Borg. *Mind and Language, 23*, 359–367.

Carston, R., Hall, A. in preparation. Pragmatic pragmatics and the extent of linguistic systematicity. Ms, UCL.

Cohen, J. L. (1971). Some remarks on Grice's views about the logical particles of natural language. In Y. Bar-Hillel (Ed.), *Pragmatics of natural language* (pp. 50–68). Dordrecht: Reidel.

Davidson, D. (1967). Truth and meaning. *Synthese, 17*, 304–323.

Fodor, J. (1983). *The modularity of mind*. Cambridge: MIT Press.

Fodor, J. (2000). *The mind doesn't work that way*. Cambridge: MIT Press.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik, 100*, 25–50.

Garcia-Carpintero, M., & Kölbel, M. (Eds.). (2008). *Relative truth*. Oxford: Oxford University Press.

Glanzberg, M. (2007). Context, content, and relativism. *Philosophical Studies, 136*, 1–29.

Grice, H. P. (1967). Logic and conversation. William James lectures. Harvard University.

Grice, H. P. (1975)/(1989). Logic and conversation. In: P. Cole, J. Morgan, (Eds.), *Syntax and semantics 3: Speech Acts* (pp. 41–58). New York: Academic; reprinted in H. Paul Grice 1989, 22–40.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press.

Hall, A. (2008). Free enrichment or hidden indexicals? *Mind and Language, 23*, 426–456.

Hall, A. (2009). Free pragmatic processes and explicit utterance content. University College London PhD dissertation.

Hall, A. (2014). 'Free' enrichment and the nature of pragmatic constraints. *International Review of Pragmatics* 6.1.

Kaplan, D. (1977)/(1989). Demonstratives. In J. Almog, J. Perry, H. Wettstein (Eds.), *Themes from Kaplan* (481–563). Oxford: Oxford University Press.

King, J., & Stanley, J. (2005). Semantics, pragmatics, and the role of semantic content. In Z. G. Szabó (Ed.), *Semantics versus pragmatics* (pp. 111–164). Oxford: Oxford University Press.

Kölbel, M. (2008). Introduction: Motivations for relativism. In M. Garcia-Carpintero, M. Kölbel (Eds.), (1–38).

Korta, K., & Perry, J. (2006). Three demonstrations and a funeral. *Mind and Language, 21*, 166–186.

Korta, K., & Perry, J. (2008). The pragmatic circle. *Synthese, 165*, 347–357.

Larson, R., & Segal, G. (1995). *Knowledge of meaning: An introduction to semantic theory.* Cambridge: MIT Press.

Lasersohn, P. (2005). Context dependence, disagreement and predicates of personal taste. *Linguistics and Philosophy, 28*, 643–686.

Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature.* Cambridge: MIT Press.

MacFarlane, J. (2007). Relativism and disagreement. *Philosophical Studies, 132*, 17–31.

MacFarlane, J. (2012). Relativism. In Graff Fara, D., Russell, G. (Eds.), *The Routledge companion to the philosophy of language* (pp. 132–142). New York: Routledge.

Neale, S. (1999). Coloring and composition. In K. Murasugi & R. Stainton (Eds.), *Philosophy and linguistics* (pp. 35–82). Boulder: Westview.

Neale, S. (2004). This, that, and the other. In A. Bezuidenhout & M. Reimer (Eds.), *Descriptions and beyond* (pp. 68–182). Oxford: Oxford University Press.

Neale, S. (2007). On location. In M. O'Rourke & C. Washington (Eds.), *Situating semantics: Essays on the philosophy of John Perry* (pp. 251–393). Cambridge: MIT Press.

Nunberg, G. (1993). Indexicality and deixis. *Linguistics and Philosophy, 16*, 1–43.

Nunberg, G. (1995). Transfers of meaning. *Journal of Semantics, 12*, 109–132.

Perry, J. (1986). Thought without representation. *Proceedings of the Aristotelian Society, 60*, 137–151.

Perry, J. (2001). *Reference and reflexivity.* Stanford: CSLI Publications.

Predelli, S. (1998). I am not here now. *Analysis, 58*(2), 107–115.

Pupa, F., & Troseth, E. (2011). Syntax and interpretation. *Mind and Language, 26*, 185–209.

Recanati, F. (1987). Contextual dependence and definite descriptions. *Proceedings of the Aristotelian Society, 87*, 57–73.

Recanati, F. (1993). *Direct reference: From language to thought.* Oxford: Blackwell.

Recanati, F. (1994). Contextualism and anti-contextualism in the philosophy of language. In S. Tsohatzidis (Ed.), *Foundations of speech-act theory* (pp. 156–166). London: Routledge.

Recanati, F. (1996). Domains of discourse. *Linguistics and Philosophy, 19*, 445–475.

Recanati, F. (2002). Unarticulated constituents. *Linguistics and Philosophy, 25*, 299–345.

Recanati, F. (2004). *Literal meaning.* Cambridge: Cambridge University Press.

Recanati, F. (2007). *Perspectival thought: A plea for (moderate) relativism.* Oxford: Oxford University Press.

Recanati, F. (2010). *Truth conditional pragmatics.* Oxford: Oxford University Press.

Russell, B. (1905). On denoting. *Mind, 14*, 479–493.

Sag, I. (1981). Formal semantics and extra-linguistic context. In P. Cole (Eds.), *Syntax and semantics 13. Radical pragmatics* (pp. 273–294). New York: Academic.

Salmon, N. (1991). The pragmatic fallacy. *Philosophical Studies, 42*, 37–45.

Searle, J. (1978). Literal meaning. *Erkenntnis, 13*, 207–224.

Searle, J. (1980). The background of meaning. In J. Searle, F. Kiefer, & M. Bierwisch (Eds.), *Speech act theory and pragmatics* (pp. 221–232). Dordrecht: Reidel.

Searle, J. (1992). *The rediscovery of the mind.* Cambridge: MIT Press.

Soames, S. (2002). *Beyond rigidity: The unfinished semantic agenda of naming and necessity.* Oxford: Oxford University Press.

Sperber, D., Deirdre W. (1986)/(1995). *Relevance: communication and cognition*. Oxford: Blackwell.

Sperber, D. (2002). In defense of massive modularity. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in Honor of Jacques Mehler* (pp. 47–57). Cambridge: MIT Press.

Stainton, R. J. (2006). *Words and thoughts: subsentences, ellipsis, and the philosophy of language*. Oxford: Oxford University Press.

Stalnaker, R. (1999). Indicative conditionals. In *Context and content*, Robert Stalnaker (pp. 63–77). Oxford: OUP.

Stanley, J. (2000). Context and logical form. *Linguistics and Philosophy, 23*, 391–434.

Stanley, J. (2002). Making it articulated. *Mind and Language, 17*, 149–168.

Stanley, J. (2005). Semantics in context. In G. Preyer & G. Peter (Eds.), *Contextualism in philosophy: Knowledge, meaning, and truth* (pp. 221–254). Oxford: Oxford University Press.

Stanley, J. (2007). *Language in context*. Oxford: Oxford University Press.

Stanley, J., & Szabó, Z. G. (2000). On quantifier domain restriction. *Mind and Language, 15*, 219–261.

Stephenson, T. (2007). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy, 30*, 487–525.

Stern, J. (2000). *Metaphor in context*. Cambridge: MIT Press.

Stern, J. (2006). Metaphor, literal, and literalism. *Mind and Language, 21*, 243–279.

Szabó, Z. G. (2006). Sensitivity training. *Mind and Language, 21*, 31–38.

von Fintel, K., & Gillies, A. (2011). Might made right. In E. Andy & W. Brian (Eds.), *Epistemic modality* (pp. 108–130). Oxford: Oxford University Press.

Wedgwood, D. (2007). Shared assumptions: semantic minimalism and relevance theory. *Journal of Linguistics, 43*, 647–681.

Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Advances in pragmatics* (pp. 230–260). Basingstoke: Palgrave Macmillan.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua, 87*, 53–76.

Explicatures are NOT Cancellable

Alessandro Capone

Abstract In this chapter I argue that explicatures are not cancellable on theoretical grounds. I take that explicatures are loci of pragmatic intrusion, where pragmatics mimics semantics. I attempt to differentiate explicatures from conversational implicatures on logical grounds. I answer some objections to Capone (2009) by Seymour (2010) and I also respond to Carston (2010). The crucial problem addressed in this paper is whether by cancellability of explicatures we should intend the evaporation of an explicature from an act of saying when a different context is considered. I discuss the logical problems which this view gives rise to. In this paper, I explore the consequences of considering cancellability of an explicature a language game. I conclude that the cancellability test proposed by Carston can never be unified with the other side of cancellability (explicit cancellability cannot be unified with cancellability due to an aspect of the context that cancels the inference). Furthermore, I consider that cancellability à la Carston is neither a definitional, nor a constructive nor an eliminative language game. The paper makes use of important considerations by Burton-Roberts (2013) on intentionality and also discusses some of his examples.

1 Introduction

In this chapter, I will not embark on the task of unifying various considerations on the cancellability of explicatures (or rather the lack of it) scattered in my papers on pragmatics and modularity of mind, attributive/referential and quotation. Since here I mainly want to deal with a high level of abstraction, I will not consider those data in detail (but I need to say that they appear 'prima facie' to support my own

A. Capone (✉)
University of Messina/Palermo, Messina, Italy
e-mail: alessandro.capone@istruzione.it

inclination to say that explicatures are NOT cancellable). I confine myself to theoretical considerations which are in line with those data. The positive position I will explore, support and justify is that explicatures are natural loci of the tension between semantics and pragmatics, where the tension is resolved in favor of pragmatics but the cost involved is that pragmatics becomes more and more semanticised. And this may mean that explicatures should not be cancellable if they constitute loci of the tension between semantics and pragmatics and loci where pragmatics simply aims to mimic the semantic resources of the language, that is its truth-conditional apparatus. I have already hinted at this in my paper 'On Grice's circle', even if the aim of that paper was to resolve a specific problem (the circularity of the view that explicatures take input from pragmatics and implicatures take input from explicatures), and not to address the general problem of how language mobilizes resources of a pragmatic nature to mimic semantic resources (and to amplify them).

2 A First Distinction

While I was attempting to publish Capone (2009), a referee suggested that I should distinguish between the pragmatic components of the explicature and its semantic components. Presumably, the referee thought that there was a semantic basis on which the explicature was built (the output of decoding) and that these parts ought to be distinguished, presumably because the pragmatic components of explicatures are cancellable (or are considered to be cancellable), whereas the semantic components (the entailments) are not or should not be.[1] Of course, it should be added that theorist believed that entailments cannot be cancelled without contradiction of what is said and that explicatures (or rather, conversational implicatures) are cancellable without contradiction of what is said. Yet, at some point it appeared to me that it is clear that contradiction is itself a logical notion that (possibly) requires pragmatic intrusion, because if you do not fix the references of the words and if you do not clarify that you are talking about a serious assertion, in which a speaker is committed to the explicatures, there can be no contradiction of what is said,

---

[1] An objection to this view could be the following: the explicated proposition is pragmatically inferred by the hearer. And doesn't that mean that, however the hearer recovers them, the "components" of an explicated proposition are entailments of that proposition? In other words, it's difficult to know what I mean by "pragmatic versus semantic components of the explicature." An explicature is a proposition and propositions are constituted by their entailments.

My reply is that, while obviously it is true that the explicated proposition is part of the truth-conditional content of the sentence, relevance theorists might insist that there is a semantic component (the output of linguistic decoding) and a component which is pragmatically inferred. [See Capone (2009) on this]. I agree with the objection that both components are subject to pragmatic processing and thus even if initially it makes sense to distinguish or want to distinguish between the semantic and the pragmatic components of the explicature, in the end it does n ot really make sense to make this distinction.

simply because we have not settled on the question of how to elucidate what is said properly. I think all this has been clear enough since some eminent statements by Levinson (1983), perhaps now ignored by the current theorists. Levinson, to explain deixis, made clear that a very simple deduction does not work unless we fix the references of the terms involved. So deduction occurs at the level of statements, not of sentences. Does this mean that entailments are things that belong to statements, rather than to sentences? Not necessarily. It is obvious that if I say that 'Every man is clever', then a, b, c, d which make up the domain of the quantifier are all (and each) clever. If I want to test the entailment and to do so I need the notion of contradiction, I need to test the entailment through an assertion. However, this is natural, since the aim of a semantic theory is to provide the resources necessary to make statements and it is through statements that we can test the logical properties of words (entailments) provided that we are able to separate what belongs to semantics from what belongs to pragmatics. To make an example:

> I say:
> This man is clever and happy.
> Then I say:
> This man is clever but unhappy.

If I were to unify those two thoughts, I would obtain a contradiction PRO-VIDED that I keep the reference of 'this' fixed. Although I have tested the entailments of my first sentence through a statement, I have obtained intuitions about semantics, since the pragmatics was kept fixed: in particular, I kept fixed the reference(s) of 'this man' and I have made a serious assertion in both statements.

So now, to go back to the issue of keeping separate the semantic and the pragmatic components of the explicature, I should at least say that the entailments that form part of the explicature must go through a pragmatic filter, as they must be judged part of a serious (or otherwise not serious) assertion. Thus it is not so clear that it makes sense to distinguish between the semantic and the pragmatic components of the explicature. In fact, if we consider the explicature a contribution to what is said, to the proposition which a speaker is committed to, we had better not make this distinction at all, since the entailments work only through the statement (they are active only if expressed through a statement). Considerations on cancellability also militate against making such a distinction, since the entailments, once they go through the pragmatic filter which makes them eligible for a serious assertion, cannot be cancelled in the same way in which the pragmatic components of the explicature cannot be cancelled. In fact, as everyone knows since the work by Hintikka on knowledge, entailments can be cancelled, in the context of a loose assertion, as when one says 'John knows that Mary is in Paris but she is in London'. Here the entailment of 'know' is cancelled in favor of the interpretation 'believes he knows'. Explicatures, instead, are loci where entailments cannot be cancelled at all—and that is because they are fixed through intentions and intentions exclude loose uses if the speakers do not intend to speak loosely.

3 On Intentions

A man who speaks intends the hearer to grasp the message she intends to convey through the use of words and syntactic concatenations of linguistic elements as shaped by knowledge of the language and by pragmatic principles (or principles of language use). A linguistic action is different from a non-linguistic one in that it is not only animated by an intention but makes that intention explicit through a linguistic form. So, when a man switches the light on, he is animated by the intention of switching the light on. The action reveals that intention in a non-linguistic way. Given the action, it is obvious enough that that is what he intended to do (unless he wanted to deceive us). However, given a linguistic action, it is not obvious enough what its point is, as that may but may not be revealed by the words used (serious vs. non-serious uses). Furthermore, given that through an utterance one can accomplish many actions, even if we were to find a simple way to detect the correlate intention, we could not easily find out the other intentions which are linked to the utterance.

But at least this must be clear. The speaker said u for some reason and by saying u he had the intention of doing x. So, the main task for the language users is how to grasp the intention which is behind the utterance (sometimes hidden by the utterance). So far we have been presupposing, perhaps simplistically, that intentions are a priori and fixed through utterances. And this is what some philosopher imbued with anthropology (say Duranti 1988) may want to deny, favoring the idea that intentions emerge from interactions and that hearers are instrumental in fixing such intentions. I do not deny that there are such complex cases. When a novelist discusses his novel with his readers he may very well come up with interesting remarks on the authorial intentions and he may even accept such suggestions. There is also the case of the academic writer who discusses an article with her editor who is able to maieutically extract what the author really wanted to say and helps her put that into writing. There is also the case of an intention which is 'in fieri' and, which, therefore, is likely to be modified by interaction with an audience. However, it is undeniable that there are also cases of simple *a priori* intentions, as when Mario asks his mother whether dinner is ready. It is clear enough that he intends to eat, in the context of utterance, and there is no cogent reason for doubting that he had such an intention in saying what he said. So, in this paper, I will assume that what I have to say is only applicable to the simple cases, while I accept that the special cases need deeper discussion.

And now the Deus-ex-Machina of this paper. How can an intention be executed/implemented and then be cancelled? Surely, a speaker can retract an intention if he repents saying what he has said and is willing to replace it with a different assertion. However, the very fact that the intention must be retracted means that the intention is still there, behind the previous utterance [See Burton-Roberts (2005), 2013]. Consider the case of the politician who says, in the middle of a conversation, 'That bloody negro….' and then stops because he remembers that there is a black person within the audience. He has committed an offence, and

although surely he can retract the previous thought, the procedure for doing so is quite laborious and it is not obvious at all that he can immediately cancel the offence he has produced. All he can do is to rely on the forgiveness of the hearers, but he need not expect that the offence can be eliminated so easily as it was produced. This example merely shows that intentions (once executed) in many cases are not easily retractable. In the easiest cases, however, one can pretend that he used the wrong word.

Communicative (or better, communicated) intentions are entirely transmitted through pragmatics. No (communicative) intentions can be fixed through the semantics of the language, although language is instrumental in fixing intentions. Even the law requires interpretation, and despite the fact that the law-maker tries to be as explicit as possible, there are residual interpretative ambiguities. In the end, the most rational interpretation of the law is the one which wins (Dascal 2003), but we still need interpretation, which shows that executions of intentions are pragmatic things.

Linguistically expressed intentions require a matching between what is understood and what is said. Without this matching, there can be no communication, at least ideally. According to some theorists, it is sufficient that the message understood and the message conveyed are similar enough; a strict matching is not indispensable (Wilson and Sperber 2012). I quite agree that a certain degree of approximation should be tolerated in actual communication; however, ideally communication cannot be successful unless there is a perfect match between the speaker's intentions and the message recovered by the hearer. A sound linguistic methodology will prescribe that we should not be happy unless the communication processes described by our linguistic theoretical apparatus capture this match. An ideal pragmatic theory is not one that solely deals with interpretation, but one that deals with the way intentions are communicated. The same predispositions to communicate information should work both at the level of codification and at the level of interpretation. Take for example the principle of Relevance by Sperber and Wilson (1986). According to this principle, a speaker communicates by an ostensive act a presumption of Relevance. It follows that this Presumption of Relevance should also guide interpretation. Interpretation is mainly a reconstruction of the speaker's communicative strategies. There is even a mirroring relationship between what the hearer does to understand and what the speaker does to communicate. The speaker takes into account the hearer, her needs and limits, and the hearer takes into account the predispositions of the speaker to take into account the dimension of the hearer.

4 Explicatures

I have said that explicatures are loci of the tension between semantics and pragmatics. In particular, they are loci where pragmatic inferences are hard or impossible to cancel. The reasons why they are hard to cancel may be multiple.

We may go along with Burton-Roberts (2005; 2013) and say that explicated meanings are cases of speaker's meaning, that is to say cases in which a speaker commits himself to his meaning (what he said). Since he intended to say something, such a message cannot be un-said and the intention behind it cannot be cancelled. Explicature cancellation, in Burton-Roberts's view, amounts to aborting an (executed) intention, but how can an (executed) intention be aborted if it was there in the first place (that is if it was already executed)? Another cogent reason why an explicature cannot be cancelled (or aborted) has to do with the logical structure of discourse. If an explicature is there to play a role in the logical structure of discourse, in particular in rescuing a fragment of discourse from illogicality, contradiction, and logical absurdity, then such an explicature cannot be aborted, because this would amount to returning to the problems which, in the first place, necessitated the explicature. We can, derivatively, couch this notion in Burton-Roberts' notion that intentions cannot be abrogated, provided that we are clear at this point that it's not individual intentions—arrived at through specific clues disseminated in the text—that count in this case, but the intentions that are derived through the desire to say something logical—and not illogical. So, at the basis of explicatures, we can find the general intention to be logical, from which other individual and concrete intentions can be derived. We are obviously faced with Jaszczolt's (1999) distinction between the individual and the social path of intentionality, where individual intentions have to conform to what must be the case in order to preserve the logical structure of discourse. So, the difference between Burton-Roberts and me, although minimal, is not trivial and is worth being discussed. The other difference between Burton-Roberts and myself is that I said that explicatures are loci of the tension between semantics and pragmatics. When linguistic decoding cannot independently provide a logical structure of discourse, pragmatics has to intervene and must provide a truth-conditional intrusion. In other words, there is a gap in truth-conditional meaning which is due to the insufficiency of semantics and this gap must be filled through pragmatics. There is a tension because pragmatics intervenes to fill the truth-conditional lacuna, and also because pragmatics becomes attracted by semantics and ends up playing the role of a substitute, which has at least some of the properties of the thing for which it is substituted. One of such properties is non-cancellability. So, if one side of the story on cancellability of explicatures depends on intentions, the other half of the picture depends on the structural role played by pragmatics and, in particular, by the exigency of replacing semantics and of mimicking at least some of its properties.

Before delving into the theoretical part of the paper, it may be convenient to provide and briefly discuss some examples of explicatures. The leading idea of explicature is that pragmatic intrusion contributes to the truth-conditional import of the statement (thus, it contributes to what is said). Important scholars like Carston (2002) and Levinson (2000), therefore, have been busy to show that, without pragmatic intrusion, it is not possible to calculate the full truth-conditional import of a statement. Consider the following:

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.

Semantically, the two conjuncts (conjoined by 'but') appear to be the same; but then the contrast would not be justified; however, there is a genuine contrast if pragmatic intrusion applies and the pragmatics of 'and' contributes to the full truth-conditional import of the utterance.

2. It is better to meet the love of one's life and get married than to get married and to meet the love of one's life.

One of the requirements of the comparative 'better' is that the things compared be different. Thus, we expect that the propositions compared are different. At the sentential level, however, they are the same propositions, thus we need pragmatics to arrive at the full propositional level (where temporal variables are assigned values).

3. If the children eat some of the cake, then we will eat the remainder.

Now consider the quantified expression 'some cake'. Unless a scalar implicature applies to it and it is interpreted as 'some but not all of the cake', the conditional does not work properly, as the consequent follows ONLY if the scalar implicature is actually computed and becomes part of truth-conditional meaning.

4. A: Why don't you join us for dinner?

B: Thanks, but I have already eaten.

In this example, B is clearly explicating that she has already eaten dinner; it is not enough that she has eaten, say, an ice-cream. The reply counts as an explanation for the tacit refusal. The speaker cannot accept because she has already had dinner and one cannot have dinner twice in the same day. To make the reply relevant, it is not enough that the speaker had dinner at some time in the past, but the explanation to be relevant has to be about a time interval immediately preceding the time of the invitation to dinner.

In all these cases, it does not make sense to cancel the explicature, because by cancelling it one returns to a discourse which is pointless; if an explicature is needed to cure potential contradictions or absurd speeches or the provision of trivial information, then by cancelling the explicature one returns to problems which cannot be remedied otherwise. Consider now the following.

5. You will not die (of this cut).

The mother who says 'You will not die' to her son, does not obviously mean that the child will never die, but that he will not die due to his cut. The contextual provision of an adjunct serves to make the truth-conditions of the utterance more precise.

The work on explicatures does not end here. In previous work, I have in fact shown that explicatures play an important role in the following areas: belief reports (null appositives), 'de se' attitudes, Immunity to Error through Misidentification, knowing how, quotation, referential/attributive, indirect reports, pronominal clitics, etc.

5 On the Tension Between Semantics and Pragmatics

I take a semantic theory to be a system of knowledge allowing people to communicate by using signs and structured strings of signs in order to express thoughts. A structured string of signs corresponds to a thought (I take ambiguities to be related to utterances capable of being given different structural analyses). We could very well think of thoughts as minimal structures capable of being enriched through further layers of meaning through pragmatics. Yet the minimal structure must be there for enrichment to occur. These minimal structures can be assigned basic truth-conditional meaning. One reason why I take there to be a basic tension between semantics and pragmatics is that while an undoubtedly complex semantics has been devised to deal with recurrent and culturally salient aspects of reality, this cannot completely deal with new aspects of reality which require some kind of pragmatic adaptation, or extension of the semantic system. Pragmatics serves to boost and amplify the semiotic potentialities of the system; needless to say, if a construction tends to be associated through pragmatics with a certain meaning, and such an association becomes recurrent and ends up capturing an aspect of reality which, for some reason, has now become culturally salient, then there are chances that the explicature will become semanticised through various stages of language use. A stage in which the use is relatively unstable will be followed by a stage in which the use becomes stable enough in that it has come to represent the needs of a multiplicity of users who, faced with a recurrent problem, have found a certain construction and its pragmatic explicatures useful to express a recurrent aspect of reality. Only when there is a convergence between the needs of a multiplicity of language users and the potential benefit that a construction represents in that it is capable of resolving a recurrent expressive problem, does the need for grammaticalization arise. Consider the following Searlian Principle:

Anything that can be thought can be expressed.

This principle embodies the basic tension between semantics and pragmatics, since when there is an expressive problem arising due to the fact that the semiotic resources of the language are not capable of coping with a certain area of language use, then pragmatics allows expressibility. However, I would even add that pragmatics allows thought in the absence of adequate semantic resources; so it is also an amplifier of thought, a means through which thought is capable of existing, of being articulated, of being developed through more complex structures than those that are allowed through the existing semiotic system. In other words, pragmatics is a basic tool or utility which makes thought more flexible and more complex, thus ensuring progress in those cases where the limits of semantics would invariably mean stagnation. A certain amount of creativity is introduced into language though pragmatics, which does not only boost thought and the existing semantic resources, but ensures that the evolution of thought can take place even in the absence of new linguistic resources or of 'ad hoc' creations.

6 On the Distinction Between Conversational Implicatures and Explicatures

A sound and coherent picture would be one according to which explicatures are calculated before implicatures. The rationale for this is that explicatures contribute to truth-conditional meaning while conversational implicatures can be cancelled (are cancellable in principle). In any case, conversational implicatures are normally calculated after truth-conditional meaning is calculated. It is not impossible that implicatures and explicatures can be calculated at the same time and that implicatures help determining the explicature. Even granting this logical possibility, these cases are rare. The cases of explicatures I have discussed are confined to those where explicature comes to rescue the discourse from a defect, such as illogicality, contradiction, triviality, etc. It is in such cases that it is hard to cancel the explicature. Considerations of parsimony also militate against the idea that explicatures can be cancelled, because once the cost of pragmatic inference has been incurred, some extra cognitive cost will be required to cancel the explicature. But this extra cost is not generally justified. Instead the cost of pragmatic inference in the case of explicature is justified by the need of liberating discourse from some obvious defect.

So, what's the difference between an explicature and an implicature? They are obviously generated by the same pragmatic principles and they are both generated when the discourse seems defective for some reason. In the case of conversational implicatures, there is often a defect in the flow of information and to restore the balance of the flow of information an implicature is needed. In the case of explicatures, there is a problem with the logicality of the discourse and one needs an explicature to liberate it from e.g. some obvious contradiction or absurdity. So, in any case both the implicature and the explicature can potentially deal with problems, but the problems cured by explicatures are more acute and are not confined to lack of relevance or lack of information on the part of the speaker.

The consequences of this preliminary discussion on cancellability are obvious. Explicatures are obviously not cancellable, because by cancelling them one returns to a severely ill-formed fragment of discourse. Conversational implicatures are cancellable in the sense that one can retract the intention behind them more easily. Consider the following case:

6. A: We should get rid of Berlusconi.

Properly contextualized, A's utterance could be taken as an invitation to get rid of Berlusconi physically. And in recent political discourse, there has been a controversy on whether this type of language counts as an incitement to violence. The implicature, however, could easily be denied by the speaker, who might simply say that he was speaking metaphorically (get rid of Berlusconi from the political scene).[2] Since there is a residual vagueness, intentions of this type can be

---

[2] A commentator doubted that these inferences could ever arise. Notice, however, the analogy with the utterance allegedly proffered by Henry II "Who will rid me of this meddlesome priest?"

easily retracted. This is not to say that in all cases of conversational implicatures, these can be cancelled. If there are numerous clues all leading in the direction of fixing the speaker's intention, then it will be difficult to cancel the implicature, because the intention is expressed in the form of a strong implicature. It may make sense to distinguish (as Wilson and Sperber 2012 do) between strong and weak implicatures. It is obvious that the stronger implicatures are hard to cancel.

But then, could we not count the explicatures as cases of strong implicatures? Considerations of parsimony would lead us to get rid of the cumbersome distinction between implicatures and explicatures. This is more or less what Levinson (2000) does, even if he talks of intrusive constructions. Yet, I would resist the idea of conflating explicatures and implicatures, because while surely strong intentions are present in the case of explicatures, it is structural configurations which make the intentions stronger. While in the case of strong implicatures, one might say that the implicatures are stronger because the speaker disseminated such an amount of clues in the text as to make cancellation difficult or impossible and strong intentionality depends on the speaker's intention to make his intention evident, in the case of explicatures, it is the structural configurations of discourse rather than the amount of clues disseminated which make the intention stronger and difficult to cancel.

7 The Pragmatic Cancellation Principle

According to Carston (2002, 138) all pragmatic inference is cancellable. Since explicatures are cases of pragmatic inference, it would follow that they are cancellable too. Now, these apparently innocent remarks require investigation and proper deepening. It may seem obvious that many cases of explicatures involve cases of Gricean scalar implicatures or in any case of Gricean generalized implicatures. For example, use of the connective 'and' may give rise to temporal readings (and then) or even causal readings (and therefore); and the use of the quantifier 'some' may give rise to interpretations such as 'some but not all'. So according to Carston, Grice's GCIs can be analysed as explicatures. Surely we must grant that at least some explicatures are arrived at through pragmatic inferences and, in particular, generalized implicatures. Yet, we must also recognize that explicatures are loci of the tension between semantics and pragmatics and that, even if generalized implicatures are utilized to fix an intention, at these loci of pragmatic intrusion there is more at stake than conversational implicatures. Conversational implicatures are only one ingredient of explicatures; then these

---

(Footnote 2 continued)

which were heard by some of his knights as an incitement to violence against Thomes Beckett. Of course, the political context is different and will yield different implicatures. In the Italian political scene, the incitement to violence interpretation is a bit strained, but certainly the supporters of Berlusconi argued that utterances like (6) could be interpreted in this way.

must be combined with the output of linguistic decoding, and in the case of explicatures a particular mode of pragmatic composition prevails, one that uses pragmatic inference to cure a logical problem. Thus structural considerations prevail and the conversational implicature is only a tool to be used in a complex machinery that throws out entailments rather than implicatures. So we may say that implicatures are part of the input, but the compositional machinery turns pragmatic inference into semantic inference. Thus, it is true that pragmatic inference is cancellable (e.g. potential explicatures), but once it is turned into semantic inference, it is no longer cancellable. It is the recognition of the loci of pragmatic intrusion or of the tension between semantics and pragmatics that makes Carston's Pragmatic cancellation inference quite irrelevant with respect to cancellability of explicatures. Since explicatures are logically different from implicatures, even if explicatures are made out of implicatures, Carston's Pragmatic Cancellation principle no longer applies.

There are further reasons for believing that Carston's Pragmatic Cancellation Principle is innocuous. To have full validity and generality, one should be able to contrast linguistic decoding and pragmatic information. (See Burton-Roberts 2013). Presumably, on a view such as Carston's, pragmatic inference is cancellable, while semantic inference is not. Yet, as shown in Capone (2009), the entailments that constitute the semantic layers of the explicatures also need pragmatic intrusion to rise to the level of intended meaning.[3] Only when they rise to the level of intended meaning, they are no longer cancellable. Otherwise, as insisted on by theorists such as Kent Bach (2001), the entailments are neither here nor there. We can easily suspend them or cancel them, as shown by numerous cases of ironic utterances. So, in the same way as we can distinguish between weak and strong implicatures, we can distinguish between weak and strong entailments. And it appears that entailments are strong, in the sense of not being cancellable, only when they rise to the level of speaker's intentions. In other words, it is the speaker's intentions that determine that the entailments cannot be cancelled. But if such are our conclusions, there are no strong or cogent reasons for distinguishing between semantic and pragmatic inference in terms of cancellability—while of course there are other reasons to ground the distinction, mainly having to do with compositionality.

8 Cancellability as a test for Conversational Implicature

The literature on conversational implicature has converged on the idea of using cancellability for testing conversational implicatures (in particular, for distinguishing them from entailments). All textbooks agree that cancellability is the most

---

[3] Perhaps the most clear case of cancellability of entailments is constituted by Hintikka's consideration that the entailments of 'know' can be cancelled, as in loose or parasitic uses such as 'John knew that p, but it turned out that p was false'.

important test for conversational implicature, in the sense that, if something is an implicature, then it should be cancellable (without contradiction of what is said). However, Sadock's (1978) seminal paper, at least some scholars have insinuated a doubt as to whether cancellability can be used as a(n automatic) test. Levinson (1983) despite his faith in cancellability agrees that as an autonomous test, it can sometimes fail and that cancellability must be used together with other tests (non-detachability, non-conventionality, the ability of constructing an argument that posits the implicature). The fact that we need cumulative testing shows that cancellability, after all, does not work properly as a test. And why should we need a test after all? Why should we need a diagnostics? We need testing when we are not sure; but if we are sure that something is an inference, we do not need to test it.

If this were not enough, two delicate problems besiege the theory based on cancellability as a test (or as a way of testing). There are cases of meanings which start as pragmatic and end up becoming semantic. They are cases of frozen pragmatics. It would be best to say that these are cases in which the tension between semantics and pragmatics on the one hand has used pragmatics to extend prior senses, on the other hand usage has consolidated the pragmatic innovations, accepting them as part of the praxis because of their usefulness and because of the communicative success of the innovation—a majority of speakers have felt that the innovation has been useful and has filled a hole in the system, has provided something for which there was an acute need. A use becomes consolidated when it provides a concept that is culturally salient enough. Of course, for such consolidated uses, cancellability as a test does not work—not only for the easy reason that one is mixing synchrony and diachrony, but because usage has, so to say, invested an inference with the approval of the community and has thus circulated the inference as a culturally salient use, rather than as an 'ad hoc' creation or innovation.

The other problem, which is more acute, is that even assuming that conversational implicatures are naturally—even if not uncontroversially—cancellable, we have no certainty that explicatures are cancellable. If they were to be completely identified with implicatures, then by identification, we would expect them to be cancellable. However, if there is a complex relationship, which is not necessarily one of identification, between implicatures and explicatures, then we should not expect explicatures to be cancellable on a par with implicatures.

In fact, if it is natural to say, to posit, or to argue that if there is a test for conversational implicatures, this should include cancellability (however controversial that test should be), it is not natural to argue that cancellability is a test or diagnostics for explicatures,[4] because we have seen that explicatures arise in loci of the tension between semantics and pragmatics, where pragmatics becomes a substitute for semantics and provides full truth-conditional meaning. Thus, it would be natural to expect that indeed non-cancellability should be a test or diagnostics for explicature. After all, explicatures are cases where the speaker's

---

[4] This cannot be Carston's view since she also believes that implicatures are cancellable.

intentions cannot be different from those that help rescuing a certain discourse configuration from implausibility and where the speaker's intentions render logical something which is or should be 'prima facie' illogical (contradictory or absurd). If intentions proceed along the social path of intentionality (Jaszcolt 1999), then the social intentionality would make them hard to cancel or uncancellable, because cancelling them would amount to proceeding along the path of individual intentionality, as opposed to social intentionality. The social path of intentionality ensures that an inference is not cancellable, because cancelling it would result into something which is not socially acceptable. Contradiction, in general, or absurdity is not socially acceptable, thus the intentions that promote contradictory or absurd readings cannot be tolerated.

9 Sliding from Generalized Implicatures to Explicatures

Burton-Roberts (2013) believes that the reason why Carston (2002) sticks to the idea that explicatures are cancellable is that, after all, she believes that certain conversational implicatures can be analysed as explicatures and, therefore, ends up arguing that generalized implicatures are 'ipso facto' explicatures, presumably because they can be embedded. Consider the following case:

7. If the children eat some of the cake, we will eat the remainder.

Presumably, the conditional makes sense on the understanding that the children will eat part of the cake and NOT all of it; only in this case, in fact, can the adults eat the remainder. There is no remainder if the children eat all of it. I agree that conditional constructions are loci of the tension between semantics and pragmatics, where pragmatic inferences become semanticised and can no longer be cancelled. However, I do not think that scalar implicatures (in general) are 'ipso facto' 'explicatures'. Consider, in fact, the following case:

8. I hope some students will come (to the class).

The professor who hopes that some students will come to the class, may be open to the idea that if all the students come, that is even better. Certainly, he does not hope that some students will not come, even if he may believe that some are not likely to come. A case like the one above is enough to show that in certain contexts, the implicature (potential, in fact) does not get through. It is well known that professors want their classes to be full and it is certainly not appropriate behavior for a teacher to hope that certain students will not come. In this scenario, the generalized implicature does not get through. Thus to say that conversational implicatures can be analysed as explicatures is not correct (alternatively, the claim must be qualified further to avoid generalization), as the relationship between implicature and explicature is a complex one. An explicature requires an implicature, but it also requires a locus of tension between semantics and pragmatics. Conversational implicatures do not require such loci of tension between semantics

and pragmatics. In easier terms, we could say that if there is an explicature, then there is an implicature; but if there is an implicature, there may or may not be an explicature. In other words, the implicature, to be promoted to an explicature, needs to be associated with an intrusive construction in the sense of Levinson (2000). The term 'intrusive' construction has been avoided by Relevance theorists presumably because it may indicate that the contribution of pragmatics to semantics is sporadic and not systematic. Instead, Relevance Theorists believe that semantics is radically underdetermined and that pragmatics is needed to arrive at full truth-conditional meaning. Making use of the term 'intrusive constructions' does not, however, amount to denying that the contribution of pragmatics to semantics is systematic (even if we may accept that it is more sporadic than claimed by Relevance Theorists). Intrusive constructions are pretty systematic and to recognize them (or their types) amounts to accepting that the role played by pragmatics in complementing and integrating semantics is systematic. In fact, work must be done in recognizing all types of possible intrusive constructions. To say, in a rather general manner, that the output of linguistic decoding is totally underdetermined amounts to allocating a role to pragmatics which competes with linguistic decoding; so much so that it does not make sense to start with linguistic decoding at all. Pragmatics could then very well take over. Recognizing that there is a tension between semantics and pragmatics amounts to recognizing the foundational role of semantics, which constitutes the first type of semiotic layer, and then to admitting that in certain cases, where semantics is not sufficient, pragmatics takes over. Furthermore, one also recognizes that semantics, to work, must be embedded in a pragmatic layer that allows it to work, by ensuring that speakers' intentions are serious as opposed to non-serious ones. Pragmatics constructs a certain path in which semantics can work (Higginbotham, p.c.) and the loci of tension between semantics and pragmatics are presumably the pragmatic scaffolding which is needed so that semantics can work properly. But now, if we assume that pragmatics is a sort of scaffolding on which semantics works properly, why should we take this sort of pragmatics to be cancellable? If it is pragmatics that ensures that a certain string of words has to be taken seriously, rather than say ironically or metaphorically, why should we think that pragmatics should be cancellable? The structural role played by pragmatics in doing the scaffolding is not compatible with the idea that pragmatic inference is cancellable, even if we are open, of course, to the idea that some pragmatic inference is cancellable (e.g. potential implicatures in the sense of Gazdar).

10 An Escape Route: Seymour Against Capone (2009)

One might argue against my tack on explicatures something along the lines of Seymour (2010):

Capone (2009) has argued recently that some particularized conversational implicatures were not cancellable, 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 looks as though sentence meaning were 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 become quite positive. So the fact that an implicature cannot be cancelled from a particular context of utterance does not imply that it is not cancelable. Cancelability should suppose the consideration of different contexts of use. The fact that a particular implicature cannot be cancelled from a particular context of use is compatible with its cancellability within a different context 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 conversational implicatures are cancelable. (Seymour 2010, 2871).

Notice, for the time being, that Seymour's considerations apply to implicatures (in fact, particularized ones), and not necessarily to explicatures. However, since we believe that, generally speaking, explicatures imply or require pragmatic processing, such considerations are against my general apparatus concerning conversational explicatures. I will later examine an objection by Carston to Burton-Roberts on cancellability of explicatures which is analogous to this one by Seymour. Summing up Seymour's argument, particularized conversational implicatures are 'prima facie' not explicitly cancellable, however since the very act of saying could be proffered in a different context (promoting or eliminating the possibility of such an implicature) they are contextually cancellable (which means that, in a different context, the same implicature would not arise). Presumably Seymour is writing of evaporation of explicatures, as opposed to explicit cancellation (without contradiction of what is said). The example provided by Seymour is that of Grice's reference letter in which a professor praises a candidate's handwriting without saying much about the candidate teaching abilities. That letter is clearly negative, but if the context was one in which the candidate applies for a different job, the letter might very well be positive. This I do not deny, of course. And of course, Seymour's considerations are stimulating and worth replying to. If we take Seymour seriously, it is an act of saying which, in a given context, gives rise to an implicature and the implicature might be different, depending on the context. Does the fact that the implicature might be different depending on the context amount to saying that the implicature (whatever it is) is cancellable? To cancel an implicature, the minimum we require is an act of saying and a context and the speaker's intentions. However, since the same act of saying might give rise to a different implicature in a different context, we cannot say that a different context or the implicature that arises there can cancel (or contradict) the

implicature we wanted to test with respect to cancellability. Cancellability would minimally require the implicature generated by the new context to contradict the implicature generated in the previous one. Consider the case of the handwriting reference letter. The fact that the professor in a different context might be taken to praise the student in support of whom he is writing cannot be said to cancel the negative implicature generated in the context in which the letter was intended to support a candidate for professorships because, in this other (more positive) context, there was no such implicature in the first place. There can be no contradiction between supporting a student for a secretarial job and not supporting her for an academic job. Cancellability requires contradicting a previous assumption—whether an implicature or an explicature. So, the contextual variation which Seymor has in mind does not reach the status of cancellability.

Most importantly, if we were to take Seymour's considerations seriously, we could very well model implicature contextual cancellation after deixis. In deixis too, an act of saying has different meanings in different contexts. Should we say that while we are in one context, the meaning which a deictic expression such as a pronominal (e.g. 'This man') has in another context is cancelled or cancellable? Surely nobody has proposed so far such a view of deixis—and this is compatible with the view that contextual variability is a way of cancelling possible meanings, but not of cancelling actual meanings. Now, if conversational implicatures follow the model of deixis, we could say that, since the implicatures given rise to by an act of saying are infinite, any act of saying in a definite context involves the cancellation of infinite (or a high number of) conversational implicatures.[5] This is the unpalatable consequence of Seymour's embracing of a contextual view of cancellation—and this is the obvious consequence of considering implicature cancellation not as a process relating to certain definite intentions, but as a process relating to possible intentions. If we accept Seymour's view, we would have to accept that cancelling an implicature is a trivial thing, since at the same time we would have to cancel many other possible implicatures related to the same act of saying, except for one implicature which would arise if a definite context were chosen.

Now while conversational implicatures and interpretations of deictic expressions may have a number of things in common, they are different in the way the intentions are fixed and they are also different because deictic expressions are merely related to referents, whereas conversational implicatures serve to convey full-fledged thoughts.

---

[5] Huang (2007) considers deixis fixing as a case of pragmatic intrusion. Yet this does not automatically amount to accepting that deixis fixing is determined through conversational implicature. Deixis fixing looks more like a semantic/pragmatic phenomenon determined by conventions of use. If I utter 'Today I am going to give a talk at Oxford university' I am expected by the audience to fix the date of the lecture by the day of the utterance event through a rule of usage. This is not a conversational implicature.

11 Another Escape Route: Carston (2010) against Burton-Roberts (2005)

As I said, similar arguments could be used to argue that explicatures are indeed cancellable. Carston, like Seymour, claims that we must rely not on explicit cancellation but on contextual variability. I will keep my reply short, because my considerations against Seymour are the same I can use against Carston. If we only rely on contextual variation, we are not capable of distinguishing between cases of conversational implicatures and cases of deixis. Yet, these cases, despite similarities, arguably should be kept separate. (But notice that radical pragmaticists may hold that deixis fixing is exactly a pragmatic process). Most importantly, it would be useless to use cancellability as proof that a phenomenon is inferential, because such a phenomenon could be very well assimilated to a deictic inference. That might proceed along different lines, as the intentions might be fixed by a gesture (that is to say semantically), while the intentions in an implicature are never (just) fixed semantically through a demonstration, but normally through reasoning (whether compressed or not).

Consider now Carston's (2010) statement of her ideas on cancellability as put by Burton-Roberts (2013):

TH

An explicature or implicature p of a given utterance in its context C1 is CANCELLABLE if and only if either (1) it can be cancelled explicitly (i.e. by an explicit act of the speaker) in C1 or (2) there is ANY CONCEIVABLE CONTEXT—Ca—in which p would not be explicated/implicated by an utterance of the same expression.

In addition to my own considerations, there are other reasons for believing that TH is dubious. Consider the following examples taken from Burton-Roberts' most illuminating work (Burton-Roberts, 2013):

9. Bill: Have you read any Proust?

   Anne: Yes.

10. Bill: Have you booked a table?

    Anne: Yes.

The implicature of (9)'s 'Yes' is that Anne has at least read some Proust; instead, the implicature of (10)'s 'Yes' is that Anne has booked a table.

Following Carston's considerations, we could claim that the explicature in (9) is cancellable, because, in fact, the same act of saying, in a different context, does not trigger the same explicature. As I said, such a notion of cancellability is not a diagnostics of conversational implicature/explicature, because it can apply very well to deictic terms. Second, what (9) and (10) at most can show is that the same act of saying can carry different implicatures in different contexts and this is not logically related with the notion of cancellability, because it is trivial that if we add

different elements to S, we obtain different utterances. Ideally, we should be able to relate cancellability in the ordinary sense (an inference can be cancelled in a given context) with cancellability in Carston's revised sense (Cc). (I take this suggestion from Burton-Roberts 2013). The attempted unification, however, does not work, because if Cc predicts that the explicature is cancellable in (9), any attempt to cancel the explicature in the context of (9) fails (Consider how we would take a speaker who answers the question in (9) with 'Yes' and then goes on to say 'But I did not mean that I have read some Proust).

The unification instead holds for explicit cancellability and contextual cancellability (in the sense that an explicature is cancelled by some feature of the context). The details are presented in Burton-Roberts (2013).

In addition to the considerations so far, which I take to be quite cogent, I want to ask the radical question whether Cc can work as a diagnostics of conversational implicature in the sense of explaining this notion. While generalized implicatures (potential implicatures) are cancellable in that their putative nature is put to the test by a given context, which may promote or otherwise cancel an implicature, the Cc test cannot apply to them, because in the case of generalized implicatures we do not want to know whether in one context the same act of saying promotes implicature x and in other it promotes implicature y. For potential implicatures, all we want to know is whether a context does promote or otherwise abort the implicature. Presumably Cc applies only to particularized implicatures—but whereas for generalized implicatures the diagnostics was important because it predicted that an implicature could get through or not, in the case of particularized implicatures we do not use context to cancel the implicature but to promote the implicature. Thus, the fact that a different context is able to promote a different implicature comes as no surprise and has no intuitively important theoretical weight.

A final argument against Carston's considerations on Cc (Carston's cancellability) may be the following. Carston accepts that an explicature can be cancelled by embedding the act of saying that generated the implicature in a different context. In such a context, the same act of saying no longer has the same explicature. Suppose this line of thought is entertained. Then one should also accept that, however one changes the context, the explicature is cancelled. But one could, in fact, change the context in such a way that the same act of saying still preserves the explicature. Suppose that one is patient enough to contrive a number of contexts in which the inference is preserved and a number of contexts in which the inference evaporates. Should we then say that the explicature is cancellable or not? Contextual variation, at this point, does not seem to be enough to ensure cancellability—one ought to specify those features of the context that genuinely militate against the explicature. And yet such contexts could be embedded in larger contexts that allow us to preserve the implicature/explicature of the original act of saying. At this point, since any context can be embedded in a larger context, for every context that cancels the explicature we could embed it in a larger context that preserves the explicature. Since the proponents and the opponents of the

theory of cancellability could be equally genial in enlarging the context, nobody could really win the dispute.

One further way to see that there is something faulty in Carston's argument is to translate the argument into the terminology of language games [(Wittgenstein 1953); See Carapezza and Biancini (2013) for an articulation of Wittgenstein's ideas in terms of the recent idea of pragmemes (Mey 2001)]. Language games are linguistic acts which avail themselves of context to reach their ultimate meaning. Suppose this time that cancelling an inference is a language game. To be a language game an act must follow a rule. Let us suppose that the rule required in cancelling an implicature is that the context should display elements that contradict the implicature, which are at odds with it. It makes sense to engage in the language game of cancelling an implicature in the case of generalized implicatures, because these inferences are devised in such a way that if everything goes well, they get through, otherwise they do not. Cancelling an inference is surely costly, but the cost must be offset by a congruous number of contextual effects. The language game of cancelling an implicature of the generalized type involves engaging in an act of communication in which the utterance most of the times hooks into contexts which fit it and were made for it. Cancelling the inference is thus recognizing that the context we are in does not fit the act of saying.

When particularized implicatures are concerned, instead one could play a different language game, that is one could try to change the context of the utterance and see how the same act of saying acquires a different shade of meaning (implicature). Since the contexts are potentially infinite, have we got any reason for saying that this, rather than that implicature is cancelled, when another context is encountered? What kind of language game would this be, if, after all, we have no more reason to say that this inference, rather than that inference, is cancelled? When we are dealing with generalized implicatures we know which inference is cancelled and when. But with particularized implicatures, it makes no sense to say: this inference is cancelled because in that other context another meaning accrues to another utterance of the same sentence. We could very well say the same thing of the implicature which arises in that other context. So, which implicature is cancelled? All and none, one could very well answer. And one now finally notices that this language game is impossible, because I do not know where to start and where to end the language game. Furthermore, I do not know what the purpose of the language game is. With generalized implicatures, the language game was to tell when an implicature arises and when it does not. Here we cannot say when an implicature arises and when it does not, since in different contexts different implicatures would arise. What benefits do we have in cancelling the implicature? None. There are no benefits to anyone. In fact, since the things being compared are different, it is impossible to say that one implicature cancels the potential which the sentence in another context would have of generating a certain implicature. In fact, the embedding of a sentence in a context does not at all interfere with the way the same sentence would behave in another context and with the implicature it would trigger. The language game we are embarking on is neither definitional, since we have already said that in this way we cannot distinguish implicature from

deixis, nor constructional, as we are not constructing anything at all. The language game, furthermore, is not even eliminative, since by saying that an act of saying has a certain implicature in a certain context, I am not eliminating completely the possibility that the same act of saying carry the same implicature in a different context, since we are always capable of embedding a context in a different one.

What emerges clearly is that, even if we were to accept Carston's considerations on cancellability, we would have to have two language games, one for generalized implicatures and one for particularized implicatures (I am sure Carston would want to insist that the language game for cancelling generalized implicatures can be partly utilized in the case of particularized implicatures). Now, supposing that we have two different language games, we still would not know how to unify them. Like Burton-Roberts, I believe that unification is impossible.

12 Conclusion

It appears to me that only theoretical, rather than empirically-oriented considerations, can guide or orient our philosophical investigations on the pragmatics of language and on the usefulness of the notion of cancellability (of implicatures/ explicatures). Should we find out that cancellability is of considerable use, we should try to explain why. I doubt that cancellability is of use in determining whether an inference is an implicature—because we intuitively know that when an inference is not driven by semantics it is an implicature. Instead, it is of use in the case of generalized implicature (potential implicatures) because it defines the kind of role which context can play in shaping meaning—namely a negative role. Since in the case of particularized implicatures, context does not have a negative role to play—as the inference is not potential, but must be singled out by the interplay of sentential meaning and context—it can only have a positive role to play. Hence we expect that cancellability has no utility for particularized implicatures.

For explicatures, lack of cancellability, rather than cancellability, seems to be a crucial diagnostics of it, contrary to what is assumed in the literature.

References

Bach, K. (2001). Semantically speaking. In I. Kenesei, & R. M. Harnish (Eds.), *Perspectives on semantics, pragmatics and discourse. A festschrift for Ferenc Kiefer* (pp. 146–170). Amsterdam: John Benjamins.

Binkley, T. (1979). The Principle of Expressibility. *Philosophy and Phenomenological Research, 19*(3), 307–325.

Burton-Roberts, N. (2005). Robyn Carston on semantics, pragmatics and encoding. (Review article of Carston 2002). *Journal of Linguistics, 41*, 389–407.

Burton-Roberts. (2013). Grice and cancellation. *Journal of Pragmatics, 48*(1), 17–28.

Capone, A. (2006). On Grice's circle (further considerations on the semantics/pragmatics debate. *Journal of Pragmatics, 38*, 645–669.

Capone, A. (2009). Are explicatures cancellable? Toward a theory of the speaker's intentionality. *Intercultural Pragmatics, 6*(1), 55–83.

Capone, A. (2010). Between Scylla and Charibdis: The semantics and pragmatics of attitudes 'de se'. *Journal of Intercultural Pragmatics, 7*(3), 471–503.

Capone, A. (2011). The attributive/referential distinction, pragmatics, modularity of mind and modularization. *Australian Journal of Linguistics* 31 (2), 153–186.

Capone, A. (2012). Consequences of the pragmatics of 'de se' attitudes. In N. Feit & A. Capone (Eds.), *Attitudes 'de se': linguistics, epistemology, metaphysics*. Stanford: CSLI publications.

Capone, A (2013). The pragmatics of quotation, explicatures and modularity of mind. *Pragmatics and Society*.

Carapezza, M, Biancini, P (2013). Language game: calcolus or pragmatic act? In A. Capone, F. Lo Piparo, M. Carapezza (Eds.), *Perspectives of pragmatics and philosophy*, Dordrecht, Springer.

Carston, R. (2002). *Thoughts and utterances*. The pragmatics of explicit communication. Oxford: Blackwell.

Carston, R (2010). Explicit communication and free pragmatic enrichment. In Explicit communication: Robyn Carston's Pragmatics, eds. Soria and Romero, (pp. 217–87). Basingstoke: Palgrave Macmillan.

Dascal, M. (2003). *Interpretation and understanding*. Amsterdam: John Benjamins.

Duranti, A. (1988). Intentions, language and social action in a Samoan context. *Journal of Pragmatics, 12*, 13–33.

Gazdar, G. (1979). *Pragmatics: Implicature, presuppositions and logical form*. New York/London: Academic Press.

Hintikka, J. (1962). *Knowledge and belief*. Cornell: Cornell University Press.

Huang, Y. (2007). *Pragmatics*. Oxford: OUP.

Jaszczolt, K. (1999). *Discourse, beliefs and intentions*. Oxford: Elsevier.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: CUP.

Levinson, S. C. (2000). *Presumptive meanings*. Cambridge: MIT Press.

Mey, J. (2001). *Pragmatics*. Oxford: Blackwell.

Sadock, J. (1978). On testing for conversational implicature. In P. Cole (Ed.), *Syntax and semantics 9: Pragmatics* (pp. 281–298). New York: Academic Press.

Searle, J. (1970). *Speech acts*. Cambridge: CUP.

Seymour, M. (2010). Speech act pluralism, minimal content and pragmemes. In Pragmemes, ed. A. Capone. *Special issue of the Journal of Pragmatics* 42, 2870–2881.

Sperber, D., & Wilson, D. (1986). *Relevance*. Oxford: Blackwell.

Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge: CUP.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

The Pragmatics of Indirect Reports and Slurring

Alessandro Capone

Abstract According to Volosinov (1971) there is a tension between two indirect discourse practices; one in which the reported message's integrity is preserved and the boundaries between the main message and the embedded reported message are formally marked and one in which such boundaries are dissolved as the reporting context allows the reporting speaker to intrude to a greater extent and transform the message by stylistic interpolations. This tension is clearly resolved, in the context of my paper on indirect reports, through the recognition of pragmatic principles which assign default interpretations (according to which the boundaries between the reporting message and the reported message are clearly visible and the reported speaker's voice prevails at least within the embedded message), while allowing context to create priorities which override the default interpretations and make the otherwise costly violations of the pragmatic principles worthwhile thanks to the facilitation and subordination of the information flow to the exigencies of the embedding context (Of course, this tension is clearly instantiated in language (it is not only a theoretical problem). As a referee points out, we are focusing on a case in which two practices are in tension. The resolution of a tension between two different, possibly opposite, practices clearly depends on practical considerations leading the language users to prefer one to the other. Deviation from a practice that conforms to ideal principles of use must always involve a cost that needs to be offset by practical advantages. One of these advantages could be the facilitation of the recognition of a referent. Another possible advantage could be, as happens in many cases, the simultaneous utterance of a speech report and a criticism).

A. Capone (✉)
University of Messina/Palermo, San Francesco P 107 98051 Barcellona, Italy
e-mail: alessandro.capone@istruzione.it

1 Introduction

The practice of indirect reporting involves a mixture of serious and non-serious use, as this practice, on the one hand, involves transformations in the sense of Goffman (1974),[1] on the other hand it involves using language in the context of a serious activity, such as describing what another person said. The practice of indirect reporting is sensitive to contextual information and, thus, it goes without saying that the richer the cues and clues allowing speakers to interpret transformations (see Dascal and Weizman 1987), the more complex are the transformations involved in the indirect reports. And the more complex the transformations are, the greater the need for a decoupling principle along the lines of Clark and Gerrig (1990):

> Speakers intend their addressees to recognize different aspects of their quotations as depictive, supportive, and annotative.
> *Mutatis mutandis*, we can apply the Decoupling Principle to indirect reports:

> Speakers intend their addressees to recognize different voices belonging to the indirect report and, in particular, to separate voices attributing them to the original source, the current speaker (the indirect reporter) or some other person involved in context. They also intend addressees to recognize supportive and annotative aspects.

To make the considerations above less cryptic, I note that *supportive* aspects are those which in one way or the other allow the speaker to make the indirect report. For example, the reporter may use English to report a Latin utterance. This use of English is clearly supportive and NOT depictive (of course, hearers should have pragmatic ways to decouple such aspects). *Annotative* aspects are those which are noted, without serving a principal purpose in the practice of reporting (for example I can note that the original speaker was giggling while using a certain word). Depictive aspects concern the words actually proffered.

I have now already departed to some extent from the standard practice to consider indirect and direct reports neatly differentiated. Clark and Gerrig themselves consider the two practices to be neatly separated, because quotation prevalently makes use of depictive aspects of language use while indirect reports make use of descriptions. Presumably, using Clark and Gerrig's terminology, there are other reasons for keeping the two practices distinct. Clark and Gerrig (p. 771) note that quotation involves both serious and non-serious language use. It involves serious language use in that the quoted item is syntactically an NP; it involves non-serious language use in that the quoted item is syntactically a sentence (S) and, thus, depictive elements prevail if the item is considered a sentence.[2] Presumably,

---

[1] For example, shifts from serious to non-serious or depictive uses.

[2] Presumably, Clark and Gerrig seem to accept that an NP is presuppositional, thus expresses an extensional object; a sentence embedded in a verb of propositional attitude or in a quotative structure can express a non-extensional object. Simple cases that can illustrate what Clark and Gerrig have in mind could be the following: "I want that car" (or "I want the car"). Here it is plausible that the NP following 'want' identifies an extensional object and not ONLY an

by Clark and Gerrig's standards, indirect reports should only involve serious uses of language, since only NPs are involved here, rather than sentences intended in their depictive sense. However, we all know that indirect reporting is very often a **polyphonic** practice where the hearer's main task is to separate voices attributing them to different actors. Even if we stay within Clark and Gerrig's terminology, it is universally recognized that there are what are often called 'mixed quotations', that is to say cases of indirect reports in which some segments are quoted. Mixed quotations are considered relatively rare cases—while the point of my discussion is that they should be considered as prototypical cases of indirect reports and that indirect reports in general should be modeled after mixed quotations (see Capone 2010a).

I have already said that indirect reports are interpreted in context. Here, however, the term 'context' is ambiguous, because, strictly speaking at least two types of context should be relevant to the interpretation of indirect reports: the context of utterance (of the original speaker) and the context of utterance (of the indirect reporter). There is often an interesting interplay between the two. We should note from the beginning that chronological considerations are important in ranking the two contexts and that the context of utterance (of the reporting speaker) is the departure point from which interpretation starts. It is often useful, therefore, to bear in mind what the purpose of the indirect report is or might be.

Indirect reports are cases in which you transmit knowledge of what another person said and what another person said is the only way or one of the ways in which you can gain knowledge about a certain situation or event **s**. The situation is clearly different from that of perception, where the only mediating elements are the perception system and certain a priori principles of knowledge. In indirect reports, the situation **s** is transformed two times[3]: once by the original speaker and then by

---

(Footnote 2 continued)

intentional one. There are cases where we might object to the equation of an NP with an extensional object, as in the case "John wants to sell his cello". Here 'His cello' could either escape the scope of the modal 'wants' or it could still be under its scope. Despite these controversial criteria, there are syntactic positions correlating with extensional/non-extensional, such as 'want NP', especially if the NP is definite. Anyway, I quite agree with a referee that the criteria by Clark and Gerrig are not uncontroversial.

[3] The situation described by an indirect report is usually an utterance by an original speaker who, in his speech act, described or brought about a situation (in the case of a non-assertive speech act). A situation is a state or event with possible participants in it. When I say that a situation is transformed through an indirect report, I mean that the reporter uses NPs that are not neutral, but may express his/her point of view and, in particular, a critical attitude (for example if the reporter makes use of epithets). The situation may be transformed in another way, as the reporter may avoid using NPs actually used by the original speaker, but may use different NPs to make sure that the Hearer can identify the referent in question. I used the term 'transformations' but I could have used the term 'modifications'. However, 'transformations' refers to an operation effected linguistically and possibly in a systematic way. In fact, it might be predictable and therefore systematic that if the hearer cannot identify the referent through an NP used by the original speaker, the reporter must use a different NP, one that allows the hearer to identify the referent. Transformations are rather systematic practices. When you are confronted with an indirect report

the reporting speaker. So, the task of the hearer is clearly an inferential task; how to delete possible transformations and how to get (back) to **s** without the interference of possible transformations. This is clearly an inferential task requiring pragmatics. Now, if the hearer of the indirect report is interested in the indirect report mainly because it allows her to have access to **s**, the reporting speaker knows this and may very well take this into account in her treatment of the information concerning the original utterance. So we may grant that at least part of the transformations may be shaped by the desire to meet the interests of H in knowing about **s**. Other transformations may be independent of the interests of H or may conflict with it. Just to mention a case, consider the reporter who said: 'John said that the bus for Oxford is on the left when you get out of the airport'. It is crucial, in this interpretation process, that the perspective be the same. And that must be: passenger getting out of the airport. If the perspective adopted in the indirect report was different from that adopted in the original utterance, confusion would ensue. Thus we exclude that the perspective could be: relatives waiting for the passenger out of the airport. If, for some reason, the indirect reporter transformed the utterance without taking into consideration the hearer's interests, an uninterpretable utterance would result (or to be more correct an utterance providing misleading information would result). As upshot of this, the purpose of the indirect report must feature prominently among the factors to take into account in the interpretation as well as in the production of indirect reports. Let us consider, provisionally, the basic structural elements that go into an indirect report.

Context 1 (original speaker; original Hearer)
Context 2 (reporting speaker; reporting speaker's Hearer)

Decoupling Principle
Separate the original speaker's from the reported speaker's voice. Establish which portions of the text have a directly pictorial function.[4]

    Separate those parts which have a supportive or an annotative function.

Purpose 1 (original speaker)
Purpose 2 (reporting speaker)
Purpose 3 (addressee).

(Footnote 3 continued)

that makes use of epithets (that bastard), you may be pretty sure that the speaker is using language in a critical way and thus a systematic effect on the hearer is the desire to know the difference between the NP used by the reporter and the one used by the original speaker. These transformations are systematic also in the sense that it might be possible to spot them and to go back to the original utterance via reflective processes.

[4] Perhaps the best example of the pictorial function is the following: John has SEEN Mary in the BATHROOM. There are cases, like the one above, in which language is used to express the form (boldface, for example) of an utterance. In this case we have a visual dimension, but sometimes we have an aural dimension, as in the case of a speaker who imitates the voice of another speaker (imitation) (or the style). Normally, however, by 'pictorial' Clark and Gerrig mean that a speaker depicts the actual words employed in a certain utterance.

Point of view 1 (Original speaker)
Point of view 2 (reporting speaker)
Point of view 3 (addressee).

Now that the structural components of the practice of indirect reports are in place, we can expect that a theory of indirect reports could be built objectively on this basis, perhaps on the basis of recursive operations that take into account the basic components.

1. Indirect reports as language games

In my previous work on indirect reports I have focused on reports as language games (Capone 2010a, 2012). Language games are activities produced through speech in conformity to social rules determining what can count as what (in the speech situation). A language game is a strip of social activity (of social life) where language (speech acts) play an important role in the execution and deco-dification (and interpretation) of the activity.[5] A language game is a form of life, the individual being able through it to participate in a social form of life (being integrated in a social dimension and coordinated through action with other members of the group). Now, while there may be differences between Goffman's terminology as used in the previous section and the terminology of language games, it is also clear that there is substantial overlap. Goffman presumably saw the continuum of social practices as segmented [or 'framed' (Goffman 1974)]. Each segment was to be recognizable as there had to be boundaries between outside and inside activities. An example of Goffmanian analysis that is well known is that of the lecture (Goffman 1981). The lecture is a bounded activity, which has its own rules. Participants know well and in advance how to behave in this segmented area, they know that there is little space for interruptions, they know that lectures have a forthcoming segment reserved for questions and answers (by the lecturer). Clearly, the lecture is also a language game, because it is structured, it has rules, it is part of societal activities, it is sufficiently differentiated from other language games. So, substantially, Goffman's theory of frames and forms of talk must coincide in broad lines with a theory of language games—or at least it must be possible to explore interconnections and overlapping territory.

But why should we want to deal with indirect reports—activities confined to small segments of interaction—in terms of language games, which are usually activities that unfold for some time and occur at some place which is substantially involved in the language game and even serves to characterize it (for example, court procedures)? And now my answer is that even if indirect reports are not normally really extended in time as language activities, they involve embeddings such as those described in the Introduction, and which we may illustrate sche-matically as in the following:

Indirect report C (indirect reporter)

---

[5] In conformity with Wittgenstein's Philosophical Investigations.

Original speaker C (original speaker)
Addressee C (addressee).

Although the language game is temporally limited, if we consider the dimension of the linguistic activity that unfolds in a temporal succession, the temporal embeddings obtained by reconstructing the original speaker's situation are potentially manifold and complex. Complexity is introduced when we see connection with other language games such as the following.

Consider a child game, which almost everyone practiced in childhood or adolescence:

There are, say, 20 boys (or girls) in a room. Each whispers to the next person in the line (or circle) what was whispered into his hear previously. The aim of the game is to show that, although, ideally, the initial and the last utterance have to be the same, the initial utterance is so transformed that the last utterance can hardly be heard to bear any meaningful relation to it.

This might be a game pointing to a practice which is quite standard in society and is based on reliable methods for transmitting and preserving information during the transmission process. The previous game dramatically illustrates the problems inherent in the game 'reporting information' or 'reporting an utterance'.

Consider another game such as the dumb-show.

A dumb-show was one of our favorite games in childhood. We practiced it, I presume, as a form of preparation in view of more serious or important societal language games. In a dumb-show you must depict information by avoiding words. You usually use gestures, even if you can point to words which happen to be written on a blackboard or on a poster. Now, since depicting occurs so heavily in direct reporting and, also in indirect reports, this is clearly a case in which we consider 'depicting' an important part of language games, a component which may be shared by different language games.

Another language game which is crucial for the understanding of indirect reports is a theatrical performance. In a theatrical performance we usually pretend, we are not using language in a serious way. An actor does not talk for herself, but on behalf of a character. This is more or less what happens in direct reports, but also what happens in indirect reports, if we consider them as polyphonic activities (see also the problematic case of mixed quotation in indirect reports).

Another case of language game that is deeply rooted in society is testimony in court. Here it might be important to be able to report what another person said on a certain occasion. This may well be an extreme case, where there is little freedom for transformations and where one needs to separate one's voice from that of the original speaker by formal markings. This practice diverges from the daily practice, to a great extent. Here a reporter may be asked by the prosecutor to reflect on the words used, to make an effort to separate her own voice from that of the original speaker. This practice may well involve a meta-representational component, as one is interested in the meanings as well, in the connotations as well as in the denotations. The reporter may well be turned into an analyst of her own speech.

(Do you exactly mean that …?). This is clearly a case where reporting is considered a meta-linguistic activity.

So, although it is true that indirect reports are small segments of talk or small strips of social behavior, they nevertheless have many features in common with other strips of behavior which we are less reluctant to call 'language games'. This may be enough to see that the connection between language games and indirect reports is well justified.

Dascal et al. consider that the notion of language game by Wittgenstein involves a shift from phenomenalism to physicalism, language games being primarily intended to create social reality. Can the language game of indirect reporting be so intended? If we follow Tannen (1989), indirect reports can, indeed, be considered as actions serving to construct social reality. An indirect report can have effects on deliberation, on action, in that it can present a piece of information that can be integrated into the argumentative structure of practical reasonings. Seen in this light, an indirect report can become a 'form of life' (Wittgenstein 1953).

Another feature of language games, according to Dascal et al. (1996) is that they are cooperative [they are constructed jointly by different speakers (or the speaker and the hearer)]. Can this be a characteristic of indirect reports (such as language games)? My reply in Capone (2012) was that the recognition of the role of the Hearer (or addressee) in the amount of transformations required in the practice of indirect reporting amounts to a recognition of the cooperative nature of indirect reports. Indirect reports—like other language games—involve an altruistic stance towards the addressee, which is instantiated in important linguistic choices that can be seen as transformations.

In Capone (2012) I specifically discussed indirect reports as language games, in the light of considerations by Dascal et al. (1996) on language games. Here I cannot expand that discussion, but I confine myself to extrapolating the most important points. Dascal et al. consider polyphony a specific language game—now, while surely indirect reports are interesting also for other features, such as representational ones, it is clear that polyphony is a language game that is embedded in the practice of indirect reporting. The game also consists in the way clues and cues are utilized to separate the voices of the participants.

The language game 'polyphony' aims at the integration of different voices (expressing different points of view). Integration does not mean summation, but an interaction between two points of view such that one is, often, a commentary on the other. One of the problems we encounter in the description of indirect reports is, in fact, that an apparently single utterance contains different voices/points of view (thus it is polyphonic), apparently making it difficult for the hearer to separate them. However, the problem is not only how to separate points of view, but how to see the interaction between them. In an indirect report, we do not only have a neutral presentation of points of view, but normally the point of view of the indirect reporter is the main filter through which we hear other voices. Thus, it happens characteristically that there may be a relationship of criticism or otherwise affiliation between the point of view of the indirect reporter and the point of view of the reported speaker. Polyphony, as a language game, therefore does not consist in a mere

summation of voices, but in the integration of them within a relationship of commentary between the voices (one voice being a comment on the other). Polyphony involves commentary in that the reporting speaker can present the reported speaker's voice in a derogatory way (he may shout, speak sardonically, express rage, etc. when expressing the reported speaker's voice). We may very well speak of a concert of voices, which are however, regimented by the reporting speaker and by the inferences of the hearer. The hearer is capable of using inferences to differentiate voices, but also to notice if some element is added illegitimately by the reporting speaker. So, the game is not only one which has the reporting and the reported speaker as its main participants, but one where the hearer is an important judge, who can add things not said or subtract unnecessary elements.

2. Davidson on indirect reports

In this paper, I am not after the logical form of indirect reports. I am mainly after a pragmatic treatment based on the notion of the language game. However, I will briefly mention Donald Davidson's (1968) treatment of the logical form of indirect reports because it is the treatment that best accords with my view of indirect reports as language games. According to Davidson a sentence such as:

1. John said that Mary is in Paris

is to be accounted for, truth-conditionally, by the following logical form:

John said that. Mary is in Paris.

In other words, Davidson asks us to consider a proposal according to which the complementizer 'that' disappears from logical form, being replaced by the pronominal 'that'. A propos of this, Davidson briefly mentions historical considerations on the development of the complementizer 'that' from the pronominal 'that'. Now, I am aware that there is a strand of research that builds on Davidson's proposal (sometimes aiming to ameliorate it, sometimes aiming to destroy it; see Rumfitt 1993). But as in this paper I am mainly interested in the language game 'indirect report' and in the pragmatics of indirect reporting, I will skip such discussions. I will nevertheless rehearse some considerations by Davidson, which are now very popular in philosophy:

> We tried to bring the flavor of the analysis to which we have returned by rewording our favorite sentence as "Galileo uttered a sentence that meant in his mouth what 'The earth moves' means now in mine". We should not think ill of this verbose version of "Galileo said that the earth moves" because of apparent reference to a meaning ("What the earth moves means"), this expression is not treated as a singular term in the theory. We are indeed asked to make sense of the judgment of synonymy between utterances, but not as foundations of a theory of Language, merely as an unanalyzed part of the content of the familiar idiom of indirect discourse. The idea that underlies our awkward paraphrase is that of same saying: when I say that Galileo said that the earth moves, I represent Galileo and myself as same sayers. (Davidson 1968, 140).

Now, by extrapolating this excerpt, I want to emphasize that for Davidson it was clear that *oratio obliqua* is a discourse involving multiple voices. The mouths

uttering the words belong to different persons. The sentences actually uttered, as far as Davidson is aware, may very well be different provided that the two utterances are semantically equivalent, that is to say their imports are truth-conditionally equivalent. There are two voices, two points of view involved, and an indirect report is, obviously, a transformation of the original utterance. Baldwin (1982, 273) claims that one defect standardly attributed to Davidon's formulation of the theory is that it seems to imply that there is one more utterance besides the utterance 'The earth moves'. This, which from a philosophical point of view, counts as a defect (which could be remedied anyway, if we follow the discussion in Baldwin), is not necessarily a defect from a linguistic point of view as it makes us see that the case of indirect reports (and its logic) depends on the tension between the reported speaker's voice and the reporter's voice. It is no surprise that there may be two utterances, whose content is fundamentally the same, although parts of it, those parts which do not count for the provision of an extensional semantic theory of indirect reports, need not be the same.

The considerations by Davidson on p. 143 are not equally famous, but in my opinion they lead to a view of indirect reports as language games, in the study of which pragmatics is prevalently or at least substantially involved:

> We would do better, in coping with this subject, to talk of inscriptions and utterances and speech acts, and avoid reference to sentences. For what an utterance of "Galileo said that" does is announce a further utterance. Like any utterance, this first may be serious or silly, assertive or playful, but if it is true, it must be followed by an utterance synonymous with some other. The second utterance, the introduced act, may also be true or false, done in the mode of assertion or play. But if it is as announced, it must serve at least the purpose of conveying the content of what someone said. (Davidson 1968, 143).

At this point we notice that Davidson has touched on a deep issue—the content of indirect reports may be determined pragmatically. So, it is possible that the utterance x following "Galileo said that" may be synonymous with an utterance which is not truth-conditionally equivalent to x, but can be made pragmatically equivalent to x, say through pragmatic intrusion. (In other words, we should consider the explicatures as truth-conditionally equivalent). In general, the excerpt above raised the important question that the purpose and the speech act communicated by the indirect report may prominently figure when we try to establish whether the reporting utterance and the original utterance match in content. I discussed, however, briefly this notion in Capone (2010a). For the sake of this discussion, it is important to point out that Davidson thinks we must separate truth-conditional content and pragmatic content. Even if Davidson does not move towards a radical pragmatic view of indirect reports, it is clear that the notion of pragmatic equivalence is what is at stake when we say that the original utterance and the reporting utterance match in content. Suppose, for example, that the original utterance is:

2. Mario is really brave

and the reporter, whether accurately or not, transforms (2) by uttering (3) (with a view that (3) and (2) match in content).

3. John said that Mario is a lion.

Should we say that the indirect report matches in content the original utterance? For some purposes, we may be persuaded to answer positively, even if the locutionary forces of these utterances (clearly) do not match. It may be argued that (2) and (3) cannot be taken to convey the same content, as metaphorical meaning is in some sense non-conventional (as referee 2 says, in order to guarantee sameness of content the metaphor would have to be conventional. But then it would be well on its way to being a lexical sense of the expression). I am not persuaded by the idea that two utterances match in content only if the conventional meanings match, but of course I agree with referee 2 that a metaphorical expression conveys (usually) much more than the conventional expression it was used to replace. There are effects in terms of poetry, force, rhetoric which are not expressed by a non-metaphorical expression. But granting some differences, think now of the following language game. We can report thoughts by using certain cards, on each of which a certain word is printed. We do not have cards for every word. So we must do what we can to express our thoughts, and our readers must accept the approximations which we can use. Now suppose we have a card for 'lion' but not one for 'brave'. Could we engage in the language game of reporting the speech act, nevertheless? The answer, in the context of this language game, with its obvious limitations, is positive. To report 'John said that Mario is a lion' is certainly better than nothing and our readers will have to put up with the limitations of our language game. However, even in a different context, a reporting speaker may want to modify somewhat the original utterance, to convey something which John did not say but probably wanted to say (or would have said in different circumstances). Perhaps the reporting speaker is judging that his indirect report is more faithful to the speaker's intentions than the original speaker's words. Perhaps the reporting speaker is relying on clues which are not available to the hearer (or reader) and is reconstructing the speaker's intentions to the best of his own abilities. After all, are we not allowed to infer and voice someone's intentions, even if that person was not capable of fully expressing them? Now, this argument, clearly, has taken me some way from the considerations by referee 2.

In ending this section, I want to remind readers that the initial Davidsonian formulation of indirect discourse was criticized because it was immune to intentionality (Baldwin 1982, 272) and was thus later replaced by a better analysis which was completely extensional (Baldwin 1982, 273):

Galileo said x iff (∃ y) [Galileo uttered y and Same in content (x, y)].

3. Capone (2010) and indirect reports

In Capone (2010a) I advanced a number of ideas on how to capture constraints on replacements of co-referential NPs in the context of indirect reporting (and, in particular, in the complement that-clause). The explanation may be parallel, but not identical with the one I gave on the issue of belief reports in Capone (2008). Such an explanation rests on the idea that replacements of co-referential NPs should not alter the speech act which the indirect report aims to report (or describe)

and that the original speaker would like to see herself reported in such a way that it does not attribute her offenses, impoliteness, rudeness, obscenity, and also slurring. In other words, reporting must be done in a way that the voice of the reporter is separated from the voice of the reported speaker or, if this separation is not possible, in such a way that the original speaker's voice is prevalent. Why should the reported speaker's and NOT the reporting speaker's voice be prevalent? I assume that it is a matter of relevance. Since we are dealing with the verb 'say', we are happy to primarily express the original speaker's voice and then the reporting speaker's voice, but only if this is possible. I now succinctly sum up the main points of Capone (2010a).

The practice of indirect reports rests on the following principles:

Paraphrasis Principle[6]
The that-clause embedded in the verb 'say' is a paraphrasis of what Y said that meets the following constraint: should Y hear what X said he (Y) had said, he would not take issue with it, but would approve of it as a **fair paraphrasis** of his original utterance.

The following is a precisification of the previous Principle, which remedies some of its defects, as it does not only take content into account, but also makes reference to form.

Paraphrasis/Form Principle
The that-clause embedded in the verb 'say' is a paraphrasis of what Y said, and meets the following constraints: should Y hear what X said he (Y) had said, he would not take issue with it, as to content, but would approve of it as a fair paraphrasis of his original utterance. Furthermore, he would not object to vocalizing the assertion made out of the words following the complementizer 'that' on account of its form/style.

In my paper I also discussed possible objections to the Paraphrasis/Form principle. Since this discussion will be amplified in the present paper, I present some of the original discussion in this section.

Depending on the context, I needn't be beholden to the original speaker's 'approval' of my paraphasis as fair, nor need I avoid manners of speech which the original speaker would shy away from. In such contexts, if John said of a person x that he will be coming to the party, my report to that effect is true whether I refer to person x politely, as John would approve of, or impolitely, as (let us imagine) my hearer would approve of. John may, upon hearing my report, demur: "Well, I don't know why you'd call x a jerk but, yes, I did say he was coming to the party". The Paraphrasis Principle and the author's other remarks are intended to rule out

---

[6] This position is somewhat reminiscent of Seymour's (1994) treatment of indirect reports, in which reference to a translation of the reported sentence is explicitly incorporated in the semantics of indirect reports.

contexts of indirect reporting that seem to allow this type of license with the original speaker's words.

As I said in Capone (2010a), I am quite open to the possibility that in suitable contexts[7] one should be able to replace an NP with a coreferential expression in the that-clause of an indirect report. However, I stick to the proposal that, in the absence of abundant contextual clues and cues allowing us to separate the original speaker's voice from that of the reporter, the default interpretation of the utterance conforms to the paraphrasis rules stated above.

4. Some considerations on Wieland on indirect reports

Wieland (2013) considers that most theories on indirect reports conclude that the practice of indirect reporting must be studied essentially from a pragmatic point of view. Wieland, however, refuses to accept that one cannot say something systematic and of general import about the practice of indirect reporting. She is adamant in considering the case of indirect reporting distinct from the case of quotation and the case of belief reports. Now, if such propositions are accepted, it goes without saying that indirect reports allow a certain amount of substitution (of NPs having identical referents) and thus it is not to be taken for granted that they are characterized by opacity. Since they are not expressions of belief, the attitude of the original speaker need not interfere with substitution of NPs having identical reference. Now, I do not want to dispute these propositions, as there is obviously some truth in them. But it is possible that the inferential step from these propositions to the lack of opacity exhibited (according to Wieland) by indirect reports is not necessary or needed; in other words, it may distract us from some obvious connections between a theory of quotation and a theory of indirect reports. And the most obvious link between the two theories is that in both cases we need to establish which voices belong to the various segments making up the utterance. Indirect reporting (as made clear by Cappelen and Lepore 2005b) involves mixed quotation, at least in some cases. So the only way to make the two issues separate now is to insist on quotation as being characterized strictly by opacity and indirect reports as not being characterized by opacity (or in being characterized less strictly by it). However, if we grant that indirect reports can contain quoted segments, it is less clear that opacity and lack of opacity can be used to distinguish the two cases. In my article on quotation (Capone 2013)[8] I insisted that inverted commas need not always be used to signal the quotative function, as they are often absent in the oral language. Rather we need pragmatic ways of signaling that certain segments are being quoted. But if this is the case, then it goes without saying that implicitly many segments of indirect reports can come out as being quoted, at least through

---

[7] One of the most typical contexts allowing substitutions of coextensive NPs is one where the NP used in the original speech act would not allow the hearer (of the indirect report) to identify the referent, and thus the reporter deems it necessary to use an NP which does indeed allow the hearer to recognize the referent (See also Capone 2008). On the role played by context in inferential processes see Capone (2010b).

[8] See also Saka (1998) for a discussion of quotation in philosophy of language.

some pragmatic means. These differences of opinion between Wieland and myself do not prevent me from seeing the importance of her other considerations on indirect reports. And it is on these crucial considerations—which I should say are both important and controversial—that I want to concentrate now.

Somehow departing from my considerations in Capone (2010a), Wieland argues that in some contexts, when the reporting speaker has a purpose which serves to advance the communication process—rather than impeding it through the use of an NP whose semantic import is not known to the hearer—it is licit to inter-substitute co-referential terms.[9] Consider this co-referential substitution:

4. A: My favourite *tapa* is *patatas bravas.*

   B: A said that her favorite *tapa* is the third item on your menu.
   Wieland says:

   > 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 (Wieland 2013).

And I agree that in reporting the original utterance by transforming an NP in this way allowing the hearer to get to the referent in a quicker way, a speaker has a practical purpose. This practical purpose does not completely transform the original utterance, in ways that might give rise to complaints by the original speaker. Furthermore, this is clearly a case in which the NP used to transform the original NP is quite neutral; and most importantly, by using it, a hearer can have access to the thought entertained by the original speaker (in saying whatever he said), as the NP which was used as a replacement will eventually, albeit not immediately now during the indirect report, but once the report has been heard in its entirety, allow the hearer to reconstruct the item that is momentarily missing. I propose to use a technical term for items such as 'the third item on the menu'—these are sort of pro-forms, but unlike pronominals, which point to objects, they are **quotative pro-forms**, as they point to locutionary segments of the talk. (Obviously they refer to types, rather than tokens).

There are other interesting transformations which Wieland draws our attention to. Consider the following, from her paper:

5. A: I went to the taco stand and bought a soda.

   B: A said that she went to the taco stand.

B's utterance is clearly obtained by conjunction elimination. Now apparently, this is the case of an innocuous, even innocent transformation. However, there are doubts that this transformation can be effected without consequences when conjunction is involved in an explicature, as in the famous examples by Carston (2002). So, suppose that Churchill said (6)

---

[9] This is in line with Wieland (2010).

6. The Germans raided London and we fought them back

    or

7. The Germans raided London but we fought them back.

There may be explicatures or conventional implicatures (see Potts 2005) attached to a certain conjunction (and as a consequence, we fought them back). Thus eliminating a conjunct from an indirect report in such cases, gives us the impression that part of the original meaning is lost. So if (7) is reported as (8)

8. Churchill said that we fought the Germans back

We have partially reported the utterance. It is a partial report. Could a partial report be felicitous? There are contexts in which it might and contexts in which it might not be felicitous. So, it is not straightforward that conjunction elimination is an operation that can be used always felicitously in indirect reporting.[10]

Consider now modifier elimination. It might be thought that modifier elimination is an innocuous logical operation in indirect reports, simply because it is supported by logical/semantic entailments:

If NP [VP ADVB V NP], then it must be the case that NP [VP V NP].

So, if I met a beautiful woman at the party, it must be the case that I met a woman at the party. And if John says:

9. I met a beautiful woman at the party

    it could be claimed that one could report felicitously:

    (9) John said that he met a woman at the party.

But now suppose that on a different occasion John said of the same woman, unaware that she was that woman:

10. That woman is horrible.

Now we could conjoin (9) with (10), since after all John was talking about the same woman and obtain:

11. John said he met a woman, who was horrible, at the party.

---

[10] A case for the potential infelicity of partial indirect reports. Mrs Savatta was the headmistress in a high school in Italy. At a meeting with the teachers, she said 'Suppose I say that Mr Buccheri is an idiot." Of course, she said that in a context, and her context was provided in part by her previous utterances. There was a rhetorical relationship between this utterance and the previous ones—she was presumably using this utterance as part of a (complex) argument. However, the teacher was offended by this and a long legal quarrel followed. The secretary of the meeting reported just this utterance but completely omitted the previous utterances, thus making it appear as if the headmistress was completely mad. Granting that there was something amiss in this linguistic contribution, however, there was something completely amiss in the report of her utterance, because it was a partial report. Cutting an utterance and reporting just part of it can make things appear in the wrong light, as the function of an utterance in a sequence of speech (in particular the rhetorical connections) seems to have been lost and the immediate result is that the speaker can be presented as a deranged person.

So the problem I see in modifier elimination is that it will allow us to conjoin a report of what John said on some occasion with a report of what he said on another occasion which contradicted what he said before.[11] The contradiction passes unnoticed, if we simply support the view that modifier elimination is a feasible operation in indirect reporting.

I should notice that Wieland adds a little later that "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." I quite agree with these considerations, even if I would take side with a more general position in which partial indirect reports are always less informative than exhaustive indirect reports and thus they require a context that justifies the extra cognitive effort required in the logical operation of the reporting (since reducing involves an extra logical operation). This may well be in line with the general position by Sperber and Wilson (1986) according to which Relevance is a balance of positive rewards (effects) and cognitive efforts.

Another important consideration by Wieland is that the logical operation **inference** can be incorporated into indirect reports. She felicitously calls this case: inferential indirect report. An example of this practice might be the following (always from Wieland 2013):

12. A: I didn't fail any students.

    B: Professor A said Maryanne passed her exam.

Wieland says: "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 inter-substitutability 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". Now there is something weird about this case. Suppose Professor B is universally known as passing only very good students (he fails those who are passable for other professors). Then, given what is known about Professor B's beliefs, it could be claimed that Professor B said that Maryanne was a very good student. Then suppose it is well known that professor B believes that all his good students will become University Professors. Then it will be held that Professor B said that Maryanne will become a University Professor. But it is not clear that Professor B said all these things. Now, while in

---

[11] The problem is more or less of the same type as noted by Igor Douven in connection with the pragmatics of belief. One should not make inferences that are likely to deceive one's future self. Now, while clearly the inferences Douven has in mind are pragmatic, here we have a logical operation of modifier elimination. But the result is similar as one's future self may be misled by being allowed to make other logical operations (such as conjunction). By the way, I am not thinking that modifier elimination and conjunction are related things. However, there is a danger in using both operations, sometimes.

my own examples, the problem might derive from identifying the words said with the beliefs normally associated with those words, in Wieland's case the problem is even worse, because professor B is said to have said something without even believing it, as he never had any beliefs about Maryanne (suppose the examination was carried out on papers marked by a code, to make them anonymous). There might be interminable discussions on points such as these—and it is good that these discussions should be undertaken. My intuition is that we are at a point in which it is not easy to distinguish between legitimate cases of indirect reports and cases that are parasitic on them. It is possible that this might be a loose usage. But even if a loose usage, it is still an indirect report, and thus Wieland does well to point out that inference may play an element in reporting. (Given that it may play a role in establishing the truth of a report, I propose that we give great consideration to Wieland's case).

The case just discussed reminds me of cases in which pragmatic inferential augmentations are banned by Igor Douven's (2010) the Pragmatics of belief and, in particular, by his Epistemic Hygienics.

Igor Douven proposes that when we store a belief (in the form of an assertion or a sentence or a thought), we avoid storing it together with inferential augmentations which may lead us later to remember something which was not the case. This is called Epistemic Hygienics. A vivid example which comes from that paper is the reference to Gettier's problem. Suppose I know that p. Then, even if I can infer 'p or q' from 'p', it will not do to store in memory 'p or q' if that is going to create trouble later, leading me to believe something that is false or unjustified. We may remember that what creates havoc in Gettier's problem is the shift from 'p' to 'p or q'. Keeping in memory 'p or q' when one believes 'p' may possibly create trouble, as that may lead to an apparently justified belief which happens to be true.

The Principle which will avoid us many problems in the future is the following:

Epistemic Hygienics (EH): Do not accept sentences that could mislead your future selves.[12]

Other interesting examples by Douven are the following:

---

[12] A referee makes an important consideration and says that virtually it could be possible for any sentence to mislead one's future self (so we need a story about what it is that makes a sentence a candidate to mislead). Well, consider the sentence: 'John went to the cinema'. I may utter it having in mind the referent 'The Apollo', but if I memorize the sentence without associating the referent 'The Apollo' to the NP 'the cinema', I may end up in the future using the sentence to refer to 'The Odeon'. After all could not mistakes of this sort happen? To avoid the over-generation of entailments, we would probably have to keep in mind that we need to memorize not only abstract sentences, but sentences uttered in context, hence complete thoughts. Pragmatic intrusion is a good way to avoid the over-generation of possible entailments. Having done so, we still have to avoid those entailments which are likely to mislead our future selves. Of course, I should note that for Douven the problem is not an entailment 'per se', but the fact that when we commit things to memory, we could keep the entailments separate from the sentences that generated them and we could even end up, in extreme cases, admittedly, forgetting the sentences which generated those entailments, while retaining the entailments in question. And this is quite bad, because we will end up remembering things which are likely to mislead us and have negative consequences on action.

13. Peggy's car is blue;
14. Peggy's car is bluish.

Now, it is clear that if Peggy's car is blue, it is also bluish (blue being a stronger gradation of bluish). However, if one commits to memory 'Peggy's car is bluish' when one believes that it is blue, one will commit to memory a piece of information which may possibly mislead one's future self (Suppose that, in a couple of days, the same person is asked by Mary to say the color of Peggy's car; he says that it is bluish; then Mary is not able to identify Peggy's car in the office's garage. Some trouble has ensued). Douven compares memorizing or committing to memory to writing notes (e.g. Turn off the gas) which will be of use to our future selves. If memories are like notes, we should avoid writing notes that mislead our future selves.

Igor Douven's paper is of great importance to epistemology but also to pragmatics. He shows that pragmatics and epistemology are intimately connected. While Igor Douven's story can be interpreted in the light of more general principles of cognition (a memory that is misleading obviously is a case in which a believed assumption is more costly than beneficial in terms of cognitive effects; positive cognitive effects being those which put me in touch with reality, not those which drive me away from it), I cannot do this in this paper.

Now, to return to Wieland's case. How can we deal with it in terms of the pragmatics of belief by Igor Douven? If we accept:

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

it is clear that creating indirect reports by resorting to inferential steps that can mislead our future selves is illicit.

So, if on the basis of (12a), I make the indirect report (in (12b), I will be entitled in the future to expect that, on meeting Maryanne, Professor A will recognize her and say 'Hello' to her.[13] But this may never happen, if he passed her only by marking an anonymous paper. Nor should we expect that, being really impressed by her paper, on seeing a paper by Maryanne in the Journal of Philosophy, he will be able to connect this paper to his past positive experience (commenting "Oh, this

---

[13] A referee said that 12b does not entitle anyone to expect recognition. And, of course I agree that Professor A is not expected to recognize Marianne. But, given 12b, is not a hearer led to believe somehow that Professor A has someone in mind (possibly just the name and the thought that a person with that name has passed the exam)? However, minimal, this thought seems to have been conveyed. Of course, one could adjust the context somehow. Suppose everyone knows that Professor A does not look at the names on the papers, he just covers them (or asks his secretary to do that) with colored sellotape. Then he marks the papers and gives them to his secretary who assigns marks to individual students. This habit is so remarkable that professor A has become famous for this. Then, in this (heavily contrived) context, the utterance 'Professor A said that Marianne passed the exam' could be interpreted as 'Professor A passed Marianne's paper'. But, even with all this contextual adjustment, we have a feeling that 'Professor A passed Marianne's paper' and 'Professor A said that Marianne passed her exam' are very different utterances, as the latter implies somehow that Professor A said something of Marianne, that he had her or her name in mind, at some point.

is another paper by Maryanne). But all this makes sense, if we are aware that there is something strange in the practice allowing us to go from the first step of (12) to its second step.

The last case discussed by Wieland that is of considerable interest (presumably based on some cases I myself pointed out in Capone (2010a), as kindly noted by Wieland) is whether we should consider the literal or the metaphorical/indirect/ironic level as the basic level of content of an indirect report. Wieland seems to opt for the view that the content of an indirect report should be constituted by interpreted and not by literal segments of speech. Thus an utterance of (15)

15. Mary is a lioness

   should be reported as:

16. John said that Mary is brave.

However, I notice that it is not cases of metaphors that are particularly thorny, because here by reporting the literal level of meaning, one allows the hearer nevertheless to compute the indirect or not literal level of meaning. The most problematic cases are those of irony, because the context of the original utterance is missing (or may be missing) and thus the hearer cannot move from the literal to the ironic (or echoic) meaning. Thus the transition from (17) to (18) is not easy:

17. The talk was very good.
18. He said that the talk was really bad and he didn't like it much.

It appears that Wieland is uncontroversially moving towards a view of indirect reports in which the content of the indirect report is only the intended meaning, rather than the (possibly unintended) literal meaning[14] of the original utterance. Now, if such a view is accepted, indirect reports could NOT be used as Cappelen and Lepore (2005a) do as tests for literal meaning or minimal semantics. My impression is that in context we must settle whether an indirect report is a literal or a non-literal report. There is evidence in favor of both views. Given the fact that it is possible to use direct quotation, when we want to mention the words used, the use of an indirect report for the same purpose would ultimately obtain the same effects of a quotation, but with great processing efforts (as one will ultimately compare the quotative construction with the indirect report). However, given that indirect reports are often mixed with quotative segments and given that quotation is (as I claimed in Capone 2013) a radically pragmatic operation, it is possible in theory that an indirect report might overlap with a quotative structure (see also Burton-Roberts 2006)—which is what happens in the most thorny examples by Cappelen and Lepore. I will stop the discussion here, as I do not want it to slide into a discussion of Semantic Minimalism. In this paper, I am mainly interested in

---

[14] Of course, a literal meaning can be intended, in which case I think Wieland would have to accept that it is the content of an indirect report. If the literal meaning is not intended, in the sense that it is superseded by non-literal elements which are speaker-meant, then Wieland will not accept it as part of the content of the indirect report.

the polyphonic structure of indirect reports and it is this aim I have in mind throughout. The overlap between quotation and indirect reports amply attests to this polyphonic structure.

A case not discussed by Wieland, which certainly fits the typology of examples she proposes, is that of how to report an ungrammatical utterance. Surely we should ask ourselves whether correcting an ungrammatical original statement by proposing an indirect report from which the error has been removed (abiding by the Principle of Charity) results in altering drastically what the original speaker said and in such a way that s/he would not approve of the indirect report. And can indirect reporting with correction result in opacity, in that the indirect report purges the original speaker's thought of something that was essential to the thought? In other words, we want to establish whether opacity only rests on the impossibility of intersubstituting co-referential NPs or whether it also rests on the impossibility of intersubstituting coreferential sentences one of which is syntactically incorrect. Paradoxically, the case is not of importance for the illiterate speaker, who attaches little importance to grammar and who may even be unaware of the substitution. However, consider what happens when the original speaker is a grammarian and the original utterance is reported through an indirect report whose grammar exhibits an element with which the original speaker may take issue. Fidelity to the grammar of the original statement may well depend on the context. If we are in a context in which we have to assign marks depending on the grammatical correctness of what the original speaker said (suppose we are marking students' papers), even slightly improving the grammar of her original sentence in an indirect report may be considered unacceptable. In this case mixed quotation may be deemed necessary.

5. Indirect reports and quotation

While scholars are generally adamant that there is a clear-cut distinction between quotation and indirect reports, this paper is, in fact, blurring these two practices. And the result of blurring the two practices fits in with the idea that opacity is a phenomenon to be found both in quotations and in indirect reports. In fact, the Davidsonian treatment of indirect reports also involved the blurring of quotation and indirect reports, as the complementizer 'that' for Davidson was a demonstrative pronominal and the thing which followed the demonstrative pronominal could be easily assimilated to a quotation (which explained where the opacity came from) (See Baldwin's 1982 important considerations, which agree with this[15]). Current scholars try to keep apart indirect reports and quotation—and perhaps their practice is correct up to a point. However, doing so in a rigid manner

---

[15] Baldwin (1982, 273) writes: "Davidson argues against such quotational theories and thereby implies that his paratactic theory is not a quotational one. But he treats quotation as abbreviated spelling out, and if, more sensibly, one treats quotation marks as a demonstrative device, and one treats the symbols within the quotation marks as a display of that which is referred to by the demonstrative, then the difference between paratactic and quotational theories becomes one largely of notation."

would prevent us from understanding where opacity comes from in such cases. Even if we grant that indirect reports are not always opaque, we surely must concede that they are preponderantly opaque. And even if they were only sometimes opaque, we would still have the trouble of explaining where the opacity comes from. And of course, the opacity of indirect reports comes from the fact that quotation and indirect reports are similar to some extent, as invariably proven by the practice of mixed quotation (in indirect reports). I want to believe that mixed quotation is not just a quirk, something that occurs sometimes, but is something that occurs frequently, since I have accepted (Capone 2013) that quotation both in the oral and in the written language can dispense with quotation marks and can resort to pragmatic marking. Given that any segment of an indirect discourse could be marked pragmatically as being mixed quoted, it is clear that the analogies between quotation and indirect reports are quite striking.

Suppose that we accept what I said in Capone (2009) on cancellability of explicatures (namely that explicatures are NOT cancellable). Then if we have pragmatic clues leading us to interpret a linguistic item as enveloped in inverted commas, the quotational interpretation cannot be cancelled, but will amount to a pragmatic intrusion into truth-conditional meaning. From this it follows that it will not do to replace that linguistic item with another coextensive item, because, otherwise, the speaker's commitment to having uttered that thought will evaporate. But this is exactly what opacity amounts to. We cannot replace a linguistic item with a coextensive one, without expressing a different thought. However, we have obtained opacity through some pragmatic means. It is not exactly semantic opacity we are writing about (to be more precise).

But now I want to pursue this line of reasoning further. Consider taboo words, usually relating to sexual organs, etc. Scholars have insisted that, despite the fact that a speaker takes great pains to distance herself from the use of a taboo word, thanks to quotation, she cannot really manage to do so, and for some strange reason, still to be explained adequately, the taboo word is assigned to her voice as well. So, consider the following example:

19. Mary said that '….T…..'.

(Where T stands for a taboo word inserted within a sentential frame …..). Regardless of the framing device of quotation, the responsibility for the taboo word is assigned equally to Mary and the (direct) reporter. Now, we would expect the matter to be different in indirect reports. Given that 'that' is not a demonstrative pronominal (as the Davidsonian analysis has it), but only a complementizer, the that-clause should come from the perspective of the indirect reporter. Thus we could expect, if there was a real difference between direct quotation and indirect reports, that only the reporter would be responsible for the taboo word in the following utterance type:

20. Mary said that …..T….

But this expectation is not born out. We equally attribute the T word to the reporter and to the original speaker.[16] And we possibly attribute the gaffe to the original speaker to the same extent as to the reporter.[17] So things stand exactly in the same way, as far as obscenities and other taboo words are concerned. Now, given that we are willing to give similar analyses of the indirect reports and of the direct reports in these cases, it is clear that neither quotation marks nor the complementizer can prevent responsibility from being assigned to the reporter. The two different functions of the complementizer and of quotation marks would lead us to expect that quotation marks could be more protective for the reporter, but this is not the case. The presence of the complementizer in indirect reports would lead us to expect that the complementizer could be more **protective** for the original speaker, but this is not the case. And why not? The truth is that if quotation and concealed mixed quotation in indirect reports are triggered and interpreted pragmatically, then we have a pragmatic machinery capable of explaining why the responsibility of a certain segment of talk is assigned to the original speaker, or both to the original speaker and the reporter.

Now, at this point, we can go on using the machinery of indirect reports for direct quotations as well.

Paraphrasis Principle

The that-clause embedded in the verb 'say' is a paraphrasis of what Y said that meets the following constraint: should Y hear what X said he (Y) had said, he

---

[16] The reason for this is that the reporter could have chosen a different word (a more descriptive strategy), but he did not do so (thus he is guilty and responsible for the words reported). The original speaker is attributed the words by default, because the statement is about what he said, and the partial responsibility of the reporter cannot eradicate the topicality of the indirect report, as that is ABOUT the utterance proffered by the original speaker. If the original speaker did not utter those words, why should we report them, allowing the hearer to possibly attribute the words to the original speaker? A report that focuses on slurring or taboo words in case the original speaker did not use those words is infelicitous and irrelevant, since, do not forget, the purpose of an indirect report is to focus on what the original speaker said, NOT on what he did not say.

[17] Referee 2 says that we can easily imagine a reporter indirectly reporting an original utterance that contains a taboo word using a euphemism instead or pointing out heavily that the taboo word attaches explicitly to the original speaker in some way. The tendency of reporters to use descriptive euphemisms for taboo words, e.g. 'The S-word' for 'Shit' or the 'F-word' for 'fuck' suggests we attribute the taboo word to reporters more than to the original speaker. I think these considerations are illuminating. I certainly agree that in reporting a speaker is sensitive to certain rules (a prohibition against using or even mentioning taboo words). But the fact that the reporter is dissociating himself from the use of those words clearly is an indication that in the world shared by the reporter and the reported speaker using those words is prohibited. And if merely mentioning those words is prohibited, we can easily imagine that there is an even stronger prohibition against USING those words (and presumably those words were used and not only mentioned by the original speaker). So I agree with the referee only up to a point, as I take the fact that the reporter is distancing himself from the T-words to reveal a judgment against the use of those words (in the original utterance) in the first place.

would not take issue with it, but would approve of it as a **fair paraphrasis** of her original utterance.

Paraphrasis/Form Principle

The that-clause embedded in the verb 'say' is a paraphrasis of what Y said, and meets the following constraints: should Y hear what X said he (Y) had said, he would not take issue with it, as to content, but would approve of it as a fair paraphrasis of his original utterance. Furthermore, he would not object to vocalizing the assertion made out of the words following the complementizer 'that' on account of its form/style.

According to these two principles, we can explain why the responsibility for the obscenity is assigned to the original speaker both in the case of indirect reports and in the case of direct quotation. Of course the case of indirect reports flows easily and directly from the principles above. In the case of quotation, we need a D-tour. It is pragmatics that assigns the obscenity to the original speaker, by marking a segment as being quoted, since the point of the quotation is to assign her those words. It follows that if the pragmatics of quotation is ok, the original speaker would approve of the utterance that is being attributed to her.

Now, why is it that the reporter (both the direct and the indirect reporter) is guilty of obscenity? Why is it that the quotation marks do not protect her? And the answer is obvious. The reporter could have avoided reporting the locution and could have found ways of expressing the content in such a way that the content as well the obscenity could be perceived, without **depicting** the obscenity but by describing it.[18] In this way, she would have dissociated herself (her voice) from the voicing of the obscenity. Now, in the indirect report, the original speaker is guilty of the obscenity to a greater extent because a segment of the indirect report is being mixed-quoted through the pragmatic machinery. Nevertheless, the reporter is responsible for the obscenity—even if to a smaller extent—because she could have reported the content by describing the obscenity rather than by depicting it. Since she preferred depicting to describing, he must be deemed guilty of not sparing the hearer the embarrassment of hearing the obscenity.

5.1 Michel Seymour (1994) on indirect discourse and quotation

My approach to indirect reports is reminiscent of the ideas expressed by Seymour (1994), which is a unique and, in my view, important paper on the close connections between indirect reports and quotation. Seymour is ambivalent between quotation proper and a domesticated view of quotation in which the quoted sentence describes an act of saying in the direct sense, but translates it according to the conceptual scheme of the reporter's translational manual. Now, if I am correct, Seymour allows a mixture of elements which reflect the quoted person's voice and elements which reflect the reporter's conceptual translation manual. So, if the (English) reporter reports 'She said that Mary went to Rome', it

---

[18] See also referee 2's considerations, voiced in footnote 16.

is possible that the original speaker used 'Maria' and 'Roma' in her utterance, but these are translated as 'Mary' and 'Rome'. The basic structure and content of the quoted item is the same, but certain interpolations were made. Now I believe that the great merit of Seymour's analysis lies in making us see that indirect reports are (normally) a blend of quotation and pure indirect reports. Pure indirect reports only represent a schematic summary made by the reporter of what the reported speaker said. The blended report couples this summary with a quotation structure, or couples a quotation structure with a use of the same sentence as if it was not quoted. In my view pure indirect reports do not exist or represent an abstract ideal, while the quotative approach to indirect reports (the approach according to which there are implicit quotation marks provided pragmatically inside the that-clause) is what I accept fully, provided that we accept that speakers and hearers rely on a pragmatic machinery allowing them to distinguish voices in the indirect report. So, does Seymour definitely abandon the Davidsonian analysis? Clearly he does not, since he blends a sentential approach (one that considers the sentence reported as if in quotation marks) with a paraphrase approach (in which content (regardless of the words used) is of paramount importance), and, furthermore, he accepts that indirect reports rest on a semantic theory based on the concept of truth, of systematicity and recursiveness. The fact that Seymour's (as well as my view) is a blend of the praraphase and sentential theories does not prevent the theory from being based on truth, since both paraphrase and quotation are structures which can be evaluated truth-conditionally. The theory is clearly systematic—being based on an abstract linguistic system that works through compositionality. And it is recursive, since it is possible to apply the same semantic rules recursively (John said that Mary said that Robert said that…).

My views, however, diverge from Michel Seymour in at least an important respect. My analysis of 'quotation' does not involve/presuppose (like his) a view based on names and is clearly based on a more developed view of quotation, say the one based on Recanati (2010) and the one I developed in Capone (2013), which is radical in claiming that pragmatics only is involved in deciding what the thing quoted is (a lexeme, a phonetic form, a written form, something somebody said, etc.). The other important difference is that I do not attach special importance to the ambiguity (whether semantic or interpretative, but I assume it makes sense to claim it is interpretative) between a sense of 'X said that' that is that of indirect reporting the content of what another person said and another sense which amounts to a special interpretation of quotation: in reporting 'X said that p', one is basically saying that there is a proposition p, such that X said 'p' and the content of 'p' is given (translated) by the sentence uttered by X.

6. **Igor Douven's point of view.**

Reacting to my paper, Igor Douven (personal communication) writes the following:

> I was wondering whether the paraphrase principles do not give too much weight to the speaker's approval. Couldn't a speaker have ulterior motives for disapproving some

paraphrase, even if an impartial third party would approve of it? Perhaps the speaker regrets what he or she said. Or the speaker has a false memory about what he/she said and is perfectly honest (though mistaken, as seen from an impartial standpoint) in disagreeing with the paraphrase.

I was also wondering whether it would be worth trying to adopt instead of the paraphrase principles a principle like the following, which would connect to the current debate about contextualism in epistemology: 'S said that p' is true iff by an assertion of that sentence the hearer comes to know what S said. As various epistemologists have argued, the standards for knowledge may vary with context. In some contexts, not much evidence is needed to gain knowledge; in other contexts, a lot of evidence is needed; and of course there are all sorts of intermediate cases. This might explain why in some contexts we think a loose paraphrase of what someone said is OK, while in others we feel that the speaker should stay very close to the original speaker's wording.

Ok. Consider the case in which disapproval comes because the speaker regrets having said what she said. We are not worried about this case, because if the speaker is honest enough, he must accept that his prior self would have approved (regardless of the reservations by his current self) the sentence reported in the indirect report, if it reflected the form and the thought he expressed in the original utterance.

The fact that a speaker disapproves what he said because she has a false memory of what she said does not worry us either—as we may confine ourselves to the case in which the original speaker remembers well what she said.

Considering the second part of Douven's comments, I am sympathetic towards a contextualist view of the matter. Presumably Douven connects 'X said p' with knowledge of what X said on the part of the hearer. Transforming the issue of indirect reports into an epistemic issue amounts to bringing in contextualism. According to Contextualists (e.g. Keith De Rose 2009) the truth of a knowledge claim may depend on the amount of evidence required to assess it. In some contexts, we need a greater amount of evidence for the truth of 'X knows that p'. In other contexts, we need an inferior amount of evidence. In high stake contexts, the evidence needed is superior than the one needed in low stakes contexts. Analogously, in high stake contexts, we could say that the Paraphrase Principle is adhered to more strictly than in low stakes contexts. But this is not the only case in which we need to depart somehow from the Paraphrase principle. I have already discussed the case in which a speaker may be interested in letting the hearer identify a referent and thus may use a mode of presentation of the reference distinct from the one used or approvable by the original speaker. This situation is not linked to contextualism in a theory of knowledge, as the mode of presentation is different regardless of whether we are in a high stake or a low stake situation. Presumably, however, Igor Douven would want to say that we are in a low stake situations and this explains why the reporter is inclined to modify the mode of presentation used by the original speaker.

Now there are cogent reasons to be sympathetic to Igor Douven's treatment, even if a modification of his way of putting things is required. I propose to modify his assertion:

'S said that p' is true iff by an assertion of that sentence the hearer comes to know what S said.

I prefer, instead:

an assertion of 'S said that p' is felicitous iff by an assertion of that sentence the hearer comes to know what S said.

Should the Paraphrase principle be abandoned then? Perhaps a reformulation is needed that links it to high stakes contexts. Alternatively, one could opt for the position that assertions of 'X said that p' which depart from the Paraphrase Principle are parasitic or loose uses. This would give greater legitimacy to the Paraphrase Principle while admitting that in some contexts we may depart from it somehow.

7. **Slurring**

If the considerations above on taboo words relating to the sexual sphere are correct, we would expect an analogy to work between taboo words in general and slurring. Slurring—to take up ideas by Lepore and Anderson (2013) amounts to using words that are derogatory and offend vast categories of people (usually minorities) such as Jews, Chinese (in USA), black people, homosexuals, etc. Our problem is not slurring *per se*, but what effects does slurring have on quotation and on indirect reports. Lepore and Anderson mainly deal with indirect reports—which use plugs such as the verb 'say'—but it is clear that indirect reports and quotations work in a parallel way when slurring is embedded in the quotation or indirect report structure. Lepore and Anderson reject the view that slurring persists in indirect reports (in that the reporter is being assigned responsibility for the slurring, rather than the original speaker) because of a conventional implicature (Williamson 2007) or because of a presupposition (see Williamson 2007 for discussion). Presuppositions usually do not escape verbs of saying, which are called 'plugs' because they tend to block presuppositions (see also Levinson 1983). But then slurs behave unlike presuppositions because they can survive embedding in plugs (even if they often survive embedding in negation, if-clauses, etc. like most presuppositions). Of course Lepore and Anderson do not consider a pragmatic view of presupposition (along the lines of Simons 2013), according to which, at least in several cases, presuppositions are projected through conversational implicatures (but then, in this case they are not presuppositions but conversational implicatures). We know how Lepore and Anderson would reply to a possible objection by Simons. If the persistence of the slurring is due to a conversational implicature, first of all we should account for the implicature through a pragmatic story. Second, the implicature would have to be cancellable, at least in some contexts. And yet we see that the implicature can hardly be cancelled, although it may be mitigated to some extent say in scientific contexts in which the writer makes it absolutely clear that her purpose in dealing with the prohibited word is scientific. If only mitigation is obtained through contextual variation, then it is hardly the case of a conversational implicature. The case against conventional implicature is more thorny. As usual, we are interested in cases of plugs, such as:

21. John said that Mary is obstinate but brave (however, I do not personally see any contrast between being obstinate and being brave).

Plugs do not make the conventional implicature disappear completely, as the speaker of (21) presumably accepts that for someone it must be true that there is a contrast between being obstinate and being brave. However, they demote it from the epistemic commitments of the speaker. Thus, Lepore and Anderson are justified in holding that slurring cannot be a matter of conventional implicature.

One of the properties that characterises slurring is its persistence despite self-correction. This shows that the entailments of the slurring word cannot be un-said; and in this respect the slurring words are different from other words, which allow self-corrections. I could say 'Mary is tall. Oh, sorry, I meant short, I got confused'. These corrections are put up with in the oral language (less in written texts). But with slurring words, this is not the case, and no replacement or correction can repair the slurring which was caused by using a slurring word. Consider, in fact, the following:

22. Look at what that negro is doing—oh, I mean that black gentleman.

A repair like the one in (22) seems to make things worse, because it tends to add an ironic interpretation.[19]

Lepore and Anderson discuss at length the word 'Negro'—but they do not discuss—not even *en passant*—that in the past the word 'Negro' seemed to be acceptable or usable in American English. Consider for instance the 'I have a dream' speech by M. L. King. I was myself perplexed by such uses. Would they count as uses involving camaraderie among blacks or are they echoic uses to be wrapped in inverted commas? (It is possible to oscillate between the two views). It is difficult to answer this question in the context of this paper, as it involves diachronic considerations too. However, if there is at least one such context in which the slurring word, wrapped up by quotation marks, does not count as slurring, one could opt for a conversational implicature. So, the only cards on the table are the following: a conversational implicature and a rule of use. The rule of use view has been advocated by Lepore and Anderson. They claim that there is a prohibition against using slurring words. Of course, this prohibition works for the groups outside the potentially slurred groups. So, there is no prohibition for members of the slurred groups against using a slurring word. This could explain well why the contextual variation has such powerful transformative effects on the slurring potential of the work 'negro' or 'queer'. The conversational implicature view would no longer be needed—or could count as an alternative view having more or less the same explanatory power. But what would the conversational implicature view amount to? Without going into details, it would have to say that

---

[19] Kennedy (2002, 19) writes about the word 'negro': 'nigger' is an ugly, evil, irredeemable word. He cites someone considering the word "the nuclear bomb of racial epithets" (p. 61).

certain words are slurs in ordinary contexts where the speaker speaks for herself (and no direct report or quotative structure is involved), and they are slurs presumably because there is a societal rule against the use of these words. Then it would have to explain, on the basis of this general prohibition, why inverted commas or indirect reports do not rescind the responsibility of the indirect reporter from that of the original speaker who presumably is responsible for slurring. But now the conversational implicature view is parasitic on the rule of use advocated by Lepore and Anderson. So, it would be simpler to hold that the rule of use based on a societal Prohibition works both for the original speaker and the reporter. But if it was a rule of use, how can we explain the fact that quotation marks do not rescind the responsibility of the reporter from that of the original speaker? After all, it is commonly held that quotation involves mentioning (at least in semantic textbooks such as Lyons 1977). If it involves mentioning, why should a rule of use be applicable to the reporter? Clearly indirect reports do not pose a serious threat to Lepore and Anderson because it might be claimed by theorists that the complemetizer 'that' need not work like a demonstrative pronominal and the indirect reporter can be considered as one who uses the words in the that-clause, at least partially. What I have said before about the parallel considerations on quotation and indirect reports discourage us from this Pyrrhic victory, so cheaply obtained. I claimed that in indirect reports too the hearer is faced with the thorny task of separating the original speaker's from the reporter's voice. Thus, it is not impossible, especially in the presence of appropriate clues, to consider the slurring words of the indirect report as being embedded in inverted commas (in this case the original speaker would have to accept responsibility for the slurring). So the problem raised by quotation is not trivial. The rule of use advocated by Lepore and Anderson does not seem to work well, first of all because quotation structures as well as indirect reports intended as having a quotative structure do not allow us to pass the theory based on a rule of use (a prohibition), as the original speaker could be assigned major responsibility for uttering the slurring words. Second, we need to note that contrary to Lepore and Anderson, who claim that indirect reports containing slurring words assign greater responsibility for the slurring to the reporter than to the original speaker,[20] I claim that, if anything, a pragmatic theory like the one voiced in Capone (2010a) makes it the case that the original speaker has responsibility too. So, we need a pragmatic machinery like the one expressed in:

---

[20] Lepore and Anderson (2013) write that "Indirect reports and other attitudinal inscriptions fail to attribute slurring to whomever they report since the offense of the reporter "screens off", so to speak, the offense of whoever is being reported. This position is interesting, but needless to say, it would need greater justification.

Paraphrasis Principle[21]

The that-clause embedded in the verb 'say' is a paraphrasis of what Y said that meets the following constraint: should Y hear what X said he (Y) had said, he would not take issue with it, but would approve of it as a **fair paraphrasis** of her original utterance.

Paraphrasis/Form Principle

The that-clause embedded in the verb 'say' is a paraphrasis of what Y said, and meets the following constraints: should Y hear what X said he (Y) had said, he would not take issue with it, as to content, but would approve of it as a fair paraphrasis of his original utterance. Furthermore, he would not object to vocalizing the assertion made out of the words following the complementizer 'that' on account of its form/style.

Now, these principles would allow us to assign the original speaker the principal responsibility for the slurring, taking for granted or presupposing Lepore and Anderson's rule of use (or prohibition). The reporting speaker, given such a use, is guilty of not having used an alternative word[22] or a description, rather than a segment which has depictive properties. Given that she has not avoided the slurring word, when she obviously could do so, she herself becomes responsible for the slurring. But now we have explained why the pragmatic explanation, despite being parasitic on Lepore and Anderson's rule of use, does more work than the original explanation by Lepore and Anderson. Thus, it could be recommended by Modified Occam's Razor, because even if Lepore and Anderson's view appears to be simpler, it cannot explain what the conversational implicature view—which is more complex—does explain.

Objection. Why should the reporter have to use some form of substitution of the slurring in question, if after all the devices of quoting and of mix-quoting in indirect reports allow her to avoid responsibility, since after all quoting does not amount to using a certain expression. The reply is simple. It is true that the reporter is not using the slurring in question and, therefore, cannot be accused of having **used** a slurring word. However, in depicting the slurring, rather than describing it by a suitable transformation and by some descriptive phrase alluding to the

---

[21] This position is somewhat reminiscent of Seymour's (1994) treatment of in direct reports, in which reference to a translation of the reported sentence is explicitly incorporated in the semantics of in direct reports.

[22] It is difficult to suggest which alternative is more neutral than a slurring term. Baugh (1991) frankly points out that Americans find it difficult to find a term which is not insulting or less insulting that 'negro'. The term 'black' used to be offensive in the past, but no longer is. The term 'coloured' used to be acceptable in the past but is now offensive. Presumably the least connotative is 'African American' a term which the Reverend Jesse Jackson managed to introduce into American's public life. However, as Du Bois (1928) stresses, if hatred and despise target a certain social group, then it will survive despite the fact that new names replaced the old ones associated with negative connotations. However, I want to point out that the term 'African American' is destined to be successful because it avoids all reference to the color of the skin. So, ideologically, it is much better than many other names. It voices the desire NOT to be classified by color and a rejection of the old stereotype that people should be classified by color.

slurring character of the original phrase, the reporter is signalling some **complicity** since she is not distancing herself from the trespasser (the original speaker). Since using depictive elements involves taking the shortest route in the description process, when there is an alternative route which by embarking on a transformation involves greater processing efforts (and production efforts), it is clear that the avoidance of greater processing costs is taken as a sign of complicity, while the more costly transformation is taken (or would be taken) as a way of signalling that one is distancing oneself from the offensive segment of talk. We could consider 'complicity' a language game, in which two voices blend in case they share the same point of view. While in the normal case in which two speakers have different points of view, they tend to differentiate their voices, in the case of complicity two voices are presented as undifferentiated. Indirect reports are prototypical cases in which an utterance gives expression to two voices, the original speaker and the reporter. Thus, it goes without saying that an indirect report should present two slots in case the original speaker's voice and the reporter's voice are differentiated and only one slot in case the two voices blend (being undifferentiated). The presence of just one slot, instead of two slots clearly exhibits the complicity between the two voices. Of course, readers may ask, how can we have two or just one slot for voices in indirect reports? Is this a semantic or a pragmatic matter? The natural answer is that the slots are provided pragmatically and should be considered not as distinct syntactic positions but as portions of text where we can reveal implicit quotation marks. It is the rich structure of cues and clues which will point towards two or, rather, one slot capable of expressing point of view.

2 Conclusion

In this chapter, I have deepened my view that indirect reports are cases of language games. I have mainly explored the analogies between quotation and indirect reports, and I have maintained that such analogies allow a parallel pragmatic treatment. In the end, I have concentrated on slurring and I have explained why both taboo words and slurring words cannot be embedded in quotation structures without losing their anti-social status. It is clear that slurring too involves the task of separating voices and of accepting the essentially polyphonic structure of discourse. Essentially the problem, in our case, is how it comes about that when someone reports a slurring expression, there are in fact at least two people—and not just one—doing the slurring. This is a complicated but interesting question, which puts to the test both the theory of quotation and that of indirect reports, throwing light on parallel problems about polyphony and the way it is supported by conversational implicature.

References

Baldwin, T. (1982). Prior and Davidson on indirect reports. *Philosophical Studies, 42*(2), 255–282.

Baugh, J. (1991). The politicization of changing terms of self-reference among American slave descendants. *American Speech, 66*(2), 133–146.

Burton-Roberts, N. (2006). Cancellation and intention. Newcastle University, School of English, Mn. Published in the Newcastle working papers in linguistics. In E. Romero & B. Soria (Eds.), *Explicit communication: Robyn Carston's pragmatics* (2010) (pp. 138–155). London: Palgrave-Macmillan.

Capone, A. (2008). Belief reports and pragmatic intrusion: The case of null appositives. *Journal of Pragmatics, 40*, 1019–1040.

Capone, A. (2009). Are explicatures cancellable? Toward a theory of the speaker's intentionality. *Intercultural Pragmatics, 6*(1), 55–83.

Capone, A. (2010a). On the social practice of indirect reports (further advances in the theory of pragmemes). *Journal of Pragmatics, 42*, 377–391.

Capone, A. (2010b). What can modularity of mind tell us about the semantics/pragmatics debate. *Australian Journal of Linguistics, 30*(4), 497–520.

Capone, A. (2012). Indirect reports as language games. *Pragmatics and Cognition*, *20*(3), 593–613.

Capone, A. (2013). The pragmatics of quotation. *Pragmatics and Society*, *4*(3), 259–284

Cappelen, H., & Lepore, E. (2005a). *Insensitive semantics: A defence of semantic minimalism and speech act pluralism*. Oxford: Blackwell.

Cappelen, H., & Lepore, E. (2005b). Varieties of quotation revisited. *Belgian Journal of Philosophy, 17*, 51–76.

Carston, R. (2002). *Thoughts and utterances. The pragmatics of explicit communication*. Oxford: Blackwell.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language, 66*(4), 764–805.

Dascal, M., & Weizman, E. (1987). Contextual exploitation of interpretation clues in text understanding: An integrated model. In J. Verschueren & M. Bertuccelli-Papi (Eds.), *The pragmatic perspective* (pp. 31–46). Amsterdam: John Benjamins.

Dascal, M., Hintikka, J., & Lorenz, K. (1996). Games in language. In M. Dascal, D. Gerhardus & K. Lorenz (Eds.), *Sprachphilosophie* (vol. II, pp. 1371–1392). Mouton De Gruyter.

Davidson, D. (1968). On saying that. *Synthese, 19*, 130–146.

De Rose, K. (2009). *The case for contextualism*. Oxford: OUP.

Douven, I. (2010). The pragmatics of belief. *Journal of Pragmatics, 42*, 35–47.

Du Bois, W. E. B. (1928). The name 'Negro'. *The Crisis, 35*, 96–97.

Goffman, E. (1974). *Frame analysis*. New York: Harper & Row.

Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Kennedy, R. (2002). *'Nigger': The strange career of a troublesome word*. New York: Pantheon.

Lepore, E., Anderson, L. (2013). Slurring words. *Noûs, 47*(1), 25–48.

Levinson, S. (1983). *Pragmatics*. Cambridge: CUP.

Lyons, J. (1977). *Semantics, 1-2*. Cambridge: CUP.

Potts, C. (2005). *The logic of conventional implicatures*. Cambridge: CUP.

Recanati, F. (2010). *Truth-conditional pragmatics*. Oxford: Oxford University Press.

Rumfitt, I. (1993). Content and context: The paratactic theory revisited and revised. *Mind, 102*(407), 429–454.

Saka, P. (1998). Quotation and the use-mention distinction. *Mind, 107*(425), 113–135.

Seymour, M. (1994). Indirect discourse and quotation. *Philosophical Studies, 74*(1), 1–38.

Simons, M. (2013). Mandy Simons, on the conversational basis of some presuppositions. In A. Capone, F. Lo Piparo & M. Carapezza (Eds.), *Perspectives on pragmatics and philosophy*. Dordrecht: Springer.

Sperber, D., Wilson, D. (1986). *Relevance theory*. Oxford: Blackwell. Reprinted with postface in 1995.

Tannen, D. (1989). *Talking voices. Repetition, dialogue, and imagery in conversational discourse*. Cambridge: CUP.

Vološinov, V. N. (1971). Reported speech. In L. Matejka & K. Pomorsko (Eds.), *Readings in Russian poetics* (pp. 149–175). Cambridge, Mass.: MIT Press.

Wieland, N. (2010). Minimal propositions and real world utterances. *Philosophical Studies, 148*, 401–412.

Wieland, N. (2013). Indirect reports and pragmatics. In A. Capone, F. Piparo & M. Carapezza (Eds.), *Perspectives on pragmatics and Philosophy*. Dordrecht: Springer.

Williamson, T. (2007). Reference, inference and the semantics of pejoratives. In J. Almog & P. Leonardi (Eds.), *Festschrift for David Kaplan*. Oxford: OUP.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.

Grammars as Processes for Interactive Language Use: Incrementality and the Emergence of Joint Intentionality

Eleni Gregoromichelaki and Ruth Kempson

Abstract Recent research in the formal modelling of dialogue has led to the conclusion that bifurcations like language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation are all based on an arbitrary and ultimately mistaken dichotomy, one that obscures the unitary nature of the phenomena because it insists on a view of *grammar* that ignores essential features of natural language (NL) processing. The subsequent radical shift towards a conception of NL grammars as procedures for enabling interaction in context (Kempson et al. 2009a, b) now raises a host of psychological and philosophical issues: The ability of dialogue participants to take on or hand over utterances mid-sentence raises doubts as to the constitutive status of Gricean intention-recognition as a fundamental mechanism in communication. Instead, the view that emerges, rather than relying on mind-reading and cognitive state metarepresentational capacities, entails a reconsideration of the notion of communication and a non-individualistic view on meaning. Coordination/alignment/intersubjectivity among dialogue participants is now seen as relying on low-level mechanisms like the grammar (appropriately conceived).

1 Introduction

Following Chomsky ([1965](#)), there has been a widespread perception, until recently, that formal accounts of natural language (NL) grammars must be grounded in the description of sentence-strings without any reflection of the

E. Gregoromichelaki (✉) · R. Kempson
King's College London, London, UK
e-mail: elenigregor@kcl.ac.uk

R. Kempson
e-mail: ruth.kempson@kcl.ac.uk

dynamics of language performance. Departures from this anti-functionalist methodology were rejected on the basis that language use is often disfluent and disorderly, hence presumed to preclude rigorous systematization, a stance independently propounded by the antiformalist approach of Ordinary Language philosophy (Austin 1975) and followed up by many theoretical approaches to pragmatics. However, structural, formal accounts consistent with performance considerations are now being considered (see e.g. Newmeyer 2010), as witness the huge growth in context-modelling and information update in formal semantics since the development of DRT and related frameworks. However, when required to interface with standard grammar formalisms, these developments in formal semantics/pragmatics are now beginning to show that the standard methodological dichotomies, e.g. language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation seem problematic. This is because all phenomena of NL context-dependency are explainable only by bifurcating them into grammar-internal versus grammar-external/discourse processes. This is because NL grammars are, on the one hand, taken to be limited to phenomena occurring within sentence boundaries but, on the other, unable to reflect the incremental word-by-word comprehension and production at the subsentential domain. However, context-dependency phenomena—anaphora, ellipsis, tense-construal, quantification, etc.—all allow unified ways of resolving how they are to be understood within and across sentence boundaries and even across distinct interlocutor turns in dialogue (Purver et al. 2009; Gregoromichelaki et al. 2011). And these update mechanisms are constrained at all levels by the incremental nature of processing. Hence, in this chapter, we suggest that these bifurcations -language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation- are all based on an arbitrary and ultimately mistaken dichotomy of phenomena, one that obscures their unitary nature because it insists on a view of grammar that ignores essential features of NL processing like incremental update.

As a response to such considerations, grammatical models have recently begun to appear that reflect aspects of performance to varying degrees (e.g. Purver 2006; Fernandez 2006; Ginzburg 2012; Gregoromichelaki et al. 2011; Hawkins 2004; Phillips 1996; Sturt and Lombardo 2005; Ginzburg and Cooper 2004; Kempson et al. 2001; Cann et al. 2005). One such model, *Dynamic Syntax* (DS), has the distinctive characteristic of taking a fundamental feature of real-time processing—the concept of underspecification and incremental goal-directed update—as the basis for grammar formulation. This shift of perspective has enabled the modelling of core syntactic phenomena as well as phenomena at the syntax-semantics-pragmatics interface in a unified and hence explanatory way (see e.g. Kempson et al. 2001; Cann et al. 2005; Kempson et al. 2011b). Moreover, instead of ignoring dialogue data as beyond the remit of grammars, DS takes the view that joint-construal of meaning in dialogue is fundamentally based on the same mechanisms underlying language structure: structure is built through incremental procedures, that integrate context in every step, and this provides principled explanations for the syntactic properties of linguistic signals; but, in addition, since

the grammar licenses partial, incrementally constructed structures, speakers can start an utterance without a fully formed intention/plan as to how it will develop relying on feedback from the hearer to shape their utterance and its construal and this provides the basis for the joint derivation of structures, meaning and action in dialogue. Thus, with grammar mechanisms defined as inducing growth of information and sustaining interactivity, the availability of derivations for genuine dialogue phenomena from within the grammar shows how core dialogue activities can take place without any other-party meta-representation at all. From this point of view then, communication is not definitionally the full-blooded intention-recognising activity presumed by Gricean and post-Gricean accounts. This then leads to questions regarding fundamental notions in philosophy and pragmatics, namely, the status of notions like intentions, common ground and linguistic versus extra-linguistic knowledge and their role in communication. We turn to examine those questions next.

2 Rethinking Intentionalism[1] in Communication

2.1 Intentions, Common Ground and Communication

The noted discrepancies between the representations delivered by the *grammar*, i.e. syntax/semantics mappings ('sentence meaning' or encoded content), and 'speaker meaning' (conveyed content) led to Grice's account of *meaning$_{NN}$*, (Grice 1975) to become the point of departure for many subsequent pragmatic models (see Levinson 1983; Bach 1997; Bach and Harnish 1982; Cohen et al. 1990, Searle 1969, 1983 a.o.).[2] From this point of view, it has been seen as necessary that, beyond some modular linguistic knowledge, communication should essentially involve notions of rationality and cooperation. In certain versions, this is displayed by the requirement that interpretation must be guided by reasoning about mental states: speaker's meaning, whose recovery is elevated as the fundamental criterion for successful communication, involves the speaker, at minimum, (a) having the intention of producing a response (e.g. belief) in the addressee (i.e. having a thought about the addressee's thoughts) and (b) also having a higher order intention regarding the addressee's belief about the speaker's second order thought (in order to capture the presumed fulfilment of the communicative intention by means of its recognition). Under this definition, speakers must, in order to

---

[1] The term is from Levinson (1995: 228) denoting the view that any kind of interaction involves an attribution of meaning or intention to the other.

[2] Note that our arguments here do not necessarily concern Grice's philosophical account, in so far as it is seen by some as just normative, but its employment in subsequent (psychological/computational) models of communication/pragmatics.

communicate, have (at least) fourth order thoughts and hearers must recover the speaker's meaning through reasoning about these thoughts.

   Millikan (1984: Chap. 3, 2005) argues that the standard Gricean view, with its heavy emphasis on mind-reading (see Cummings, this volume) over-intellectualises communication. Unlike the Gricean conception of meaning$_{NN}$ which rules out causal effects on the audience, e.g. involuntary responses in the hearer, Millikan's account, to the contrary, examines language and communication on the basis of phenomena studied by evolutionary biology, with linguistic understanding seen as analogous to direct perception rather than reasoning (see also McDowell 1980)[3]: Objects of ordinary perception, e.g. vision, are no less abstract than linguistic meanings, both requiring contextual enrichment through processing of the incoming data in order to be comprehended. Yet, in the case of ordinary perception, this processing does not require any consideration of someone else's intention. An analogous assumption can then be made as regards linguistic understanding, so that the resolution of underspecified input in context does not require considering interlocutors' mental states as a necessary ingredient. Millikan then provides an account of linguistic meaning in a continuum with natural meaning based on the *function* that linguistic devices have been selected to perform (their survival value). These functions are defined through what linguistic entities are supposed to do (not what they normally do or are disposed to do) so that "function", in Millikan's sense, becomes a normative notion. Norms of language, "conventions", are uses that had survival value, and meaning is thus equated with function. In contrast then to accounts of intentional action which see the structures involved as distinctive of rational agents, distinguishing them from entities exhibiting merely purposive behaviour (see, e.g. Bratman 1999: 5), in Millikan's naturalistic perspective, function, i.e. meaning, does not depend upon speaker intentions. Nonetheless, speakers indeed can be conceived as behaving purposefully in producing tokens of linguistic devices (as hearts and kidneys behave purposefully) but without representing hearers' mental states or having intentions about hearers' mental states (see also Csibra and Gergely 1998; Csibra 2008). Similarly, hearers understand speech through direct perception of what the speech is about without necessary reflection on speaker intentions.[4,5]

---

[3] The strict dichotomy between "meaning$_{NN}$" and "showing" has also been disputed within Relevance Theory (see, e.g., Wharton 2003).

[4] Of course, adults can, and often do, use reflections about the interlocutor's mental states; but the point is that this is not a necessary ingredient for meaningful interaction. Gricean mechanisms, that is, can be invoked but only as derivative or in cases of failure of the normal functioning of the primary mechanisms involved in the recovery of meaning, such as deception, specialised domains of discourse etc.

[5] An alternative account of communication combining Gricean and Millikanesque perspectives is that of Recanati (2004), which makes Gricean higher-order intention recognition a prerequisite only for implicature reconstruction. For what he terms "primary processes", on the other hand, Recanati adopts Millikan's account of understanding-as-direct-perception for the pragmatic processes that are involved in the determination of the truth-conditional content of an underspecified linguistic signal. These processes are blind and mechanical relying on

Early on, philosophers like Strawson (1964) and Schiffer (1972) severally presented scenarios where the criterion of higher-order intention recognition was satisfied even though this still was not sufficient for the cases to be characterised as instances of "communication" (as opposed to covert manipulation, "sneaky intentions" etc.). This led to the postulation of successively higher levels of intention recognition as a prerequisite for communication, and an attendant concept of "mutual knowledge" of speaker's intentions, both of which were recognised as facing a charge of infinite regress (see e.g. Sperber and Wilson 1995: 256–77). Although in applications of this account in psychological implementations it is not necessary to assume that explicit reasoning takes place online, nevertheless, an inferentially-driven account of communication on this basis has to provide a model that explicates the concept of 'understanding' as effectively analysed through an inferential system that implements these assumptions (see e.g. Allott 2005). So, even though such a system can be based on heuristics that short-circuit complex chains of inference (Grice 2001: 17), the logical structure of the derivation of an output has to be transparent if the implementation of that model is to be appropriately faithful (see e.g. Grice 1981: 187 on the 'calculability' of implicatures). Agents that are not capable of grasping this logical structure independently cannot be taken to be motivated by such computations, except as an idealisation pending a more explicit account. On the other hand, ignoring in principle the actual mechanisms that implement such a system as a competence/performance issue, or an issue involving Marr's (Marr 1982) computational versus the algorithmic and implementational levels of analysis (see e.g. Stone 2005, 2004; Geurts 2010) does not shield one from charges of psychological implausibility: if the same effects can be accounted for with standard psychological mechanisms, without appeal to the complex model, then, by Occam's razor, such an account would be preferable, especially if subtle divergent predictions can be uncovered (as in e.g. Horton and Gerrig 2005).

In this respect then, a range of psycholinguistic research suggests that recognition of intentions is an unduly strong psychological condition to impose as a prerequisite to effective communication. First, there is the problem of autism and related disorders. Autism, despite being reliably associated with inability (or at least markedly reduced capacity) to envisage other people's mental states, is not a syndrome precluding first-language learning in high-functioning individuals (Glüer and Pagin 2003). Secondly, language acquisition across children is established well before the onset of ability to recognise higher-order intentions (Wellman et al. 2001), as evidenced by the so-called 'false-belief task' which necessitates the child distinguishing what they believe from what others believe (Perner 1991). Given that language-learning takes place very largely through the

---

(Footnote 5 continued)

'accessibility' so that no inference or reflection of speaker's intentions and beliefs is required. It is only at a second stage, for the derivation of implicatures, that genuine reasoning about mental states comes into play.

medium of conversational dialogue, these results appear to show that at least communication with and by children cannot rely on higher-order intention recognition.

Such evidence has led to a move within Relevance Theory (RT) (Sperber and Wilson 1995) weakening further its Gricean assumptions (Breheny 2006). The RT view of communication is that the content of an utterance is established by a hearer relative to what the speaker could have intended (relative also to a concept of 'mutual manifestness' of background assumptions). This explanation involves meta-representation of other people's thoughts, but the process of understanding is effected by a mental module enabling hypothesis construction about speaker intentions. As noted by RT researchers, along with the communicated propositions, the context for interpretation falls under the speaker's communicative intention and the hearer selects it (in the form of a set of conceptual representations) on this basis. So, even though, unlike common ground, mutual manifestness of assumptions is in principle computable by conversational participants, and the interpretation process is not a "rational" one in the sense of Grice (cf. Allott 2008), it still remains the case that speaker meaning and intention are the guiding interpretive criteria which are implemented on mechanisms that have evolved to effect mind-reading. For this reason, Breheny argues that children in the initial stages of language acquisition communicate relative to a weaker 'naive-optimism' strategy in which some context-established interpretation is simply presumed to match the speaker's intention, only coming to communicate in the full sense substantially later (see also Tomasello 2008). In effect, this presents a non-unitary view of communication, which, based on the occasional sophistication that adult communicators exhibit radically separates the abilities of adult communicators from those of children and high-functioning autistic adults.

But there is also very considerable independent evidence that even though adults are able to think about other people's perspectives, they are significantly influenced by their own point of view (*egocentrism*) (Keysar 2007). This suggests that the complex hypotheses required by Gricean reasoning in communication may not reliably be constructed by adults either.[6] This is corroborated by an increasingly large body of research demonstrating that Gricean "common ground" is not a necessary building block in achieving coordinative communicative success: speakers regularly violate shared knowledge at first pass in the use of anaphoric and referential expressions which supposedly demonstrate the necessity of established common ground (Keysar 2007, a.o.).[7] Given this type of observation, checking in parsing or producing utterances that information is jointly held by the dialogue participants—the perceived *common ground* (see Allan, this volume)—

---

[6] Indeed, it is useful to note that even adults fail the false belief task, if it is a bit more complex (Birch and Bloom 2007).

[7] Though 'audience design' and coordination effects are regularly observed in experiments (see e.g. Hanna et al. 2003), these can be shown to result from general memory-retrieval mechanisms rather than as based on some common ground calculation based on metarepresentation or reasoning (see Horton and Gerrig 2005; Pickering and Garrod 2004).

cannot be a necessary condition on such activities. And there is psycholinguistic evidence that such neglect of common ground does not significantly impede successful communication and is not even detected by participants (Engelhardt et al. 2006, a.o.). Moreover, if such data are set aside as exceptional or unsuccessful acts of communication, one is left without an account of how people manage to understand what each other has said in these cases. But it is now well-documented that "miscommunication" phenomena not only provide vital insights as to how language and communication operate (Schegloff 1979), but also facilitate coordination: as Healey (2008) shows, the local processes involved in the detection and resolution of misalignments during interaction lead to significantly more positive effects on measures of successful interactional outcomes (see also Brennan and Schober 2001; Barr 1998). In addition, these localised procedures lead to more gradual, group-level modifications, which in turn account for language change. It seems then from this perspective that the Gricean and neo-Gricean focus on detecting speaker meaning as the sole criterion of communicative success misrepresents the goals of human interaction: miscommunication (which is an inevitable ingredient in the interaction of interlocutors that do not share a priori common ground) and the specialised repair procedures made available by the structured linguistic and interactional resources available are the main means that can guarantee intersubjectivity and coordination; and, as Saxton (1997) shows, in addition, such mechanisms, in the form of negative evidence and embedded repairs (see also Clark and Lappin 2011), crucially mediate language acquisition (see also Goodwin 1981: 170–171).

2.2 Joint Intentions, Planning and Dialogue Modelling

More recently, work in philosophy has started exploring notions of joint agency/ joint action/joint intentions (see e.g. Searle 1990, 1995; Bratman 1990, 1992, 1993, 1999; Gilbert 1996, 2003; Tuomela 1995, 2005, 2007 a.o.). As the Gricean individualistic view of speaker's intention being the sole determinant of meaning underestimates the role of the hearer, current dialogue models have turned to Bratman's account of *joint intentions* to model participant coordination. The controversial notion of 'intention' as a psychological state has been explicated in terms of hierarchical planning structures (Bratman 1990), a view generally adopted in AI models of communication (see, e.g. Cohen et al. 1990). In this type of account, collective intentions are reduced to individual intentions and a network of mutual beliefs. A similar style of analysis features prominently in H. Clark's model: dialogue involves joint actions built on the coordination of (intention-driven) individual actions based on shared beliefs (*common ground*):

> What makes an action a joint one, ultimately, is the coordination of individual actions by two of more people (Clark 1996: 59).

In this respect, a strong Gricean element underlies the psycholinguistic and computational modelling of dialogue reflecting reasoning about speakers' intentions even though now supported by an account in terms of joint action and conversational structure. Thus, within psycholinguistics and (computational) semantics, the move from individualistic accounts of action, planning and intention to joint action and coordination in dialogue has seen the latter as derivative.

However, joint action seems to involve a number of lower-level cognitive phenomena that cannot be easily explicated in Gricean terms. We should distinguish here between the terms 'coordination' and 'cooperation': cooperation is taken as involving a defined shared goal between interlocutors whereas coordination is the dynamically matched behaviour of two or more agents so that it might appear that there is a joint purpose, whether there is one or not (see also Allott 2008: 15). In this respect, psycholinguistic studies on dialogue have demonstrated that when individuals engage in a joint activity, such as conversation, they become "aligned", i.e. they (unconsciously) synchronise their behaviour at a variety of different levels, e.g. bodily movements, speech patterns etc. These coordinations draw on subpersonal, synchronised mechanisms (Pickering and Garrod 2004) or emotional, sensory-motor practices that are, crucially, nonconceptual (Gallagher 2001: 81; Hutto 2004).

From this perspective, taking the individualistic conception of intention in, e.g. Bratman's analysis as the basis of conversational dialogue seems either conceptually or cognitively implausible (Tollefsen 2005; Becchio and Bertone 2004). In this connection, the Schiffer and Strawson scenarios mentioned earlier that led to a more complicated picture of utterance meaning seem to show, in fact, that Gricean assumptions are on the wrong footing as a foundation for accounts of communication: The method of generalising from these elaborate cases to cases of ordinary conversation makes it inevitable that paradoxes will be generated, e.g. the *mutual knowledge paradox* (Clark and Marshall 1981), according to which, interlocutors have to compute an infinite series of beliefs in finite time. The dilemma here is that there is plenty of evidence for *audience design* in language production, a type of (seemingly) cooperative, coordinative behaviour, posing the problem of how to model the interlocutors' abilities allowing them to achieve this during online processing. But the solution to such problems, ideally, should not replicate the problematic structure involved (as in, e.g. Clark and Marshall 1981, who assume that interlocutors carry around detailed models of the people they know which they consult when they come to interact with them). Replacing such accounts with a psychological perspective that focuses on the lower-level mechanisms involved can undercut the intractability of such solutions by invoking independently established memory mechanisms that provide explanation of how people appear to achieve "audience designed" productions without in fact constructing explicit models of the interlocutor or metarepresentations. In this respect, Horton and Gerrig (2005) show, through subtle experimental manipulations, that the ordinary retrieval of episodic memory traces during interaction predicts much better both participants' conformity but also, and more crucially, their deviations from the assumptions derived from the "common ground" idealisation.

In the same spirit, empirical Conversational Analysis (CA) accounts of the sequential coherence of conversations emphasise the importance of the turn-by-turn organisation of dialogue which allows juxtaposition of displays of participant understandings and provides structures for organised repair (see e.g. Schegloff 2007). Rather than interlocutors having to figure out each other's mental states and plans through metarepresentational means, conversational organisation provides the requisite structure for coordination through repair procedures and routines. Accordingly, as Garrod and Anderson (1987) observe, in task-oriented dialogue experiments, explicit negotiation is neither a preferential nor an effective means of coordination, as would be expected to be if reasoning about speaker plans and common ground were the primary means of coordination. Explicit negotiation, if it occurs at all, usually happens after participants have already developed some familiarity with the task. Hence, the Interactive Alignment model developed by Pickering and Garrod (2004) emphasizes the importance of tacit alignment mechanisms and implicit common ground as the primary means of coordination. The establishment of routines and the significance of repair as externalised inference are also noted by Pickering and Garrod. Further psycholinguistic experiments reported in Mills and Gregoromichelaki (2008, 2010) and Mills (2011) suggest that, by probing the process of coordination in task-oriented dialogue, it can be demonstrated that notions of joint intentions and plans emerge gradually in a regular manner, rather than guiding utterance production and interpretation throughout. The hypothesis that these implicit means, rather than intention recognition, are the primary method of coordination is probed in these experiments by inserting artificial clarifications regarding intentions (*why?*) and observing the responses they receive at initial and later stages of rounds of games. At early stages, individuals display little recognition of specific intentions/plans underpinning their own utterances and explicit negotiation is either ignored or more likely to impede (see also Mills 2007; Healey 1997). This is because participants have not yet figured out the structure of the task, hence they do not have yet developed a metalanguage involving plan and intention attribution in order to explicitly negotiate their purposes. As CA research indicates, this then implies that discursive constructs such as "intentions" need to emerge, even in such task-oriented joint projects. Initially, participants seem to follow trial-and-error strategies to figure out what the task involves and coordinate their responses. These strategies and the routines participants develop lead, at later stages of the games, to highly coordinated, efficient interaction and, at this stage, issues of "intention/plan" can be raised. These results appear to undermine both accounts of coordination that rely on an a *priori* notion of (joint) intentions and plans (e.g. Bratman 1990) and also accounts which rely on some kind of strategic negotiation/agreement to mediate coordination. This is because it seems that, even in such task-specific situations, joint intentionality is not guaranteed ab initio but rather has to evolve incrementally with the increasing expertise.

These observations seem consonant with an alternative approach to planning and intention-recognition according to which forming and recognising such constructs is a subordinated activity to the more basic processes that underlie people's

performance (see e.g. Suchman 1987/2007; Agre and Chapman 1990). Given the known intractability of notions like plan recognition and common ground/mutual knowledge computation (see, e.g. Levinson 1995), computational models of dialogue, even when based on generally Clarkian theories of common ground, have now largely been developed without explicit high-order meta-representations of other parties' beliefs or intentions except where dealing with complex dialogue domains (e.g. non-cooperative negotiation, Traum et al. 2008). With algorithmically defined concepts such as *dialogue gameboard*, *QUD*, (Ginzburg 2012; Larsson 2002) and default rules incorporating rhetorical relations (Lascarides and Asher 2009; Asher and Lascarides 2008), the necessity for rational reconstruction of inferential intention recognition is largely sidestepped (though see Lascarides and Asher 2009; Asher and Lascarides 2008 for discussion). Even models that avow to implement Gricean notions (see e.g. Stone 2005, 2004) have significantly weakened the Gricean reconstruction of the notion of "communicative intention" and meaning$_{NN}$, positing instead representations whose content does not directly reflect the logical structure (e.g. reflexive or iterative intentions) required by a genuine Gricean account.

The philosophical underpinnings of dialogue models that rely on Gricean notions are sought in accounts that explicate intentions as mental states, independent of and prior to intentional action. However, the tradition following late Wittgensteinian ideas sees 'intention' as part of a discursive practice (Anscombe 1957) rather than a term referring to an actual mental state. Accordingly, language is to be understood as action, rather than the means of allowing expression of inner, unobservable cognitive entities. Such approaches criticise standard dialogue models, e.g. H. Clark's theory, based on the claim that that these approaches retain a communication-as-transfer-between-minds view of language treating intentions and goals as pre-existing private inner states that become externalised in language (see, e.g. Hutto 2004). In contrast, philosophers like Brandom (1994) eschew the individualistic character of accounts of meaning espoused by the Gricean perspective, analysing meaning/intentionality as arising out of linguistic social practices, with meaning, beliefs and intentions all accounted for in terms of the *linguistic* game of giving and asking for reasons. This view has been adopted in the domain of computational semantics and dialogue modelling by Kibble (2006a, b) among others (e.g. Matheson et al. 2000; Walton and Krabbe 1995; Singh 1999). The guiding principle behind such social, non-intentionalist explanations of communication and dialogue understanding is to replace mentalist notions such as 'belief' with public, observable practical and propositional 'commitments', in order to resolve the problems arising for dialogue models associated with the intersubjectivity of beliefs and intentions, i.e. the fact that such private mental states are not directly observable and available to the interlocutors. A further motivation arises from the fact that it has been shown that beliefs, goals and intentions underdetermine what "rational" agents will do in conversation: social obligations or conversational rules may in fact either displace beliefs or intentions as the motivation for agents' behaviour or enter as an additional explanatory factor (e.g. the (social) obligation to answer a question might displace/modify the

"intention" not to answer it, see, e.g. Traum and Allen (1994)). Brandom's account presents an inferentialist view of communication which seeks to replace mentalist notions with public, observable practical and propositional commitments. Under this view, commitment does not imply 'belief' in the usual sense. A speaker may publicly commit to something which she does not believe. And 'intention' can be cashed out as the undertaking of a practical commitment or a reliable disposition to respond differentially to the acknowledging of certain commitments.[8]

From our point of view, the advantage of such non-individualistic, externalist accounts (see also Millikan 1984, 2005; Burge 1986) is that, in not giving supremacy to an exclusively individualist conception of psychological processes, they break apart the presumed exhaustive dichotomy between behaviourist and mentalist accounts of meaning and behaviour (see e.g. Preston 1994) or code versus inferential models of communication (see e.g. Krauss and Fussell 1996). Instead, ascribing contents to behaviours is achieved by supra-individual social or environmental structures, e.g. conventions, "functions", embodied practices, routinisations, that act as the context that guides agents' behaviour. The mode of explanation for such behaviours then does not enforce a representational component, accessible to individual agents, that analyses such behaviours in folk-psychological mentalistic terms, to be invoked as an explanatory factor in the production and interpretation of social action or behaviour. Individual agents instead can be modelled as operating through low-level mechanistic processes (see e.g. Böckler et al. 2010) without necessary rationalisation of their actions in terms of mental state ascriptions (see e.g. Barr 2004 for the establishment of conventions and Pickering and Garrod 2004 for coordination). This view is consonant with recent results in neuroscience indicating that notions like 'intentions', 'agency', 'voluntary action' etc. can be taken as post hoc "confabulations" rather than causally efficacious (work by Benjamin Libet, John Bargh and Read Montague, for a survey see Wegner 2002): according to these results, when a thought that occurs to an individual just prior to an action is seen as consistent with that action, and no salient alternative "causes" of the action are accessible, the individual will experience conscious will and ascribe agency to themselves.

Accordingly, when examining human interaction, and more specifically dialogue, notions like intentions and beliefs may enter into common sense psychological explanations that the participants themselves can invoke and manipulate, especially when the interaction does not run smoothly. As such, they do operate as resources that interlocutors can utilise explicitly to account for their own and others' behaviour. In this sense, such notions constitute part of the metalanguage participants employ to make sense of their actions in conscious, often externalised reflections (see e.g. Heritage 1984; Mills and Gregoromichelaki 2010; Healey

---

[8] An intermediate position is presented by Lascarides and Asher (2009); Asher and Lascarides (2008) who also appeal to a notion of public commitment associated with dialogue moves but which they link to a parallel cognitive modelling component based on inference about private mental states (see also Traum and Allen 1994; Poesio and Traum 1997).

2008). Cognitive models that elevate such resources to causal factors in terms of plans, goals etc. either risk not doing justice to the sub-personal, low-level mechanisms that implement the epiphenomenal effects they describe, or they frame their provided explanations as competence/computational level descriptions (see e.g. Stone 2005, 2004). The stance such models take may be seen as innocuous preliminary idealisation, but this is acceptable only in the absence of either emerging internal inconsistency or alternative explanations that subsume the phenomena under more general assumptions. For example, there are well-known empirical/conceptual problems with the reduction of agent coordination in terms of Bratman's joint intentions (Searle 1990; Gold and Sugden 2007)[9]; and there are also psychological/practical puzzles in cognitive/computational implementations in that the plan recognition problem is known to be intractable in domain-independent planning (Chapman 1987).[10] But, in addition, empirical linguistic phenomena seem to escape adequate modelling in that the assumption that speakers formulate and attempt to transmit determinate meanings in conversation seems implausible when conversational data is examined. We turn to a range of such phenomena next.

2.3 Emergent Intentions

The fundamental role of intention recognition and the primary significance of speaker meaning in dialogue has been disputed in interactional accounts of communication where intentions, instead of assuming causal/explanatory force can be characterised as "emergent" in that the participants can be taken to jointly construct the content of the interaction (Gibbs 2001; Haugh 2008; Mills and Gregoromichelaki 2010; Mills 2011). This aspect of joint action has been explicated via the assumption of the "non-summativity of dyadic cognition" (Arundale and Good 2002; Arundale 2008; Haugh 2012; Haugh and Jaszczolt 2012) or in terms of "interactive emergence" (Clark 1997; Gibbs 2001). This view gains experimental backing through the observation of the differential performance of participants versus over-hearers in conversation (Clark and Schaefer 1987; Schober and Clark 1989) and the gradual emergence of intentional explanations in task-oriented dialogue (Mills and Gregoromichelaki 2010). Standard dialogue systems, by contrast, are serial, modular and operate on complete utterances underpinned by a speaker plan and its recognition. Typically, such models include a parser responsible for syntactic and semantic analysis, an interpretation manager, a

---

[9] In addition, such accounts of coordination are not general enough in that they are discontinuous with explanations of collective actions, in e.g. crowd coordination, individuals walking past each other on the sidewalk, etc.

[10] In addition, it has been argued that use of such folk-psychological constructs are culture/occasion-specific (Du Bois 1987; Duranti 1988), hence should not be seen as underpinning general cognitive abilities.

dialogue manager and a generation module. The output of each module is the input for another with speaking and listening seen as autonomous processes. This goes against the observation that, in ordinary conversation, utterances are shaped genuinely incrementally and "opportunistically" according to feedback by the interlocutor (as already pointed out by Clark 1996) thus genuinely engendering co-constructions of utterances, structures and meanings (see e.g. Lerner 2004). In our view, the main reason for this inadequacy in dialogue modelling are methodological assumptions justified by the competence/performance distinction, separating the grammar from the parser/generator and the pragmatic modules, with the result that the grammatical models employed lack the capability to fully manipulate and integrate partial structures in an incremental manner (for recent incremental systems see Petukhova and Bunt 2011; Poesio and Rieser 2010).

2.4 Incrementality in Processing and Split Utterances

The incrementality of on-line processing is now uncontroversial. It has been established for some considerable time now that language comprehension operates incrementally; and, standardly, psycholinguistic models assume that partial interpretations are built more or less on a word-by-word basis (see e.g. Sturt and Crocker 1996). More recently, language production has also been argued to be incremental (Kempen and Hoenkamp 1987; Levelt 1989; Ferreira 1996; Bock and Levelt 2002). Guhe (2007) further argues for the incremental conceptualisation of observed events resulting in the generation of preverbal messages in an incremental manner guiding semantic and syntactic formulation. In all the interleaving of planning, conceptual structuring of the message, syntactic structure generation and articulation, psycholinguistic incremental models assume that information is processed as it becomes available, reflecting the introspective observation that the end of a sentence is not planned when one starts to utter its beginning (see e.g. Guhe et al. 2000). In accordance with this, in dialogue, evidence for radical incrementality is provided by the fact that participants incrementally "ground" each other's contribution through *back-channel* contributions like *yeah*, *mhm*, etc. (Allen et al. 2001). In addition, as shown in (1), interlocutors clarify, repair and extend each other's utterances, even in the middle of an emergent clause (*split utterances*):

1. Context: Friends of the Earth club meeting
   A: So what is that? Is that er… booklet or something?
   B: It's a book
   C: Book
   B: Just… talking about al you know alternative
   D: On erm… renewable yeah
   B: energy really I think
   A: Yeah [BNC:D97].

In fact, such completions and continuations have been viewed by Herb Clark, among others, as some of the best evidence for cooperative behaviour in dialogue (Clark 1996: 238).

But even though, indeed, such joint productions demonstrate the participants' skill to collaboratively participate in communicative exchanges, this ability to take on or hand over utterances raises the problem of the status of intention-recognition within human interaction when the aim is an explicit procedural model of how such exchanges are achieved. Firstly, on the Gricean assumption that pragmatic inference in dialogue operates on the basis of reasoning based on evidence of the interlocutor's intention, delivered by establishing the semantic propositional structure licensed by the grammar, the data in (1) cannot be easily explained, except as causing serious disruptions in normal processing, hence the view of dialogue as "degenerate" language use in formal analyses. Secondly, on the assumption that communication necessarily involves recognising the propositional content intended by the speaker, there would be an expected cost for the original hearer in having to infer or guess this content before the original sentence is complete, and for the original speaker in having to modify their original intention, replacing it with that of another in order to understand what the new speaker is offering and respond to it. But, wholly against this expectation, interlocutors very straightforwardly shift out of the parsing role and into the role of producer and vice versa as though they had been in their newly adopted role all along. Indeed, it is the case that such interruptions do sometimes occur when the respondent appears to have guessed what they think was intended by the original speaker, what have been called *collaborative completions*:

2. Conversation from A and B, to C:
   A: We're going to…
   B: Bristol, where Jo lives.
3. A: Are you left or
   B: Right-handed.

However, this is not the only possibility: as (4)–(5) show, such completions by no means need to be what the original speaker actually had in mind:

4. Morse: in any case the question was
   Suspect: a VERY good question inspector [Morse, BBC radio 7].
5. Daughter: Oh here dad, a good way to get those corners out
   Dad: is to stick yer finger inside
   Daughter: well, that's one way (from Lerner 1991).

In fact, such continuations can be completely the opposite of what the original speaker might have intended as in what we will call *hostile continuations* or *devious suggestions* which are nevertheless collaboratively constructed from a grammatical point of view:

6. (A and B arguing:)
   A: In fact what this shows is
   B: that you are an idiot.
7. (A mother, B son)
   A: This afternoon first you'll do your homework, then wash the dishes and then
   B: you'll give me £10?

Furthermore, as all of (1)–(7) show, speaker changes may occur at any point in an exchange (Purver et al. 2009), even very early, as illustrated by (8), with the clarification *Chorlton?* becoming absorbed into the final in-effect collaboratively derived content:

8. A: They X-rayed me, and took a urine sample, took a blood sample. Er, the doctor
   B: Chorlton?
   A: Chorlton, mhmm, he examined me, erm, he, he said now they were on about a slide <unclear> on my heart [BNC: KPY 1005–1008].

This phenomenon has consequences for accounts of both utterance understanding and utterance production. On the one hand, incremental comprehension cannot be based primarily on guessing speaker intentions: for instance, it is not obvious why in (4)–(7), the addressee has to have guessed the original speaker's (propositional) intention/plan before they offer their continuation.[11] On the other hand, speaker intentions need not be fully-formed before production: the assumption of fully-formed propositional intentions guiding production will predict that all the cases above where the continuation is not as expected would have to involve some kind of revision or backtracking on the part of the original speaker. But this is not a necessary assumption: as long as the speaker is licensed to operate with partial structures, they can start an utterance without a fully formed intention/plan as to how it will develop (as the psycholinguistic models in any case suggest) relying on feedback from the hearer to shape their utterance (Goodwin 1979).

While core pragmatic research has largely left on one side the phenomenon of collaborative construction of utterances, the emergence of propositional contents in dialogue has been documented over many years in Conversation Analysis (CA) (see e.g. Lerner 2004). The importance of feedback in co-constructing meaning in communication has been already documented at the propositional level (the level of speech acts, 'adjacency pairs') within CA (see e.g. Schegloff 2007). However, it

---

[11] These are cases not addressed by DeVault et al. (2009), who otherwise offer a method for getting full interpretation as early as possible. Lascarides and Asher (2009); Asher and Lascarides (2008) also define a model of dialogue that partly sidesteps many of the issues raised in intention recognition. But, in adopting the essentially suprasentential remit of SDRT, their model does not address the step-by-step incrementality needed to model split-utterance phenomena.

seems here that the same processes can operate sub-propositionally, but this can be demonstrated only relatively to models that allow the incremental, sub-sentential integration of cross-speaker productions. We turn to two such models next.

3 Grammar and Dialogue

It seems to be a standard assumption that linguistic knowledge has to be modelled as providing constraints on linguistic processing (see e.g. Bosch 2008, a.o.). In this sense linguistic knowledge is (often) characterised in abstract static terms whereas linguistic processing is argued to be characterised by three indispensable features, namely: immediacy (i.e. context-dependence), incrementality, multi-modality (see Marslen-Wilson and Tyler 1980; Altmann and Steedman 1988). However, against this view, work on linguistic phenomena, e.g. ellipsis, that cross-cut monologue and dialogue, sentence and discourse, has shown that a unified story requires all these three processor properties to be included in the theory of linguistic knowledge/grammar (see, e.g. Gargett et al. 2009; Kempson et al. 2009a, b). Otherwise, separating linguistic knowledge (grammar) from processing results in a view of dialogue as "degenerate" language use. Notably, this separation has led even dialogue-oriented psycholinguists, e.g. Clark (1996), to distinguish language$_S$ (language structure) versus language$_U$ (language-in-use).

In contrast, here we would like to argue for a reconciliation between the "language-as-action" and "language-as-product" traditions, at the same time shifting the boundaries between grammar and pragmatics. The reason for this is that the two approaches should be seen, in our view, as constituting not a dichotomy but a continuum. However, in order to substantiate such a view, linguistic knowledge has to be reconceptualised as encompassing the update dynamics of communication which crucially involves:

- representations integrating multiple sources of information
- word-by-word incrementality within the grammar system
- NL grammars as mechanisms for communicative interaction relative to context.

This is because what we see as inherent features of the grammar architecture, utilised to solve traditional grammatical puzzles (see e.g. Kempson et al. 2001; Cann et al. 2005; Kempson et al. 2011b), also underlie many features of language use in dialogue. Firstly, the function of items like inserts, repairs, hesitation markers etc. interact with the grammar at a sub-sentential level (Clark and Fox Tree 2002). Hence the grammar must be equipped to deal with those in a timely and integrated manner. In addition, the turn-taking system (see, e.g., Sacks et al. 1974) seems to rely on the grammar, based on the predictibility of (potential) turn endings; in this respect, recent experimental evidence have shown that this predictability is grounded on syntactic recognition rather than prosodic cues etc. (De Ruiter et al. 2006); and further evidence shows that people seem to exploit such predictions to manage the timing of their contributions (Henetz and Clark 2011). More

importantly for our concerns here, incremental planning in production allows the grammar to account for how the interlocutors interact sub-sententially in dialogue to derive joint meanings, actions and syntactic constructions taking in multi-modal aspects of communication and feedback, a fact claimed to be a basic characteristic of interaction (Goodwin 1979, 1981).

3.1 Modelling the Incrementality of Split Utterances

The challenge of modelling the full word-by-word incrementality required in dialogue has recently been taken up by two models which employ distinct approaches: a neo-Gricean model by Poesio and Rieser (2010) (*P&R* henceforth) and Dynamic Syntax (Kempson et al. 2001).

P&R set out a dialogue model for German, defining a thorough, fine-grained account of dialogue interactivity. Their primary aim is to model *collaborative completions*, as in (2) and (3) in cooperative task-oriented dialogues where take-over by the hearer relies on the remainder of the utterance taken to be understood or inferrable from mutual knowledge/common ground.[12] Their account is an ambitious one in that it aims at modelling the generation and realisation of joint intentions which accounts for the production and comprehension of co-operative completions. The P&R model hinges on two main points: the assumption of recognition of interlocutors' intentions according to shared joint plans (Bratman 1992), and the use of incremental grammatical processing based on LTAG. With respect to the latter, this account relies on the assumption of a string-based level of syntactic analysis, for it is this which provides the top-down, predictive element allowing the incremental integration of such continuations. However, exactly this assumption would seem to impede a more general analysis, since there are cases where split utterances cannot be seen as an extension by the second contributor of the proffered string of words/sentence:

9. Eleni: Is this *yours* or
   Yo: *Yours* [natural data].
10. with smoke coming from the kitchen:
    A: I'm afraid I burnt the kitchen ceiling
    B: But have *you*
    A: burned *myself*? Fortunately not.

In (9), the string of words (sentence) that the completion yields is not at all what either participant takes themselves to have constructed, collaboratively or other-wise. And in (10) also, even though the grammar is responsible for the dependency that licenses the reflexive anaphor *myself*, the explanation for B's continuation in the third turn of (10) cannot be string-based as then *myself* would not be locally

---

[12] Thus, notably, the P&R data involve data collected after task training.

bound (its antecedent is *you*). Moreover, in LTAG, P&R's syntactic framework, parsing relies in the presence of a head that provides the skeleton of the structure. Yet, as (1)–(10) indicate, utterance take-over can take place without a head having occurred prior to the split (see also Purver et al. 2009, Howes et al. 2011), and even across split syntactic dependencies (in (10) an antecedent-anaphor relation and in (11) between a Negative Polarity Item and its triggering environment, the question):

11. A: *Have* you mended
    B: *any* of your chairs? Not yet.

Given that such dependencies are defined grammar-internally, the grammar has to be able to license such split-participant realisations. But string-based grammars cannot account straightforwardly for many types of split utterances except by treating each part as elliptical sentences requiring reconstruction of the missing content with case-specific adjustments to guarantee grammaticality/interpretability (as is needed in (9)–(10)).

   Furthermore, if the attempt is to reconstruct speaker's intentions as the basis for the interpretation recovered, as P&R explicitly advocate, there is the additional problem that such fragments can play multiple roles at the same time (e.g. the fragments in (3) and (9) can be simultaneously taken as question/clarification/ completion/acknowledgment/answer; see also Sbisà, this volume). Notice also that co-construction at the sub-propositional level can be employed for the perfor-mance of speech acts by establishing (syntactic) conditional relevances,[13] i.e. exploiting grammatical mechanisms as a means to induce the coordination of social actions. For example, such completions might be explicitly invited by the speaker thus forming a question–answer pair:

12. A: And you're leaving at
    B: 3.00 o'clock.
13. A: And they ignored the conspirators who were …
    B: Geoff Hoon and Patricia Hewitt [radio 4, Today programme, 06/01/10]
14. Jim: The Holy Spirit is one who <pause> gives us? Unknown: Strength
    Jim: Strength. Yes, indeed. <pause> The Holy Spirit is one who gives us? <pause>
    Unknown: Comfort. [BNC HDD: 277–282]
15. George: Cos they <unclear> they used to come in here for water and bunkers you see
    Anon 1: Water and?
    George: Bunkers, coal, they all coal furnace you see,… [BNC, H5H: 59–61]

Within the P&R model, such multifunctionality would not be capturable except as a case of ambiguity or by positing hidden constituent reconstruction that has to be

---

[13] For the concept of *conditional relevance* in conversation see, e.g., Schegloff (1996).

subject to some non-monotonic build-and-revise strategy that is able to apply even within the processing of an individual utterance. But, in fact, in some contexts, invited completions have been argued to exploit the vagueness/covertness of the speech act involved to avoid overt/intrusive elicitation of information (Ferrara 1992):

16. (Lana = client; Ralph = therapist)
    Ralph: Your sponsor before…
    Lana: was a woman

Hence, the resolution of such fragments cannot be taken to rely on the determination of a specific speaker-intended speech-act (see also Sbisà, this volume).

It has to be said that the P&R account is not intended to cover such data, as the setting for their analysis is one in which participants are assigned a collaborative task with a specific joint goal, so that joint intentionality is fixed in advance and hence anticipatory computation of interlocutors' intentions can be fully determined; but such fixed joint intentionality is decidedly non-normal in dialogue (see e.g. Mills and Gregoromichelaki 2010) and leaves any uncertainty or non-determinism in participants' intentions an open challenge. Nonetheless, by employing an incremental model of grammar, the P&R account marks a significant advance in the analysis of such phenomena. Relative to any other grammatical framework, dialogue exchanges involving incremental split utterances of any type are even harder to model, given the near-universal commitment to a static performance-independent methodology. Thus, first of all, in almost all standard grammar frameworks, it is usually the sentence/proposition that is the unit of syntactic/semantic analysis. Inevitably, fragments are then assigned sentential analyses with semantics provided through ellipsis resolution involving abstraction operations as in Dalrymple et al. (1991) (see e.g. Purver 2006; Ginzburg and Cooper 2004; Fernandez 2006). The abstraction is defined over a propositional content provided by the previous context to yield appropriate functors to apply to the fragment. Of course, multiple options of appropriate "antecedents" for elliptical fragments are usually available (one for each possible abstract) resulting in multiple ambiguities which are then relegated to some performance mechanism for resolution. Such mechanisms are defined to appeal to independent pragmatic assumptions having to do with recognizing the speaker's intention in order to select a single appropriate interpretation. But the intention recognition required for disambiguation is unavailable in sub-sentential split utterances as in (1), (3), (9)–(16) in all but the most task-specific domains. This is because, in principle, attribution of recognition of the speaker's intention to convey some specific propositional content is unavailable until the appropriate propositional formula is established. This is particularly clear where an antecedent is required too early in the emergent proposition so that no appropriate abstract definable from context is available as in (8) above.

In response to the challenge that such data provide, we turn to *Dynamic Syntax* (DS: Kempson et al. 2001; Cann et al. 2005) where the correlation between parsing

and generation, as they take place in dialogue, can provide a basis for modelling recovery of interpretation in communicative exchanges without reliance on recognition of specific intentional contents.

3.2 Dynamic Syntax

DS is an action-based formalism. It models "syntax" in procedural terms as the goal-directed, incremental, stepwise transition from strings of words to meaning representations which dynamically integrate both linguistic and extra-linguistic or inferred information. These are the only representations constructed during the interpretation of utterances, hence no distinct syntactic level of representation is assumed. As in DRT and related frameworks (see also Jaszczolt 2005), semantic, truth-conditional evaluation applies solely to these contextually enriched representations, hence no semantic content is ever assigned to strings of words (sentences).

3.2.1 Radically Contextualist Representations

The examination of linguistic data seems to indicate evidence of structure underlying the linear presentation of strings. Similar types of evidence can also be found in dialogue. First of all, it has been shown both by corpus research (Fox and Jasperson 1995) and experimental results (Eshghi et al. 2010) that repair processes in dialogue target primarily 'constituents' whereas other factors like pauses, time units etc. play a secondary role. For example, Fox and Jasperson, who examine self-repairs, claim that "in turn beginnings, if repair is initiated after an auxiliary or main verb, the verb and its subject are always recycled together; the verb is never recycled by itself." (1995:110). Moreover, the use of fragments ("elliptical" utterances) during interaction, follows syntactic constraints indicating their appropriate integration in some structured representation. This is more evident in languages with rich morphology and case systems. For example, although it has been established that speakers can use fragments like the following in (17) to perform speech acts that do not presuppose the recovery of a full sentence ('non-sentential speech acts': Stainton 2005), languages like German and Greek require that the fragment bears appropriate case specifications, otherwise it is perceived as ungrammatical:

17. Context: A and B enter a room and see a woman lying on the floor:
    A to B: Schnell, den Arzt/*der Arzt (German)
    "Quick, the doctor_ACC/*the doctor_NOM"

One might take these as evidence for a separate (possibly autonomous) level of syntactic analysis. Indeed, based on similar observations, standard grammatical

models postulate an independent level of structure over strings (see e.g. Ginzburg and Cooper 2004; Ginzburg 2012) whereas categorial grammars that deny the existence of any level of independent structuring with syntactic relevance have difficulty in explaining such data. Both these types of account are not sustainable as there is also evidence that explanations for such phenomena cannot be string-based. As shown below in (18) and earlier in (9)–(10), splicing together the two partial strings gives incorrect interpretations since elements like indexicals have to switch form in order to be interpretable as intended or for grammaticality:

18. G: when you say it happens for a reason, it's like, it happened to get *you* off
    D: off *my* ass [Carsales 3 cited in Ono and Thompson (1995)]

In contrast, even though DS, like categorial grammar, takes the view that syntactic constraints and dependencies do not justify a separate level of representation for structures over stings, nevertheless, it handles such data successfully via the definition of constraints on the updates of the semantic representations induced by the processing mechanism. So the reduction in representational levels, instead of impeding the definition of syntactic licensing, allows in fact the handling of a wider range of data via the same incremental licensing mechanisms. So, instead of data such as those in (9)–(10) and (18) being problematic, use of the licensing mechanisms across interlocutors illustrates the advantages of a DS-style incremental, dynamic account over static models (for detailed analyses see Kempson et al. 2009a, b, 2011a; Purver et al. 2010, 2011; Gregoromichelaki et al. 2009, 2011; Gargett et al 2008). Given that linguistic processing has to be incrementally interleaved with processes of inference and perceptual inputs, this is essential for dialogue as not only is comprehension heavily reliant on context and multimodal input but also dialogue management issues are handled by interaction of linguistic and non-linguistic resources. For example, Goodwin (1979) suggests that in face-to-face interaction completion, extension and allocation of turns are managed through a combination of gaze and syntactic information.

3.2.2 Incrementality

Because of this procedural architecture, two features usually associated with parsers, *incrementality* and *predictivity*, are intrinsic to the DS grammar and are argued to constitute the explanatory basis for many idiosyncrasies of NLs standardly taken to pose syntactic/morphosyntactic/semantic puzzles. As can be seen in (1) above, dialogue utterances are fragmentary and subsentential. This implies that dialogue phenomena like self-repair, interruptions, corrections etc. require modelling of the incremental understanding/production and if the grammar needs to license such constructions it needs to deal with partial/non-fully-sentential constructs. Modular approaches to the grammar/pragmatics interface deny that this is an appropriate strategy. Instead they propose that the grammar delivers under-specified propositional representations as input to pragmatic processes that achieve

full interpretations and discourse integration (see e.g. Schlangen 2003, following an SDRT model). However, an essential feature of language use in dialogue is the observation that on-going interaction and feedback shapes utterances and their contents (Goodwin 1981), hence it is essential that the grammar does not have to licence whole propositional units for semantic and pragmatic evaluation to take place. And this is the strategy DS adopts as it operates with partial constructs that are fully licensed and integrated in the semantic representation immediately. This has the advantage that online syntactic processing can be taken to be implicated in the licensing of fragmentary utterances spread across interlocutors without having to consider such fragments as elliptical sentences or non well-formed in any respect. And this is essential for a realistic account of dialogue as corpus research has shown that speaker/hearer exchange of roles can occur across all syntactic dependencies (Purver et al. 2009):

19. Gardener: I shall need the mattock.
    Home-owner: **The**…
    Gardener: **mattock**. For breaking up clods of earth [BNC].
20. A: or we could just haul: a:ll the skis in [[the:]] dorms
    B: [[we could]]
    [[haul all the skis into the dorm]]
    C: [[hh uh hhuhhuh]] (1.0)
    B: **which** (0.3)
    A: **might** work
    B: might be the best [BNC].
21. Jack: I just returned
    Kathy: **from**…
    Jack: Finland [from Lerner 2004]
22. Teacher: Where was this book lub- published?
    Teacher: Macmillan publishing company **in**? (.)
    Class: New York ((mostly in unison))
    Teacher: Okay, [from Lerner 2004].
23. Therapist: What kind of work do you do?
    Mother: on food service
    Therapist: **At_**
    Mother: uh post office cafeteria downtown main point office on Redwood
    Therapist: °Okay° [Jones and Beach 1995].
24. S: You know some nights I just- (0.2) if I get bad flashes I c- I can't mo:ve.
    C: No: =
    S: So some nights he's got the baby and me:huh(.)
    C: hhhh Uh by flashes you mean flashbacks
    S: Yea:h.
    C: **To::**,
    S: To- To the bi:rth

C: To the birth itse:lf. mm.(0.2)
S: And thee uhm (.) the- the labor an' thee the week in the hospital
afterwa:rd[s.]
C: [Y]e:s. Ye:s. [from Lerner 2004]

But if the grammar is conceived as operating independently of the dialogue processes that manage turn handling and derivation of content across participants there is no way to account for the licensing, the formal properties and eventual interpretations of such fragmentary utterances (see also Morgan 1973). Instead, DS grammar constraints operate incrementally, on a word-by-word basis, thus allowing participants to progressively integrate contents and modify each other's contributions.

3.2.3 Predictivity

As we said earlier, the turn-taking system (see Sacks et al. 1974) relies heavily on the grammar via the notion of predictibility of (potential) turn endings. Fluent speaker/hearer role switch relies on participants' being able to monitor the on-going turn and project constituent completions so that they can time their exits and entries appropriately. Experimental results have shown that this ability is primarily grounded on syntactic recognition (rather than prosodic clues etc. see, e.g. De Ruiter et al. 2006). The ability of recipients to project the upcoming turn completion so that they can plan their own contribution seems to favour predictive models of processing (e.g. Sturt and Lombardo 2005) over head-driven or bottom-up parsers. DS incorporates exactly such a notion of predictivity/goal-directedness inside the grammar formalism itself in that processing (and hence licensing) is driven by the generation and fulfilment of goals and subgoals. This architectural feature of DS is fully compatible with observations in interactional accounts of conversation where it is noted that 'anticipatory planning' takes place (Arundale and Good 2002). In addition, given the format of the semantic representations employed by DS (linked trees annotated with conceptual content in functor-argument format), a second stage of composition of what has been built incrementally also occurs at constituent boundaries thus giving the opportunity for 'retroactive assessment' of the derived content (as noted again by Arundale and Good 2002).

Because DS is *bidirectional*, i.e. a model of both parsing and production mechanisms that operate concurrently in a synchronized manner, its goal-directedness/predictivity applies symmetrically *both* in parsing and generation (for predictivity in production see also Demberg-Winterfors 2010). And the consequences in this domain are welcome. Given that the grammar licenses the generator to operate with partial sub-propositional objects, speakers can be modelled as starting to articulate utterances before having planned a complete proposition. Split utterances follow as an immediate consequence of these assumptions: given the general predictivity/goal-directedness of the DS architecture, the parser/

generator is always predicting top-down structural goals to be achieved in the next steps. But such goals are also what drives the search of the lexicon ('lexical access') in generation, so a hearer who achieves a successful lexical retrieval before processing the anticipated lexical input provided by the original speaker can spontaneously become the generator and take over. As seen in all cases (1)–(15) above, the original hearer is, indeed, using such anticipation to take over and offer a completion that, even though licensed, i.e. a grammatical continuation of the initial fragment, might not necessarily be identical to the one the original speaker would have accessed had they been allowed to continue their utterance as in (7)–(9). And since the original speaker is licensed to operate with partial structures, without having a fully-formed intention/plan as to how it will develop (as the psycholinguistic models in any case suggest), they can integrate immediately such offerings without having to be modelled as necessarily revising their original intended message[14] (for detailed analyses see Kempson et al. 2009a, b; Purver et al. 2010, 2011; Gregoromichelaki et al. 2009, 2011; Gargett et al 2008).

Thus DS reflects directly and explicitly, from within the grammar itself, how the possibility arises for joint-construction of utterances, meanings and structures in dialogue and how this is achieved. And these explanations are fundamentally based on the same mechanisms underlying language structure: since the grammar licenses partial, incrementally constructed objects, speakers can start an utterance without a fully formed intention/plan as to how it will develop relying on feedback from the hearer to shape its structure and its construal. Moreover, the syntactic constraints themselves can be exploited ad hoc as a source of "conditional relevances" (Schegloff 2007) by setting up sequences (joint speech acts or 'adjacency pairs') sub-sententially (see (20)–(22) above). Thus, syntactic devices and their goal-directed, projectible nature can be manipulated by interlocutors to manage conversational organisation and perform speech acts without fully-formed propositional contents.

Given these results, in our view, the dichotomy between $language_S$ (language structure) and $language_U$ (language use) postulated in standard linguistic models does not withstand the test of application in dialogue, the primary site of language use. Instead, the grammar has to be seen as underpinning communication with, as DS suggests, the syntactic architecture viewed in dynamic terms as the crystallisation of action patterns derived from language use and wider cognitive/social considerations.

4 Conclusion

With grammar mechanisms defined as inducing incremental context-dependent growth of information and employed symmetrically in both parsing and generation, the availability of derivations for genuine dialogue phenomena, like split

---

[14] But, of course, this is not excluded either.

utterances, from within the grammar, shows how core dialogue activities can take place without any other-party meta-representation at all.[15] On this view, as we emphasised earlier, communication is not at base the intention-recognising activity presumed by Gricean and post-Gricean accounts. Rather, speakers can be modelled as able to air propositional and other structures with no more than the vaguest of planning and commitments as to what they are going to say, expecting feedback to fully ground the significance of their utterance, to fully specify their intentions (see e.g. Wittgenstein 1953: 337). Hearers, similarly, do not have to reconstruct the intentions of their interlocutor as a filter on how to interpret the provided signal; instead, they are expected to provide evidence of how *they* perceive the utterance in order to arrive at a joint interpretation. This view of dialogue, though not uncontentious, is one that has been extensively argued for, under distinct assumptions, in the CA literature. According to the proposed DS model of this insight, the core mechanism is incremental, context-dependent processing, implemented by a grammar architecture that reconstructs "syntax" as a goal-directed activity, able to seamlessly integrate with the joint activities people engage in.

This then enables a new perspective on the relation between linguistic ability and the use of language, constituting a position intermediate between the philosophical stances of Millikan and Brandom, and one which is close to that of Recanati (2004). Linguistic ability is grounded in the control of (sub-personal, low-level) mechanisms (see e.g. Böckler et al. 2010) which enable the progressive construction of structured representations to pair with the overt signals of the language. The content of these representations is ascribed, negotiated and accounted for in context, via the interaction among interlocutors and their environment. From this perspective, constructing representations of the other participants' mental states, rational deliberation and planning, though a possible means of securing communication, is seen as by no means necessary.

---

[15] Though, of course, use of reasoning over mental states is not precluded either; such richer contexts and consequent derived implications are modelled via the construction of appropriately linked representations, whose mechanisms for construction are independently available in DS, see Purver et al (2011).

References

Agre, P. E., & Chapman, D. (1990). What are plans for? *Robotics and Autonomous Systems, 6*(1–2), 17–34.

Allen, J., Ferguson, G., & Stent, A. (2001). *An architecture for more realistic conversational systems. In Proceedings of the 2001 International Conference on Intelligent User Interfaces (IUI)*, January 2001.

Allott, N. E. (2005). Paul Grice, reasoning and pragmatics. *UCL Working Papers in Linguistics, 17*, 217–243.

Allott, N. E. (2008). *Pragmatics and rationality*. (PhD thesis, University of London).

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191–238.

Anscombe, G. E. M. (1957). *Intention*. Oxford: Basil Blackwell.

Arundale, R. B. (2008). Against (Gricean) intentions at the heart of human interaction. *Intercultural Pragmatics, 5*(2), 229–258.

Arundale, R. B., & Good D. (2002). Boundaries and sequences in studying conversation In A. Fetzer & C. Meierkord (Eds.), *Rethinking sequentiality. Linguistics meets conversational interaction* (pp. 121–150). Amsterdam: John Benjamins.

Asher, N., & Lascarides, A. (2008). Commitments, beliefs and intentions in dialogue. *Proceedings of LonDial*, 35–42.

Austin, J. L. (1975). *How to do things with words*. Oxford: Oxford University Press.

Bach, K. (1997). The semantics-pragmatics distinction: What it is and why it matters. *Linguistische Berichte, 8*(1997), 33–50.

Bach K., & Harnish R. M. (1982). *Linguistic communication and speech acts*. MIT Press.

Barr, DJ. (1998). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In S. Santi, I. Guaitella, C. Cave & G. Konopczynski (Eds.), *Oralité et gestualité. Interactions et comportements multimodaux dans la communication* (pp. 597–600). Paris: L'Harmattan.

Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science, 28*(6), 937–962.

Becchio, C., & Bertone, C. (2004). Wittgenstein running: Neural mechanisms of collective intentionality and we-mode. *Consciousness and Cognition, 13*, 123–133.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382.

Bock, K., & Levelt, W. (2002). Language production: Grammatical encoding. In G. T. M. Altmann (Ed.), *Psycholinguistics* (pp. 405–452). London: Routledge (Critical concepts in psychology, 5).

Böckler, A., Knoblich, G., & Sebanz, N. (2010). Socializing Cognition. *Towards a Theory of Thinking*, 233–250.

Bosch, P. (2008). Processing definite determiners. In: P. Bosch, D. Gabelaia & J Lang (Eds.) *Proceedings of the 7th International Tbilisi Symposium on Language, Logic and Computation*.

Brandom, R. B. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard University Press.

Bratman, M. E. (1990). What is intention? In P. R. Cohen, J. Morgan, & E. M. Pollack (Eds.) *Intentions in communication*. Cambridge: MIT Press.

Bratman, M. E. (1992). Shared cooperative activity. *Philosophical Review, 101*, 327–341.

Bratman, M. E. (1993). Shared intention. *Ethics, 104*, 97–113.

Bratman, M. E. (1999). *Faces of intention: Selected essays on intention and agency*. Cambridge: Cambridge University Press.

Breheny, R. (2006). Communication and folk psychology. *Mind & Language, 21*(1), 74–107.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*(2), 274–296.

Burge, T. (1986). Individualism and psychology. *The Philosophical Review, 95*(1), 3–45.

Cann, R., Kempson, R., & Marten, L. (2005). *The dynamics of language*. Oxford: Elsevier.

Chapman, D. (1987). Planning for conjunctive goals. *Artificial Intelligence, 32*(3), 333–377.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, A. (1997). *Being There.* Cambridge: MIT Press.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speech. *Cognition, 84*, 73–111.

Clark, A., & Lappin, S. (2011). *Linguistic nativism and the poverty of the stimulus*. London: Wiley-Blackwell.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.

Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes, 2*(1), 19–41.

Cohen, P. R., Morgan, J., & Pollack, M. E. (1990). *Intentions in communication*. Cambridge: MIT Press.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705–717.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science, 1*(2), 255–259.

Dalrymple, M., Shieber, S. M., & Pereira, F. C. N. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy, 14*(4), 399–452.

Demberg-Winterfors, V. (2010). *A broad-coverage model of prediction in human sentence processing*. (PhD thesis. University of Edinburgh).

DeVault, D., Sagae, K., & Traum, DR. (2009). *Can I finish? In Proceedings of the SIGDIAL 2009 Conference*, pp. 11–20, London.

Du Bois, J. W. (1987). Meaning without intention: lessons from divination. *Papers in Pragmatics*, *1*(2).

Duranti, A. (1988). Intentions, language, and social action in a Samoan context. *Journal of Pragmatics, 12*(1), 13–33.

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of quantity? *Journal of Memory and Language, 54*(4), 554–573.

Eshghi, A., Healey, P. G. T., Purver M., Howes, C., Gregoromichelaki, E., & Kempson, R. (2010). *Incremental turn processing in dialogue. In Proceedings of the 16th Annual Conference on Architectures and Mechanisms for Language Processing (AmLAP)*, New York, Sep 2010.

Fernandez, R. (2006). *Non-sentential utterances in dialogue: classification, resolution and use*. (PhD thesis, King's College London, University of London).

Ferrara, K. (1992). The interactive achievement of a sentence: Joint productions in therapeutic discourse. *Discourse Processes, 15*(2), 207–228.

Ferreira, V. S. (1996). Is it better to give than to donate: syntactic flexibility in language production. *Journal of Memory and Language, 35*, 724–755.

Fox, B. A., & Jasperson, R. (1995). A syntactic exploration of repair in English conversation. In P. W. Davis (Ed.), *Descriptive and theoretical modes in the alternative linguistics* (pp. 77–134). Amsterdam: John Benjamins.

Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies, 8*, 83–108.

Gargett, A., Gregoromichelaki, E., Howes, C., & Sato, Y. (2008). *Dialogue-grammar correspondence in dynamic syntax. In Proceedings of the 12th SemDial (LonDial).*

Gargett, A., Gregoromichelaki, E., Kempson, R., Purver, M. & Sato, Y. (2009). Grammar resources for modelling dialogue dynamically. *Journal of Cognitive Neurodynamics, 3*(4).

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition, 27*, 181–218.

Geurts, B. (2010). *Quantity implicatures*. Cambridge: Cambridge University Press.

Gibbs, Jr., & Raymond, W. (2001). Intentions as emergent products of social interactions. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality* (pp. 105–122). Cambridge: MIT Press.

Gilbert, M. (1996). *Living together: Rationality, sociality, and obligation*. Lanham: Rowman & Littlefield.

Gilbert, M. (2003). The structure of the social atom: Joint commitment as the foundation of human social behavior. In F. F. Schmitt (Ed.), *Socializing metaphysics: The nature of social reality* (pp. 39–64). Lanham: Rowman & Littlefield.

Ginzburg, J. (2012). *The interactive stance: Meaning for Conversation*. Oxford University Press.

Ginzburg, J., & Cooper, R. (2004). Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and Philosophy, 27*(3), 297–365.

Glüer, K., & Pagin, P. (2003). Meaning theory and autistic speakers. *Mind and Language, 18*(1), 23–51.

Gold, N., & Sugden, R. (2007). Collective intentions and team agency. *Journal of Philosophy, 104*(3), 109–137.

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97–121). New York: Irvington Publishers.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Gregoromichelaki, E., Sato, Y., Kempson, R., Gargett, A., Howes, C. (2009). *Dialogue modelling and the remit of core grammar. In: Proceedings of IWCS 2009*.

Gregoromichelaki, E., Kempson, R., Purver, M., Mills, G. J., Cann, R., Meyer-Viol, W., Healey, PGT. (2011). Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse 2*(1), 199:233, special issue on Incremental Processing in Dialogue.

Grice, P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts* (pp. 41–58). New York: Academic Press.

Grice, P. (1981). Presupposition and implicature. In P. Cole (Ed.), *Radical* pragmatics (pp. 183–198). New York: Academic Press. Reprinted in studies in the way of words. Cambridge, MA: Harvard University Press, 1989, pp. 269–82.

Grice, P. (2001). *Aspects of reason*. Oxford: Clarendon Press.

Guhe, M. (2007). *Incremental conceptualization for language production*. New Jeresy: Lawrence Erlbaum Associates.

Guhe, M., Habel, C., & Tappe, H. (2000). Incremental event conceptualization and natural language generation in monitoring environments. In *Proceedings of the 1st international conference on natural language generation*-(Vol. 14, pp. 85–92). Association for Computational Linguistics.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*(1), 43–61.

Haugh, M. (2008). Intention in pragmatics. *Intercultural Pragmatics, 5*(2), 99–110.

Haugh, M. (2012). Conversational interaction. In K. M. Jaszczolt & K. Allan (Eds.), *Cambridge handbook of pragmatics*. Cambridge: Cambridge University Press.

Haugh, M., & Jaszczolt, K. M. (2012). Speaker intentions and intentionality. In K. M. Jaszczolt & K. Allan (Eds.),*Cambridge Handbook of Pragmatics* (pp.87–112). Cambridge: Cambridge University Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Healey, P. G. T. (1997). *Expertise or expert-ese: The emergence of task-oriented sub-languages. In P. Langley & M. G. Shafto (Eds.), Proceedings of the 19th Annual Conference of the Cognitive Science Society* (pp. 301–306). Stanford: Stanford University, Aug 1997.

Healey, P. G. T. (2008). Interactive misalignment: The role of repair in the development of group sub-languages. In R. Cooper & R. Kempson (Eds.), *Language in Flux*. College Publications.

Henetz, T., & Clark, H. H. (2011). *Managing delays in speaking*. Paper presented at the IPrA meeting, Manchester, July 2011.

Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes, 40*(1), 1–35.

Howes, C., Purver, M., Healey, P. G. T., Mills, G. J. (2011). On incrementality in dialogue: Evidence from compound contributions dialogue and discourse. (Vol 2, No 1). Special issue on Incremental Processing in Dialogue.

Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind and Language, 19*(5), 548–573.

Jaszczolt, K. M. (2005). *Default semantics: Foundations of a compositional theory of acts of communication.* Oxford: Oxford University Press.

Jones, C. M., & Beach, W. A. (1995). Therapists' techniques for responding to unsolicited contributions by family members. In G. H. Morris & R. J. Chenail (Eds.), *The talk of the clinic: Explorations in the analysis of medical and therapeutic discourse* (pp. 49–70). Hillsdale: Lawrence-Erlbaum Associates.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*(2), 201–258.

Kempson, R., Meyer-Viol, W., & Gabbay, D. (2001). *Dynamic syntax: The flow of langage understanding*. Oxford: Blackwell.

Kempson, R., Gregoromichelaki, E., & Sato, Y. (2009a). *Incrementality, speaker/hearer switching and the disambiguation challenge. In Proceedings of European Association of Computational Linguistics (EACL Athens 2009).*

Kempson, R., Gregoromichelaki, E., Purver, M., Mills, G. J., Gargett, A., Howes, C. (2009b). *How mechanistic can accounts of interaction be? In: Proceedings of Diaholmia, the 13th workshop on the semantics and pragmatics of dialogue.*

Kempson, R., Gregoromichelaki, E., Meyer-Viol, W., Purver, M., White, G., & Cann, R. (2011a). *Natural-language syntax as procedures for interpretation: the dynamics of ellipsis construal. In Proceedings of the PRELUDE Workshop on Games, Dialogues and Interaction*. Lecture Notes in Computer Science. Springer.

Kempson, R., Gregoromichelaki, E., & Howes, C. (2011b). *The dynamics of lexical interfaces*. Stanford: CSLI Publications.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics, 4*(1), 71–84.

Kibble, R. (2006a). Reasoning about propositional commitments in dialogue. *Research on Language and Computation, 4*(2), 179–202.

Kibble, R. (2006b). Speech acts, commitment and multi-agent communication. *Computational and Mathematical Organisation Theory*, *12*, 127–145 Special issue on normative multi-agent systems.

Krauss, R. M., & Fussell, S. R. (1996). Social and psychological models of interpersonal communication. In Kruglanski, A W. & Higgins, E. T. (Eds.), *Social psychology: Handbook of basic principles* (655–701). Guildford.

Larsson, S. (2002). *Issue-based dialogue management*. (PhD thesis, Göteborg University). Also published as Gothenburg Monographs in Linguistics 21.

Lascarides, A., & Asher, N. (2009). Agreement, disputes and commitments in dialogue. *Journal of Semantics, 26*(2), 109.

Lerner, G. H. (1991). On the syntax of sentences-in-progress. *Language in Society, 20*, 441–458.

Lerner, G H. (2004). Collaborative turn sequences. In G. H. Lerner (Ed.) *Conversation analysis: Studies from the first generation* (pp. 225–256). Amsterdam: John Benjamins.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

Levinson, S. C. (1983). *Pragmatics. Cambridge textbooks in linguistics*. Cambridge: Cambridge University Press.

Levinson, S. C. (1995). Interactional biases in human thinking. In E. N. Goody (Ed.), *Social intelligence and interaction* (pp. 221–260). Cambridge: Cambridge University Press.

Matheson, C., Poesio, M., Traum, D. (2000). *Modelling grounding and discourse obligations using update rules. In: Proceedings of NAACL* (pp. 1–8).

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Henry Holt and Co., Inc.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*, 1–71.

McDowell, J. (1980). Meaning, communication, and knowledge. In Z. van Straaten (Ed.), *Philosophical subjects: Essays presented to PF Strawson*. Oxford: Clarendon Press.

Millikan, R. G. (1984). *Language, thought and other biological categories*. Cambridge: MIT Press.

Millikan, R. G. (2005). *Language: A biological model*. USA: Oxford University Press.

Mills, G. J. (2007). Semantic co-ordination in dialogue: the role of direct interaction. PhD thesis, Queen Mary University of London.

Mills, G. J. (2011). *The emergence of procedural conventions in dialogue. In: Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Boston.

Mills, G. J. & Gregoromichelaki, E. (2008/in prep.) *Coordinating on joint projects*. Paper based on talk given at the Coordination of Agents Workshop, Nov 2008, KCL.

Mills, G. J. & Gregoromichelaki, E. (2010). *Establishing coherence in dialogue: sequentiality, intentions and negotiation. In Proceedings of the 14th SemDial*, PozDial.

Morgan, J. (1973). Sentence fragments and the notion 'sentence'. In B. Jashru, R. Lees, Y. Malkiel, A. Pietrangeli, & S. Saporta (Eds.), *Issues in linguistics* (pp. 719–751). Urbana: University of Illinois Press.

Newmeyer, F. J. (2010). Formalism and functionalism in linguistics. *Wiley Interdisciplinary Reviews*: Cognitive Science, *1*(3), 301–307, May/June 2010.

Ono, T., & Thompson, S. A. (1995). What can conversation tell us about syntax? In P. W. Davis (Ed.), *Descriptive and theoretical modes in the alternative linguistics* (pp. 213–271). Amsterdam: John Benjamins.

Perner, J. (1991). *Understanding the representational mind*. Cambridge: MIT Press.

Petukhova, V., & Bunt, H. (2011). *Incremental dialogue act understanding. In: IWCS '11 Proceedings of the 9th International Conference on Computational Semantics*.

Phillips, C. (1996). Order and structure. PhD dissertation, Massachusetts Institute of Technology.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169–226.

Poesio, M., & Rieser, H. (2010). Completions, coordination, and alignment in dialogue. *Dialogue and Discourse, 1*(1), 1–89.

Poesio, M., & Traum, D. R. (1997). Conversational actions and discourse situations. *Computational Intelligence*, *13*(3), 309–347.

Preston, B. (1994). Behaviorism and mentalism: Is there a third alternative? *Synthese, 100*(2), 167–196.

Purver, M. (2006). The theory and use of clarification requests in dialogue. PhD thesis, University of London.

Purver, M., Howes, C., Gregoromichelaki, E., & Healey, P. G. T. (2009). *Split utterances in dialogue: a corpus study. In: Proceedings of SigDial London*, Sep 2009.

Purver, M., Gregoromichelaki, E., Meyer-Viol, W., & Cann, R. (2010). *Splitting the i's and crossing the you's: Context, speech acts and grammar. In Proceedings of SemDial 2010 (PozDial)*, Poznan, June 2010.

Purver, M., Eshghi, A., & Hough, J. (2011). *Incremental semantic construction in a dialogue system. In: Proceedings of 9th International Conference on Computational Semantics (IWCS)*, Oxford, Jan 2011.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Ruiter De, J. P., Holger Mitterer, & Enfield, N. J. (2006). Projecting the end of a speakers turn: A cognitive cornerstone of conversation. *Language, 82*(3), 515–535.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.

Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language, 24*(01), 139–161.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. *Syntax and semantics, 12*, 261–286.

Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52–133). Cambridge: Cambridge University Press.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis.* Cambridge: Cambridge University Press.

Schiffer, S. R. (1972). *Meaning*. USA: Oxford University Press.

Schlangen, D. (2003). A coherence-based approach to the interpretation of non-sentential utterances in dialogue. (PhD thesis, University of Edinburgh).

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhears. *Cognitive Psychology, 21*, 211–232.

Searle, J. R. (1969). *Speech acts, an essay in the philosophy of language*. Cambridge: Cambridge University Press.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 401–415). Cambridge: MIT Press.

Searle, J. R. (1995). *The social construction of reality*. New York: Free Press.

Singh, M. P. (1999). *A social semantics for agent communication languages. In: Proceedings IJCAI'99 Workshop on Agent Communication Languages* (pp. 75–88). (Issues in agent communication lecture notes in computer science, 2000, Vol. 1916/2000, pp. 31–45, DOI: 10.1007/10722777_3).

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell. 2nd edn.

Stainton, R. J. (2005). In defense of non-sentential assertion. In Z. G. Szabo (Ed.), *Semantics versus Pragmatics* (pp. 383–457). Oxford: OUP.

Stone, M. (2004). Intention, interpretation and the computational structure of language. *Cognitive Science, 28*(5), 781–809.

Stone, M. (2005). Communicative intentions and conversational processes in human–human and human-computer dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use* (pp. 39–70). Cambridge: MIT Press.

Strawson, P. F. (1964). Intention and convention in speech acts. *The Philosophical Review, 73*(4), 439–460.

Sturt, P., & Crocker, M. W. (1996). Monotonic syntactic processing: A cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes, 11*, 448–494.

Sturt, P., & Lombardo, V. (2005). Processing coordinate structures: Incrementality and connectedness. *Cognitive Science, 29*, 291–305.

Suchman, L. A. (1987/2007). *Plans and situated actions: The problem of human-machine communication.* Cambridge: Cambridge University Press.

Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences, 35*, 75–97.

Tomasello, M. (2008). *Origins of human communication*. Cambridge: MIT Press.

Traum, D. R. & Allen, J. F. (1994). Towards a formal theory of repair in plan execution and plan recognition. *Proceedings of UK planning and scheduling special interest group*.

Traum, D. R., Marsella, S. C., Gratch, J., Lee, J., & Hartholt, A. (2008). *Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In 8th International Conference on Intelligent Virtual Agents*.

Tuomela, R. (1995). *The importance of us: A philosophical study of basic social notions*. Palo Alto: Stanford University Press.

Tuomela, R. (2005). We-intentions revisited. *Philosophical Studies, 125*(3), 327–369.

Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. Oxford: Oxford University Press.

Walton, D. N., & Krabbe, E. C. W. (1995). *Commitment in dialogue*. Albany: State University of New York Press.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge: MIT Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684.

Wharton, T. (2003). *Pragmatics and the showing–saying distinction*. (PhD thesis, UCL).

Wittgenstein, L. (1953). *Philosophical investigations, trans GEM Anscombe*. Oxford: Blackwell.

Logophoricity and Neo-Gricean Truth-Conditional Pragmatics

Yan Huang

Abstract Logophoricity refers to the phenomenon whereby the 'perspective' or 'point of view' of an internal protagonist of a sentence or discourse, as opposed to that of the current, external speaker, is being reported by using some morphological and/or syntactic means. The term 'perspective' or 'point of view' is used here in a technical sense and is intended to encompass words, thoughts, knowledge, emotion, perception and space-location [e.g. Huang (2000a: 173, 2001, 2002: 213–224, 2006/2009: 18–25, 2010a: 75–101)]. The aim of this article is threefold. In the first place, I shall provide a cross-linguistic, descriptive analysis of the phenomenology of logophoricity. Secondly, I shall present a pragmatic account of logophoricity and the related use of regular expressions/pronouns in terms of conversational implicature, utilizing the revised neo-Gricean pragmatic theory of anaphora developed by Huang (1991, 1994/2007, 2000a, 2000b, 2004, 2007, 2010a: 75–101, 2010b: 33–37)] [see also e.g. Levinson (2000)]. Thirdly and finally, I shall argue that (1) the neo-Gricean pragmatic analysis of logophoricity and the related use of regular expressions/pronouns in terms of pragmatic intrusion made here provides further evidence in support of the thesis that contrary to the classical Gricean position, pragmatics does 'intrude' or enter into the conventional, truth-conditional content of a sentence uttered, (2) pragmatic intrusion into

---

Y. Huang (✉)
University of Auckland, Auckland, New Zealand
e-mail: yan.huang@auckland.ac.nz

---

logophoricity is a conversational implicature rather than an explicature/pragmatically enrich said/impliciture, and (3) it involves 'pre'-semantic neo-Gricean pragmatics.

1 Introduction

Logophoricity refers to the phenomenon whereby the 'perspective' or 'point of view' of an internal protagonist of a sentence or discourse, as opposed to that of the current, external speaker, is being reported by using some morphological and/or syntactic means.[1] The term 'perspective' or 'point of view' is used here in a technical sense and is intended to encompass words, thoughts, knowledge, emotion, perception and space-location (e.g. Huang 2000a: 173, 2001, 2002, 2006/2009, 2010a, c). The concept of logophoricity was introduced in the analysis of African, especially West African languages like Donno Sɔ, Ewe and Tuburi, where there is a separate paradigm of logophoric pronouns, which is employed for such a purpose (e.g. Hagège 1974; Clements 1975). As an illustrating example, consider (1) taken from Donno Sɔ (Culy 1994).[2]

(1)		Logophoric pronouns: free form					
	a.	Oumar	Anta	inyemeñ	waa	be	gi.
		Oumar	Anta	LOG-ACC	seen	AUX	said
		'Oumar$_1$ said that Anta$_2$ had seen him$_1$.'					
	b.	Oumar	Anta	woñ	waa	be	gi.
		Oumar	Anta	3SG-ACC	seen	AUX	said
		'Oumar$_1$ said that Anta$_2$ had seen him$_3$.'					

In (1a) the use of the logophoric pronoun encodes a coreferential reading between it and the matrix subject. By contrast, in (1b) the employment of the regular pronoun indicates a disjoint reference. Described in this way, logophoricity can be regarded as a special case of anaphora.

The organization of this paper is as follows. I shall first discuss cross-linguistic marking of logophoricity in Sect. 2. Section 3 presents a number of implictional universals for logophoricity. Next in Sect. 4, I shall provide a pragmatic analysis of logophoricity and the related use of regular expressions/pronouns in terms of conversational implicature, utilizing the revised neo-Gricean pragmatic theory of anaphora developed by Huang (1991, 1994/2007, 2000a, 2004, 2007, 2010a) (see also e.g. Levinson 2000). Finally, in Sect. 5, I shall argue that (1) the neo-Gricean pragmatic analysis of logophoricity and the related use of regular expressions/

pronouns made here provides further evidence in support of the thesis that contrary to Grice, pragmatics does 'intrude' or enter into the conventional, truth-conditional content of a sentence uttered, (2) pragmatic intrusion into logophoricity is a conversational implicature rather than an explicature/pragmatically enriched said/impliciture, and (3) it involves 'pre-'semantic neo-Gricean pragmatics.

2 Cross-Linguistic Logophoric Marking

Cross-linguistically, logophoricity may be morphologically and/or syntactically accomplished by one or more of the following mechanisms: (1) logophoric

(2)		Logophoric pronouns: cliticised to the verb		
		(Ewe, Clements 1975)		
	a.	Kofi	be	ye-dzo.
		Kofi	say	LOG-leave
		'Kofi$_1$ said that he$_1$ left.'		
	b.	Kofi	be	e-dzo.
		Kofi	say	3SG-leave
		'Kofi$_1$ said that he$_2$ left.'		

(3)		Logophoric addressee pronouns						
		(Mapun, Frajzyngier 1985)						
	a.	n-	sat	n-wur	taji	gwar	dim	n	Kaano.
		I	say	BEN-3SG	not	ADDR	go	to	Kano
		'I told him$_1$ that he$_1$ may not go to Kano.'							
	b.	n-	sat	n-wur	taji	wur	dim	n	Kaano.
		I	say	BEN-3SG	not	3SG	go	to	Kano
		'I told him$_1$ that he$_2$ may not go to Kano.'							

(4)		Logophoric cross-referencing		
		(Akɔɔse, Curnow 2002)		
	a.	a-hɔbe	a	mə-kag
		he-said	RP	LOG-should go
		'He$_1$ said that he$_1$ should go.'		
	b.	a-hɔbe	a	a-kag
		he-said	RP	he-should go
		'He$_1$ said that he$_2$ should go.'		

---

(5)		Logophoric verbal affixes		
		(Eleme, Bond 2006)		
	a.	a-gbi	kɔ	a-dɔ-ɛ.
		3 AP-think	that	3 AP-fall-LOG
		'He$_1$ thought that he$_1$ fell.'		
	b.	a-gbi	kɔ	a-dɔ.
		3 AP-think	that	3 AP-fall
		'He$_1$ thought that he$_2$ fell.'		

---

(6)	First person logophoric marking				
	(Karimojong, quoted in Curow 2002)				
	abu-papa	tolim	ɛbe	alozi	iŋez	moroto.
	AUX-father	say	that	1S-go	3S	Moroto
	'The father said that he was going to Moroto'					

---

(7)		Long-distance reflexives				
		(Chinese)				
	(a)	Xiaoming	yiwei	Xiaohua	xihuan	ziji.
		Xiaoming	think	Xiaohua	like	self
		'Xiaoming$_1$ thinks Xiaohua$_2$ like self$_{1/2}$.'				
	(b)	Xiaoming	yiwei	Xiaohua	xihuan	ta.
		Xiaoming	think	Xiaohua	like	3SG
		'Xiaoming$_1$ thinks Xiaohua$_2$ like him$_{1/3}$.'				

---

pronouns, which may take free forms, as in (1a) above or be cliticised to the verb, as in (2a), (2) logophoric addressee pronouns, as in (3a), (3) logophoric cross-referencing, as in (4a), (4) logophoric verbal affixes, as in (5a), (5) first-person logophoric marking, as in (6), and (6) long-distance reflexives, as in (7a).

Out of these logophoric marking devices, (1), (2), (3) and (7) represent a (pro)nominal strategy, and logophoricity is marked overtly in syntax. By contrast, (4) and (5) display a verbal strategy, and logophoricity is indicated morphologically.[3] Notice that the verbal strategy represents a violation of categorical iconicity, because the function of reference-tracking is indicated on the verb rather than on the noun (e.g. Huang 2000a: 175). A further point of interest is that a language may use a combination of the logophoric marking mechanisms, mentioned above, to encode logophoricity. For example, Mapun has both logoporic

---

(8)	Hierarchy of logophoric marking mechanisms	
	a.	Logophoric expressions
	b.	Long-distance reflexives

---

pronouns and logophoric addressee pronouns (Frajzyngier 1985). Donno Sɔ is a language which contains both logophoric pronouns and first-person logophoric marking. In Moru, there are both logophoric pronouns and logophoric cross-referencing (Curow 2002). Finally, as pointed in Huang (2000a: 226), cross-linguistically logophoric marking is done according to the hierarchy in (8).

What (8) basically says is this: for logophoric marking, a logophoric expression will be used if there is one; otherwise, a long-distance reflexive will be used. For example, Gokana is a language that has logophoric forms, i.e. logophoric verbal suffixes or logophoric clitics. In this language, logophoricity is encoded by a logophoric verbal suffix or clitic. On the other hand, Modern Greek has no logophoric expressions, and consequently uses long-distance reflexives to mark logophoricity (Chiou and Huang 2010). The hierarchy in (8) also predicts that if a language has both logophoric forms and reflexives, logophoric expressions and reflexives will have different syntactic distributions. This prediction is borne out, for example, by Ewe (Clements 1975) and Moru (Andersen and Goyvaerts 1986).[4]

3 Some Implicational Universals with Respect to Logophoricity

Following Hyman and Comrie (1981), I have proposed a number of implicational universals for logophoricity in Huang (2000a, 2001, 2002, 2006/2009, 2010a). The first of these is concerned with the person distinction of logophoric expressions.

---

(9)   Person hierarchy for logophoric pronouns
       3 > 2 > 1
       First-person logophoric expressions imply second-person logophoric expressions,
       and second-person logophoric expressions imply third-person logophoric expresssions.

Given (9), it is predicted that in all languages with logophoric expressions, logophoric expressions can be third-person; in some, they can also be identified as second-person; in a few, they can be distinguished on first-person as well. For example, the logophoric expression in Donno Sɔ can be third-person only (Culy 1994). The same is true of the logophoric forms in Kresh, Noni, Sango, Togo Kã (e.g. Huang 2000a) and Karimojong and Lotuko (Curow 2002). By contrast, in Mundani, the logophoric pronoun is used for third- and second-, but not for first-person (Parker 1986). Other African languages whose logophoric expressions can be third- and second, but not first-person include Akɔɔse, Moru (Anderson and Goyvaerts 1986), and Ngbaka. Finally, in languages like Lele (Wiesemann 1986), and Yag Dii, logophoric marking can be done in all three persons (see also von Roncador 1992 for a two-way marking system on person based on the argument that some languages such as Ewe exhibit syncretism for second-and third-persons with regard to logophoric pronouns).

(10)	a.	aè	kɔ	aè	dɔ-ɛ.
		he	said	he	fell-
					LOG

'He₁ said that he₁ fell.'

	b.	aè	kɔ	aè	dɔ.
		he	said	he	fell

'He₁ said that he₂ fell.'

(11)	a.	oò	kɔ	oò	dɔ-ɛ.
		you	said	you	fell-
					LOG

'You said that you fell.'

	b.	oò	kɔ	oò	dɔ.
		you	said	you	fell

'You said that you fell.'

(12)	a.	mm	kɔ	mm	dɔ-ɛ.
		I	said	I	fell

'I said that I fell.'

	b.	mm	kɔ	mm	dɔ.
		I	said	I	fell

'I said that I fell.'

A further piece of evidence in favor of (9) comes from Gokana. In this language, while third-person logophoric marking is obligatory, second-person logophoric marking is optional but preferred, and first-person logophoric marking is optional but dispreferred (Hyman and Comrie 1981).

This pattern of person distinction holds also for long-distance reflexives. As pointed out in Huang (2000a, 2001, 2002, 2006/2009, 2010a), there is a functional/pragmatic explanation for (9). For referential disambiguity, the non-deictic third-person distinction is the most, and the deictic first-person distinction, the least useful, with the deictic second-person distinction in between, since third-person is closer to nonperson than either first- or second-person. It follows, therefore, that the fact that first-person logophoric expressions are very rare, if not non-existent, in natural languages, is hardly surprising, given that logophoric forms are one of the (most common) devices the current, external speaker (which is encoded usually in terms of a first-person pronoun) utilizes in reflecting the perspective of anyone else (usually an internal protagonist) but him- or herself.

(13)		Number hierarchy for logophoric pronouns
		Singulars > plurals
		Plural logophoric expressions imply singular logophoric expessions.

Secondly, there is the implicational universal for the number specification of logophoric expressions.

While all languages with logophoric expressions allow singular logophoric expressions, only some permit plural logophoric expressions as well. Mundang, for example, is a language which has only singular logophoric pronouns (Hagège 1974). Consequently, when reference is made to a set of internal protagonists of a reported event, the plural form of a first-person regular, nonlogophoric pronoun is used instead. Other African languages which have only singular logophoric forms include Babungo, Igbo, and Songhai. By contrast, Ewe, Gbandili, and Ngwo are languages whose logophoric pronouns have both singular and plural forms (e.g. Huang 2000a). This pattern of number specification is true also for long-distance reflexives. Again, from the viewpoint of referential disambiguity, singulars are more important than plurals.

A further point of interest is that a plural logophoric form can be used for a singular antecedent, provided that the antecedent is properly included in the set denoted by the plural logophoric expression (and that the singular antecedent and the plural logophoric form accord to the universal for conjunction of different persons, i.e. $1 + 1$, $1 + 2$, $1 + 3 =$ 1plural.; $2 + 2$, $2 + 3 =$ 2plural; $3 + 3 =$ 3plural) (e.g. Hyman and Comrie 1981). In contrast, the use of a plural regular pronoun in general does not include the matrix subject. This is illustrated in (14). The same pattern is found in Ewe, Donno Sɔ, Gokana, Lele, and Mapun (see e.g. Huang 2000a for examples).

(14)	(Eleme, Bond 2006)			
a.	aɲɛ	lama-mi	kɔ	ɛ-ba-dɔ-ba.
	3SG	tell-1SG	COMP	3-3PL-fall-LOG
	'He₁ told me that they$_{(1 + 2)}$ fell.'			
b.	aɲɛ	lama-mi	kɔ	ɛ-ba-dɔ.
	3SG	tell-1SG	COMP	3-3PL-fall
	'He₁ told me that they₂ fell.'			

We move next to the grammatical functions a logophoric expression can take. Again, there seems to be an implicational universal here, namely:

(15)	Grammatical function hierarchy for logophoric expressions
	Non-possessives > possessives
	Possessive forms/functions imply non-possessive forms/functions.

In Igbo, for example, the logophoric pronoun has only one form, and can occur only as subject. But no such restriction is imposed on logophoric expressions in the majority of African logophoric languages. As already mentioned, Mundang distinguishes between personal and possessive forms of the logophoric pronoun. In Zande, *u* is the logophoric form for third-person singular subject, *rus*, the logophoric form for third-person singular object, *ami*, the logophoric form for third-person plural subject, and *ra*, the logophoric form for third-person plural object. In a similar way, Mundani also uses different forms for different grammatical functions such as subject, object, possessor, and emphatic (e.g. Huang 2000a). The same implicational universal holds for the grammatical function distinction of long-distance reflexives.

Next, mention should be made of logocentric triggers, namely those NPs that can act as an antecedent for a logophoric expression and a long-distance reflexive. First, logocentric triggers are generally constrained to be a core-argument of the logocentric predicate of the matrix clause. Secondly, they are typically subjects. In other words, a logophoric expression and a long-distance reflexive are canonically subject-oriented. Contrariwise, a regular pronoun is not. This contrast is illustrated in (16) and (17).

---

(16)	(Tuburi, Wiesemann 1986a, b: 448–449)					
	a.	Pɔl	riŋ	Jaŋ	gá	sɛ	lɛ'ɛ.
		Paul	said	to-John	that	LOG	fell
		'Paul$_1$ said to John$_2$ that he$_1$ fell.'					
	b.	Pɔl	riŋ	Jaŋ	gá	a	lɛ'ɛ .
		Paul	said	to-John	that	he	fell
		'Paul$_1$ said to John2 that he$_{2/3}$ fell.'					

---

(17)	(Japanese)					
	Takasi-ga	Taroo-ni	Yoshiko-ga	zibun-o	hihansita	to	itta.
	Takasi-NOM	Taroo-DA	Yoshiko-NOM	self-ACC	criticized	that	said
	'John$_1$ told Bill$_2$ that Mary$_3$ criticized self$_{1/*2/3}$.'						

However, logocentric triggers can also be some other, non-subject argument, provided that this argument represents the 'source' of the proposition or the 'experience' of the mental state that is being reported. Two types of construction are particularly common. The first involves the predicate 'hear from', as in (18) and (19). Similar examples can be found in Donno Sɔ, Gokana, and Tuburi and many other long-distance reflexivization languages.

---

(18)	(Ewe, Clements 1975)					
	Ama	se	tso	Kofi	gbɔ	be	yè-xɔ	nunana.
	Ama	hear	from	Kofi	side	COMP	LOG-receive	gift
	'Ama$_1$ heard from Kofi$_2$ that she$_1$/he$_2$ had received a gift.'						

(19)	(Korean)				
	Kim-un	Lee-loputhe	caki-ka	tayhak	iphaksihem-ey
	Kim-TOP	Lee-from	self-NOM	college	entrance examination-at
	hapkyekhayssta-nun	iyaki-lul		tulessta.	
	passed	-that story-ACC		heard	
	'Kim$_1$ heard from Lee$_2$ that self$_{1/2}$ passed the college entrance examination.'				

The second involves 'psychological' predicates expressing emotional states and attitudes, of which the 'experiencer' frequently acts as direct object or object of preposition. This is the case in (20), (21), and (22). The same pattern can be found in Gokana, Ewe, and Mundani, and numerous long-distance reflexivization languages

(20)	(Tuburi, Wiesemann 1986a, b: 449)					
	hééné	jɔŋ	Pɔl	gá	sɛ̀	lɛ̀'	cegè.
	fear	has	Paul	that	LOG	fall	sick
	'Fear grips Paul$_1$ that he$_1$ will fall sick.'					

(21)	(Eleme, Bond 2006)			
	a-waa	osaro	kɔ	a-dɔ -ɛ̀.
	3AP-anger	Osaro	that	3AP-fall-LOG
	'It angered Osaro$_1$ that he$_1$ fell.'			

(22)	(Tamil, Amritavalli 2000)				
	taan	toottadu	kumaare	romba	paadiccadu.
	self	defeat-PAST-it	Kumar-ACC	much	affect-PAST-it
	'That self$_1$ was defeated affected Kumar$_1$ very much.'				

We have the hierarchy for logocentric triggers/antecedent in (23).

(23)	Hierarchy for logocentric triggers
	Surface structure: subject > object > others
	Semantic role: agent > experiencer > benefactor > others

What (23) basically says is that the higher an NP is on the hierarchy, the more likely it will function as an antecedent for a logophoric expression or a long-distance reflexive. Given that the subject of the matrix clause is typically the NP that is highest on the hierarchy (and incidentally most animate) within a sentence, it is hardly surprising that it is the typical antecedent for a logophoric expression or a long-distance reflexive. Hierarchy (23) also provides a natural explanation for the examples in (20), (21) and (22). Because the subjects in these examples are non-referential and non-human, by (23), the logophoric expressions are naturally linked to the next highest NP available on the hierarchy, namely the objects. More or less the same can be said of examples of the following kind, where the ante-

(24)	(Tuburi, Wiesemann 1986a: 449)							
	ɓil	ɓɛg	fɛh	wɛrmàngá	sɛ	ko	Jaŋ	mɔnɔ.
	stomach	his	ACCOM	happy	because	LOG	see	John
	'He was happy because he saw John.'							

cedent of the logophoric pronoun is a possessor.

Taken together, the above four hierarchies predict that the most basic, unmarked pattern of logophoric marking is one which encodes logophoricity by the use of a third-person, singular, non-possessive, logophoric expression or long-distance reflexive which refers to a human subject.

Finally it should be pointed out that logophoric expressions and long-distance reflexives usually occur in a logophoric domain, that is, a stretch of discourse in which the internal protagonist's perspective is being represented. In general, a logophoric domain starts in a clause which is subordinate to one in which the logocentric trigger is identified either explicitly or implicitly. Following Culy (1994), I shall call this part of the logophoric domain the sentential logophoric domain. In contrast, I shall call the logophoric domain which operates across clause boundaries the discourse logophoric domain. Logophoric 'binding' across sentences is found in a number of African languages including Angas, Bwamu, Donno Sɔ, Ewe, Fon, Gokana, Tuburi, and perhaps Babungo and Mundani (e.g. Huang 2000a). Likewise, 'binding' of long-distance reflexives can also be commonly across sentence boundaries into discourse.

The logophoric domain is commonly created by a logocentric licenser, which is of two types: (1) logocentric predicates, and (2) logocentric complementisers. Logocentric predicates can largely be distinguished on a semantic basis. The most common types of logocentric predicate are predicates of speech and thought. But other types of predicate such as those of mental state, knowledge, and direct perception can also trigger a logophoric domain. Languages differ in allowing precisely which type of predicate to function as a logocentric licenser. For example, while in Ewe and Mundani, the first four types of predicate mentioned above are allowed to act as a logocentric licenser, in Donno Sɔ, predicates other than those of speech and thought are in general excluded. In Mundang, only

predicates of asserting, ordering, and more rarely, thinking can license a logo-phoric domain. There are even languages where logocentric predicates are further restricted. Igbo and Mapun are just such languages; the former restricts the logophoric domain to predicates of communication, and the latter, just to predicates of speech, and perhaps only to the predicate *sat* 'say' (Frajzyngier 1985). Another African language whose logophoric domain is limited to the predicate *tèlè* 'say' seems to be Sango. Furthermore, the types of logocentric predicate may be affected by the grammatical functions the logophoric pronoun performs, as in Togo Kã (e.g. Huang 2000a).

(25)	An implicational universal for logocentric predicates
	Speech predicates > epistemic predicates > psychological predicates > knowledge predicates > perceptive predicates > unmarked directional predicates

But cross-linguistically there does seem to be an implicational universal for logophoric predicates (Stirling 1993; Culy 1994; Huang 2000a).

What (25) basically says is this: if a language allows (some) predicates of one class to establish a logophoric domain, then it will also allow (some) predicates of every class higher on the hierarchy to do the same. Thus, if a language has logophoric marking with predicates of, say, psychological state, it will then necessarily have it with predicates of thought and communication.

There is one pattern of long-distance reflexives (especially in East Asian languages and Turkish) which has not been attested for logophoric expressions in African language. This involves the use of deictically-oriented directional predicates such as 'come/go' and 'bring/take'. As can be shown by (26), while the use of 'come' in (26a) allows long-distance reflexivisation, the use of 'go' in (26b) does not. Furthermore, note that this contrast is independent of whether or not a logocentric predicate occurs in the matrix clause (see Huang 1994, 2000a for examples from Turkish, Chinese and Korean).

(26)	(Japanese, quoted in Huang 1994/2007, 2000a)							
a.	Taroo	wa	zibun	ni	ai	ni	kita	hito
	Taroo	TOP	self	to	see	to	came	people
	ni-wa	dare-demo,	syokuzi o	dasu.				
		whoever	meal	offer				
	'Taroo$_1$ offers a meal to anybody who has come to see self$_1$.'							
b.	$^?$Taroo	wa	zibun	ni	ai	ni	itta	hito
	Taroo	TOP	self	to	see	to	went	people
	ni-wa	dare-demo,	syokuzi o	dasu.				
		whoever	meal	offer				
	'Taroo$_1$ offers a meal to anybody who has gone to see self$_1$.'							

This contrast seems to be attributed to the fact that the use of 'come' in (26a) makes clear what is reported is from the space-location of the matrix subject, therefore the matrix subject is the pivot, or the relativized 'centre of deixis' in the logophoric domain. Hence the possibility of long-distance reflexivisation. In other words, 'come' must be interpreted as describing movement towards the matrix subject. On the other hand, the use of 'go' in (26b) is an indication that what is described is not from the 'camera angle' of the matrix subject, rather it indicates movement away from the matrix subject, therefore the matrix subject cannot be the pivot or the logocentric trigger. Hence long-distance reflexivisation is rather bad. The deictic nature of the pivot can further be seen by the fact that (26b), for example, can be improved given an appropriate context.

Now, given the hierarchy for logocentric predicates, one immediate question arises: why is the hierarchy ordered in the way it is. Currently, the most plausible explanation (Culy 1994) proposed that the hierarchy is associated with a notion of 'reliability'. The link between the hierarchy and reliability is threefold. Firstly, the more reliable the current, external speaker deems the situation to be, the more likely it is that the context will be marked as a logophoric one. Logophoric marking is used in this case to show that the current, external speaker (or the primary ego) is reporting on the logocentric trigger (or the secondary ego) and is not taking responsibility for what he or she is reporting. This notion of s[ituation]-reliability seems to explain why cross-linguistically speech predicates are more likely to generate logophoric marking than all the other predicates on the hierarchy. Unlike the action denoted by the other predicates, that referred to by speech predicates can be directly perceived. Having direct sensory evidence about the situation is certainly the clearest sign of s-reliability. The second factor affecting the hierarchy is reliability of the report, that is, the less reliable/objective the report is, the more likely it is that the context will be made a logophoric domain. Logophoric marking is used in this case to reflect the subjectivity of the truth, content, or linguistic characterization of the report. This is called r[eport 1]-reliability. R1-reliability appears to be responsible for the fact that cross-linguistically languages are less likely to mark logophoricity with predicates of knowledge and of direct perception than with predicates of speech and of thought. This is because knowledge and direct perception predicates are factive, in that they presuppose the truth of their complements. Thirdly and finally, with predicates of direct perception, the logocentric trigger also has direct, sensory evidence for the report. This is r[eport 2]-reliability. Consequently, the reports with knowledge and direct perception predicates are more reliable/objective than those with the other predicates, hence logophoric marking is less used with these reports (e.g. Huang 2000a, see also Dimmendaal 2001 and Speas 2004).

Of further interest is that as pointed in Speas (2004), the hierarchy for logocentric predicates in (25) can be inversely correlated with the hierarchy for evidentiality in (27).

(27)   Hierarchy for evidentiality (Willett 1988)
              Personal experience > direct (e.g. sensory) evidence > indirect (e.g. inferable) evidence > reported evidence or hearsay

The inverse correlation can be presented in (28).

---

(28)   An inverse correlation between logocentric predicates and evidentiality
         a.   Logocentric predicates
                   Speech predicates > epistemic predicates > psychological predicates >
                   knowledge predicates > perceptive predicates > unmarked directional
                       predicates
         b.   Evidentiality
                   Hearsay > indirect evidence > direct evidence > personal experience

---

What (28) basically says is that the more likely a predicate is to engender a
logophoric domain, the less likely it is to be a category in the evidential paradigm.
For example, the speech predicate 'say' is the predicate that is most likely to
trigger a logophoric domain crosslinguistically; hearsay is the least likely to be a
part of an evidential paradigm (e.g. Speas 2004).[5]

Also worth pointing out is what is called the skipping effect of logophoric
marking. This refers to the phenomenon whereby the embedding under a logo-
centric predicate of a sentence which originally does not allow logophoric marking
can render it logophoric, in that the logophoric form can skip over one or more
layers of embedding to reach up the logocentric trigger of the logocentric predi-
cate. For example, in Togo Kã the factual knowledge predicate 'know' does not
function as a logocentric predicate, hence the ungrammaticality of (29a). However,
when it is embedded under the logocentric predicate 'say', the sentence becomes
grammatical with the interpretation that the logophoric pronoun is coreferential
with the matrix subject, as in (29b). The same is true of Ewe.

---

(29)   (Culy 1994)
       a.   *Omar   Anta   ɛnɛ   ɔɛ $\overline{\text{II}}$ wɔ .
               Omar   Anta   LOG   saw   know   AUX
               'Omar$_1$ knows that Anta saw him$_1$.'
       b.   Madu   Omar   wa   Ali   ɛnɛ   laran   ɔɛ $\overline{\text{II}}$ wɔ   gi.
               Madu   Omar   SUBJ   Ali   LOG   sister   saw   know   AUX   said
               'Madu$_1$ said that Omar$_2$ knows that Ali$_3$ saw his$_{1/*2/*3}$ sister.'

Next, note that in some African languages the skipping effect can also arise with
respect to clause types. Ewe, for example, is a language which does not allow a
logophoric pronoun to be used inside a relative clause, as in (30a). But such a
restriction is lifted when the relative clause is embedded under the logocentric
predicate 'say', as in (30b).

(30)   (Ewe, Clements 1975)

	a.	*Ama	do	ŋku	nyɔ nuvi	hi	dze	yè	gbɔ	dyi.
		Ama	set	eye	girl	REL	stay	LOG	side	on

'Ama₁ remembered the girl who stayed with her₁.'

	b.	Ama	gbl	be	yè-o	ŋku	nyɔ nuvi	hi	dze	yè	gbɔ	dyi.
		Ama	say	COMP	LOG-set	eye	girl	REL	stay	LOG	side	on

'Ama₁ said that she₁ remembered the girl who stayed with her₁.'

The other common type of logocentric licenser is what Stirling (1993: 260) called 'report-opening' complementizers or what Dimmendaal (2001) labeled *de dicto* or 'reported speech markers' such as *be* in Ewe, *kɔ* in Gokana, *se* in Mundang, *ne* in Mundani and *ga* in Tuburi. These complementisers are often homophonous with the verb 'say' and are often developed historically out of it. Evidence for this evolutional pattern has been found in a wide range of languages of West Africa (e.g. Ewe), East Asian (e.g. Japanese and Korean), Southeast Asian (e.g. Lahu and Thai) and Chinese (e.g. Cantonese and Taiwanese), Kriyol, the English-based creole Krio, the Micronesian language Pingilapese, Russian and Sranan (e.g. Huang 2000a; Dimmendaal 2001). Some of them still carry the force of speech, as can be seen by the fact that a predicate of speech is frequently omitted before such a complementiser. This can be illustrated by Kana, Mundani, Tuburi, Lele and many other Chadic languages.

(31)   (Kana Bond 2006)

a(-kɔ)	kɔɔ	e-kii-e.
he-say	COMP	he-go-LOG.

'He₁ said that he₁ would leave.'

Interestingly, as pointed out in Huang (1994/2007), the logophoric domain in Chinese can also be triggered by the semi-complementiser *shuo*. This semi-complementiser is homophonous with the verb *shuo* 'say' and still carries the force of speech. This can be evidenced by the fact that (1) it cannot co-occur with the verb 'say', as in (32a) and (2) it can co-occur only with predicates of speech, as in (32b).

(32)

	a.	*Xiaoming	shuo	shuo	Xiaohua	bu	xihuan	ziji.
		Xiaoming	say	say	Xiaohua	not	like	self

'Xiaoming₁ says that Xiaohua₂ does not like self₁/₂.'

	b.	Xiaoming	gaosu	mama	shuo	Xiaohua	bu	xihuan	ziji.
		Xiaoming	tell	mum	say	Xiaohua	not	like	self

'Xiaoming₁ tells mum that Xiaohua₂ does not like self₁/₂.'

In many logophoric languages, a complementiser of this kind does seem to play an important role in logophoric marking. Thus, Clements (1975) reports that in Ewe all logophoric constructions contain the complementiser *be*. In a similar way, Hagège (1974) notes that in Tuburi, the use of the complementiser *ga* always gives rise to a logophoric domain. Essentially the same can be shown to hold for Gokana where the presence of the complementiser *kɔ* is sufficient for triggering logophoric marking. This connection between complementisers and logophoricity can also be observed in Lele; according to Wiesemann (1986), the whole system of logophoric pronouns in this language has developed from the grammaticalisation of the complementiser *na*. Similar observations have also been made of the correlation between complementisers and logophoric marking in other logophoric languages such as Akɔɔse, Banda-linda and Efik (e.g. Huang 2000a).

However, as pointed out by Culy (1994) and Huang (2000a), contrary to Dimmendaal (2001) there does not seem to be a universal correlation between complementisers and logophoric marking. On the one hand, logophoric marking can be without complementisers, as in Mundang (Hagège 1974), hence complementisers may not be necessary. On the other hand, complementisers may not result in logophoric marking. Furthermore, of particular interest is that not only are languages different with regard to requiring a complementiser to be present to activate a logophoric domain, so are individual logocentric predicates within a single language. This is the case in Donno Sɔ. In this language, some logocentric predicates require the complementiser *gɔ,* which is homophonous with a definite determiner rather than the verb 'say', while others do not (Culy 1994). There is thus clear evidence that complementisers do not in and of themselves give rise to a logophoric domain.[6]

4 A Neo-Gricean Pragmatic Account

Having provided a cross-linguistic description of logophoricity, I now turn to present a neo-Gricean pragmatic analysis of logophoricity and the related use of regular expressions/pronouns in logophoric languages. Let me start with the three neo-Gricean pragmatic principles proposed by Levinson (e.g. 2000, see also Huang 1991, 1994/2007, 2000a, 2007, 2010a, b, 2011).

---

(33)   Levinson's Q-, I-, and M-principles (simplified) (e.g. Huang 2007)

    a.   The Q-principle

        Speaker: Do not say less than is required (given I).

        Addressee: What is not said is not the case.

    b.   The I-principle

        Speaker: Do not say more than is required (given Q).

        Addressee: What is generally said is stereotypically and specifically exemplified

    c.   The M-principle

        Speaker: Do not use a marked expression without reason.

        Addressee: What is said in a marked way is not unmarked.

Each of these three principles has two sides: a speaker's maxim, which specifies what the principle enjoys a speaker to say, and a recipient's corollary, which dictates what it allows an addressee to infer.

The basic idea of the Q-principle is that the use of an expression (especially a semantically weaker one) in a set of contrastive semantic alternates Q-implicates the negation of the interpretation associated with the use of another expression (especially a semantically stronger one) in the same set. In other words, the effect of this pragmatic strategy is to give rise to an upper-bounding conversational implicature: a speaker, in saying '...*p*...', *ceteris paribus* conversationally implicates that (for all he or she knows) '... at most *p*...'. Seen the other way round, from the absence of a semantically stronger expression, we infer that the interpretation associated with the use of that expression does not hold. Schematically (I use the symbol +> to indicate 'conversationally implicate'):

---

(34) Q-scale: <all, some>

Most of his teachers were persecuted during Mao's 'Cultural Revolution' in China.

+> Not all of his teachers were persecuted during Mao's 'Cultural Revolution' in China

---

Next, the basic idea of the I-principle is that the use of a semantically general linguistic expression I-implicates a semantically specific interpretation. In other words, the I-principle is an upper-bounding pragmatic law which may be (and systematically is) exploited to invite lower-bounding conversational implicatures: a speaker, in saying '...*p*...', conversationally implicates that (for all he or she knows) '... more than *p*...'. More accurately the operation of the I-principle induces an inference to a proposition that is best in keeping with the most stereotypical and explanatory expectation given world knowledge. Schematically:

---

(35)   I-scale: [x,y]

y +>$_I$ x

---

(36)   (Conjunction buttressing)

p and q +> p and then q

+> p therefore q

+> p in order to cause q

John turned the key and the drawer opened.

+> John first turned the key and then the drawer opened.

+> John turned the key and therefore the drawer opened.

+> John turned the key in order to cause the drawer to open.

Finally, the basic idea of the M-principle is that the use of a marked expression M-implicates the negation of the interpretation associated with the use of an alternative, unmarked expression in the same set. In other words, from the use of a

marked expression, we infer that the stereotypical interpretation associated with the use of an alternative, unmarked expression does not hold. Schematically:

---

(37)   M-scale: {x,y}
                y +>M ∼ x

---

(38)   a.   The bus comes frequently.
                          +> The bus comes, say, every ten minutes.
                b.   The bus comes not infrequently.
                          +> The bus comes not as frequently as the uttering of (a)
                              suggests, say, every half an hour

---

Taken together, the I-, and M-principles give rise to complementary interpretations: the use of an unmarked expression tends to convey an unmarked message, whereas the use of a marked expression, a marked message. Furthermore, inconsistencies arising from the Q-, I-, and M-principles are resolved by an ordered set of precedence.

---

(39)   Levinson's resolution schema
                          a.   Level of genus: Q > M > I
                          b.   Level of species: e.g. Q-clausal > Q-scalar

---

This amounts to saying that genuine Q-implicatures tend to precede I-implicatures, but otherwise I-implicatures take precedence until the use of a marked expression triggers a complementary M-implicature to the negation of the applicability of the pertinent I-implicature (see also e.g. Huang 1991, 1994/2007, 2000a, 2000b, 2007, 2010b, 2011 for further discussion).

We move next to the neo-Gricean pragmatic theory of anaphora advanced by Huang (1991, 1994/2007, 2000a, 2004, 2007, 2010a) (see also e.g. Levinson 2000). The underlying idea is that the interpretation of certain patterns of anaphora can be made using pragmatic enrichment, parasitic on a language user's knowledge of the range of options available in the grammar and of the systematic use or avoidance of particular linguistic expressions or structures on particular occasions.

Applying the Q-, I- and M-principles, sketched above, to the domain of anaphoric reference, we can derive a general pragmatic apparatus for the interpretation of zero anaphors, pronouns, reflexives, and lexical NPs in (40).

(40) A revised neo-Gricean pragmatic apparatus for anaphora (Huang 1991, 1994/2007, 2000a, 2004, 2007, 2010a)

    (a) Interpretation principles

        (1) The use of an anaphoric expression x I-implicates a local coreferential interpretation, unless (2) or (3).

        (2) There is an anaphoric Q-scale $\langle x, y \rangle$, in which case, the use of y Q-implicates the complement of the I-implicature associated with the use of x, in terms of reference.

        (3) There is an anaphoric M-scale $\{x, y\}$, in which case, the use of y M implicates the complement of the I-implicature associated with the use of x, in terms of either reference or expectedness.

    (b) Consistency constraints

        Any interpretation implicated by (a) is subject to the requirement of consistency with

        (1) The revised DRP.

        (2) Information saliency, so that

            (a) implicatures due to matrix constructions may take precedence over implicatures due to subordinate constructions, and

            (b) implicatures to coreference may be preferred according to the saliency of antecedent in line with the following hierarchy: topic > subject > object, etc.; and

        (3) General implicature constraints, namely,

            (a) background assumptions,

            (b) contextual factors

            (c) meaning-nn, and

            (d) semantic entailments.

(41) The revised disjoint reference presumption (DRP) (see e.g. Huang 2000a, 2004, 2007)

    The co-arguments of a predicate are intended to be disjoint, unless one of them is reflexive-marked.

Coming back to the use of logophoric expressions and long-distance reflexives to encode logophoricity, the interpretation of them and their associated regular expressions/pronouns can be determined by (40).[7]

Let me start with logophoric expressions in African languages. Notice that referentially, the use of logophoric forms in these languages is in complementary distribution with that of regular expressions/pronouns, and that logophoric expressions are the only option available in the grammar of these languages to encode coreference in a logophoric domain. This is shown in (1)–(7) above.

As a result, any speaker of these languages who intends coreference will also have to use a logophoric expression. This has the consequence that if a logophoric form is not employed but a regular expression/pronoun is used instead, a Q-implicature will arise, namely neither the marking of logophoricity nor coreference is intended. In other words, we have a Q-scale $\langle$logophoricexpression, regular expression/pronoun$\rangle$ here, such that the use of the semantically weaker regular expression/pronoun Q-implicates that the use of the semantically stronger logophoric expression cannot be truthfully entertained, that

is to say, both the logophoric reading and the coreferential interpretation which are associated with the use of a logophoric form should be avoided. Schematically:

---

(42)   <logophoric expression [+ logophoric, +coreference], regular expression/pronoun
               [-logophoric, -coreference]>

             regular expression/pronoun +>Q $\sim$ logophoric expression

---

We move next to the use of long-distance reflexives to mark logoophoricity in a language without logophoric expressions. Note that unlike in African languages, in these languages, long-distance reflexives are not in referential complementary distribution with regular pronouns. In other words, there is usually a referential overlap between long-distance reflexives and regular pronouns, as the Chinese example in (7) above and the Bangla example in (43) indicates.

---

(43)   (Bangla, Sengupta 2000)

a.	babli	bolo	nije	kaj-Ta	korbe.
	Babli-NOM	said	self	work-CL-ACC	will do

'Babli$_1$ said self$_1$ would do the job.'

b.	babli	bolo	se	kaj-Ta	korbe.
	Babli-NOM	said	she	work-CL-ACC	will do

'Babli$_1$ said she$_{1/2}$ would do the job.'

While a long-distance reflexive is used for the purpose of encoding logophoricity, for coreference, a regular pronoun can be employed. Put another way, whereas the use of a long-distance reflexive encodes both logophoricity and coreference, the use of a regular pronoun may or may not encode coreference, but not logophoricity. This is sufficient enough to form a Q-scale⟨long-distance reflexive, regular pronoun⟩ to the effect that the unavailability of the semantically stronger long-distance reflexive will Q-implicate the speaker's intention to avoid at least one of the features associated with its use, namely logophoricity. Long-distance reflexives are semantically stronger than regular pronouns in that (1) syntactically they usually require to be somewhat 'bound', even in a discourse, and (2) semantically they typically have to be referentially dependent. Schematically:

---

(44)   <long-distance reflexive [+ logophoricity, +coreference], regular pronoun
                       [-logophoricty, ±coreference]>
                     regular pronoun +>Q $\sim$ long-distance reflexive

---

Alternatively, the use of long-distance reflexives and regular pronouns in languages without logopohoric expressions can be accounted for in terms of the systematic interaction between the I- and M-principles. Since the grammar of these languages allows the unmarked regular pronoun to be used to mark coreference, a speaker will use it if such an interpretation is intended. On the other hand, if the unmarked regular pronoun is not used, but the marked (morphologically more complex) long-distance reflexive is used instead, then an M-implicature is created, namely not only coreference but logophoricity as well is intended. Schematically:

(45)	{regular pronoun [-logophoric, $\pm$ coreference], long-distance reflexive [+ logophoric, +coreference]} long-distance reflexive +>M $\sim$ regular pronoun

Also worth noting is that if relevant, the choice between long-distance reflexives on the one hand and regular pronouns, on the other, is correlated with that between subjunctive and indicative mood: it is common for the use of a long-distance reflexive to go in tandem with subjunctive mood, and for the employment of a regular pronoun to go with indicative mood, as (46) shows.

(46)	(Icelandic, quoted in Huang 2000a, b)					
	a.	Jón	segir	að	María	elski	sig.
		John	says-INDIC	that	Mary	loves-SBJV	self
		'John$_1$ says that Mary loves self$_1$.'					
	b.	Jón	veit	að	María	elskar	hann.
		John	knows-INDIC	that	Mary	loves-INDIC	him
		'John$_1$ knows that Mary loves him$_1$.'					

Once again, the correlation seems to be a reflection of a semantic/pragmatic choice made by the external speaker about the responsibility he or she assumes for the truthfulness of what he or she is reporting. If a regular pronoun and indicative mood are used, as in (4bb), it shows that the speaker asserts that the report is true. He or she cannot go on to deny it because doing so will give rise to Moore's paradox. If on the other hand, a long-distance reflexive and subjunctive mood are deployed, as in (46a), it indicates that the speaker does not take the responsibility for the truth of the report. He or she can then go on to deny it. Thus, the optionality of long-distance reflexives/regular pronouns and of subjunctives/indicatives provides the speaker with a useful means of expressing his or her attitudes toward the truth of what he or she is reporting, or more broadly, of expressing evidentiality.

5 Pragmatic Intrusion into Logophoricity: Explicature, Pragmatically Enriched Said, ImplicIture or ImplicAture?

On a classical Gricean account of meaning and communication (e.g. Grice 1989), a distinction is made between what is said and what is conversationally implicated. Given Grice's definition, what is said is generally taken to be (1) the conventional meaning of a sentence uttered with the exclusion of any conventional implicature, and (2) the truth-conditional, propositional content of the sentence uttered. However, according to Grice, the truth-conditional, propositional content of what is said is not fully worked out until reference is identified, deixis is interpreted and ambiguity is resolved. How can all this be done? Grice (1989: 25) seemed to take the recovery of the truth-conditional, propositional content as largely the outcome of linguistic and contextual decoding.

It turns out however, that contrary to this classical Gricean position, the determination of indexicality and the related area of what is said involves pragmatic enrichment of some kind. It is now generally acknowledged that in working out the classical Gricean characterization of what is said, there is pragmatic intrusion of some sort, namely the pragmatically inferred content, into the truth-conditional, propositional content of the sentence uttered. This is further evidenced by the neo-Gricean pragmatic analysis of logophoricity and the related use of regular expressions/pronouns in terms of pragmatic intrusion offered here as a special case of reference determination.

The question that arises next is what pragmatic intrusion under consideration is. Currently, two approaches can be identified: the non-conversational implicature and the conversational implicature approaches. Within the first camp, three analyses are of particular interest. First, in relevance theory, pragmatic intrusion into what is said is analyzed as explicature—an inferential development of one of the linguistically-given incomplete conceptual representation or logical forms of a sentence uttered (e.g. Sperber and Wilson 1995). Defined thus, it is a pragmatically inferred component of the Gricean notion of what is said (though what is said is abandoned in relevance theory). Secondly, somewhat similar to the relevance-theoretic view is the position taken by Recanati (e.g. 2004). According to Recanati, pragmatic intrusion under discussion is part of pragmatically enriched said. Finally, a third approach is due to Bach (e.g. 2004). On Bach's view, there is no pragmatic intrusion into what is said. This is because certain communicative contents do not need to be recognized as either part of what is said or part of what is conversationally implicated. Rather, they constitute a middle ground between what is said and what is conversationally implicated. Bach dubbed this middle level of speaker-meaning conversational impliciture or impliciture for short (see also Huang 2010d).

On the other hand, within the neo-Gricean pragmatic framework, Levinson (2000: 195–196) is of the view that pragmatic intrusion into what is said is neither an explicature; nor the pragmatically enriched said; nor an impliciture. Rather, it is

the same beast as a neo-Gricean conversational implicature. In my view, the reason why it is a conversational implicature is threefold. Firstly, so-called explicature/the pragmatically enriched said/impliciture is engendered largely by the same Gricean pragmatic mechanism that yields a conversational implicature. Secondly, Recanati (1993) put forward two tests, i.e. the availability principle and the scope principle, to differentiate explicature/the pragmatically enriched said/impliciture from conversational implicature. But as I argued in Huang (2007), neither of Recanati's tests really works from a theoretical point of view (but see Capone 2009). This is also the case with work carried out in experimental pragmatics. I do not think that there is any experiment that can differentiate explicature/pragmatically enriched said/impliciture from conversational implicature. Therefore, currently there is no failsafe test (both conceptual and experimental) that can be employed to distinguish alleged explicature/the pragmatically enriched said/impliciture from conversational implicature on a principled basis. Thirdly, other things being equal, given the metatheoretical principle known as 'Occam's razor' ('theoretical entities are not to be multiplied beyond necessity'), the implicature analysis is theoretically and methodologically preferable, because it postulates less representational levels in the interpretation of an utterance than the explicature/pragmatically enriched said/impliciture account because the latter has also to include conversational implicature analysis. If neo-Gricean conversational implicature can intrude on to truth-conditional content of an utterance, then a problem known as Grice's circle arises, namely, how what is conversationally implicated can be defined in contrast to, and calculated on the basis of what is said, given that what is said seems to both determine and to be determined by what is conversationally implicated (e.g. Levinson 2000; Huang 2007). Levinson's proposal was that one should reject the 'received' view of the pragmatics-semantics interface, namely, the view that the output of semantics is the input to pragmatics, and allow implicatures to play a systematic role in 'pre'-semantics, that is, to help determine the truth-conditional, propositional content of an utterance (Levinson 2000, see also Huang 2007). Putting it slightly differently, in order to avoid Grice's circle, one needs both 'pre'-semantic pragmatics and 'post'-semantic pragmatics or what Korta and Perry (2008) called near-side and far-side pragmatics.

In conclusion, both logophoric expressions and long-distance reflexive are the anaphoric-linking or reference-tracking devices that are used mainly to encode logophoricity. The neo-Gricean pragmatic analysis of logophoricity and the related use of regular expressions/pronouns in terms of pragmatic intrusion made here provides further evidence in support of the thesis that contrary to the classical Gricean position, pragmatics does 'intrude' or enter into the conventional, truth-conditional content of a sentence uttered. Pragmatic intrusion into logophoricity is a conversational implicature rather than an explicature/pragmatically enrich said/impliciture, and it involves 'pre'-semantic neo-Gricean pragmatics.

6 Notes

1. There is at least one other, wider definition for logophoricity in the generative syntax literature. Under this definition, any NP, in particular, any anaphor in the Chomskyan sense which cannot be bound in its local domain either in the sense of minimal configuration or in the sense of minimal predicate is taken to be a logophoric expression. There is, however, one major problem attendant to this definition of logophoricity, namely the problem of circularity. Unless the notion of logophoricity can be defined independently of binding theory, as in our case, to say that an anaphor is a logophoric form simply because it violates binding theory is circular.

2. Abbreviations: ACC, accusative; ADDR, addressee pronoun; AUX, auxiliary; COMP, complementizer: DAT, dative; INDIC, indicative;LOG, logophoric; NOM, nominative; REL, relative clause; RP, report particle; SBJV, subjunctive; SUBJ, subject; TOP, topic.

3. In some cases, the distinction is not clear-cut. Curnow (2002), for example, is of the view that logophoric pronouns that are cliticised to verbs are intermediate between logophoric pronouns and logophoric cross-referencing.

4. Following suggestions made by von Roncador (1992) and Culy (1994), I grouped languages into three types with respect to logophoricity: (1) full or pure logophoric languages, languages which have special morphological and/or syntactic forms that are employed only in logophoric domains, be the forms logophoric pronouns, logophoric addressee pronouns, logophoric cross-referencing forms, logophoric verbal affixes and/or first-person logophoric marking devices (e.g. Babungo, Pero, and Ekepeye); (2) non-logophoric languages, languages which have no such special morphological and/or syntactic forms (e.g. Arabic, English and perhaps Abrom, Agni, Bargu, Mambar, and Moore) (von Roncador 1992); and (3) semi or mixed logophoric languages, languages which allow either logophoric expressions to be used for non-logophoric purposes (e.g. Igbo, Idoma and Yoruba) or the extended use of reflexives in logophoric contexts (e.g. Italian, Malay, and Northern Pomo). For a list of full or pure logophoric languages, see Huang (2000a, b). See also Cunrow (2002) for a different view.

   Interestingly enough, as Culy observes, while logophoric languages are found in many places throughout the world, full/pure logophoric languages seem to be found only in Africa. Furthermore, while full/pure logophoric languages are not in a contiguous area, logophoric languages as a whole are in a contiguous area. This geographic distribution of logophoric languages is fascinating as well as surprising, and for the time being, remains unexplained.

5. Given my formulation of the implicational universal for logocentric predicates, the matching can only be partial.

6. Logophoric domains in African languages can be extended to syntactic constructions which do not seem to be directly related to the reporting of an

internal protagonist's perspective. Again, languages vary greatly here and the extension may have to be stipulated on a language-by-language basis. But there do appear to be some common patterns cross-linguistically, two of which are (1) purpose clauses, as has been found in Babungo, Donno Sɔ, Gokana, Lele and Yag Dii; and (2) relative clauses, as has been attested in Mundang, Donno Sɔ, Gokana, Mundani, and Tuburi. In East Asian languages, logophoric domain can be extended to other types of syntactic construction such as the topic construction, and the relative construction.

7. A point to be borne in mind is that logophoricity and coreference are two distinct, though intimately related notions; logophoricity entails coreference, but not vice versa.

References

Amritavalli, R. (2000). Lexical anaphors and pronouns in Tamil. In Lust et al. (Eds.), (pp. 169–216).

Andersen, T., & Goyvaerts, D. L. (1986). Reflexivity and logophoricity in Moru-Madi. *Folia Linguistica, 20*, 297–318.

Bach, K. (2004). Pragmatics and the philosophy of language. In L. R. Horn & G. Ward (Eds.), (pp. 463–87).

Bond, O. (2006). A broader perspective on point of view: Logophoricity in Ogonoid languages. In J. Mugane, J. Hutchison & D. Worman (Eds.), *Selected Proceedings of the 35th Annual Conference on African Linguistics* (pp. 234–244).

Capone, A. (2009). Are explicatures cancellable? *Intercultural Pragmatics, 6*, 55–83.

Chiou, M., & Huang, Y. (2010). NP-anaphora in modern Greek: a partial neo-Gricean pragmatic approach. *Journal of Pragmatics, 42*, 2036–2057.

Clements, G. N. (1975). The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages, 2*, 141–177.

Culy, C. (1994). Aspects of logophoric marking. *Linguistics, 32*, 1055–1094.

Curnow, T. (2002). Three types of verbal logophoricity in African languages. *Studies in African Linguistics, 31*, 1–25.

Dimmendaal, G. (2001). Logophoric marking and represented speech in African languages as evidential hedging strategies. *Australian Journal of Linguistics, 21*, 131–157.

Frajzyngier, Z. (1985). Logophoric systems in Chadic. *Journal of African Languages and Linguistics, 7*, 23–37.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Hagège, C. (1974). Les pronoms logophoriques. *Bulletin de la Société de Linguistique de Paris, 69*, 287–310.

Huang, Y. (1991). A neo-Gricean pragmatic theory of anaphora. *Journal of Linguistics, 27*, 301–335.

Huang, Y. (1994/2007). *The syntax and pragmatics of anaphora: A study with special reference to Chinese*. Cambridge: Cambridge University Press.

Huang, Y. (2000a). *Anaphora: A cross-linguistic study*. Oxford: Oxford University Press.

Huang, Y. (2000b). Discourse anaphora: Four theoretical models. *Journal of Pragmatics, 32*, 151–176.

Huang, Y. (2001). Marking of logophoricity in West African, East and South Asian languages. In *Proceedings of the 6th International Conference on the Languages of East, Southeast Asia and West Africa* (pp. 224–238).

Huang, Y. (2002). Logophoric marking in East Asian languages. In T. Güldemann & M. von Roncador (Eds.), *Reported discourse: A meeting ground for different linguistic domains* (pp. 213–224). Amsterdam: John Benjamins.

Huang, Y. (2004). Anaphora and the pragmatics-syntax interface. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 288–314). Oxford: Blackwell.

Huang, Y. (2006/2009). Anaphora, cataphora, exophora, logophoricity. In K. Brown (Ed.), *The encyclopaedia of language and linguistics* (2nd ed., vol. 1, pp. 231–238). New York: Elsevier Science. (Reprinted in *Concise encyclopaedia of semantics*, pp. 18–25, by A. Keith, Ed., New York: Elsevier Science).

Huang, Y. (2007). *Pragmatics*. Oxford: Oxford University Press.

Huang, Y. (2010a). Switch-reference in Amele and logophoric verbal suffix in Gokana: A generalized neo-Gricean pragmatic analysis. In D. F. Shu & K. Turner (Eds.), *Contrasting meaning in languages of the East and West* (pp. 75–101). Berlin: Peter Lang.

Huang, Y. (2010b). Neo-Gricean pragmatic theory of conversational implicature. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 607–631). Oxford: Oxford University Press.

Huang, Y. (2010c). The pragmatics of anaphora. In L. Cummings (Ed.), *The pragmatics encyclopaedia* (pp. 33–37). London: Routledge.

Huang, Y. (2010d). Impliciture. In L. Cummings (Ed.), *The pragmatics encyclopaedia* (pp. 238–240). London: Routledge.

Huang, Y. (2011). Types of inference: entailment, presupposition, and implicature. In W. Bublitz & N. Norrick (Eds.), *Foundations of pragmatics* (pp. 397–421). Berlin: Mouton de Gruyter.

Hyman, L., & Comrie, B. (1981). Logophoric reference in Gokana. *Journal of African Languages and Linguistics, 3*, 19–37.

Korta, K., & Perry, J. (2008). The pragmatic circle. *Syntheses, 165*, 347–357.

Levinson, S. (2000). *Presumptive meanings*. Cambridge, MA: The MIT Press.

Lust, B. C., Wali, K., Gair, J. W., & Subbarao, K. V. (Eds.). (2000). *Lexical anaphors and pronouns in selected South Asian languages*. Berlin: Mouton de Gruyter.

Parker, E. (1986). Mundani pronouns. In Wiesemann (Ed.), (pp. 131–166).

Recanati, F. (1993). *Direct reference*. Oxford: Blackwell.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Sengupta, G. (2000). Lexical anaphors and pronouns in Bangla. In Lust et al. (Eds.), pp. (277–332).

Speas, M. (2004). Evidentiality, logphoricity and the syntactic representation of pragmatic features. *Lingua, 114*, 255–276.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell.

Stirling, L. (1993). *Switch-reference and discourse representation*. Cambridge: Cambridge University Press.

von Roncador, M. (1992). Types of logophoric marking in African languages. *Journal of African Languages and Linguistics, 13*, 163–182.

Wiesemann, U. (1986a). Grammaticalized coreference. In Wiesemann (Ed.), (pp. 437–464).

Wiesemann, U. (Ed.). (1986b). *Pronominal systems*. Tübingen: Narr.

Willett, T. (1988). A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language, 12*, 51–97.

Some Notes on Points of View

Eros Corazza

Abstract I'll argue that the notion of viewpoint plays central stage in our understanding and interpretation of many utterances. I'll claim that such a notion is best characterized on the background of indexical reference; yet it cannot be reduced to it. I'll thus show how points of view can be unarticulated (roughly, unmentioned) and yet play an important role in our linguistic practice inasmuch as the understanding of some utterances rests on the grasping of the point of view associated with them. Finally, I'll mention how the notion of viewpoint (as an unarticulated linguistic phenomenon) plays an essential role in the understanding and interpretation of utterances containing anaphoric reflexive pronouns.

E. Corazza (✉)
ILCLI, UPV-EHU, Carlos Santamaria Zentroa, Elhuyar Plaza 2,
20018 Donostia San-Sebastian, Spain
e-mail: eros_corazza@yahoo.com

E. Corazza
IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

E. Corazza
Carleton University, Philosophy, 3A41 Paterson Hall, Colonel By Dr. 1125,
Ottawa, ON K1S 5B6, Canada

1 The Importance of Points of View

In our everyday life, down to some of the most basic activities in which we engage, points of view play a crucial role. To understand what someone says or does we often take on board her viewpoint. We can hardly interpret someone's action, let alone her intentions, without considering her point of view. We often face sentences like:

(1) Hugo Chavez's and Alvaro Uribe's viewpoints on Washington's foreign policies differ.
(2) Chomsky's point of view on the mind/body divide differs from that of Descartes.
(3) If you take Anya's point of view, you can easily understand why she left Bob.
(4) On this particular issue Anya and John have similar viewpoints.
(5) G.W. Bush and Tony Blair failed to appreciate the viewpoint of the Arab world.
(6) On this subject Anya has no particular point of view.
(7) If you understand my viewpoint you will not criticize me.

In these sentences points of view can, roughly, be defined as the general perspective one has on something. The latter seems to involve, among other things, the set of beliefs, dispositions, etc. one has on a given issue/problem/object/event/… This, though, is not exactly what I have in mind when I talk about points of view. What I am interested in is a more modest and narrow view on points of view. In what follows I shall confine myself to the notion of point of view as it is linked to one's agency, in particular to one's perceptual or sensory apparatus. That is, the notion of a point of view I am interested in is the one that would be expressed by sentences like:

(8) From my viewpoint I cannot see the shop near that building.
(9) Ian and Anya heard the same noise because they shared a viewpoint.
(10) Anya's point of view is the best; she can admire the entire scene.

In (8)–(10) the notion of point of view comes close to the notion of the location, place, position, perspective, etc. from which one apprehends and can act on one's surroundings. And one usually apprehends and acts on one's surroundings from a given place, at a given time and with one's own sensory apparatus. A way to characterize the notion of point of view I have in mind is to employ the analogy of the camera. A point of view is analogous to the point from which the camera films the action in a movie. As the camera can move around and register an event from different points of view, an agent can move around and perceive an event from different viewpoints.

Along this line a point of view can be cashed out using what came to be known (after Perry 1979) as *essential indexicals*: 'I', 'here', and 'now'. The latter are characterized in epistemological terms. Take 'I', for instance. It has a *cognitive impact* insofar as it triggers self-centered behaviors. The same thing holds for the indexicals 'now' and 'here'; they trigger self-centered behaviors as well. Essential

indexicals cannot be explained away or replaced by co-referring terms without destroying the cognitive impact their use conveys (see Castañeda 1966, 1967, 1968 and 1979). Privatus may know that Privatus is a war hero without knowing (being amnesiac, for instance) that he himself is a war hero and, thus, without behaving appropriately. As a first approximation we can summarize the notion of point of view as follows:

- *Point of View*

A point of view is the perspective from which one interacts with one's own surroundings and from which one can perform a given action. As such one can characterize one's own viewpoint using the indexicals 'I', 'here' and 'now'.[1]

Actually, if one were asked to express one's own viewpoint one would end up articulating it using 'I', 'now', and 'here'. As such, the notion of point of view I have in mind is intrinsically linked to the notion of perspective, insofar as a point of view encapsulates the egocentric perspective one has on some contextual salient aspects of one's surrounding, be it an event, an object, and individual or what you have. For this reason, the notion of point of view is best explained against the background of indexical reference. This doesn't mean, though, that viewpoints reduce to indexical reference. I shall argue that points of view can work as the non-conceptual or subdoxastic perspective one entertains on one's surrounding. The occurrence of an indexical, on the other hand, explicitly expresses a conceptual perspective one entertains on a give item of discourse and/or thought. And this, we shall see is the main difference between indexical expressions and viewpoints. The notion of point of view I have in mind should capture the way one cognizes one's surrounding without being constrained by conceptual representations.

From Kaplan (1977, 1989) we learned that indexical expressions have a linguistic meaning (character) which can be represented as a function taking as argument the context and giving as value the content (or referent). In short, a semantic account of indexicality must take on board the following notions: (1) the indexical linguistic meaning (character) and (2) the context on which the character operates whose parameters are: the *agent(s)*, *time*, *location*, *demonstratum (demonstrata)*, and *possible world*. As we'll now see, viewpoints don't have a linguistic meaning (character) operating on some aspect of context to deliver a content.

---

[1] The notion of action I have in mind is an intuitive one. It need not involve the notion of responsibility, for instance. The agent performing an action is the one engaging in behavioral movements, she is not the one who may influence, force, induce, etc. … someone else to do something. I also ignore cases where one can be said to perform an action at a distance. E.g.: when one leaves a will giving instructions to perform some actions after one's death, or when one leaves some instructions on an answering machine.

2 Beyond Indexicals: Points of View

When two people engage in a communicative interaction they do so from different perspectives. Anya sees the world with *her* eyes and talks about it with *her* mouth, while Ian sees it with *his* eyes and talks about it with *his* mouth. A difference in points of view is particularly salient when two people perceive something from a different perspective or angle. While one sees one side of a figure someone may see the other side and thus have a different view of the same object.

To illustrate the importance of points of view in the understanding of some utterances, consider a face-to-face communication like:

(11)   Anya to Ian: "The book is to the left of the pen"

In that exchange, for Ian to understand Anya's utterance he has to grasp Anya's viewpoint, i.e. he must understand that the book is on the left of the pen *relative* to Anya's position. If the relevant book is placed between Anya and Ian it is on the left *vis-à-vis* Anya and on the right *vis-à-vis* Ian. If Anya and Ian's communicative interaction were by telephone, for instance, the relevant point of view could be different. Suppose that Anya forgot her book at home and calls Ian to ask him to find the book and bring it to her. In her phone conversation Anya guides Ian to the location of the book (e.g., in her room on the desk under the window, etc.) and then utters (11). In such a situation the relevant viewpoint is Ian's, not Anya's. It is Ian's viewpoint insofar as Anya jumps so to speck, into Ian's shoes, i.e. she assumes Ian's perspective.[2]

Further examples involving the notion of a point of view can be furnished by so-called *contextuals* like: 'local', 'foreigner', 'enemy', 'national', etc. (cf. Vallée 2003). Consider:

(12) Anya: "Ian is a foreigner"

Anya is likely to be claiming that Ian is a foreigner regarding *her own* nationality or, if Anya is a foreigner *vis-à-vis* the country in which she produces her utterance Anya is likely to be claiming that Ian is a foreigner regarding the *location* of the utterance. For, one is a foreigner *vis-à-vis* some people and/or some places. Were Anya to communicate that Ian is a foreigner *vis-à-vis* someone else or some other location but not a foreigner *vis-à-vis* herself or the location in which she produces her utterance, she is likely to make that explicit, either by overtly articulating it or by relying on some information surrounding the discourse situation. If, for instance, Anya and her audience are discussing Ian's planned travels to Afghanistan, Anya can express the worry that Ian is in danger. Her friend asks why and Anya replies with (12). In that case the relevant nationality *vis-à-vis* which Ian is a foreigner is neither Anya's nationality nor the place of the utterance,

---

[2] The capacity we have to assume others perspective—what Vendler (1984) characterizes as transference—underlies most of our linguistic exchanges and joint activities.

but Afghanistan.[3] In uttering (12) Anya doesn't explicitly refer to a specific viewpoint. Furthermore, Anya need not have a conceptual representation of the relevant point view *vis-à-vis* which Ian is a foreigner.

As I have already hinted, the notion of a point of view I am interested in is the one tied to the notion of the agent's perspective. Every utterance is an event produced by someone. As such an utterance is intrinsically linked to a point of view. Hence, like the performance of an action, a speech act is fundamentally linked to *who*, *where*, and *when* it is produced. It's chiefly for this reason that if one were to express one's viewpoint one would end up using the essential indexicals 'I', 'now', and 'here'. Yet when one acts (e.g. kicks a ball), as when one produces an utterance, one doesn't usually mention, let alone think about, one's own egocentric perspective. There is no action lacking an agent, a location and a time. Furthermore, one cannot influence an event like picking up a glass which happened yesterday (time traveling is out of our power) or in some other location. If one is in New York one cannot pick up a glass in, say, Paris, let alone drink the burgundy it contains.[4]

Points of view may not be relevant in the production and understanding of an utterance. If one utters "$2 + 2 = 4$" or (as Galileo once famously said) "The Earth moves",[5] for instance, the point of view accompanying these utterances does not play a particular role in their processing and understanding. That is, whether these utterances are produced by Claire, Ian or Anya, at a time t, t^1, or t^2 in location l, l^1, or l^2 they would express the same content (say the same thing) and convey the very same information.[6] To grasp the latter one need not grasp the producer's points of view, i.e. by whom, where and when it has been produced. In short, in saying that a point of view is intrinsically tied to an utterance I am not suggesting that the understanding of the latter necessarily rests on the conceptualization of the accompanying point of view. On the one hand, there are viewpoint-free utterances and, on the other hand, there are viewpoint dependent utterances that one can

---

[3] This characterization should be neutral on whether or not we consider contextuals like 'foreigner' to work on the model of indexicality because they present an argument place (a hidden indexical) in their underlying grammatical structure—this would be the view favoured by so called *Indexicalist*s (e.g. Stanley 2000). If one were to embrace Indexicalism one could claim that the relevant contextual parameters fixing the nationality *vis-à-vis* which Ian is judged to be a foreigner are furnished by a point of view contextually selected by a hidden indexical or argument place. Notice, though, that one could be a Indexicalist concerning contextuals terms without endorsing Indexicalism for all the utterances that happen to be viewpoint dependent.

[4] I invite you to restrain your science fiction imagination. It may help philosophers to state some thesis, but it does not help us to understand what is going on in everyday life and in situations like the ones I am describing and interested in.

[5] What he actually said is: "Eppur si muove".

[6] For the sake of simplicity I'm ignoring the difference between what is literally (semantically) expressed and what is communicated. One could indeed defend the view that what is conveyed or communicated transcends what is literally expressed or said. This would be for instance the view presented by minimalists inspired by Cappelen and Lepore (2005). These distinctions, as interesting as they may be, shouldn't affect the main argument of this chapter.

successfully process without conceptualizing the relevant viewpoint. If I am right, points of view are usually neither something one explicitly refers to nor something one conceptualize. Points of view aren't something ending up in the proposition expressed. When one produces an indexical utterance like "Today I meet the Dean", the day referred to by the occurrence of 'today' ends up in the proposition expressed. When one utters a viewpoint dependent utterance like "The salt is on the left of the pepper" the relevant point of view fixing the perspective *vis-à-vis* which the salt can be judged to be on the left doesn't enter the proposition expressed. The speaker didn't say "The salt is to the left of the pepper from my viewpoint". And in uttering this sentence the speaker need not represent her own perspective when thinking that the salt is to the left of the pepper. This is, I reckon, one of the main difference between indexical reference and viewpoints. In other words, when one entertains a thought one would express by uttering "Now I must go to meet Jane" one comes to entertain a representation of the relevant time. One thinks of it as *now*. The same with thoughts expressed or grasped by utterances like "Here is cold" or "Today I must go to my office" one entertains indexical thoughts representing the relevant day as *today*, the relevant location as *here* and oneself as *I* and *me*. In claiming that points of view need not be conceptualized I mean that a speaker and her audience need not represent the relevant point of view. The structure of the situation in which their linguistic interchange occurs may raise to salience the relevant viewpoint without the speaker and her audience having to represent it. This does not mean, though, that points of view never get conceptualized. It simply means that in many cases a point of view need not be the constituent of a thought. In short, when one produces an utterance one need not represent, the point of view from which the utterance is made. To borrow Perry's (1986) terminology we could say that a point of view can be an *unarticulated constituent* of both of the utterance and the accompanying thought. To highlight this phenomenon let us consider utterances of sentences like:

(13) It's raining

      and

(14) It's 3:00 p.m.

These sentences are context-sensitive—if uttered in London (13) may be true, while if uttered in New York it may be false. Yet there is no indexical expression appearing in them designating the relevant location and time zone. No specific element in the utterance operates on context to designate a particular item.[7] As Perry argues, the relevant location and the relevant time zone in (13) and (14) are *unarticulated constituents* of the propositions expressed (see Perry 1986, 2001). In an utterance of "It is raining" the relevant location, *qua* unarticulated

---

[7] Unless one defends the view that there's a hidden indexical (or argument) in the logical form of the sentence singling out a determinate location and/or time zone. This would be the position advocated by so-called Indexicalists (e.g. Stanley 2000). If this were the case "It's raining" and "It's raining here" would differ only at the superficial, grammatical, level.

constituent, is picked out, following Perry's suggestion, by the utterance as a whole.[8] What does it exactly mean, though, to claim that an utterance *as a whole* can single out a given location? The notion of a point of view helps us here. It is because our relevant utterances are made from, or concern a specific point of view that their full truth conditions also depend on a specific location/time/etc. In other words, it is the implicit point of view accompanying an utterance which contributes in making an utterance like (13) to concern a specific location and an utterance of (14) to concern a specific time zone. (13) and (14) can thus be considered as typical examples of viewpoint-dependent utterances.

The question many philosophers and linguists raised is whether the speaker (and the audience) of utterances like (13) and (14) ought to represent the relevant place and time zone of the utterance. While many would not contest that the truth value of these utterances also depends on the relevant time zone and location, opinions diverge on whether or not the utterances and the corresponding thoughts ought to represent the time zone and the location.

Friends of so-called "Indexicalism" would argue that the relevant location and time zone are represented both in the utterance and in the corresponding though. And they are so represented because at the level of the logical form there is a hidden indexical (or implicit argument) selecting the relevant time zone and location. Given that the notion of logical form corresponds to the level of syntactic representation which represents the properties relevant for semantic interpretation, the relevant location and time zone are somewhat represented in the thoughts associate with (13) and (14).

One of the chief arguments put forward by indexicalists (see for instance Stanley) is the so-called binding argument. It runs as follows: a sentence like (13), for instance, can be encapsulated into a quantificational sentence like:

(15)   Every time I lit a cigarette it is raining

meaning, roughly, that it is raining where I happen to be when litting a cigarette. While in (13) the relevant place is provided by the location where the utterance occurs, in (15) it depends upon (and varies with) the domain of the quantifier. We are told that the natural way to understand utterances such as (13) and (15) is to posit a hidden argument place for a location, so that the implicit argument place for the verb 'to rain' in (13) works like a free variable, while in (15) it works as a variable bound by the quantifier. As far as I know, the first person to suggest that we have to postulate an argument place for the alleged unarticulated constituent when binding is possible is Partee (1989). For more on the argument from binding and the way it suggests the presence of tacit arguments at the level of LF see Stanley (2000).

---

[8] However, the relevant location and time zone need not, *pace* Perry, enter the proposition expressed. One could argue that the utterance concerns them insofar as the (minimal) proposition expressed is situated. The time zone and location belong to the situation or circumstance of evaluation. This debate, as interesting as it may be, transcends the scope of the present chapter.

The same argument can be run for viewpoint-dependent utterances. Consider:

(16)   Every client thought that the salt was to the left of the pepper

which could mean that the salt is to the left of the pepper regarding different orientations depending on the different values of the quantifier. If one embrace the binding argument one is likely to consider points of view as a kind of indexical reference.

Recanati (2002) and Cappelen and Lepore (2004) propose what I consider a convincing argument against the binding argument. Recanati's main argument is that it forces unwelcome consequences. In particular, it forces us to postulate the presence of argument places where, intuitively, there is none. Recanati invites us to consider an intransitive verbs like 'to eat' which denotes the property of eating. In that case, he argues, the contextually provided constituent results from free enrichment and not from the semantics of the verb, for in its intransitive reading 'eat' is not a two-places predicate. But in a sentence like "Jon ate" binding can occur:

(17)   Jon is anorexic, but whenever his father cooks mushrooms, he eats.

The intuitive way to understand it is that Jon eats *them*, i.e., *the mushrooms his father has cooked*. Examples like this seem to prove that intuitive binding, *per se*, does not entail the presence in the logical form of an argument place and, therefore, that the argument from binding is not compelling. Cappelen and Lepore (2004) propose the following reduction of the argument from binding:

(18)   Everywhere I go, $2 + 2 = 4$

Here is the Binding Argument applied to (18). Intuitively, (18) says that, for every place I go, $2 + 2 = 4$ at that place. So we should present the logical form of (17) along the following lines:

(19)   For all places x, if I go to x, then $2 + 2 = 4$ at x.

The quantifier phrase 'Everywhere I go' is binding a place variable in the logical form of "$2 + 2 = 4$"—otherwise, there would be nothing for the quantifier phrase to bind. This establishes that the logical form of the sentence '$2 + 2 = 4$' has a freely occurring place variable.

Like Recanati and Cappelen & Lepore I don't think that the binding argument forces us to posit hidden indexicals or argument places in logic position when dealing with underdetermined utterances like (13) and (14). The same story can be told about viewpoint-dependent utterances.

Concerning the general phenomena involving points of view, I am sympathetic to Perry's view that although the full truth conditions of the utterances and the accompanying thoughts involve the relevant time zone and location, the agent need not represent them. It is a matter of an external, contextual, relation that makes the thought concerning a relevant location and time zone. As an analogy, think of the mental representation one, John, has when perceiving a given individual, say Anya. The fact that John's perceptual representation is about Anya and

not her identical twin need not be represented by John. It is because Anya and not her twin is in John's perceptual field that John's representation is about Anya and not her twin. If Anya's twin, instead of Anya, were in John's perceptual field John's representation would be about Anya's twin and not Anya. If John expresses his thought by uttering "That woman looks anxious" his thought would be about Anya if Anya happens to be in John's perceptual field and it would be true iff Anya looks anxious at the time John perceives her. Were Anya's twin in John's perceptual field the very same representation would be true iff Anya's twin looks anxious. If one were to build into the relevant representation the time and perceptual relation, one would commit oneself with the view that in both scenarios John would entertain different thoughts insofar as the relevant representations concern different individuals.[9] In short, the picture I have in mind goes as follows. The very same mental representation can be about different objects/events/… It is the situation in which the representation occurs that links that representation to an object/events/… and this contextual link need not be represented by the agent entertaining that representation. As far as I understand, this comes close to Perry's idea that the relevant location in (13) and time zone in (14) is picked up by the utterance as a whole and that they are unarticulated constituents entering the full truth conditions. Furthermore, the relevant location and time zone need not be represented, i.e. we can have, to borrow Perry's terminology, thoughts without representation for the time zone and location.[10]

   To further highlight this point thinks of a child who is unaware of time zones. Our child can utter "It's 3:00 p.m." and, in so doing, succeed in passing along some relevant information and engage in a successful linguistic interchange. Our child's speech act can be successful even if the speaker (and the hearer) do not represent the relevant time zone, e.g. without them having to think that it is, say, 3:00 p.m. *Pacific Time*. Little-John and little-Jane can decide to meet at the playground at 3:00 p.m. without them having to represent the relevant time zone. They lack the cognitive resources to represent time zones. In cases like this, all the relevant parameters granting the success of the linguistic interaction are *fully* provided by the structure of the situation in which the exchange occurs.[11]

---

[9] This would be the view advocated among others by some neo-Fregeans (e.g. Evans 1982 and McDowell 1984) who defend the existence of *de re* senses, i.e. modes of presentations involving the objet itself. A sense, as a thought constituent, would thus vary with a change of the object involved.

[10] Corazza (2007) and Corazza and Dokic (2007, 2010) claim that alleged unarticulated constituents need not end up, *pace* Perry, in the proposition expressed but can remain in the situation *vis-à-vis* which the proposition is assessed to be true or false.

[11] Following Barwise and Perry (1983) seminal work a situation can be characterized as a partial possible world: "Reality consists of situations—individuals having properties and standing in relation at various spatiotemporal locations. We are always in situations; we see them, cause them to come about, and have attitudes toward them" (Barwise and Perry 1983: 7). A real situation comports infinitely many aspects. Yet we can cognize only parts of it; what we cognize depends on many factors such as our interests, activities, practices, etc. The very same individuals and properties may appear in different situations. As such they are uniformities. Locations and time

This seems to obey a general cognitive principle of economy, i.e. that people (like many organisms) tend to minimize internal representational resources. Since the situation fixes all that needs to be fixed, the speaker and her audience need not represent what their discourse concerns.[12] In a nutshell, what is fixed by a situation need not be fixed by the agent's representational system. As I just pointed out, this reflects a principle of cognitive economy.[13] If, for instance, a group of people living on a small island never travel and never have contact with the outside world (e.g. they do not observe reports coming from abroad and their telephone line does not extend outside their small island), they need not have representations for time zones. The situation in which they utter, say, "It's 3:00 p.m." provides all that is needed for them to get the time right and their actions are automatically attuned to it. Yet their time-utterances (and thoughts) concern a given time zone. These islanders, though, don't need any mental effort to distinguish various time-zones. They are simply unaware of the existence of time zones. Since we are often aware of time zones and since we sometimes communicate with people in other parts of the world and thus in different time zones we need some cognitive capacity to keep track of various time zones. If Ian from San Francisco calls Anya in New York and tells her "I'll call you back tomorrow at 3:00 p.m." Anya ought *to know* whether Ian will call at 3:00 p.m. Pacific Time or 3:00 p.m. Eastern Time. There are different ways, though, in which Anya can know about the relevant time that Ian will call. It may be the case that when calling Anya, Ian always refers to Anya's time zone, i.e. the Eastern Time zone. In that case, based on past practice, Anya need not think about the relevant time zone. She takes it for granted that Ian will call at 3:00 p.m. Eastern Time. This kind of knowledge can be stored in Anya's (and Ian's) long-term memory. As such, it need not be articulated in Anya's and Ian's cognitive system. That is to say, it may be stored in their memory without them having to activate it. It can thus be considered as background knowledge and belief stored in the situation in which their time-utterances and thoughts occur. It's for this very reason that it need not enter Anya's and Ian's thoughts during their telephone exchange. Just as some relevant information can be stored in the external world (e.g.: we store telephone numbers, birthdays, meeting schedules, etc. in our notebooks), some information can be stored in one's memory. Stored memory information isn't part of one's working memory. Because of that it can be

---

(Footnote 11 continued)

are uniformities as well insofar as different things can happen in the same location at different times and various things can be going on at the same time in different locations.

[12] See Perry (1986)'s Z-land story where the inhabitants of a little island, Z-land, never travel and don't have telephone communications and broadcasting information coming from the external world. When a Z-lander utters "It's raining" her talk concerns Z-land (it rains in Z-land). Yet she doesn't (and need not) represent the location where it is raining.

[13] Cf. Clark 007 principle: "[E]volved creatures will neither store nor process information in costly ways when they can use the structure of the environment and their operations upon it as a convenient stand-in for the information-processing operations concerned" (Clark 1989: 64).

classified as tacit and/or dispositional knowledge, a sort of knowing-how or procedural knowledge.[14]

As a way of an analogy we can mention the implicit knowledge one has of the grammatical rules one follows when computing sentences. The knowledge at work in these cases can be characterized as non-representational, procedural, knowledge insofar as one is not aware of these rules. This kind of knowledge can only be deployed in one's understanding of the language. The conceptual apparatus that the linguists use in describing and characterizing syntactic rules does not belong to the conceptual stock of an ordinary competent speaker.[15]

3 Points of View and Understanding

From a third person perspective, i.e. from the audience's viewpoint, in order to understand utterances like (13) [It's raining] or (14) [it's 3:00 p.m.] one needs to grasp the intrinsic point of point of view the utterances concerns. It is only when one "knows" the relevant location (13) concerns, for instance, that one understands it.[16]

I claimed that if one were to articulate one's own viewpoint one would likely use an essential indexical ('I', 'here', 'now'). In some cases the point of view is explicitly articulated in the utterance itself. This happens when one utters one of these indexicals (or other indexical expressions).

---

[14] No doubt, more should be said about the way information can be stored in memory and how it works in our thinking and linguistic interchanges without being actualized and/or articulated. For a detailed review article on this issue and on how shared memories work, see Sutton (2009). Following Clark and Chalmers (1998) there is no principled difference between information and beliefs stored in memory and information and beliefs stored in one's notebook. Someone may reliably believe that the meeting start at 1:00 p.m. because they wrote it down in their notebook (see Clark and Charmers' case of Otto who, suffering from Alzheimer's, cannot store in his biological memory relevant information and, as a consequence, reliably stores it and successfully retrieves it from his notebook).

[15] This comes close to Cussins (1990) when he argues that an account of experiential content is best understood in terms of an organism's abilities to act upon the perceived environment, rather than in terms of truth and truth conditions. Along this line we can argue that what an individual perceives when uttering viewpoint dependent sentences is a structured environment or situation in terms of the possibilities it affords for action. Cussins' conception of an ability-based notion of content provides a clear distinction between a level of what Dummett's (1986) characterizes as proto-thoughts that can successful trigger some actions and a level of truth apt full-fledged thoughts.

[16] As I previously mentioned, though, the knowledge at issue here need not be *explicit* knowledge. One may be said to tacitly know or grasp a viewpoint inasmuch as one's action is consonant with the relevant viewpoint. If Anya, looking out of the window, says "It's raining" John's grasping of the relevant viewpoint (location) needn't rest on John coming to entertain a thought he would express by "It's raining *here*" or "It's raining *in London*". For John to grasp the relevant viewpoint, it suffices that his action is attuned to it: e.g. he picks up an umbrella before going out, he renounces going out watering the garden, etc.

It's also interesting to note that sometimes grammar forces us to explicitly articulate someone else's viewpoint. Consider, for instance, Roger Federer's utterance:

(20) I hope to win Wimbledon

One can faithfully report what Federer said with:

(21) Roger Federer hopes to win Wimbledon

(21) captures Federer's viewpoint, i.e. the viewpoint he expressed using 'I'. (21) can but have a *de se* reading (see Chierchia 1989). As such, its underlying form can be represented as:

(22) Roger Federer$_1$ hopes [PRO_1 to win Wimbledon]

where the unpronounced subject of the report (PRO[17]) attributes to the agent of the attitude, Roger Federer, an 'I'-thought and, as such, captures Federer's viewpoint, i.e. the viewpoint he explicitly expressed using the first person pronoun in (20). Here we have syntactic evidence favoring the view that in some of our linguistic activities we explicitly convey someone's viewpoint. Another way to capture Federer's viewpoint as he manifests it in (20) would be in using what Castañeda (1966, 1967, 1968) characterizes as a quasi-indicator. The paradigmatic examples of quasi-indicators are the anaphoric 's/he (her/himself)', i.e. an anaphoric pronoun attributing to the referent of its antecedent an 'I'-thought. We could thus have:

(23) *Roger Federer$_1$* said that *he (himself)$_1$* hopes to win Wimbledon

In (18) the narrator also expresses Federer's viewpoint. That is, on top of referring to Federer the narrator also attributes to Federer a specific viewpoint. The anaphoric pronoun 'he (himself)' in (23) must be understood as a pronoun which allows the reporter to capture someone else's, in our example Roger Federer's, viewpoint.

In favor of this interpretation, viz. that we often represent someone else's viewpoint, we can also mention some cross-linguistic data. In some natural languages (so-called logophoric languages) logophoric pronouns are used to attribute a point of view explicitly. This is, for example, the case of '*sɛ́*' in Tabury (see Hagège 1974):

(24) a. á Dík lí māy mà:gā *à* kó n sú: mònò
      (He$_1$ thinks of the young girl that he$_1$ saw yesterday)

    b. á Dík lí māy mà:gā *sɛ́* kó n sú: mònò
      (He$_1$ thinks of the young girl that he (himself)$_1$ saw yesterday)

---

[17] PRO represents the null pronominal element acting as the syntactic subject of infinitives and gerunds. In other words, PRO is the null analogue of lexical pronouns.

In (24b) '*sɛ́*' explicitly attributes an 'I'-thought, thus a viewpoint, to the referent of the antecedent it is coindexed with. The same with 'yè' (singular) 'yèwo' (plural) in Ewe (cf. Clements 1975):

(25) a. Kofi be yé-dzo
       [Kofi say LOG-leave]
       (Kofi said that he (himself) left)
     b. Kofi be me-dzo
       [Kofi say I-leave]
       (Kofi said that I leave)
     c. Kofi be e-dzo
       [Kofi say s/he-leave]
       (Kofi$_1$ said that she/he$_2$ leave)

In Ewe and Tabury we thus have pronouns whose specificity is to capture someone else's viewpoint. In using these pronouns the narrator explicitly attributes a point of view, in our examples, an 'I'-thought. Castañeda (1966, 1967, 1968) created an artificial pronoun, 'she*/he*/it*', to represent in an attitude ascription the use (maybe only implicitly) of the first-person pronoun. "Sue says that she* is rich" represents Sue as saying "I am rich". These artificial pronouns are called 'quasi-indicators' and, Castañeda claims, are the only mechanism enabling the attribution of indexical reference from the third-person perspective. They are, therefore, the only tools which allow us to capture the cognitive impact conveyed by the essential indexicals—'she*' captures the cognitive impact conveyed by 'I', 'then*' the cognitive impact conveyed by 'now' and 'there*' the one conveyed by 'here'. It is an accident of English that a single pronoun 'she/he/it' can be used to perform very different speech acts.[18]

Quasi-indicators, *qua* logophoric pronouns, help one to capture someone else's viewpoints. Yet in specifying someone's viewpoint we need not ascribe the attribute a specific mental representation of her viewpoint. What we are ascribing may be best understood as a capacity to act in the appropriate way in the circumstance our attribute happens to be. For this reason we can ascribe viewpoint to non-linguistic infants and (some) other non-linguistic animals.

Further linguistic evidence highlighting the importance of points of view in our interpretation of utterances is furnished by so-called picture noun phrases (see Pollard and Sag 1992). In such cases the notion of point of view takes center stage when we attempt to explain how people process and understand utterances containing picture-noun phrases. Consider:

---

[18] "It is a mere accident of grammar that the same physical objects are used in different logical roles. The underlying rationale is this: Indicators are a primary means of referring to particulars, but the references made with them are personal and ephemeral; quasi-indicators are the derivative means of making an indexical reference both interpersonal and enduring, yet preserving it intact" (Castañeda 1967: 207).

(26) *John₁* was going to get even with Anya. That picture of *himself₁* in the paper
      would really annoy her, as would the other stunts he had planned.

In (21) the reflexive 'himself' appears in another clause. As such it cannot be
linked to its antecedent via some syntactic principles or rule. In particular, it is not
c-commanded by its antecedent.[19] In cases like these, that Pollard and Sag call
exempt-anaphors because they are not constrained by the grammatical rules
controlling ordinary reflexives, the reflexives' resolution cannot be determined
by syntax alone. Since intersentential anaphora does not obey principle A of
Government and Binding Theory whichever way one spells it out, other consid-
erations governing their use and interpretation must be considered.[20] The notion of
a point of view comes to our rescue here. The reflexive pronoun is coindexed with
an antecedent whose point of view is being reported. In (26) the narrator is
expressing John's viewpoint and the reflexive 'himself' is coindexed (and thus
coreferential) with 'John'. To highlight the importance of the notion of points of
view in our interpretation of anaphoric relation further, let us consider the
ungrammaticality of:

(27) a. *Anya was quite taken aback by the publicity *John₁* was receiving. That
          picture of *himself₁* in the paper had really annoyed her, and there was not
          much she could do about it.[21]

---

[19] A pronoun is *bound* iff it is c-commanded by a coindexed element, while a pronoun is *free* iff
it is not c-commanded by a coindexed element. The notion of c-command is defined as:
    •*C-command*
    Node A c-commands node B iff:
1. A does not dominate B and B does not dominate A; and
2. the first branching node dominating A also dominates B.
    The notion of dominance characterizes the vertical relation in a tree and can be defined as:
    •*Dominance*
    Node A dominates node B iff A is higher in the tree than B and if you can trace a line from A
    to B going only downwards.

[20] A way to state Principle A is as follows (see Pollard and Sag 1992: 263):
    Every anaphor must be coindexed with a NP in an appropriately defined command relation,
within an appropriately defined minimal syntactic domain.
    The main questions (and disagreements) focus on how the command relation and the minimal
syntactic domain should be specified. This debate, however, transcends the scope of my paper. It
is also worth stressing that the (traditional) notion of anaphor I am relying on here is not the same
as the syntactic GB notion, for Principle A never covers intersentential coreference.

[21] Tom Baldwin suggested to me that a picture noun phrase like 'picture of him/her-self' should
be read as 'self portrait'. Thus, if we replace 'that picture of himself' with 'that self-portrait',
(22a) is grammatical. I do not know whether this constitutes the default reading of a picture noun
phrase. The important point here is to compare sentences like (22a) and (22b) and to understand
why one is grammatical while the other is not. Furthermore, if in a sentence like (22a) a 'picture
of himself' means 'self-portrait', the sentence would be ambiguous on whether the relevant
picture represents Anya or John, i.e. whether it is a self-portrait of Anya or of John. Besides, and
more importantly, 'that picture of himself' cannot be automatically replaced by 'self-portrait'.
For it is not contradictory to say "I'm having a picture of myself taken by John", while it would
be contradictory to say "I'm having a self-portrait taken by John".

(27a) is ungrammatical because the narrator expresses Anya's viewpoint. For this reason the reflexive 'himself' cannot be coindexed with 'John'. The ungrammaticality is generated by a conflict of viewpoint. That is, while the narrator (with the first clause) expresses Anya's viewpoint, the reflexive 'himself' expresses John's viewpoint. If, on the other hand, the reflexive were coindexed with 'Anya' we would have the grammatical:

(27)   b. *Anya$_1$* was quite taken aback by the publicity John was receiving. That picture of *herself$_1$* in the paper had really annoyed her, and there was not much she could do about it.

Since the narrator expresses Anya's viewpoint the reflexive can be linked to 'Anya'. In that case we do not have a conflict in viewpoint. In other words, in an example like this the reflexive can only be linked to the antecedent standing for the agent whose point of view is being represented. Hence, sentences like:

(28) Ian was quite taken aback by the publicity John was receiving. That picture of himself in the paper had really annoyed him, and there was not much he could do about it.

must be represented as:

(28)   a. *Ian$_1$* was quite taken aback by the publicity John was receiving. That picture of *himself$_1$* in the paper had really annoyed him, and there was not much he could do about it.

If 'himself' is coindexed with 'John' we generate ungrammaticality:

(28)   b. * Ian was quite taken aback by the publicity *John$_1$* was receiving. That picture of *himself$_1$* in the paper had really annoyed him, and there was not much he could do about it.

Once again the ungrammaticality is triggered by the conflicting viewpoint expressed, i.e. a sentence like this cannot express both Ian's and John's point of view.[22]

Furthermore, psychological verbs such as 'bother' make evident how the notion of viewpoint is crucial in determining the antecedent of an anaphora. With 'bother' it is natural to assume that the agent whose viewpoint is being reported is the direct object of the verb:

(29)   a. The picture of *himself$_1$* in the paper bothered *Ian$_1$*

       b. *The picture of *himself$_1$* in the paper bothered *Ian$_1$'s* mother

---

[22] The notion of point of view also helps us to understand the ungrammaticality of a sentence like: * Speaking of Roger Federer, I expect himself to win Wimbledon. The ungrammaticality can easily be explained by the fact that the narrator uses 'I' and, because of this very fact, she represents her own viewpoint and, thus, she cannot represent Roger Federer's viewpoint as the reflexive 'himself' suggests she should do. The ungrammaticality is thus explained by a conflict in viewpoints.

The ungrammaticality of (29b) is explained by the fact that the viewpoint represented is that of Ian's mother, rather than Ian's; thus 'himself' cannot be coindexed with 'Ian'. We thus have a conflict of viewpoints. While the NP 'the picture of himself' brings to the fore Ian's viewpoint, the VP 'bothered Ian's mother' suggests that the viewpoint represented is Ian's mother's. The ungrammaticality is thus generated by the conflict between Ian's and his mother's viewpoints. Consider now:

(30) a. The picture of $himself_1$ in the paper dominated Ian_1's thoughts

    b. The picture of $himself_1$ in the paper made Ian_1's day

Although (30a–b) are structurally equivalent to the ungrammatical (29b), they are grammatical insofar as they bring to the fore only Ian's viewpoint. As such, unlike in (29b), there is no conflict of viewpoint.

The importance of points of view in our understanding and interpretation of sentences is further highlighted by sentences like:

(31) a. Ian_1 and his_1 father saw the game

    b. * Ian_1's father and he_1 saw the game[23]

(31a) is grammatical insofar as the only point of view represented is Ian's. (31b) is ungrammatical because two conflicting points of view are represented, i.e. Ian's and Ian's father's. The same with sentences like:

(32) a. I met $Anya_1$ and her_1 spouse

    b. ?? $Anya_1$ and her_1 spouse met me.

Since 'to meet' is mutual, i.e. *a* met/is meeting/will meet *b* iff *b* met/is meeting/will meet *a*, (31a) and (32b) are logically equivalent. Yet (32b), if not ungrammatical, is awkward because the presence of 'Anya' in subject position brings to relevance Anya's viewpoint, which ends up conflicting with the narrator viewpoint represented by 'me'. As we saw, essential indexicals explicitly articulate the narrator's viewpoint. The presence of the essential indexical 'I/me' in (31a)/(32b) explicitly represents the narrator's viewpoint. Since 'Anya' in (32a) is not in subject position Anya's viewpoint is not raised to salience. We thus do not have conflict in viewpoints, so (32a) is grammatical.

We have further cross-linguistic evidence about the importance of points of view in the understanding of sentences containing reflexives. In Icelandic (cf. Sells 1987), in multiple embedded sentences any of the subjects can be the antecedent of the reflexive. In:

(33) Jón segir að María viti að Haraldur vilji að Billi heimsæki sig
    John says that Anya knows that Harold wants that Billy visit self

---

[23] For more examples along these lines and further discussion about them, see Kuno 2004.

The antecedent of the reflexive 'sig' can be either 'Jón', 'María', or 'Haraldur'. Hence, (33) is multiple ambiguous. To dissolve the ambiguity the interpreter must decide whose point of view is being reflected. If the narrator brings to salience Jón's viewpoint, 'sig' is coindexed with 'Jón', while if the narrator reflects Anya's or Billi's viewpoint 'sig' is coreferential with 'Anya' or 'Billi', etc. Furthermore, in Icelandic (see Sells 1987) the antecedent of the reflexive can operate across clauses. In that case the point of view which helps determine the reflexive's antecedent is transmitted through the discourse:

(34) Formaδurinn₁ varδ δskaplega reiδur. Tillgan væri avívirδileg.
The-chairman₁ became furiously angry. The-proposal was outrageous.
Væri henni beint gegn sér₁ persónulega.
Was it aimed at self₁ personally.

4 Conclusion

If the story I have told comes close to being accurate, points of view must be considered among the main features when we come to the task of explaining how context-sensitivity can affect our linguistic interchanges. If I am right, context sensitivity expands behind indexicality. Yet unlike the latter, the context sensitivity conveyed by the notion of points of view need not be linguistically and mentally represented. It is an open question whether the context sensitivity conveyed by a viewpoint-dependent utterance affects the (literal) content of what one ends up expressing. Yet this kind of context sensitivity cannot be ignored when we come to explain how people manage to understand viewpoint dependent utterances. As we saw, our capacity for grasping someone else's viewpoint is crucial in our understanding of utterances of: "It is raining", "Anya is an enemy", "John is a foreigner", etc. I suggested that when a viewpoint is not explicitly mentioned it can be understood as an unarticulated constituent. As such it need not be represented either by an utterance or by the agent of the utterance. As I attempted to show, an agent's viewpoint is linked to an appropriate disposition to act. Although a disposition to act is causally grounded, the agent need not represent this grounding. It's a matter of nature, it's a given by nature, that agents act and behave from a given viewpoint. To borrow a famous terminology, we can say that it is because of a pre-established harmony that when we produce viewpoint-dependent utterances we need not represent the relevant viewpoint. Yet since we're able to assume others' perspective, i.e. to assume someone else's viewpoint, in our interpretation and understanding of utterances we sometimes come to represent the relevant viewpoint. This is particularly important when we process (anaphoric) reflexive expressions. Since in an utterance one can represent someone else's viewpoint, the anaphoric links of reflexive pronouns are often determined regarding the viewpoint represented.

References

Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge: MIT Press.

Cappelen, H., & Lepore, E. (2004). Unarticulated constituents and hidden indexicals: An abuse of context in semantics'. In M. O'Rourke & C. Washington (Eds.), *Situating semantics: Essays on the philosophy of John Perry*. Cambridge Mass: MIT Publications.

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics*. Oxford: Blackwell.

Castañeda, H.-N. (1966). 'He': A study in the logic of self-consciousness. *Ratio, 8*, 130–157.

Castañeda H.-N. (1967). Indicators and Quasi-indicators. *American philosophical quarterly 4*:85–100. Reprinted In H.-N. Castañeda (1989). *Thinking, language, and experience* (pp. 206–231). Minneapolis: University of Minnesota Press.

Castañeda, H.-N. (1968). On the Logic of attributions of self-knowledge to others. *The Journal of Philosophy, 65*, 439–456.

Chierchia, G. (1989). Anaphora and attitudes *de se*. In R. Bartsch, J. van Bentham, & P. van Emde (Eds.), *Language in context* (pp. 1–31). Dordrecht: Doris.

Clark, A. (1989). *Microcognition: Philosophy, cognitive science and parallel distributed processing*. Cambridge: MIT Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*, 10–23.

Clements, G. (1975). The logophoric pronoun in Ewe its role on discourse. *Journal of West African Languages,* 10: 141–177.

Corazza, E. (2007). Contextualism, minimalism, and situationalism. *Pragmatics & Cognition,* 15 (1): 115–137.

Corazza, E., & Dokic, J. (2007). Sense and insensitivity: Or where minimalism meets contextualism. In G. Preyer & G. Peter (Eds.), *Context-sensitivity and semantic minimalism: Essays in semantics and pragmatics* (pp. 169–193). Oxford: Oxford University Press.

Corazza, E., & Dokic, J. (2010). Situated minimalism versus free enrichment. *Synthese, 184*(2), 179–198.

Cussins, A. (1990). The connectionist construction of concepts. In M. Boden (Ed.), *The philosophy of artificial intelligence* (pp. 368–440). Oxford: Oxford University Press.

Dummett, M. (1986). *Origins of analytical philosophy*. Cambridge: Harvard University Press.

Evans, G. (1982). *Varieties of reference*. Oxford: Oxford University Press.

Hagège, G. N. (1974). Les Pronoms Logophoriques. *Bulletin de la Société de Linguistique de Paris, 69*, 287–310.

Kaplan, D. (1977). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press. (1989).

Kaplan, D. (1989). Afterthoughts. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 565–614). Oxford: Oxford University Press.

Kuno, S. (2004). Empathy and direct discourse perspective. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 315–343). Oxford: Blackwell.

McDowell, J. (1984). *De re* senses. In C. Wright (Ed.), *Tradition and influence* (pp. 98–109). Oxford: Blackwell.

Partee, B. (1989). Binding implicit variables in quantified contexts. In *Papers from the 25th regional meeting of the Chicago Linguistic Society* (pp. 342–365). Chicago: University of Chicago Press.

Perry, J. (1979). The problem of the essential indexical. *Noûs* (13), 3–21. Reprinted In J. Perry (2000). *The problem of the essential indexical and other essays* (pp. 27-44). Palo Alto: CSLI Publication.

Perry, J. (1986). Thoughts without representation. *Proceeding of the Aristotelian Society 60*, 137–52. Reprinted In J. Perry (2000). *The problem of the essential indexical and other essays* (pp. 171–188). Palo Alto: CSLI Publications.

Perry, J. (2001). *Reference and reflexivity*. Palo Alto: CSLI Publications.

Pollard, C., & Sag, I. (1992). Anaphors in english and the scope of binding theory. *Linguistic Inquiry, 23*, 261–303.

Recanati, F. (2002). Unarticulated constituents. *Linguistics and Philosophy, 25*, 299–345.
Sells, P. (1987). Aspect of logophoricity. *Linguistic Inquiry, 18*, 445–479.
Stanley, J. (2000). Context and logical form. *Linguistics and Philosophy, 23*, 391–434.
Sutton, J. (2009). Remembering. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 217–235). Cambridge: Cambridge University Press.
Vallée, R. (2003). Context-sensitivity beyond indexicality. *Dialogue, 42*, 79–106.
Vendler, Z. (1984). *The matter of mind*. Oxford: Clarendon Press.

Referring to 'What Counts as the Referent': A View from Linguistics

Keith Allan

Abstract As defined here, a speaker's act of referring is the speaker's use of a language expression in the course of talking about its denotatum. This pragmatic definition of reference is defended against more traditional usage that contrasts "referring", "denoting", "describing", "alluding", "attributing", etc. It is proposed that the various differences in meaning supposedly captured by the different applications of these terms are better dealt with in other ways that can make sharper distinctions. What the hearer recognizes as the speaker's referent necessarily only 'counts as the referent' because it is on many occasions not identical to what the speaker identifies, indeed the speaker and hearer might even have entirely contradictory conceptions of the referent and yet the language expression used by the speaker can be said to successfully refer. Consider some examples. In *President Clinton was a baby in 1946* the speaker refers to (on my definition) two temporally distinct manifestations of Bill Clinton. If Sue says to Ed *My husband's having an affair with his boss* it is perfectly possible for Ed (and us) to understand which two persons are being referred to in such a way as to distinguish them in subsequent discourse, even though neither Ed nor us have ever met either of them. Sue's referent for "my husband" will not be identical with Ed's referent, though the referent for each of speaker and hearer counts as the same for the given occasion of talk. If the Archbishop of Canterbury says to Richard Dawkins *I will offer proof of the existence of God* and Dawkins replies *But God does not exist*, the

K. Allan (✉)
Monash University, Clayton, Australia
e-mail: keith.allan@monash.edu

deity that they are both referring to only counts as the same referent, because for the Archbishop God exists and for the author of *The God Delusion* God does not; in fact they have almost contradictory conceptions of the referent. This essay argues that an expression *e* frequently cannot identify exactly the same referent r for speaker and hearer, and that it is in fact unnecessary for it to do so; all that is required is that the referent counts as the same referent for the purpose of the communication. This is why mistaken reference like *Who's the teetotaller with the glass of water?* spoken of a man quaffing a glass of vodka can often successfully communicate who it is that is being spoken of; and attributives like the subject NP of *The person who designed Stonehenge was a genius* refers to whomsoever the designer was just as efficiently as *The architect of La Sagrada Família was a genius* refers, implicitly, to Antoni Gaudí.

1 Preliminary Remarks

In order to have a consistent means for speaking about what language users do with language, in this essay I define reference much more liberally than most philosophers and many linguists. For me **reference is a speaker's use of a language expression in the course of talking about (referring to) its denotatum.**[1] In short, my topic is **a speaker's act of referring**. This is a pragmatic conception of reference completely at odds with, say, Kaplan's conception of semantic reference (Kaplan 1989b: 491 n.13). For me, a referent is something the speaker (or writer or signer) talks about on a given occasion and so a referent can be many different types of entity: a particular, a universal, a proposition, an existent, a hypothetical entity or situation, even a non-existent—although reference to these

---

[1] Immediately we have a terminological problem with the term *denotatum*. Loosely distinguished: a speaker refers, the language denotes. For me a language expression *e* denotes (designates, if you will) something in a world—mostly outside of language, e.g. a knife, an act of killing, a state of desperation, a manner of moving, but also (within language) a noun, a predicate. A speaker uses *e* to refer to something that falls within the domain of what *e* can denote, though sometimes pushing the envelope. That which is denoted is the denotatum. As will become clear, it is not the case for me that a speaker *refers* to individuals and *denotes* general concepts.

last two categories is achieved only implicitly.[2] Many philosophers will say that linguistic representations of most of these cannot function as referring expressions, which instead may be "allusive", "attributive" or "descriptive".

There are many different views and definitions of *reference* (see for example Abbott 2010; Almog et al. 1989; Bach 1987, 2008; Gundel and Hedberg 2008; MacBride 2006; Sullivan 2006, 2012). Among philosophers there is considerable variation but the most uncontroversial "referring expressions" are demonstratives (*this, that*), proper names (*London, Peter Strawson*), personal pronouns (*she, it, them*), and definite NPs (*the computer*) such as "can occur as the subject of what would traditionally be regarded as singular subject-predicate sentence" (Strawson 1950: 320). One might generalize to say that the usual conception of reference limits it to constant individuated concepts to which a speaker (or the language expression S uses) draws to audience attention; on this view, general terms (e.g. *mats*) denote and don't refer. Strawson wrote:

> I have explained identifying reference—or the central case of identifying reference—as essentially involving a presumption, on the speaker's part, of the possession by the audience of identifying knowledge of a particular item. Identifying knowledge is knowledge of the existence of a particular item distinguished, in one or another sense, by the audience from any other (Strawson 1964: 101).

A little further on Strawson insists that the audience already knows of the "existence and uniqueness" of the referent and it is "no part of the speaker's intention […] to inform the audience of the existence of" it. Abbott 2010: 9 writes: "philosophical and linguistic research has yielded no clear-cut, obviously correct criterion for identifying either those NPs which encode the possibility for referential use, or those NPs which can be said to have a referent (in such-and-such an utterance)." In the course of this essay I hope to show that the meaningful distinctions often attributed to the terms *refer* and *denote* need not be abandoned on my conception of reference, but some more explicit and less controversial labelling will be required.

---

[2] Abbott 2010: 3 allows that "[O]n the semantic conception [of reference] most kinds of linguistic expressions might be considered to have reference—not only NPs but verbs and verb phrases (VPs), adjectives and adverbs, etc." I don't see why this cannot just as well apply to the pragmatic conception of reference. Incidentally, my definition of *reference* is close to lay usage, as demonstrated by the following four examples from corpora. (1, D17 3385 ACE Corpus) **Referring** to the eclipse or corruption of religion, he wrote: "Should the lamp of religion be obscured, chaos and confusion will ensue, and the lights of fairness, of justice, of tranquillity and peace cease to shine". (2, A06 170 LOB Corpus) **Referring** to previous negotiations, Mr Macmillan looked towards Mr Reginald Maudling. (3, E17 30 LOB Corpus) Well, Polish coach Felix Stam, **referring** to the omission of such stars as Pietrzykowski, Adamski, Drogosz and Pazdior in Belgrade, declared—"They are too old". (4, G17 0360 BROWN Corpus) But I suspect that the old Roman was **referring** to change made under military occupation—the sort of change which Tacitus was talking about when he said, "They make a desert, and call it peace" ("Solitudinem faciunt, pacem appellant").

2 The Nature of Reference

> "Mentioning", or "referring", is not something an expression does; it is something that someone can use an expression to do. Mentioning, or referring to, something is a characteristic of a *use* of an expression, just as "being about" something, and truth-or-falsity, are characteristics of *a use* of a sentence (Strawson 1950: 326).

Where I differ from Strawson and many others is in a wider interpretation of "identifying knowledge of […] the existence of a particular item"; and I allow that it can be part of a speaker's intention to inform the hearer of the existence of whatever the speaker is referring to. Bach 2008: 16 writes: "Speaker reference is a four-place relation, between a speaker, an expression, an audience, and a referent: you use an expression to refer someone to something". This I agree with, though I have a much wider interpretation of the term *reference* than Bach does. I strongly adhere to Strawson's dictum that referring is characteristic of the use of an expression; and I claim that, if it were to make any sense at all to say that "an expression refers", this is a function of the fact that it is either typically or on occasion used by speakers to refer. Referring is very obviously a pragmatic act: it is situated in a particular context (of both utterance and world spoken of) and "the rules of language and of society synergize in determining meaning, intended as a socially recognized object sensitive to social expectations about the situation in which the utterance to be interpreted is embedded" (Capone 2005: 1357; see also Wettstein 1989: 432). What is less obvious is that a hearer's recognition of the speaker's reference is also a pragmatic act of interpretation that uses common ground (which I will describe in a moment) to make sense of the utterance. Literary criticism, academic and legal argument (or the like) flourish because the "same" text can be interpreted differently by different hearers and readers, surely proving that interpretation is active not passive. In the ensuing discussion, we shall see some of what is involved in this process.

To start the argument, if Joe utters (1) then Joe refers to someone called Saddam Hussein whom he supposes his audience is able to identify from common ground.

(1) Saddam Hussein is dead.

Common ground is constituted from discourse context, situation of utterance, and input from relevant encyclopaedic knowledge.[3] On my definition of reference,

---

[3] See Allan 2013; Clark 1996; Clark et al. 1983; Lee 2001; Stalnaker 1973, 1974, 2002. Common ground for any community K of two or more people that include speaker and hearer is that:

(a) every member, or almost every member, of K knows or believes some fact or set of facts F; and
(b) a member is presumed to know or believe F by (almost) every other member of K; and
(c) a member of K knows that both (a) and (b) are true.
  When a member of K applies knowledge of F in order to interpret P, a state of affairs or something said, s/he can presume that others in the community will also apply knowledge of F in order to interpret P. The existence of F, P, and the application of knowledge of F to interpreting P is common ground for members of the community K. Once attended to, P becomes part of F, incrementing the common ground.

the speaker of (1) is talking about Saddam's death at some unspecified time in the past and thus referring to that. Furthermore, in ordinary lay language, a speaker would be said to be referring to Saddam's death in (1); to exclude it from "referring" by stipulation is inappropriate without rational grounds being given, and I have seen none. This analysis takes reference beyond certain constituents of propositions to propositions themselves—or, more precisely, to a proposition used by a speaker on a given occasion.[4] Just like Saddam Hussein himself, the fact of his death is a singular entity, and so ought to fall within the philosopher's canon of referring expressions. Any conceptual difference between identifying an individual such as Saddam and identifying the purported fact of this death correlate directly with what Systemic Functional Grammar calls the lexicogrammar, and investigations of such differences (along with those for the different kinds of reference/ denotation of different types of verbs and adjectives, adverbs, etc.) require a finer tool than a controversial distinction between the application of such terms of analysis as *refer, describe, denote, allude*, etc. There is no reason to exclude the potential of the predicate in (1) to refer on grounds that it is non-corporeal (after all, Saddam himself is non-corporeal today). If the speaker were intending to remind the hearer of Saddam's death, the statement falls within Strawson's criterion of "knowledge of the existence of a particular item" (Strawson 1964). Of course, it is more likely that Joe believes he is informing the hearer that Saddam is dead, a function which Strawson disallows as an act of referring. But I can see no value in Strawson's stricture. In short, speakers may refer to (purported) facts. It happens that it is true that Saddam is dead: Joe's utterance of (1) states a true fact. Had Ed said *Saddam Hussein is alive and well in 2011* he would also have been referring to a purported fact, but this time it is false—Ed was either ignorant of the truth or deliberately lying for some reason. But Ed is nonetheless referring. Reference is no guarantee of truthfulness or accuracy; those must be judged on other grounds.

For some people (e.g. Reimer 2003) reference is naming. The verb *name* is ambiguous (cf. Kaplan 1989a: 602): *I name this ship 'Sea Nymph'* is an act of baptism in the sense of Kripke 1972; I am here using the post-baptism sense of *name* as 'picking out a particular name bearer'. Referring and naming are closely correlated, and in the context of this paper it is true that to name is to refer; but the relationship is asymmetric because to refer is not to name. Although it is arguable that in *A great architect designed this church* one of the referents is 'named a church', for convenience I shall here restrict (the term) *naming* to the use of a proper name to identify a unique individual such as la Sagrada Família or a unique set as in *the Rockies* or *the Grateful Dead* (see Allan 2001; Lehrer 2006; Reimer 2006). Because naming is a type of referring, naming is (by logical transitivity) a pragmatic act. The hearer's recognition of the name-bearer is also, therefore, a pragmatic act.

---

[4] A proposition is the denotation of a sentence, cf. Abbott 2010: 7f; Castañeda 1989.
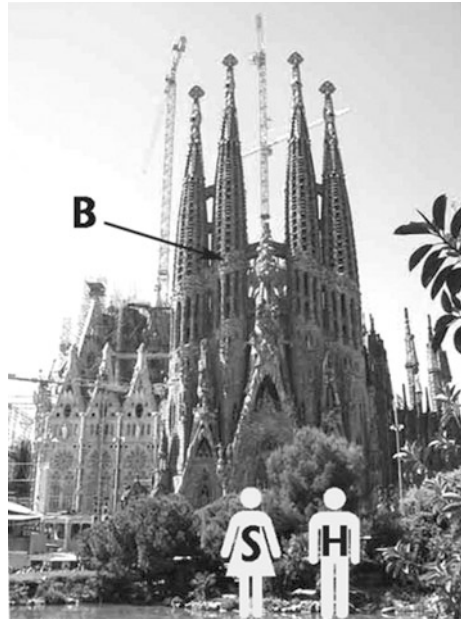
A speaker's act of referring is the speaker's use of a language expression in the
course of talking about its (purported) denotatum to an audience. What the hearer
recognizes as the speaker's referent necessarily only **counts as the referent**
because it is on many occasions not identical to what the speaker identifies (cf.
Colston 2008: 173); indeed the speaker and hearer might have substantially dif-
ferent, even contradictory, conceptions of the referent—and yet the language
expression used by the speaker can be said to successfully refer such that the
hearer recognizes the referent well enough for the communicative act to be judged
successful by both speaker and hearer. The hearer recognizes the referent well
enough if s/he is able to speak about that referent cogently and, if required, ascribe
certain properties to it; the hearer does not need to be able to identify the referent
as a physical entity.[5] For a philosopher, reference appears to be restricted to
identifying a particular (mostly singular) entity in the real world or more generally
entities that have extension in worlds and times accessible from the real world. For
the linguist, however, it is more important to be able to identify what the speaker is
(apparently) talking about when addressing the hearer in a given context, in order
to tie the way that is achieved to the language expressions used.

Consider (2).

(2) A great architect designed this church.

In order to clarify the significance of context on the interpretation of (2), I
propose that the following condition be placed on it: in Fig. 1, S (she) utters (2) to

---

[5] Wettstein 1989: 423, 439 says something similar, citing Kripke as an authority.

H (him) when standing in front of building B. Anyone who has been to El Temple de la Sagrada Família will be aware that it is visually dominant such that an utterance of (2) made when standing in front of it will unambiguously refer to la Sagrada Família unless the speaker is very obviously NOT referring to it because, for instance, s/he is pointing to a picture or description of another church (such as a guide book entry for the Basilica di San Marco).

Given my definition of reference, the speaker of (2) can legitimately be reported as referring to Antoni Gaudí, la Sagrada Família, and the fact that Gaudí designed la Sagrada Família. However, this interpretation depends on considerable inferencing from contextual and encyclopaedic data. In (2) uttered by S to H in the context described by Fig. 1, "this church" refers to B (because of its visual salience) without any recourse being necessary to the name of the church. In other words, the name of B is necessarily an additional inference from common ground. "La Sagrada Família" serves to identify the referent by naming it appropriately, but the name itself is not any part of (2). There are two things of interest here: how the reference to B is established, and how the referent is correlated with the name *Sagrada Família*. The subject NP of (2), *A great architect*, refers to the architect of B; the architect is not named and S does not necessarily know his name—she may simply be impressed by the architectural brilliance of B. If either S or H can name the architect it will be sourced from encyclopaedic knowledge (see Allan 2001, 2006) through identifying the church, probably, though not necessarily, by name.

According to Bach 2008:16, using an indefinite such as "A great architect" in (2) the speaker at best "alludes" to somebody if s/he has someone specific in mind, but s/he does not "refer" to that person. In my paragraph just above it is suggested that the speaker is talking about whoever the architect may be: Bach takes such attributives to "describe" what is spoken of (see also Abbott 2010: 263–270 which uses "speak of"). "Neither alluding to an individual nor singling one out descriptively counts as referring to it—you are not expressing a singular proposition about it" (*ibid.* 19).[6] As a linguist, I cannot see the value in these distinctions between "referring", "alluding", and "describing" or "speaking of", nor the special status awarded to expressing a singular proposition—what about all the other things we do with language? To subsume such speaker "allusion" and "description" to the speaker's act of referring, as I do, does not obscure the differences in meaning among those types of expression which (presumably) underlie Bach's distinctions.[7]

---

[6] Bach's position is basically similar to that of Russell 1905. A singular proposition is the content of a sentence containing an indicator that makes direct reference in Kaplan's sense. "The directly referential term goes directly to its referent, *directly* in the sense that it does not first pass through the proposition" (Kaplan 1989a: 569). Thus a singular proposition such as *Socrates was curious* is a statement about the man himself, not the name "Socrates" (see Abbott 2010: 34; Castañeda 1989: 114).

[7] I am not suggesting that the differences between any of e.g. universals vs particulars, definites vs indefinites, proper names vs descriptive names, etc. are uninteresting or irrelevant, just that to label some "referring terms" and others "allusive", "attributive" or "descriptive" is not the optimal way to differentiate them with respect to meaning.

The correspondence between reference and extension is complicated. Given the context of utterance defined by (2), "this church" has extension in whatever world S and H inhabit and is contemporaneous with them. S and H are recognizably fictitious characters restricted to Fig. 1 and the discussion presented by the author of this essay. For readers of the essay, "this church" has extension in Fig. 1 and, because the Figure depicts a church that exists in the world that my readers also inhabit, the church has extension in their real world. The reference to a great architect[8] also identifies an entity which has extension in whatever world S and H inhabit and the world which readers of this essay inhabit; but whereas la Sagrada Família still has extension in our world, its architect no longer does (he died in 1926).

Let's consider some things that can go wrong with (2). Ordinarily, the complex demonstrative *this church* refers to the most salient church in the foreground of attention[9]; it counts as what Reimer 2003 calls a "standard" reference. It identifies a necessary condition of the pragmeme for a particular kind of pragmatic act of reference that is clearly generalizable to a wide variety of occasions, as we shall see. There are several reasons for thinking that, in the situation of utterance described by Fig. 1, to refer using "this church" would be unambiguous. One is that S will know at least roughly where she is and under most circumstances, so will H. Even if S and H had not planned to be in front of la Sagrada Família and merely happened upon it, they would see[10] that (based on encyclopaedic knowledge) B looks like a church. They might also know that there is a church called (El Temple de) la Sagrada Família (or a translation of that name into another language). Ordinarily, but not necessarily, this will be strengthened by additional information such as knowledge of its approximate location and appearance, and/or some idea of its history and who designed it. It is conceivable that S is not aware of the name of the church; nevertheless, (2) would still be a credible utterance and so would (3), in which S accesses her encyclopaedic knowledge.

(3) This church must be the one that was designed by Gaudí.

Suppose S utters (2) in the context given by Fig. 1 but S mistakenly believes the church is La Seu (la Catedral de Santa Eulàlia), so that she could honestly comment on (2) by saying (4).

(4) In saying *a great architect designed this church* I am talking about La Seu.

---

[8] Or, if you prefer, to 'some x such that x is a great architect'.

[9] The phrase 'this church' is not a complex demonstrative when introducing the referent into the foreground of attention as in *Yesterday I came across this church with a gold-plated roof. It was just such an amazing sight.*

[10] I'm assuming neither is blind. Blindness would complicate matters, but not invalidate the general argument.

Fig. 2 *H* utters (5) to *S* as they stand before *B*
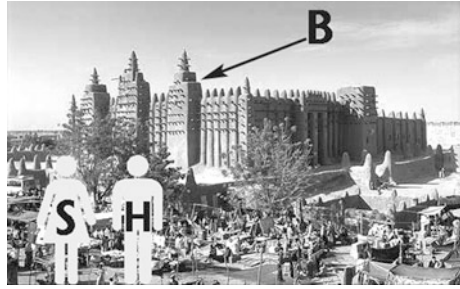
It is clear that there is nothing wrong with (2) on account of (4). The referential pract[11] is successful: S could honestly believe that a great architect designed B, referred to using "this church". The error that (4) reveals is the naming of B: in the context obtaining, (4) is referentially incorrect only because B is wrongly named "La Seu".

Let's now take the case of (5) as uttered by H to S in the context described by Fig. 2.

(5) A great architect designed this church.

H's referent for "this church" is readily identified as B, which is church-like even though it is in fact a mosque. Whether or not H (the speaker of (5)) can name the building as the Great Mosque of Djenné is irrelevant to S's successful interpretation of the utterance in (5) as referring to B. Once again, reference can be successful because of appropriate use of the pract: a certain kind of act is performed (namely, referring) in a certain kind of context—before B, which is an appropriate possible referent. Another example of misattribution that can refer successfully is suggested by Donnellan 1966: 287: the complex demonstrative in the question *Who's that teetotaller with a glass of water?* spoken of someone quaffing a glass of neat vodka can successfully refer insofar as the same person is recognized by both speaker and hearer as the one being spoken of, no matter what s/he is drinking. Mistaken reference will obviously be unsuccessful more frequently than 'correct' reference, but it doesn't cease to be reference (on my definition).

It is clear from (2), (4), and (5), that successful reference to B using "this church" is independent of the correct naming of B and even independent of whether B is literally a church, which it isn't in Fig. 2. This success is a function of

---

[11] Mey 2001: 221 writes: 'The theory of pragmatic acts [… focuses] on the environment in which both speaker and hearer find their affordances, such that the entire situation is brought to bear on what can be said in the situation, as well as on what is actually being said. […T]he emphasis is not on conditions and rules for an individual (or an individual's) speech act, but on characterizing a general situational prototype, capable of being executed in the situation; such a generalized pragmatic act I will call a *pragmeme*. The instantiated individual pragmatic acts, […] *practs*, refer to a particular pragmeme in its realizations'.

the pragmeme that appropriately combines the language material with situational data, drawing attention to B by employing the complex demonstrative phrase *this church*. Consequently, the default interpretation of the constituent "this church" merely needs to be 'the most salient church or church-like building in the foreground of attention'. I take this to be the default referent of "this church" in (2), (4), and (5).

Let's return now to Fig. 1. If H had attributed the correct name to the church referred to in (2), the resulting nonmonotonic inference, what Levinson 2000 calls an I-implicature, would be as shown in (6). If H wrongly believed that the church referred to is la Catedral de Santa Eulàlia, the implicature would be as shown in (7).

(6) H understands that S says 'a great architect designed the most salient church(-like thing) in the foreground of attention' +> a great architect designed La Sagrada Família.
(7) H understands that S says 'a great architect designed the most salient church(-like thing) in the foreground of attention' +> a great architect designed La Seu.

(6) is an appropriate interpretation of (2) in which S's reference to the church in (2) uttered under the conditions specified in Fig. 1 achieves success just because it instantiates the proper pragmeme. The S refers by means of a complex demonstrative to an entity in the world spoken of, namely B, that is readily identified. In (7) pragmemic integrity ensures that the referent of "this church" was recognized correctly as B; the fault is that the wrong name was (perhaps temporarily) assigned to B.

Suppose S were to follow up (2) with (8):

(8) In saying *A great architect designed this church*, I(S) meant that whoever [it may have been that] designed la Sagrada Família (B) was a great architect.

Let's make the default assumption that S is speaking felicitously, that is, she has genuine aesthetic grounds for stating her opinion that the design of B is such that it must be the work of "a great architect". It is notable that this may be a so-called "attributive" usage (as per Donnellan 1966) where S cannot name the architect. However (8) is also appropriate when S is able to correctly name the architect; her judgment of his skill is based on this building alone. In such a case, the optimal phraseology (without actually naming Gaudí) would be (9).

(9) In saying *A great architect designed this church*, I(S) meant that the person who designed la Sagrada Família (B) was a great architect.

However, (9) is ambiguous between what Donnellan called "attributive" and "referential" uses: "the person who designed la Sagrada Família" can be "attributive" (refer to whosoever the architect was) or, alternatively, refer to Gaudí. As an "attributive", the role of the senses of the indefinite description is direct or, the better to avoid misconstrual, **explicit**; when (2) or (9) is referring to Gaudí, the role of the sense is to make implicit reference. To be explicitly referring

to Gaudí, the speaker would need to name him in the utterance, as in *Antoni Gaudí designed la Sagrada Família* [(18) below]. **The locution typically identifies a speaker's explicit reference; the implicit reference is achieved via monotonic or nonmonotonic inference**.

At this stage it behoves me to explain why I have been putting quotes around the word *attributive* when discussing attributive uses of definite descriptions. It is because the attributive (from which I'm now dropping the quotes) is often contrasted with the referential, but I would insist that, given my definition of referring, in using attributives a speaker nonetheless refers. What characterizes an attributive such as "The person who designed Stonehenge" in (10) is that the identity of the referent is unknown and never likely to be known, but the speaker is nonetheless referring to that person who existed in prehistoric times in order to predicate a compliment of him (or, less probably, her).[12]

(10) The person who designed Stonehenge was very accomplished.

The difference between the referent of the attributive interpretation of "A great architect" in (2) given in (8) or (9) and that of the attributive in (10) is that the identity of the former is known (if not to S or H, then to us) whereas the identity of the latter is not. The true identity of a referent is not necessarily crucial for communicative success. In (11) the final "it" refers implicitly to an as yet unidentified member of the set of chocolates offered, one that will never be identifiable if the offer is refused such that the prediction fails to materialize. The reference, then, is to an entity in a hypothetical irrealis world.

(11) Take one of these lovely chocolates. I'm sure you'll enjoy it.
(12) Eat this chocolate. I'm sure you'll enjoy it.

In both (11) and (12), "it" satisfies the Bach description, quoted earlier, of "a four-place relation, between a speaker, an expression, an audience, and a referent". In both, the final pronoun "it" refers implicitly via the proposition in which it occurs and it correlates with an antecedent. Whereas the antecedent in (11), "one of these lovely chocolates", also refers implicitly[13] via the proposition in which it occurs to an unspecified member of a bounded set, the antecedent in (12) refers explicitly to the demonstrated chocolate, specified as a particular.

In uttering (13), the speaker refers to the universal (set of) "all spiders":

(13) All spiders have a cephalothorax and an abdomen.

A nominalist will dispute the realist's claim that universals exist; but that is of no concern to most speakers of English who are aware that, whether or not there is

---

[12] If it turned out that there was more than one designer of Stonehenge, (10) would be taken to refer to all of them. This is the situation with respect to the name *Homer* as author of the *Iliad* and *Odyssey* which—as they have come down to us—were composed by more rhapsodists than Homer, yet it is found convenient to refer to their author as, simply, *Homer*. Our understanding is not increased by pedantically recasting *Homer* as a collective noun.

[13] In terms of Bach 2008, this describes instead of refers.

an extension for the universal *all spiders* there is no question about the existence of spiders in this world and ordinary speakers have a conception of the meaning of the quantifier *all* ranging over spiders. Given my definition of reference, this is all that is required for a speaker to refer in using "all spiders" in (13). A speaker can also implicitly refer to something that is known not to exist by means of referring explicitly to the fact of its nonexistence as in (14). In (15) there is implicit reference to something whose existential status is uncertain. In (16) explicit reference is made to an entity that is fictional.

(14) *No human* has walked on Mars.
(15) Is there life on Mars?
(16) (Raymond Chandler's) *Philip Marlowe* is my favourite shamus.

In (14) the speaker refers to the planet Mars, and to a certain fact asserted about Mars. Many philosophers will dispute that it is possible to refer to a non-entity as opposed to, say, having it in mind. Whether or not the NP "no human" refers to a non-entity (*nullus*), it is clear that the speaker of (14) refers to a (negative) fact about Mars which can be assessed for its truth. The speaker of (15) also refers to Mars and questions the possible fact of existence of life on that planet—thus referring implicitly to (hypothetical) life on Mars. The speaker of (16) refers to a certain fact (personal judgment) about a fictional character, i.e. a person that exists in several works of fiction by Raymond Chandler that themselves exist in the real world. In my view there is explicit reference here to Philip Marlowe.[14]

To return to our investigation of meanings of (2) in the context described by Fig. 1, a further possibility is that S follows up (2) with (17):

(17) In saying *a great architect designed this church*, I(S) meant that Antoni Gaudí designed la Sagrada Família (B).

In (2), the reference to Gaudí and la Sagrada Família is implicit. To be explicit the speaker must utter (18), which is—of course—included as the explanatory clause in (17).

(18) Antoni Gaudí designed la Sagrada Família.

(2) and (18) satisfy different discursive functions and expectations, just as do the names *Hesperus* and *Phosphorus*. Notably, (18) offers no evaluation of Gaudí's prowess as an architect. The relevant practs for (2) and (18) are different: the former refers without naming, the latter names and thereby refers. The referential pragmeme is the same for both (2) and (18) and that is what sanctions identity of

---

[14] For Bach this is "pseudo-reference". The fact that there is no reason to make this distinction is attested by the following quotation: "For twenty-three years now I've been floating rivers. Always downstream, the easy and natural way. The way Huck Finn and Jim did it, La Salle and Marquette, the mountain men, Major Powell, a few hundred others." ('Preliminary Notes' to *Down the River*, Abbey 1982: 1). The author and reader recognize the reference to Twain's fictional Huckleberry Finn and Jim in the same manner as they recognize the historical persons of La Salle, Marquette, and Powell.

reference for the distinct constituents. What S said in (2) is most accurately glossed in (19) [the reasons were given above when discussing (9), (17), and (18)].

(19) In saying *a great architect designed this church*, S meant that a great architect designed the most salient church (-like building) in the foreground of attention.

There is a further possible interpretation of (2): in a third scenario S might explain her utterance (2) (in the context described by Fig. 1) by saying (20).

(20) In saying *a great architect designed this church*, I(S) meant that Frank Gehry designed la Sagrada Família (B).

First of all let's assume that "Frank Gehry" is not somehow a mismatch between tongue and brain such that the speaker had in mind Antoni Gaudí but misnamed him. In other words, S believes as she utters (2) that Frank Gehry designed B and that he deserves the accolade "a great architect". This erroneous but intended attribution in no way destroys the comprehensibility of (2) which still has the meaning that I attribute to it in (19). Only additional discourse will reveal S's mistake or perhaps, even though she may recognize her own error, it may never be explicitly corrected.

Let's consider some other quirks of reference. The speaker of (21) refers to a true fact.

(21) President Clinton was a baby in 1946.

In (21) the speaker also refers to the fact of something being a baby (the predication of babyhood) in 1946 and identifies this referent with Bill Clinton, the man who became President of the United States of America. Clearly, this state of babyhood applied to an individual entity vastly different from the one that was 42nd President of the United States from 1993 to 2001. Nonetheless, the two references count as identical in the sense that the speaker of (21) refers to what language users think of as the same individual at different times in his life. It is well recognized that a referent undergoes changes over time. This is specifically described in Heim 1983, 1988 as updating the file in any two successive references to an entity. For instance:

(22) Catch $[a\ chicken_1]$. Kill $[it_2]$. Pluck $[it_3]$. Draw $[it_4]$. Cut $[it_5]$ up. Marinade $[it_6]$. Roast $[it_7]$. When you've eaten $[it_8]$, put $[the\ bones_9]$ in the compost.

The speaker of (22) uses all nine subscripted NPs to refer (by my definition) to the creature identified as a live chicken in "$a\ chicken_1$", a nonspecific member of an unbounded set.[15] By 2 it is dead, by 3 featherless, by 5 dismembered, by 7 roasted, and by 8 eaten. 9 refers to the chicken's bones after the flesh has been

---

[15] Following the lead of Karttunen 1976, Bach 2008: 30 says that such "discourse reference" isn't reference, but he doesn't say what it is instead. Bezuidenhout 2013 might agree with me that this is discourse reference, I'm not sure.

stripped from them. Thus 7, for instance, refers not to the chicken in 1, but to the caught, killed, plucked, drawn, cut up, and marinaded pieces of that chicken. These successive states of the chicken are presented as changes in the world–time pair spoken of: although the world stays constant throughout (22), each clause corresponds to a temporal change: $time_1$, $time_2$, … $time_9$. Similarly, London (England) was very different in 1966 from London in 1666, but it was at both times named *London* and its spatial location is partly identical for the two periods. Reference to London on occasions three hundred years apart is normally taken to refer to 'the same city' even though language users recognize the differences that time has worked: we specify a temporal index to differentiate the different manifestations of the referent of *London* just as we differentiated President Clinton from the baby known in 1946 as *Billy Blythe* (William Jefferson Blythe III) who adopted the name *Bill Clinton* around 1960. The name changes that occur over time [see (23), where ⇒ can be glossed 'became' and "$t_j \succ t_i$" means 't_j succeeds t_i'] present temporally different manifestations of the referent for which different truths obtain.

(23) Billy Blythe [at t_i] ⇒ Bill Clinton [at $t_j \succ t_i$]

Byzantion [at t_i] ⇒ Kōnstantinoupolis [at $t_j \succ t_i$] ⇒ Kostantiniyye [at $t_k \succ t_j$] ⇒ İstanbul [at $t_l \succ t_k$]
Norma Jeane Mortenson [at t_i] ⇒ Norma Jeane Baker [at $t_j \succ t_i$] ⇒ Marilyn Monroe [at $t_k \succ t_j$]

There are other effects too. Compare (24) with (25).

(24) Marilyn Monroe starred in *Some Like it Hot*.
(25) Norma Jeane Baker starred in *Some Like it Hot*.

Although one can reasonably claim that "Marilyn Monroe" and "Norma Jeane Baker" have the same referent, (24) is true but (25) is not true—in the least, it is not true in the same sense that (24) is true. The speaker of (25) errs in not identifying **the appropriate manifestation of the referent** because it uses her baptismal name rather than her stage name. This fact about the appropriate manifestation of the referent is more important because more basic to the understanding of (25) than arguing over whether (25) is merely infelicitous or whether it is also false.

The speaker of (26) refers to Marilyn Monroe, her age (had she lived), a date, and another true fact.

(26) Marilyn Monroe would have been 74 on June 1, 2000.

Although Marilyn Monroe died in 1962 we can imagine a possible world of June 1, 2000 at which she was still alive and, given that she was born June 1, 1926, she would indeed be 74. Reference to things that no longer exist, reference to hypotheticals, reference to fictions, even reference to impossibilities is possible; we have already seen some examples in (11)–(16), and (27) refers to a true fact about an impossible entity.

(27) There is no largest prime number.

Impossible entities are alike to one another in being impossible, but the phrases *largest prime number* and *round square* are, nonetheless, recognisably distinct for the typical speaker of English: in fact their impossibility stems from a proper understanding of their constituent parts whose senses and the intensions conflict, e.g. something which is square cannot concomitantly be round.

If Sue says (28) to Ed, it is perfectly possible for Ed (and us) to understand that two persons are being referred to in such a way as to sufficiently distinguish them in subsequent discourse, even though neither Ed nor we have ever met either of them. Reference does not necessarily require that a hearer can physically pick out the referent; merely that it can be distinguished from distractors within the context of the particular communication.[16]

(28) My husband's having an affair with his boss.

Sue's referent for "my husband" will not be identical with Ed's referent, though **the referent for each of speaker and hearer will count as the same for this given occasion of talk**. Where the audience does not know the persons involved, the principal referent in (28) is the purported fact about Sue's husband's behaviour. Given the cooperative principle, it counts as a truth until disproved. Because it is the purported fact rather than the personae that are significant, it doesn't really matter that the boss turns out to be male rather than female because this additional fact has no bearing at the time of its utterance on the respective references in (28) to Sue's husband and her husband's boss.

Very similar are the references to the tree in (29), where the tree is not in view.

(29) LAYMAN: My elm tree is looking sick.
     ARBORIST: Is that ulmus procera or ulmus parvifolia?
     LAYMAN: I have no idea.

Both interlocutors refer to what counts as the same tree, the layman's tree, but they clearly have different conceptions of it. It is in no way infelicitous for the layman to single out this referent even if he has several elm trees on his property; none of the others is relevant to the interchange. It is sufficiently identified for him as the one that he thinks is looking sick. The arborist can make a finer distinction by identifying the subspecies of elm: it is often the case that a referent can be more precisely characterized by one interlocutor—which would, of course, be Sue's situation with respect to her husband in (28).

Hilary Putnam imagined a Twin-Earth that has counterparts to everything on Earth, including English. The sole difference is that Twin-Earth water$_T$ is not H_2O but XYZ, otherwise water$_T$ has all the properties that water$_E$ on Earth has—such as its potability and its being found in lakes (Putnam 1975: 232f). Putnam's question was that when Oscar$_E$ on Earth uses his term *water* and his Twin-Earth counterpart

---

[16] This is comparable with Putnam's example of the layman unable to distinguish an elm from a beech tree yet knowing that *elm* and *beech* denote different species of tree (Putnam 1975).

Oscar$_T$ uses his term *water* do they refer to the same thing? Putnam concludes that operationally they do but, by definition, the actual substances referred to are distinct. Suppose Oscar$_E$ visits Oscar$_T$ and asks (30).

(30) Can I have a glass of water?

I believe the reference for each of them counts as the same. It is only if a chemical analysis of the referent of *water* is at issue that the difference between water$_E$ and water$_T$ becomes critical. In this regard it should not be forgotten that on Earth (31) may be used of going into the sea, a lake, or a swimming pool and in each of those locations the referent of *water* is differently constituted. The appropriate referent will be contextually determined in accord with common ground.

(31) Can I go into the water, Mum?

At this point I will compare my account of "what counts as reference" with discussions of two-dimensional semantics. Unfortunately there are many differing, even conflicting, accounts of two-dimensional semantics e.g. in García-Carpintero and Macià 2006. One account that is comparatively appealing to me is that of Chalmers 2006a, b. To illustrate the theory take the two-dimensional account of the water on Earth and XYZ on Twin-Earth.

All water in the universe could be H_2O, which is the case on Earth, or it could be XYZ, which is the case on Twin-Earth.

Table 1 Two-dimensional account of the water on Earth and XYZ on Twin-Earth

	Earth	Twin-Earth
Earth water$_E$	H_2O	H_2O
Twin-Earth water$_T$	XYZ	XYZ

Table 2 The different roles of the real world name and the stage name

	Real world	Film world
Real world	Norma Jeane	Norma Jeane
Film world	Marilyn	Marilyn

Fig. 3 *D* is Dawkins, *A* the
Archbishop, *G* is God.
Dawkins' world is the ellipse;
the Archbishop's world is the
rectangle
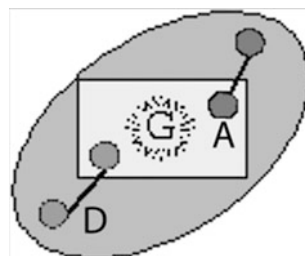
> We can say that an expression's 'diagonal intension' is a function mapping a world *w* to
> the term's extension when *w* is taken as both actual and counterfactual. (Chalmers 2006b:
> 577).

We can see in Table 1 the two pairs of diagonals that reflect the situation described: from the Earth perspective actual water is H_2O and it is counterfactual that water is XYZ (shaded cells); from the Twin-Earth perspective the situation is vice versa.

Two-dimensional semantics also seems to work with respect to the matters raised in (24) and (25). The diagonals in Table 2 do identify the different roles of the real world name and the stage name. The situation is comparable with that obtaining between the classical accounts of Phosphorous and Hesperus, both referring to Venus (Chalmers 2006a: 58–61).

Two-dimensional semantic theory works well for the Earth vs Twin-Earth scenario where there are different intensions and extensions of *water*$_E$ and *water*$_T$ and also for the different intensions of *Norma Jeane Baker* and *Marilyn Monroe* that have the same extension as demonstrated by the diagonals in Table 2. But it does not work for the different conceptions of "my husband" in (28) nor for (32) below because there is no way to get the top row and left column to match. Thus, although a version of two-dimensional semantics touches on some of the problems raised in this paper, it does not offer an account of all of them, and can be left aside.

Interlocutors may have contradictory conceptions of a referent, as in (32), uttered in 2009.

(32) ARCHBISHOP OF CANTERBURY: I will offer proof of the existence of God.

   RICHARD DAWKINS: But God does not exist.

For Rowan Williams (the Archbishop) God exists and for the author of *The God Delusion* (Dawkins 2006) God does not; so they have contradictory conceptions of the referent. Nonetheless, the deity that they are both referring to in (32) counts as the same and as overhearers we too understand them to be speaking of what counts as the same referent. The situation is represented in Fig. 3. For the Archbishop (A), he, God (G), and Dawkins (D) exist in the same world (depicted as a rectangle); for Dawkins (D), he and A exist in the same world (the ellipse), but G only exists within A's world (the rectangle). A and D have counterparts in both worlds.

3 Concluding Remarks

In this essay I have defined reference as the speaker's use of a language expression in the course of talking about its (purported) denotatum to a hearer such that the hearer should recognize what is spoken of sufficiently well for both speaker and hearer to be satisfied that the communication is successful. Referring and the recognition of what has been referred to are pragmatic acts. Thus *a referring expression* is simply a language expression that may be used by a speaker to refer. I have shown that speakers and writers can and do refer explicitly or implicitly to many different types of entity, to particulars (e.g. (1), (12)), to universals, (13), propositions [e.g. (14)], to current or former existents (18), hypotheticals, (11), and nonexistents (14), (27), (32). A referent only needs to 'count as a referent' because the nature of a referent may change over time (e.g. (21)) and because speaker and hearer may hold very different conceptions of the referent in a successful communication—as we saw in (28), (29), (30) and (32). All that is required for a speaker to successfully refer is that the hearer recognizes the referent well enough for the communicative act to be judged successful by both speaker and hearer.[17] Thus we saw in (5) that *this church* only needed to be identified with "the most salient church(-like thing) in the foreground of attention" (B, because of its visual salience). Mistaken reference was mentioned in respect of (5), and although mistaken reference will obviously be unsuccessful more frequently than 'correct' reference, it doesn't cease to be reference on that account. I distinguished explicit reference from implicit reference: *the author of 'Emma'* explicitly refers to the person who wrote 'Emma' (whoever or whatever that entity is) and only implicitly to Jane Austen—a fact determined via encyclopaedic knowledge.[18] In (2) and (5) *this church* explicitly refers to the most salient church(-like thing) in the foreground of attention. The identity between the particular church referred to and la Sagrada Família in (2) or the Great Mosque of Djenné in (5) is determined through encyclopaedic knowledge relevant to the context supplied by Figs. 1 and 2 respectively such that reference to la Sagrada Família or to the Great Mosque of Djenné is, consequently, only implicit. In (11), "one of these lovely chocolates" refers implicitly via the proposition in which it occurs to an unspecified member of a bounded set; in (14) the speaker implicitly refers to something that is known not to exist by means of referring explicitly to the fact of its nonexistence. The speaker of (15) questions the possible fact of existence of life on Mars and thereby refers

---

[17] To take a general case: a certain historical figure put to death by the Roman authorities in ancient Palestine is recognized as such in three religions, but the properties attributed to him are different: for Jews, Yeshua was just a preacher not a messiah; for Muslims, Isa was *al-Masī* the last great prophet before Mohammed but not divine; for Christians, Jesus is Christ the messiah and divine. Nevertheless, it is generally agreed that the names *Yeshua, Isa* and *Jesus* are used with the "same" referent intended—albeit with different attributed properties.

[18] If there were another lesser-known work entitled 'Emma' by Sue Flood, the implicit reference could be to her; but in this circumstance the explicit reference is the same: the speaker refers to the author of 'Emma'.

implicitly to life on that planet. The Dawkins denial of the existence of God in (32) is tricky. Arguably he implicitly refers to something whose existence he denies by means of referring explicitly to the purported fact of its nonexistence. Another account applies the kind of explanation given for the infelicity of (25), "Norma Jeane Baker starred in *Some Like it Hot*": in (32) God is manifest in the Archbishop's world wherein it is appropriate for Williams to explicitly refer to God; Dawkins can then implicitly refer to God as something manifest in the Archbishop's world[19] (comparable with a film or fictional world) but whose existence Dawkins denies by means of referring explicitly to the purported fact of its nonexistence in his world—which he takes to be the real world. A speaker of (33) refers to Philip Marlowe and while denying his real world existence implicits (see Bach 1994) his existence as a fictional persona.

(33) Raymond Chandler's LA gumshoe Philip Marlowe did not exist in the real world.

Bach 2008: 50 n.2 writes "there is a broad sense in which every expression refers (or at least every expression that has a semantic value that contributes to the propositional content of sentences in which it occurs)" and that would be my position if it is interpreted to mean 'in every expression uttered by a speaker using language normally, the speaker refers'.

Defining the pragmatic act of reference is problematic. A programmatic approach to such a definition includes the following well-known steps.

(a) The speaker S wishes to communicate with hearer H and S has an intention towards referent r. Intentionality is a property of the human mind/brain in virtue of which the mind targets a particular object of thought (see Jaszczolt 1999; Jacob 2003; Siewert 2006; Haugh and Jaszczolt 2012). The intentionality may precede the desire to communicate or vice versa.

(b) S believes that use of the language expression e_r to refer to referent r will enable H to recognize r, that is, distinguish it from potential distracters, normally, with minimum effort.[20]

(c) For r to be recognized by H from e_r, S will presume that H will make recourse to CG_i, that is, the common ground CG assumed to be shared between S and H at a time period beginning t_i, the time of anticipated interpretation. S must surmise what H will take to be in CG_i. In spoken communications CG_i is typically proximal to the time of utterance, but in written communications the time span between utterance and interpretation can be unbounded.[21]

(d) If S has the inclination and opportunity to be careful s/he will imagine him/herself in H's shoes as H seeks to recognize the referent, and S will label it accordingly. (This is standard procedure for an adroit communicator.)

---

[19] And the world of like-minded people.

[20] I intend clause (b) to encompass the reference to the man in *I saw this weird man that was screaming at passers-by on my way to work*. (Thanks to Sali Mufwene for this example.).

[21] There is some similarity here with the conclusions in Zielinska 2007: 828f.

(e) Particularly in face-to-face interaction, S can rely on H's response (verbal or nonverbal) to indicate whether or not s/he has achieved referential success (this becomes part of the common ground CG_{i+1}) and S may have the opportunity to relabel the reference using an alternative expression to help render the intended reference more amenable to H.

Clauses (b) and (c) are crucial and much has been written on these topics. For illustration consider (34), (35), and (36).

(34) Max shouted at Ed because he'd forgotten to set the alarm.
(35) Max shouted at Ed because he was in a foul mood.
(36) The vet smelled the dog's breath when she bit her.

In (34) and (35) the people referred to must normally be identifiable from common ground and in addition the "he" in (34) will most likely refer to Ed because Ed's failing to set an alarm can have unfortunate consequences which present a possible reason for Max to shout at Ed. It is less likely (but not impossible) that Max is shouting at Ed because Max himself has failed to set the alarm; but this would normally be explicitly marked as in *Max shouted at Ed although it was he himself who had forgotten to set the alarm*. In (35) "he" most likely refers to Max, because shouting at someone is evidence of being in a bad mood. Had Ed been in a foul mood, the cooperative speaker should have said something like *Max shouted at Ed for being in a foul mood*. In (36) it would be usual for "the vet" and "the dog" to be identifiable from common ground, and then knowledge of animal $\sim$ human behaviour (also a part of common ground) will identify the biter as the dog. Reporting an unusual event such as the vet biting the dog would normally demand explicit marking of agency as in *When the vet went to bite the dog she smelled its breath*.

We have seen that what counts as the referent in a successful act of communication may differ for speaker and hearer. Indeed it may differ for just the speaker if s/he is referring to different manifestations of a referent in different locations as in (30) and (31), or that have undergone mutative processes through the passage of time as exemplified in (21), or as the consequence of a series of predications as in (22). Speakers of (24) and (25) refer to different manifestations of a woman under her baptismal name and her stage name (which may constitute different legal entities). The manifestation of this same woman referred to in the counterfactual world described in (26) is yet again different. I conclude that to successfully perform a pragmatic act of reference requires astute assessment of the common ground and percipient choice of the language expression that will best point the hearer to the intended manifestation of the reference in those circumstances. Physical identification is not necessary, a hearer only needs to have a cogent grasp of what differentiates the speaker's (presumed) referent from any distractors.

References

Abbey, E. (1982). *Down the river*. New York: E.P. Dutton.

Abbott, B. K. (2010). *Reference*. Oxford: Oxford University Press.

Allan, K. (2001). *Natural language semantics*. Oxford and Malden: Blackwell.

Allan, K. (2006). Dictionaries and encyclopedias (relationship). In E. K. Brown (Ed.), *Encyclopedia of languages and linguistics* (2nd ed., Vol. 14, pp. 573–577). Oxford: Elsevier.

Allan, K. (2013). What is common ground? In A. Capone, F. L. Piparo, & M. Carapezza (Eds.), *Perspectives on linguistic pragmatics* (pp. 285–310). Cham: Springer.

Almog, J., Perry, J., & Wettstein, H. K. (Eds.). (1989). *Themes from Kaplan*. New York/Oxford: Oxford University Press.

Bach, K. (1987). *Thought and reference*. Oxford: Clarendon Press.

Bach, K. (1994). Conversational impliciture. *Mind and Language, 9*, 124–162.

Bach, K. (2008). On referring and not referring. In J. K. Gundel & N. Hedberg (Eds.), *Reference: Interdisciplinary perspectives* (pp. 13–59). New York: Oxford University Press.

Bezuidenhout, A. (2013). The (in)significance of the referential-attributive distinction. In A. Capone, F. L. Piparo, & M. Carapezza (Eds.), *Perspectives on pragmatics and philosophy* (pp. 351–366). Cham: Springer.

Capone, A. (2005). Pragmemes (a study with reference to English and Italian). *Journal of Pragmatics, 37*, 1355–1371.

Castañeda, H.-N. (1989). Direct reference, the semantics of thinking, and guise theory (constructive reflections on David Kaplan's theory of indexical reference). In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 105–144). New York/Oxford: Oxford University Press.

Chalmers, D. J. (2006a). The foundations of two-dimensional semantics. In M. García-Carpintero & J. Macià (Eds.), *Two-dimensional semantics* (pp. 55–140). Oxford: Clarendon Press.

Chalmers, D. J. (2006b). Two-dimensional semantics. In E. Lepore & B. C. Smith (Eds.), *The Oxford handbook of the philosophy of language* (pp. 574–606). Oxford: Oxford University Press.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., Schreuder, R., & Butterick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior, 22*, 245–258.

Colston, H. L. (2008). A new look at common ground: memory, egocentrism, and joint meaning. In I. Kecskes & J. L. Mey (Eds.), *Intention, common ground and the egocentric speaker-hearer* (pp. 151–187). Mouton de Gruyter: Berlin/New York.

Dawkins, R. (2006). *The god delusion*. London: Bantam Press.

Donnellan, K. S. (1966). Reference and definite descriptions. *Philosophical Review*, *75*, 281–304. Reprinted in D. D. Steinberg, L. A. Jakobovits (Eds.) (1971)., *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology* (pp. 100–114). London: Cambridge University Press.

García-Carpintero, M., & Macià, J. (Eds.). (2006). *Two-dimensional semantics*. Oxford: Clarendon Press.

Gundel, J. K., & Hedberg, N. (Eds.). (2008). *Reference: Interdisciplinary perspectives*. New York: Oxford University Press.

Haugh, M., & Jaszczolt, K. M. (2012). Speaker intentions and intentionality. In K. Allan & K. M. Jaszczolt (Eds.), *Cambridge handbook of pragmatics*. Cambridge: Cambridge University Press.

Heim, I. R. (1983). File change semantics and the familiarity theory of definiteness. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use, and the interpretation of language* (pp. 164–189). Berlin: De Gruyter.

Heim, I. R. (1988). *The semantics of definite and indefinite noun phrases*. New York: Garland.

Jacob, P. (2003). Intentionality. In N. Z. Edward (Ed.), *Stanford encyclopedia of philosophy*. Stanford: Stanford University. http://plato.stanford.edu/entries/intentionality

Jaszczolt, K. M. (1999). *Discourse, beliefs and intentions: Semantic defaults and propositional attitude ascription*. Oxford: Elsevier.

Kaplan, D. (1989a). Afterthoughts. In J. Almog, J. Perry, & H. K. Wettstein (Eds.), *Themes from Kaplan* (pp. 565–614). New York/Oxford: Oxford University Press.

Kaplan, D. (1989b). Demonstratives. In J. Almog, J. Perry, & H. K. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). New York/Oxford: Oxford University Press. (Circulated from 1977).

Karttunen, L. (1976). Discourse referents. In J. D. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground* (pp. 363–385). New York: Academic Press.

Kripke, S. (1972). Naming and necessity. In D. Davidson, & G. Harman (ed) *Semantics of natural language* (pp. 253–355). Dordrecht: Reidel. Republished separately as *Naming and Necessity*. Oxford: Blackwell. 1980.

Lee, B. P. H. (2001). Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics, 33*, 21–44.

Lehrer, A. (2006). Proper names: semantic aspects. In E. K. Brown (Ed.), *Encyclopedia of languages and linguistics* (2nd ed., Vol. 14, pp. 141–145). Oxford: Elsevier.

Levinson, S. C. (2000). *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

MacBride, F. (2006). Predicate reference. In E. Lepore & B. C. Smith (Eds.), *The Oxford handbook of philosophy of language* (pp. 422–475). Oxford: Clarendon Press.

Mey, Jacob L. (2001). *Pragmatics: An introduction* (2 (revised) ed.). Malden, MA and Oxford: Blackwell.

Putnam, H. (1975). The meaning of 'meaning'. In K. Gunderson (Ed.), *Minnesota studies in philosophy*. Minneapolis: University of Minnesota Press. Reprinted in *Mind, Language, and Reality: Philosophical Papers Vol.2* Cambridge: Cambridge University Press. 1975: 215–271.

Reimer, M. (2003). Reference. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Stanford: Stanford University. http://plato.stanford.edu/entries/reference

Reimer, M. (2006). Proper names: Philosophical aspects. In E. K. Brown (Ed.), *Encyclopedia of languages and linguistics* (2nd ed., pp. 137–141). Oxford: Elsevier.

Russell, B. (1905). On denoting. *Mind*, *14*, 479–493. Reprinted in R. C. Marsh (Ed.), (1956). *Logic and knowledge* (pp. 39–56). London: Allen and Unwin.

Siewert, C. (2006). Consciousness and intentionality. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Stanford: Stanford University. http://plato.stanford.edu/entries/consciousness-intentionality

Stalnaker, R. C. (1973). Presupposition. *Journal of philosophical logic, 2*, 77–96.

Stalnaker, R. C. (1974). Pragmatic presupposition. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and philosophy*. New York: New York University Press.

Stalnaker, R. C. (2002). Common ground. *Linguistics and Philosophy, 25*, 701–721.

Strawson, P. F. (1950). On referring. *Mind*, *59*, 320–44. Reprinted in *Readings in* J. Rosenberg, & C. Travis (Ed.,) (1971).*The philosophy of language* (pp. 175–195). Englewood Cliffs: Prentice-Hall.

Strawson, P. F. (1964). Identifying reference and truth values. *Theoria, 30*, 96–118.

Sullivan, A. (2006). Reference: Philosophical theories. In K. Brown (Ed.), *Encyclopedia of languages and linguistics* (2nd ed., Vol. 14, p. 420). Oxford: Elsevier.

Sullivan, A. (2012). Referring in discourse. In K. Allan & K. M. Jaszczolt (Eds.), *Cambridge handbook of pragmatics*. Cambridge: Cambridge University Press.

Wettstein, H. K. (1989). Cognitive significance without cognitive content. In J. Almog, J. Perry, & H. K. Wettstein (Eds.), *Themes from Kaplan* (pp. 421–454). New York/Oxford: Oxford University Press.

Zielinska, D. (2007). The selective mode of language use and the quantized communicative field. *Journal of Pragmatics, 39*, 813–830.

What is Common Ground?

Keith Allan

Abstract Language is primarily a form of social interactive behaviour in which a speaker, writer or signer (henceforth S) addresses utterances (U) to an audience (H). It requires S to make certain assumptions about H's ability to understand U. This includes choice of topic, language, language variety, style of presentation, and level of presentation (because, for instance, addressing a neophyte or a child must be differently handled from addressing a group of experts). These assumptions constitute what can conveniently be called "common ground". They have been subsumed to context (e.g. Allan 1986; Duranti 1997); and at least a part of the common ground constitutes what Lewis (1969) referred to as "common knowledge", a term adopted by Stalnaker (1973). Schiffer (1972) called it "mutual knowledge*". Prince (1981) rejected "shared knowledge", preferring "assumed familiarity". Following Grice (1981, 1989), Stalnaker (2002) named it "common ground", which he described as "presumed background information shared by participants in a conversation" [...] "what speakers [take] for granted—what they [presuppose] when they [use] certain sentences". A fatal flaw was carried over from Schiffer's definition of mutual knowledge* into Stalnaker's definition of common ground: "It is common ground that φ in a group if all members *accept* (for the purpose of the conversation) that φ, and all *believe* that all accept that φ, and all *believe* that all *believe* that all accept that φ, etc." (Stalnaker 2002: 716). The recursion within this definition would necessitate infinite processing on the part of each of S and H. This flaw has been accepted and repeated by many since (e.g. Kecskes and Zhang (2009, 2014)). Clark (1996) attempted to circumvent it but his definition includes a clause that calls itself, thus creating an endless loop. In this essay I suggest a way, inspired by Lee (2001), to characterize common ground from the points of view of both S and H and which does not admit runaway recursion. In line with Stalnaker's mingling of presupposition and common ground, it refers to the preconditions on illocutions.

K. Allan (✉)
Monash University, Melbourne, Australia
e-mail: keith.allan@monash.edu

1 Introduction

My intention in this chapter is to examine and define the notion of *common ground* (CG), which is crucial to those of us who take pragmatics seriously. I shall assume (contrary to Chomsky 1975: 56f; Chomsky 1980: 229f, 239) that human language is characteristically a form of social interactive behaviour; it may occasionally have other functions, but the motivation for its coming into existence (see Allan 2003, 2010: 233; Dunbar 1996) and by far the majority of its usage is when S (speaker, writer, signer) addresses utterance U to audience H for an unbounded number of perlocutionary and illocutionary purposes such as to establish or maintain a social relationship, to inform, question, demand, warn, apologize, and so forth. S and H are mutually aware that, normally, their interlocutor is an intelligent being. S does not need to spell out those things which are obvious to the sensory receptors of H, or such that H can very easily reason them out using the knowledge that each of us develops from birth as we experience the world around us on the basis of communicative competence (knowing the language and the conventions for its use). These constitute **common ground** (CG). Our understanding of linguistic utterances rests on an assumption of CG: e.g. when S points to something visible in the situation of utterance and says *Isn't that nice?* there is an assumption that H understands English and can also see it; saying *Let's go to Cracow* assumes that "Cracow" will be understood as referring to a certain city. Some CG is universal, e.g. knowledge of the sun as a heavenly body that is a source of light and warmth, rain as (among other things) a source of fresh water replenishing the earth, the physiological and socio-cultural differences between the sexes. Some CG is very restricted, e.g. between a couple who use *the Hobgoblin* to refer to the man's first wife. Usually S can readily assess the probable CG with H, and chooses his or her words accordingly. This requires S to make assumptions about H's capacity to understand U well enough that S's (expressed intention in the) message is, in S's opinion, more or less correctly interpreted by H (Allan 1986; Colston 2008: 173).

> The speaker designs his utterance in such a way that he has good reason to believe that the addressees can readily and uniquely compute what he meant on the basis of the utterance along with the rest of their common ground. (Clark et al. 1983: 246)

S's assumptions here are S's estimates of the CG between S and H with respect to U; this is not something S is normally conscious of except, perhaps, when communicating with a stranger—and not often then. Assumed CG is based on an

assessment of H's competence to understand U,[1] and it motivates such things as choice of language and language variety, style and level of presentation—because, for instance, addressing a neophyte or a child must be differently handled from addressing a group of experts. CG allows meaning to be underspecified by S, so that language understanding is a constructive process in which a lot of inferencing[2] is expected from H.

These are linguistic aspects of CG and there must also be assumptions about what H may know of the world, which can affect the choice of utterance topic, and even whether or not S should address H at all. H also makes assumptions about the CG with S, basing it on H's assessment of U in the context of utterance and of S as a person. S's assessment of CG with H and H's assessment of CG with S are unlikely to be identical: all that is required is that the overlap in S's and H's assessments of mutual CG enables S to be satisfied that H understands U well enough for S's communicative purpose to, in S's judgment, succeed. This will apply to each utterance in a discourse such that the relevant CG is dynamic and typically accretes. As conversation proceeds the CG develops (Stalnaker 2002: 701): if, where A, B, and C are interlocutors, A says φ and B says χ then, normally all of A, B and C (keeping score in terms of Lewis 1979) will know that A either subscribes to or purports to subscribe to φ and B to χ, whether or not the other interlocutors also subscribe to φ and χ. Furthermore, in a talk exchange, the roles of S and H will alternate among interlocutors. The situation is again complicated by the fact that, when uttering U, S will often address more than one person and so is required to assess CG with an audience of any number of people.

In the course of this essay I shall colour in this rough sketch and add some additional constituents of common ground. An oversimplifying sketch of common ground between just two people, X and Y, is given in Fig. 1. §2 of this essay discusses some alternative labels for and/or concepts closely related to what I am calling *common ground*. §3 considers some definitions of CG and seeks one without the fatal flaw of runaway recursion. §4 applies the findings of §3 to the analysis of common ground in several real language texts, both written and spoken. §5 concludes the discussion with a summary and some additional points of interest.

2 Common Ground and its Aliases

Linguists and philosophers of language who do not recognize common ground under that label often acknowledge it—at least in part—under some other name.

---

[1] Assumptions about common ground are made in any social encounter and not restricted to language, though linguistic environments are all that concern me here.

[2] Inferencing, which may arise from spreading activation within an associative network, includes enrichment of implicitures and implicatures, disambiguation and the like.
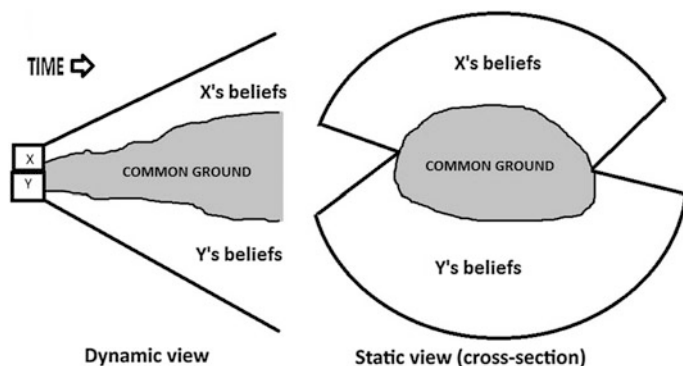
Fig. 1 An over-simplified sketch of common ground between X and Y

For instance Allan (1986) subsumes it to the recognition of aspects of context such as recognition of elements in the context of the situation of utterance and S's presentation of and H's recognition of the world spoken of in U. The world and time spoken of (or written of, or signed) is a mental model of the world which we construct in order to be able to produce or understand a phrase, a sentence, or a much longer text. In the course of interpreting any text, H must construct a model of the world and time spoken of. For instance, to interpret a declarative sentence such as (1), H models a world in which it is day-time and the sun is (mostly) shining and there is (at least) one person mowing a lawn.

1. It's a sunny day and someone is mowing a lawn.

Typically, the world and time spoken of contain people and things H knows or knows of; thus it is a contextualization of the states of affairs referred[3] to by S in terms of place, objects, and participants, etc. It can be (a reconstruction of) the real world, or some other possible world that can be imagined, desired, or supposed.[4]

Duranti 1997: 27ff includes sensitivity to cultural and procedural knowledge as aspects of "contextualization" (cf. Gumperz 1982: 131) pointing out that

> *speakers design their speech according to their on-going evaluation of their recipient as a member of a particular group or class. [… And] speakers change the content of what they say depending on whom they identify as their primary recipient.* (Duranti 1997: 299f. Sic.)

This is a reaction to S's perception of CG with H. Duranti (1997: 294) quotes Phillips 1972: 377:

---

[3] I use the term *refer* deliberately. A speaker's act of referring is the speaker's use of a language expression in the course of talking about its (purported) denotatum to an audience. For justification see Allan 2001, 2013.

[4] Occasionally people speak of logical impossibilities such as *the largest prime number*, but this does not affect the discussion here.

> Teachers use different participant structures, or ways of arranging verbal interactions with students, for communicating different types of educational material, and for providing variation in the presentation of the same material to hold children's interest.

The same applies, *ceteris paribus*, to any S addressing any audience: S must be capable of presenting different material to different audiences according to the task to which U is put in such a way as to hold audience interest. S will often modify CG by presenting (acting out) a persona taking a certain stance on what is being said (revealing S's footing, in terms of Goffman 1981: 128) and may present as not only the animator, but also the author and principal—or not—of what is said in U (*ibid.* 145, 167, 229). This is one means by which S can manipulate the CG.

A part of the CG constitutes what Lewis 1969: 56ff referred to as "common knowledge" within a population, a term adopted by Stalnaker 1973 who comments: "this background of knowledge or beliefs purportedly shared by the speaker and his audience constitute the presuppositions which define the context" (p. 448). By "context" Stalnaker apparently means both the situation of utterance and the world spoken of. Abbott 2008 offers some significant objections to Stalnaker's notion of presupposition, one being that items may be introduced into common ground if they are noncontroversial, as in Grice 1981: 190:

> For instance, it is quite natural to say to somebody, when we are discussing some concert, My aunt's cousin went to that concert, when one knows perfectly well that the person one is talking to is very likely not even to know that one had an aunt, let alone that one's aunt had a cousin. So the supposition must not be that it is common knowledge but rather that it is noncontroversial, in the sense that it is something that you would expect the hearer to take from you (if he does not already know).

Lewis 1979: 340 spoke of H "accommodating" to such introductions. I would see this as a predictable part of a personal relations frame or schema (compare that waitpersons, tables, and food are part of a restaurant frame and script; see Allan 2001; Mazzone 2011; Schank and Abelson 1977). Abbott also cites the "informative-presupposition *it*-clefts" identified by Ellen Prince, whose purpose is "to inform the hearer of that very information" (Prince 1978: 898). For example:

> The leaders of the militant homophile movement in America generally have been young people. IT WAS THEY WHO FOUGHT BACK DURING A VIOLENT POLICE RAID ON A GREENWICH VILLAGE BAR IN 1969, AN INCIDENT FROM WHICH MANY GAYS DATE THE BIRTH OF THE MODERN CRUSADE FOR HOMOSEXUAL RIGHTS. ([*Pennsylvania Gazette* February 1977: 16, referring to the Stonewall Riot, June 27–29 1969], Prince 1978: 898, ex. 41b. *Sic.*)

Another kind of informative example is *If you're going into the bedroom, would you mind bringing back the big bag of potato chips that I left on the bed?* (Abbott 2008: 531, citing Birner and Ward 1994: 93). And there are dozens more. So Abbott disputes that presupposition can be more or less equated with CG. Presuppositions are relevant to CG: S might be said to presuppose some CG with H and vice versa; but for the time being we continue with the verb *assume* in preference to *presuppose* because I hold to the view that presuppositions are the preconditions (preparatory conditions) on illocutions, subject to the maxim of quality (see Allan 2001: 204ff for discussion). Although presuppositions constitute

a part of CG, there is more to CG than them alone. I shall return to preconditions on illocutions, and hence presupposition, when defining CG.

Lewis's "common knowledge" is essentially similar to Schiffer's "mutual knowledge*",[5] described as follows:

> For example, all "normal" people know that snow is white, know that all normal people know that snow is white, know that all normal people know that all normal people know that snow is white, and so on *ad infinitum*. (Likewise, I should think, for all or most of our common general knowledge; so if *S* and *A* mutually know* that each is "normal", all of the general knowledge that each has in virtue of being a "normal" person will also be mutually known* by them.) (Schiffer 1972: 32)

As Schiffer recognized, an intractable problem with "mutual knowledge*" is that it is infinite; in other words both S and H will be processing an utterance *ad infinitum*—which is obviously contrary to fact. This problem has been resolved on an ad hoc basis by arbitrarily limiting the number of recursions to three or four. But, as we shall see in §3 and §4, it is also possible to dispense with the problem altogether.

Just as Lewis 1969 assumes that "common knowledge" implicitly defines a community (group) wherein the knowledge is common, so does Schiffer 1972:131 explicitly assign this property to "mutual knowledge*". The identification of community is a relevant aspect of CG; as Enfield 2008: 235 writes: "The management of common ground is directly implicated in our perpetual attendance to managing personal relationships within our social networks."

"Shared knowledge refers to knowledge that is possessed by all interlocutors; whether the interactants are aware of each other's awareness of this state is not relevant" (Holtgraves 2002: 125). However, Prince 1981: 230 does not agree:

> [G]ivenness in the sense of 'shared knowledge' may be described as follows: Givenness$_k$: The speaker assumes that the hearer 'knows,' assumes, or can infer a particular thing (but is not necessarily thinking about it).

This pertains to the "informative-presupposition *it*-clefts" discussed in Prince 1978: 898ff. In a remark that applies also to mutual knowledge* with its runaway recursion, Prince then rejects "shared knowledge" on grounds that it "is taking the position of an omniscient observer and is not considering what ordinary, non-clairvoyant humans do when they interact verbally" (Prince 1981: 232); she prefers "Assumed Familiarity" (233) which she further divides into seven categories that are not germane to this essay.

Following Grice 1981: 190 (= Grice 1989: 65) Stalnaker referred to "common ground", described as "presumed background information shared by participants in a conversation" (Stalnaker 2002: 701) or "what speakers [take] for granted—

---

[5] Schiffer uses the asterisk to mark the phrase a term of art. Schiffer 1972 makes explicit the essential similarity between "common knowledge" and "mutual knowledge*" though it is sometimes claimed (e.g. on http://www.gametheory.net/dictionary/MutualKnowledge.html) that "common knowledge" lacks the recursive aspect of "mutual knowledge*" described immediately below.

what they [presuppose] when they [use] certain sentences" (*ibid.* 702). What seems abundantly clear is that although one might nit-pick differences among them (see Lee 2001), the terms *common knowledge, mutual knowledge*, shared knowledge, assumed familiarity*, *presumed background information* and *common ground* are describing essentially the same thing, and it is what defines the pragmatic constituent of communicative competence: the knowledge and application of how and when to use utterances appropriately that combines with grammatical knowledge (of semantics, syntax, morphology, phonology) in the production of utterances to generate a coherent text comprehensible to its intended audience.

3 Towards a Definition of Common Ground

Stalnaker describes CG as follows:

> The common beliefs of the parties to a conversation are the beliefs they share, and that they recognize that they share: a proposition φ is common belief of a group of believers if and only if all in the group believe that φ, all believe that all believe it, all believe that all believe that all believe it, etc. (Stalnaker 2002: 704)

Stalnaker rightly adds temporary assumptions, probable presumptions, and pretended beliefs to what is mutually known as a potential part of CG. He points out that X may believe of Y that Y mistakenly believes that φ is a common belief, while X takes φ to be an uncommon belief (*ibid.* 708). Colston 2008: 160 allows for false beliefs where because φ is said to have been previously spoken of or done in X's presence, X may come to falsely believe that φ occurred, even though it did not.

Stalnaker defines CG thus:

> It is common ground that φ in a group if all members *accept* (for the purpose of the conversation) that φ, and all *believe* that all accept that φ, and all *believe* that all *believe* that all accept that φ, etc. (*ibid.* 716. *Sic*)

("To accept a proposition is to treat it as true for some reason" *ibid.*) Here again is the unsatisfactory infinitely recursive definition taken over from Schiffer's "mutual knowledge*". It is adopted with minor changes by Kecskes and Zhang 2009 without amending or mitigating this fatal flaw—fatal, because (as I have said) it requires that for both S and H an utterance is processed *ad infinitum*, which is contrary to fact.

Clark 1996 Ch.4 describes and defines CG, basing it on a presumption of awareness (p. 93f): X is aware of the world around him/her and is aware of being aware of it, i.e. X is not asleep, in a coma, stoned out of his/her mind, or the like. X is also aware that an interlocutor Y is aware in a similar way to X; likewise for Y in respect of X. The shared basis for mutual awareness is not a sharing of exactly

identical facts and suppositions, but there will be a substantial overlap. Clark's account of shared CG (Clark 1996: 94f) can be paraphrased as in (2).

2. (i) Every member of community K is aware that B (the basis for believing a set of propositions) holds true.
(ii) B indicates to every member of K that every other member of K is aware of B.
(iii) B indicates to every member of K that φ.
(iv) φ is common ground in K.
(v) [Reflexive common ground:] Every member of K has the information that φ and that (v).

The awareness condition is applicable to all human interactive behaviour and not especially relevant to the linguist's definition of CG. Although (2)(v) appears to side-step the recursion *ad infinitum* which vitiates Schiffer's "mutual knowledge*" and Stalnaker's definition of CG (as well as that of followers like Kesckes and Zhang), it nonetheless creates an endless loop by recalling itself.[6] Clark's "principle of justification" seeks to avoid the recursion and looping problems with CG:

> In practice, people take a proposition to be common ground in a community only when they believe they have a proper shared basis for the proposition in that community. (Clark 1996: 96)

However, this fails to specify the grounds for belief in a shared basis for φ, which take it back to the procedure in (2). The only recourse is to revise (2).

Allan 2001: 21 suggested the revised definition of CG in (3), which I have here ameliorated slightly by specifically mentioning S and H.

3. Common ground for any community K of two or more people that include S and H is that:
(a) every member, or almost every member, of K knows or believes some fact or set of facts F; and
(b) a member is presumed to know or believe F by (almost) every other member of K; and
(c) a member of K knows that both (a) and (b) are true.

When a member of K applies knowledge of F in order to interpret P, a state of affairs or something said, s/he can presume that others in the community will also apply (or be able to apply) knowledge of F in order to interpret P. The existence of F, P, and the application of knowledge of F to interpreting P is common ground for members of the community K. Once attended to, P becomes part of F, incrementing the common ground.

---

[6] Clark (pc, July 2011) disputes this objection claiming that 2(v) is simply self-referring on a par with *This sentence contains five words*. *Pace* Herb, but I maintain that 2(v) calls itself such that it generates unbounded recursion: Every member of K has the information that φ and that every member of K has the information that φ and that every member of K has the information that φ and that every member of K has the information that φ and that every member ….

The "community K" referred to in (3) may consist merely of people who chance to be interlocutors on a given occasion having in common an intention to communicate with each other using a language of which neither of them have to be fluent speakers but which entails some cultural and historical heritage to which they have access; most often, K satisfies the usual conditions for *community* in that its members share a common language of which they are native or native-like speakers, they share a common cultural and historical heritage, and are located in a specific locality.

"Consensus is fundamental to defining cultural communities" (Clark 1996: 105) and (3)(c) owes something to Lewis's definition of convention (Lewis 1969: 78). Each of (3)(a) and (3)(b) is common knowledge in K, and therefore so is (3)(c).[7] F includes not only behaviours but also manifest facts such as what can be seen, heard, smelt, etc. by the interlocutors. Included among F are "schemata" (Bartlett 1932; Mazzone 2011), "frames" (Minsky 1977; Fillmore 1982), "scripts" (Schank and Abelson 1977; Schank 1984), "scenarios" (Sanford and Garrod 1981), and "Assumed Familiarity" (Prince 1981)—all of which capture the fact that our brains look for, detect, and store structured patterns of information that constitute part of "common knowledge" in the sense of Lewis 1969 and Schiffer's "mutual knowledge*". On most if not all occasions P is effable: it can be expressed in a proposition or set of propositions φ. I have less confidence that F is always effable, but mostly it is. Note that (3)(a) and (3)(b) allow for a member M_i of K to not know or not believe F, permitting miscommunication to arise. For instance, if X says *I've just been talking to Harry* and Y responds *Harry who?* then X is expected to explain who Harry is. Sometimes S assumes something is not in CG with H, when in fact it is. As a rule H corrects S as in (4).

4. B: I guess he buys the books for uh something called Borders which is a bookstore that car-

   A: Yes we have it here too. (Horton and Gerrig 2005: 24)

Lee 2001 describes empirical data from the Map Task of Brown 1995 in which two interlocutors are given slightly different maps (which the other cannot see), one being identified as an update of the other. Lee discusses what seems to be taken as CG when X guides Y through a route on the map. A plausible story for CG between two "ordinary, nonclairvoyant humans" (to quote Prince again) is shown in (5), adapted from Lee 2001: 38.

5. | | Y | Y/X | Y/X/Y |
|---|---|---|---|
| X: you start below the palm beach right | CG | CG | |
| Y: right | CG | CG | CG |

---

[7] (3) does not invoke the notion of "collective belief" as described by Gilbert 1987, 1989—I am referring to what a member of K assumes about the beliefs of other members of K—and more particularly H. In my view, to convert this to what a member of K assumes to be a (collective) belief in K would be inaccurate.

The topmost CG under Y indicates that Y assumes "the palm beach" is on Y's map, Y can identify the unique "palm beach" accurately because it is in plain sight. The sister CG under Y/X assumes that Y recognizes that X assumes the palm beach (B) is on Y's map, because that is entailed by X's utterance. Y's answer "right" confirms that B is on Y's map and thereby confirms that Y recognizes that X assumes B is on Y's map. Additionally the CG under Y/X/Y indicates that Y recognizes that X believes that Y recognizes that B is on Y's map. No further recursion seems warranted. Based on many instances of data like (5), Lee notes that Brown 1995: 227 cannot see grounds for requiring more than just the three steps of recursion demonstrated in (5), and introspection surely confirms the accuracy of this. I note that there is a parallel here with Grice's notion of "reflexive intention" (Grice 1957, 1968, 1969); although Grice later modified this first to an "iterative intention" and then a "sneaky intention" (in Grice 1982) he had no need to do so, it was always satisfactory (see Bach 2012). Reflexive intention can be characterized as S's intention to have H recognize that when uttering U in context C, S intends to have a certain effect on H partly caused by H recognizing that S has the intention to communicate with S by means of U. The analogy with respect to CG is that when S mentions φ in U, S intends H to recognize φ and, furthermore, that H comprehend that S intends that H recognize φ as a result of S uttering U in context C. Hence in (5), Y's "right" means 'I start below the palm beach that you, X, have mentioned as probably being on my map and that I recognize as in fact being on my map as you, X, anticipated (which is why you, X, mentioned it)'.

How does this analysis square with the definition of CG in (3)? In (5) community K consists of just X and Y. Among the set of facts F are that each is aware that the other has a map before them and that these maps are different in that one has been identified as an update of the other. The task is a map-reading scenario with a journey script evoked. P in (3) is realized by each clause of each utterance in (5).

Missing from (5) is what X has in mind; after all, CG is what is common to both Y and X. So let's tackle the analysis differently from Lee 2001. What X says ("you start below the palm beach") pragmatically entails[8] that (a) X believes that Y has B on Y's map and that Y can confirm (or deny) this supposition ("X: … right") and (b) Y will assume that X believes that Y has B on Y's map and that Y can confirm (or deny) this supposition. What Y says pragmatically entails that (a) Y confirms that B is on Y's map and that X's supposition was correct and (b) X may assume that X's supposition was correct because B is on Y's map and that Y can therefore "start below the palm beach". No further recursion is needed.

Generalizing over this analysis we get (6):

6. i. X says φ to Y pragmatically entails (a) X believes that φ and (b) Y has some reason to believe that X believes that φ.

[8] If A pragmatically entails B, B cannot (given A) be denied without creating a paradox, absurdity, or contradiction.

ii. If in saying φ to Y, X refers to α, this act of referring to α pragmatically entails that (a) X believes Y can identify α (knows who or what α is) and (b) Y recognizes that X believes Y can identify α. Typically, when Y cannot identify α, Y asks X for further information.

The description in (6) is reminiscent of the preconditions on a felicitous statement (see e.g. Allan 2001, 2006; Austin 1962; Bach and Harnish 1979; Karttunen and Peters 1979; Searle 1969) which correspond to speaker presupposition and justify to some degree the views of e.g. Kecskes and Zhang 2013; Stalnaker 1974, 1978, 2002. There is more to CG than speaker presupposition, however, because CG must also take the beliefs of H into account.

4 Analysing Common Ground

In this section I analyse CG in one invented text, one fictional text, one nonfictional text, and one spoken text. The invented text is (7).

7. A male colleague, X, turns up late for a meeting and on entry immediately says *I'm sorry, my car broke down.*

In terms of the definition in (3), the words uttered (in italics) constitute P and the stage direction constitutes part of F. The following aspects of CG in (7) also comprise components of F.

(7A) In most Anglo societies (and many others) the situation, turning up late at a meeting, is impolite and so typically demands an apology (part of a meetings script). $K_{(7)}$ is such a community. "I'm sorry" satisfies this social constraint in $K_{(7)}$.

(7B) To explain away an offense mitigates face-loss and that is why X will normally be understood to be apologizing for being late, not for the fact that his car broke down. That is, mention of the car break-down is intended to explain away X's being late on the ground that in $K_{(7)}$ car-break-downs disrupt journey schedules by extending journey times (sometimes blocking journey completions altogether). This is part of a car-break-down script.

(7C) Even if none of X's colleagues knew he was coming by car, X does not have to spell this premise out because in most Anglo societies travel by car is a common means of transport; knowledge of this fact is part of CG within that community. Let's assume that (7) is felicitous because it is located in such a community ($K_{(7)}$).

The mundane enrichment of what is said on the basis of CG rests upon knowledge of social and cultural conventions and the cognitive principles that govern our thinking, all of which need to be accounted for in a linguistic theory of utterance meaning.

with the rather tongue-in-cheek "net curtains that needed laundering and the boy's size library table with the venerable magazines on it to give the place a professional touch". The final sentence of the first paragraph "I always left this [door] unlocked, in case I had a client, and the client cared to sit down and wait" hints that clients did not always care to sit down and wait. The description of Mrs Regan is of a smart, well-dressed woman. The reference to her legs recalls Marlowe's earlier impression of her, in Chap. 3 "I sat down on the edge of a deep soft chair and looked at Mrs. Regan. She was worth a stare. She was trouble. She was stretched out on a modernistic chaise-longue with her slippers off, so I stared at her legs in the sheerest silk stockings. They seemed to be arranged to stare at. They were visible to the knee and one of them well beyond. The knees were dimpled, not bony and sharp. The calves were beautiful, the ankles long and slim and with enough melodic line for a tone poem. She was tall and rangy and strong-looking." Again, author Chandler is establishing CG with a reader: the description of Mrs Regan renders her attractive and handsome but not sexy as in a romance—there is no reference to her face or bosom as one might find in a Mills and Boon novel. Consequently, the casting of Lauren Bacall to play Mrs Regan in the film of *The Big Sleep*[9] was perfect; Marilyn Monroe would have been entirely inappropriate (though MM could have been appropriately cast as Vivian's sister Carmen). These comments about the film of the novel are only tangentially relevant to the CG in (8): they are relevant as one reader's development on that part of the CG in (8) which established the appearance and character of Mrs Regan. H may legitimately increment what S presents as part of the CG in U beyond what is shared with S. This appears to directly conflict with a claim made by Clark and Carlson 1981: 328 that "the comprehension process must keep track of common ground, and its performance will be optimal if it limits its access to that common ground. Whether its design is actually optimal in this respect is a question that can only be answered empirically." As shown here, the empirical answer is surely negative, a conclusion that does not conflict with the view of Clark et al. 1983 that the search for *referents* is restricted to entities in CG. That is correct, as we shall see in the discussion in (9) below of the reference to Proust in (8).

The description of Mrs Regan's hat ($50 in 1938 would be close to $800 today) shows that Marlowe thought it hugely overpriced and perhaps ugly ("looked as if you could have made it with one hand out of a desk blotter"). Her reaction to his dingy office, "wrinkling her nose", is thereby put into context. Common ground with the reader thus far is along the following lines: a private detective is expected to have some kind of office; Marlowe's business is not flourishing and this is consistent with his office being small and dingy. Mrs Regan is wealthy and elegant; although slightly troubled ("She looked a little pale and strained"), she has earlier been described as "strong-looking" and is here designated "like a girl who

---

[9] Directed by Howard Hawkes, starring Humphrey Bogart and Lauren Bacall. 1946, Warner Bros. In the film, scripted by William Faulkner, Leigh Brackett and Jules Furthman, Vivian Regan was renamed Vivian Rutledge. Her sister Carmen was played by Martha Vickers.

could function under a strain". The banter with Marlowe should therefore be no surprise to a reader. Mrs Regan has been kept waiting and so chides Marlowe for being late into the office: "Well, you *do* get up [….] I was beginning to think perhaps you worked in bed, like Marcel Proust." This is a jibe and a display of culture which is apparently lost on Marlowe: a lack of CG between them. She rectifies it with a slightly dismissive identification of Proust as "a connoisseur in degenerates", and her put-down of Marlowe in "You wouldn't know him." This last is interesting: had she earlier supposed that Marlowe would not know of Proust, why did she bother to name him? We might (correctly) say that Vivian Regan is a pawn of the author and it is Raymond Chandler who names Proust to engineer this display of social disparity between Mrs Regan and Marlowe. However, it would also be in character for Mrs Regan to name Proust simply because of his reputation for working from his bed. When Marlowe reveals his ignorance of Proust it allows her to insult him such that "You wouldn't know him" has the force of *I should have known you wouldn't know who Proust was*. Marlowe's response is to tut—but at what? His own ignorance or the "connoisseur in degenerates"? It remains ambiguous. But his flippant "Come into my boudoir" in place of *Come into my office* is primed by the mention of "French" and, perhaps, the noun "degenerates" which is part of the CG with his addressee. He is understood by Mrs Regan (R) who accedes without comment. As I have said, the two final paragraphs demonstrate remembered common ground: R offers a rapprochement by admitting to have perhaps been rude; bringing it up voluntarily counts as an implicit apology based on social custom (part of CG). Marlowe's (M's) response accepts the attempt at reconciliation by admitting his own fault. Although this is fiction there is clearly a limitation on recursion. For convenient discussion, part of the conversation between R and M is repeated in (9).

9. R_1: … Marcel Proust [α]
   M_1: Who's he?
   R_2: A French writer, a connoisseur in degenerates.
   M_2: Tut, tut.

(9)R_1 pragmatically entails that (a) R believes α is known to M and (b) M recognizes that R assumes α is known to M. (9)M_1 pragmatically entails that (a) M does not know α and asks R to identify α and (b) R recognizes α is not known to M and M assumes R can identify α for M. (9)R_2 pragmatically entails that (a) R assumes that M knows nothing of α and so R identifies α for M in order to establish that α is known to M (despite the follow-up "You wouldn't know him"[10]); and (b) M is informed that α is a French writer who is a connoisseur in degenerates and M assumes (based on the cooperative principle) that R is speaking truthfully about α. (9)M_2 pragmatically entails that (a) M accepts (9)R_2 and comments on it (or his ignorance) disapprovingly and (b) R is informed that M accepts what she has said

---

[10] See the comments above.

about α and that M has commented upon it (or M's ignorance) disapprovingly. It is plain to see that there are no grounds here for positing runaway recursion in CG.

In the analysis of (9) there is a question. (10) generalizes over questions and other forms of request (though I shall not discuss these further here, see Allan 2006).

10. X asks Y φ pragmatically entails (a) X believes Y may be able to do φ and expects Y to accede or refuse (b) Y recognizes that X believes Y may be able to do φ and Y needs to decide whether to accede or refuse.

In (9) there is an instance of the situation generalized over in (6)ii., namely "If in saying φ to Y, X refers to α, this act of referring to α pragmatically entails that (a) X believes Y can identify α (knows who or what α is) and (b) Y recognizes that X believes Y can identify α. Typically, when Y cannot identify α, Y asks X for further information." We see this played out in lines (9)R_1 and (9)M_1. (9)M_1 and (9)R_2 operate via (10): (9)R_2 has Mrs Regan in the role of Y, acceding to the request in (9)M_1.

Now consider the very different, non-fiction, text in (11). In terms of the definition in (3) each sentence in (11) constitutes a P that becomes a part of F once processed by the reader; described below are other parts of F.

11. For twenty-three years now I've been floating rivers. Always downstream, the easy and natural way. The way Huck Finn and Jim did it, La Salle and Marquette, the mountain men, Major Powell, a few hundred others. ('Preliminary Notes' to *Down the River*, Abbey 1982: 1)

This is the opening paragraph of the first essay in the late American author Edward Abbey's *Down the River*. We expect knowledge of the book's title to be indicative of the book's content and to establish CG with the reader (H); but even without recourse to the book's title, the first sentence in (11) introduces the topic of river running and the author's experience of floating down rivers. The implication is that the canoes, kayaks, rafts, dories or inflatables used were powered by oars, paddles, or poles and not motors. This presumption is strengthened, though not confirmed, by the second sentence: the river's current was used as a power source. A basic knowledge of the dynamics of rivers and the running of rivers is invoked as CG with H (the reader); but H does not have to know much. The mention of 23 years suggests (based on CG with H of human life spans and contextually relevant experiences) that the author might be middle-aged (he was in fact 54) and so had considerable experience of river running. That fact is not relevant to the understanding of (11), but it is relevant to a reader's appreciation of the book that follows.

Abbey was somewhat egocentric in his assumptions about CG with the reader. Although he had lived in Europe and visited Australia, Abbey wrote primarily for

Americans with similar views to his own.[11] In part this is evident from the final sentence in (11). Many readers all over the Anglo world will recognize "Huck Finn", but the other people named are less likely to be familiar, especially to non-Americans. It is clear from context that all of them are people who floated downstream on rivers, a speculation confirmed to any reader who remembers the adventures of Huckleberry Finn and his Negro companion Jim as they floated down the Mississippi on a raft (Twain and Clemens 1884). Readers who recall the character of Huck Finn may also recollect his rebelling against oppression, searching for freedom and adventure, and revelling in life in the open—traits shared with Abbey himself and many of his books. Abbey probably anticipated this would be CG with a large number of his readers. This CG is coherent with the sense of adventure and exploration of the natural environment shared with the seventeenth century explorers René-Robert Cavelier, Sieur de La Salle, and Father Jacques Marquette SJ who, like Huck and Jim, travelled parts of the Mississippi; likewise with "the mountain men"[12]—nineteenth century fur trappers and traders in the American West—and with John Wesley Powell. Major Powell, who lost half an arm in the American Civil War, was celebrated for, and wrote eloquently about, his 1869 three-month river trip down the Green and Colorado rivers, making the first recorded passage through the Grand Canyon. A reader of (11) will typically know at least a few of these facts (F, in terms of (3)) and could readily ascertain the rest from encyclopaedias during Abbey's lifetime and, today, from the web. In other words, the interested reader (H) can readily increment CG with author Edward Abbey (S) should s/he wish to do so.

The text in (11) raises two important aspects of CG: the tendency for S to be egocentric and the importance of memory within CG. I don't doubt the presence of egocentricity as an antidote to cooperativeness, but I do believe that Keysar and Henly 2002; Keysar 2007; Kecskes and Zhang 2009 overstate the case for egocentrism in communicative discourse. Egocentrism is a function of what is severally salient to S and H whereas to seek common ground is an effortful process employing cognitive resources to incorporate beliefs about the knowledge and perspectives of other interlocutors. This view assumes conscious effort on S's or H's part but I would predict that, given the near constant exposure to language interchange during the waking hours of most human beings, under most circumstances S and H automatically assume that for S to get a message across in U and for H to understand U one has to put oneself into the interlocutor's shoes; consequently, this is what we automatically do in language exchange. For instance, it enables us to correctly interpret utterances in unfamiliar accents through a sort of analysis-by-synthesis: "It seems as if listeners sometimes perceive an utterance by reference to their own motor activities. When we listen to speech, we may be considering, in some way, what we would have to do in order to make similar

---

[11] "Speaking for myself, I write to entertain my friends and to exasperate our enemies" ('A Writer's Credo' Abbey 1988: 177). See also Solheim and Levin 1989.

[12] . http://www.mtmen.org/

sounds" (Ladefoged 1982: 104). Linguistic communication in general is a matter of putting oneself into the interlocutor's shoes and, because this behaviour is the norm, it very quickly becomes automatic except perhaps in those with autism spectrum disorders, or those who are severely narcissistic or very deeply depressed. Otherwise neither S nor H needs to consciously accommodate themselves to the needs of an interlocutor; it is automatic and takes no noticeable processing effort. In the words of Horton 2008: 202: "automatic commonality assessment provides one possible basis upon which language users may generate inferences about common ground."

Nonetheless, it is necessarily the case that what S utters is based on S's own knowledge and perspectives (egocentricity) and these may not match H's even though S is desirous of communicating with H. We find this illustrated in (11). Even so, S will make the effort to communicate effectively with H, as we see in (12) and (13).

12. Everything is disorganized that's why the lights are constantly going out and the transportation is just eh- but anyway he Don Ward lives through all this. (Horton and Gerrig 2005: 29)

13. A: Really that's what Lawrence and one of his friends that's what he did when he was in the service. Because he just uh you know the eh the you know how the tops of the tanks have those kind of ball bearing things
    B: mhm
    A: He just made sure that those ran right. (Horton and Gerrig 2005: 30)

In (12) S utters "he" but then immediately clarifies who s/he is referring to, naming "Don Ward" who is presumably in CG with H. In (13) A makes the effort to check that B has understood his reference to "those kind of ball bearing things" on the tops of tanks.

Another aspect of (11), especially its third sentence, "The way Huck Finn and Jim did it, La Salle and Marquette, the mountain men, Major Powell, a few hundred others", is the invocation of memory. Lewis 1979: 346 writes of mental representations of a conversational scoreboard and obviously S's apprehension of CG with H depends on S's long-term and short-term memory and, equally, H's understanding of U typically relies on H's short-term and long-term memory (Horton and Gerrig 2005). This is very evident in (14) and (15) below. In (14) S can't remember if the information is in CG already but thinks it isn't; and in (15) A has forgotten that the information is already in CG and can be recalled by B.

14. I got you. Yeah I've got another buddy who, uh, is a marine pilot. I'm trying to think if you had ever met this guy. I don't think so. (Horton and Gerrig 2005: 14)

15. A: My nephew's name is Jeff McDougal
    B: yeah
    A: He made his vows in the Jesuits a year ago.
    B: You said that.
    A: I I couldn't remember if I did or not. (Horton and Gerrig 2005: 19)

As Horton says:

[T]he claim is simply that conversational phenomena like audience design can, in many circumstances, be mediated through domain-general memory processes. Indeed, there are many situations in which relatively strategic considerations of commonality would be expected to occur, either because of the need to keep track of what information is shared or not, or because feedback from the partner triggers the need for possible monitoring and error correction. (Horton 2008: 217)

CG does rely on memory, including memories of schemata, frames, scenarios and scripts (Bartlett 1932; Mazzone 2011; Minsky 1977; Fillmore 1982; Sanford and Garrod 1981; Schank and Abelson 1977; Schank 1984), but CG is nonetheless a valid concept within the analysis of communication because there is more to CG than memory alone.

For the last of these textual analyses consider CG within a lengthy spoken text; this time a kind of horror story.

16. (A):208 On you Oh did you hear about um X and Y's train train journey back to
    (A):210 Far out man You have not such it's not really action packed but just the most horrific train journey I have ever heard
    (A):213 It was unbelievable They went from Ingham to Brisbane and it took them I think it took them twelve hours longer than it should have'cos
    (B):217 No
    (B):218 And that's a pretty long way anyway
    (A):220 Yeah and um oh the they were going really slowly and it
    (A):221 Oh the train usually left slowly on a long distance anyway
    (A):222 And I don't know where they were going past
    (A):223 I can't remember
    (A):224 This was on the news You would've seen it on the news
    (A):225 You know those two kids that got run over by the train
    (A):226 You didn't see it on the news There was a little two year old boy who wandered out of his out of his backyard onto the train tracks which was behind his house and he was with a friend and they he got hit by a train and he got killed
    (A):231 They were on the train and they're they were over in the back carriage and they they said that um they felt two bumps and the train stopped and they were whingeing because the train stopped for about an hour and then and a guy this woman came in and she goes
    (A):237 It's not funny but it's just the way that they told
    (A):238 A woman came in from the kitchen and she said we're serving dinner in the dining room car now and we've just hit a two year boy
    (A):241 That's how she said it
    (A):242 And Y said she started laughing'cos she thought that she said we're serving dinner in the dining room and we're having a two year old boy
    (A):245 Mmm oh god it's so oh she said it was just gross and that they had to take the carriage the front carriage thing off and and take it back to wherever

and get a new train or something
(B):250
(B):251 Why
(A):252 Because it's something superstitious They reckon or like it was pretty messed up
(A):253 just like like blood all over the the engine and stuff so they had to move it anyway
(A):255 Then they said they were going along and some guy had a heart attack and was lying in between carriages and people and this guy got up and tried to kick him out of the way and I went nah
(A):259 Then theyum and then the traum someone blew up the tracks
(A):261 Like there was a derailment and the train's tracks blew up or something so they had to stop the train for 6 h and replace the tracks
(A):263 So it took them forever to get back
(B):265 Mmm
(B):266 Oh no
(A):267 They were just spewing I would've said oh god thanks
(A):268 I think I'll take a bus
(B):270 Nuh Flying is the best way
(ICE-AUS Corpus, S1A-036)

Consider some aspects of CG in this story. First there is the opening gambit in (A):208 "Oh did you hear about um X and Y's train train journey back to …": A is checking whether this story is news to B (whether it is already among the F that is part of CG, cf. (3)); usually it is a rhetorical question. Here it headlines the story of two people, X and Y, known to both A and B who had a horrific train journey. X and Y are in the CG shared by A and B and are introduced into ours. A train journey script is CG for A, B, and us too, as overhearers. The initial "Far out man You have not such it's not really action packed" of (A):210 promises a story so unusual it is well worth the telling; this is an "evaluation" in terms of Labov and Waletzky 1967. The remainder of (A):210–213 functions as an abstract (Labov and Waletzky 1967) or synopsis: "just the most horrific train journey I have ever heard … it took them twelve hours longer than it should have". This grounds (puts into CG) the expectation of a horror story. (B):218 shows some shared knowledge that a train journey from Ingham to Brisbane takes a long time, based, presumably, on a vague idea that this journey of about 1500 kilometres normally takes nearly 26 h. (A):222–223 admits that A cannot remember the exact location of what turns out to be a gruesome accident so he appeals to B's memory in (A):224. I'll offer a more detailed analysis of (A):224–226.

(A):224 "This was on the news You would've seen it on the news" (in terms of (3), P recalling something in F) pragmatically entails that (a) A believes B has probably seen a (TV) news item about the train accident he is recounting and (b) B will assume that A has seen a (TV) news item involving the train on which X and Y were travelling which will somehow throw light on the cause of the train's extended travel time and that A believes B will have seen the same or a similar

news item. (A):225 "You know those two kids that got run over by the train" pragmatically entails that (a) A purportedly reminds B, and thereby informs B should B be unaware of it, of the topic of the news item, namely that two young children got run over by a train[13]; furthermore A knows this train to have been the one on which X and Y were travelling. (b) B is informed that two young children got run over by a train which, under the cooperative principle of relation is most probably the train on which X and Y are travelling and A is implying this. (A):226 "You didn't see it on the news …" pragmatically entails that (a) B has indicated to A (presumably kinesically) that B did not see the news item A has been speaking of and A is confirming this with B. (b) B is informed that A believes that B did not see the news item A has been speaking of and B should confirm the truth of this. Presumably B once again did this kinesically, since there is no verbal record of it, and A proceeds to describe what happened. "There was a little two year old boy who wandered out of his out of his backyard onto the train tracks which was behind his house and he was with a friend and they he got hit by a train and he got killed". It is clear from CG concerning "the most horrific train journey" that the killer train must be the one on which X and Y were travelling. There is also some reliance on knowledge (memory) of other stories in which two year old children wander from home into danger.

(A):231 adds some ghastly detail, "they felt two bumps and the train stopped", where we are to believe the bumps are caused by the carriage wheels passing over the children's bodies (or, as it turns out, one child's body) but that X and Y did not know this at the time. They were complaining at the train being "stopped for about an hour"—which again evokes CG (memory) of typical passenger behaviour when a train is held up for a while. Then the announcement of the accident is introduced, but delayed by the parenthetical apology for a funny juxtaposition of the mundane and the horrific in (A):237–241. The report of Y's laughter in (A):242 is almost light relief after the shock of (A):237–241: we might speculate (perhaps along with A and B) that Y misheard because she couldn't immediately believe what had been said. The removal of the engine and front carriage because they were covered in blood, (A):245–(A):253, appears consistent with what might be required following such an accident; i.e. knowledge in common ground about vehicle accidents renders this part of the story plausible. Two additional horrible events follow: (A):255 there is a passenger having a heart-attack whom a fellow passenger tried to kick out of the way; and (A):259–261 there is the rather confused and confusing tale of the blown up tracks and derailment. The story proper ends with (A):263 "So it took them forever to get back". B shows appropriate sympathy in (B):265–266. Then A comments on their friends' story "They were just spewing I would've said oh god thanks/I think I'll take a bus" and B joins in with a preference for yet another form of transport, flying ((A):267–(B):270). The relevant CG here is that if one form of transport causes problems, consider an alternative.

---

[13] Prince's term "informative-presupposition" seems appropriate here.

Discussion of (16) raises the question of overhearers and CG. I have been discussing CG in (16) as an overhearer and not a participant; I am nonetheless confident that my account of supposed common ground is correct and leave it to readers of this essay to agree or disagree. Where an overhearer differs from the participants in (16) is in knowing something about A, B, X, and Y. We know from the text that the action takes place in Queensland Australia and that Y is female, but we cannot be certain what sex/gender the others are, never mind any other details. This situation with respect to overhearers is, I believe, typical: overhearers customarily share less common ground with S than H does, but they can usually understand most of what is said.

5 Conclusion

In this essay I have sought to establish the means by which S identifies the supposed common ground with H. H also makes assumptions about the CG with S based on H's assessment of U in the context of utterance and of S as a person. S needs to be satisfied that H understands U well enough for S's communicative purpose to, in S's judgment, succeed. S must be capable of presenting different material to different audiences according to the task to which U is put in such a way as to hold audience interest. The initial assumption is that, normally, common ground is quite readily identified by S and recognized by H. When it is not, H typically requests clarification (where circumstances allow). S and H may come to feel they are speaking at cross-purposes and consequently seek to re-assess the common ground. At worst S fails to communicate the intended message and, because of a degree of incomprehension, H may be bored or feel insulted by S's use of language in U.

The notion of common ground necessitates a community, K, that observes social norms such as that S and H are mutually aware that, normally, their interlocutor is an intelligent and aware being. In other words each interlocutor believes of him/herself and fellow interlocutors that they are intelligent and aware beings and believes of fellow interlocutors that they too believe themselves and fellow interlocutors (including him/herself) to be intelligent and aware beings. There is a concomitant assumption of communicative competence: the knowledge and application of how and when to use utterances appropriately that combines with grammatical knowledge (of semantics, syntax, morphology, phonology) in the production of utterances in order to create a coherent text comprehensible to its intended audience. Normal use of language goes unremarked, but abnormal use may indicate a person living with autism, schizophrenia, or the like. Age, social status, educational level, and cultural background, etc. of both self and other will affect the assessment of an interlocutor's use of language and probable range of comprehension. When a member of community K applies knowledge of a set of facts F in order to interpret P, a state of affairs or something said, s/he can presume that others in K will also be able to apply knowledge of F in order to interpret P.

The existence of F, P, and the application of knowledge of F to interpreting P is CG for members of the community K. Once attended to, P becomes part of F, incrementing the common ground.

Common ground is dynamic. In conversation it is constantly developing and as themes change so does CG. We see this everywhere. For instance at the end of (16) there is a new theme introduced in (A):267–268, which B develops in (B):270:

(A):267 They were just spewing I would've said oh god thanks
(A):268 I think I'll take a bus
(B):270 Nuh Flying is the best way

Travellers X and Y were very annoyed and upset by their horrific journey ("just spewing") but (A):268 suggests that if A had endured such a journey on the train s/he would in future "take a bus", a theme of alternative means of transport which B develops by dismissing A's choice with a preference for air travel. For another example, reconsider part of (8), reproduced here as (17).

17. 'Well, you *do* get up,' she said, wrinkling her nose at the faded red settee, the two odd semi-easy chairs, the net curtains that needed laundering and the boy's size library table with the venerable magazines on it to give the place a professional touch. 'I was beginning to think perhaps you worked in bed, like Marcel Proust.'
    'Who's he?' I put a cigarette in my mouth and stared at her. She looked a little pale and strained, but she looked like a girl who could function under a strain. 'A French writer, a connoisseur in degenerates. You wouldn't know him.'
    'Tut, tut,' I said. 'Come into my boudoir.'
    She stood up and said 'We didn't get along very well yesterday. Perhaps I was rude.'
    'We were both rude,' I said. I unlocked the communicating door and held it for her.
    (*The Big Sleep*, Chandler 1939, Chap. 11)

The dingy office theme is developed for readers in the description of the outer-office furnishings. The jibe about Marlowe possibly working in bed like Marcel Proust is introduced as a tongue-in-cheek explanation for the detective's late arrival. It founders because Marlowe does not know who Mrs Regan is referring to and CG is developed as she explains and he accepts the explanation. There is then a change of theme when Marlowe invites her into his inner-office and another when Mrs Regan introduces the conciliatory admission of having been rude on their previous encounter; this aspect of CG is politely acknowledged by Marlowe as he holds the door for her to enter the inner-office.

Because of the ubiquity of language interaction among human beings it is most probable that our cognitive and social behaviour in language exchange is largely automatic and rarely consciously and deliberately planned. S and H automatically assume that the optimal means for S to get a message across in U and, concomitantly, for H to understand U is for each interlocutor to put themself into the other's shoes. Hence, even though it is necessarily the case that what S utters is based on S's own knowledge and perspectives there is normally no effortful,

cognitively costly process of accommodation to the knowledge and perspectives of the interlocutor.

I have described what common ground is by recourse to what others have said on the matter, by introspection, and by analytical explication of some longer instances of real language data, both written and spoken texts. In §2 I reviewed some near synonyms of *common ground*, including *common knowledge, mutual knowledge*, shared knowledge, assumed familiarity*, and *presumed background information*. All of these are to some extent relevant to the defining of CG. I drew attention to the significant flaw carried over from Schiffer's definition of mutual knowledge* into Stalnaker's definition of common ground: "It is common ground that φ in a group if all members *accept* (for the purpose of the conversation) that φ, and all *believe* that all accept that φ, and all *believe* that all *believe* that all accept that φ, etc." (Stalnaker 2002: 716, quoted earlier). The runaway recursion would necessitate infinite processing on the part of each of S and H. This flaw has been accepted and repeated by many since, ignoring Prince's caveat to consider what ordinary, nonclairvoyant humans do when they interact verbally (Prince 1981: 232). Clark 1996: 95 attempted to circumvent it (see (2) above) but his definition includes a clause that calls itself, thus creating an endless loop, which would also dictate infinite processing on the part of each of S and H. Some more realistic description was needed.

On the basis of analysing real language data, Lee 2001, adopting a proposal of Brown 1995, reduces the number of steps in the recursive process to three. Bach and Harnish 1979: 267ff had proposed a similar limitation on the basis of a proposal by Scheff 1967 to just three levels of agreement that are required to achieve consensus. The Bach and Harnish 1979: 269 proposal for mutual knowledge ceases at level (iii):

Members of group *G* believe

1. that *p*,
2. that the members of *G* believe that *p*, and
3. that the members of *G* believe that the members of *G* believe that *p*.

They do not say why the sequence should not continue with *iv:

4. *that the members of *G* believe that the members of *G* believe that the members of *G* believe that *p* … etc.

The explanation is that there is no reason to do so.

More directly relevant to the definition of CG is the proposal of Garfinkel 1964: 33 (which he attributes to Alfred Schutz 1899–1959): "the person assumes, assumes the other person assumes as well, and assumes that as he assumes it of the other person the other person assumes it of him". This is essentially similar to part of my definition of common ground in (3), viz.: (a) every member, or almost every member, of community K knows or believes some fact or set of facts F; and (b) a member is presumed to know or believe F by (almost) every other member of K; and (c) a member of K knows that both (a) and (b) are true.

A limitation of the analysis of CG in Lee 2001 is that he does not fully explore the commonality of CG, but looks at it from the point of view of just one participant. On three occasions in my analyses of CG in §4 I rectified this by examining the presumed beliefs of each participant in the uttering of U by S and in the understanding of U by H. On each occasion the analysis was specific to the text at hand, but the general picture is that captured in (6) and (10), collated in (18).

18. (1) X says φ to Y pragmatically entails (a) X believes that φ and (b) Y has some reason to believe that X believes that φ.
    (2) If in saying φ to Y, X refers to α, this act of referring to α pragmatically entails that (a) X believes Y can identify α (knows who or what α is) and (b) Y recognizes that X believes Y can identify α. Typically, when Y cannot identify α, Y asks X for further information.
    (3) X asks Y φ pragmatically entails (a) X believes Y may be able to do φ and expects Y to accede or refuse (b) Y recognizes that X believes Y may be able to do φ and Y needs to decide whether to accede or refuse.

It seems probable that other illocutionary types may give rise to additional patterns corresponding to the preconditions of those illocutions, but I will leave examining them for another day.

References

Abbey, E. (1982). *Down the river*. New York: E.P. Dutton.
Abbey, E. (1988). *One life at a time, please*. New York: Henry Holt.
Abbott, B. (2008). Presuppositions and common ground. *Linguistics and Philosophy, 21*, 523–538.
Allan, K. (1986). *Linguistic meaning* (Vol. 2). London: Routledge and Kegan Paul. (reprint edn, Beijing: World Publishing Corporation 1991).
Allan, K. (2001). *Natural language semantics*. Oxford: Blackwell.
Allan, K. (2003). Linguistic metatheory. *Language Sciences 25*, 533–560. Doi: http://dx.doi.org/10.1016/S0388-0001(03)00032-9.
Allan, K. (2006). Clause-type, primary illocution, and mood-like operators in English. *Language Sciences, 28*, 1–50.
Allan, K. (2010). *The western classical tradition in linguistics (second expanded edition)*. London: Equinox. (First edn 2007).
Allan, K. (2013). Referring to 'what counts as the referent': A view from linguistics. In C. Alessandro, L. P. Franco, & C. Marco (Eds.), *Perspectives on linguistic pragmatics*. Cham: Springer.
Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon Press.
Bach, K. (2012). Saying, meaning, and implicating. In Keith Allan & K. M. Jaszczolt (Eds.), *Cambridge handbook of pragmatics*. Cambridge: Cambridge University Press.
Bach, K., & Harnish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge: MIT Press.

Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University Press.

Birner, B.J. & Ward, G. (1994). Uniqueness, familiarity, and the definite article in English. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society* (pp. 93–102), Berkeley CA: Berkeley Linguistics Society.

Brown, G. (1995). *Speakers, listeners and communication: Explorations in discourse analysis*. Cambridge: Cambridge University Press.

Chandler, R. (1939). *The big sleep*. New York: Alfred A. Knopf.

Chomsky, N. (1975). *Reflections on language*. New York: Pantheon Books.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Carlson, T. B. (1981). Context for comprehension. In L. John & B. Alan (Eds.), *Attention and performance IX* (pp. 313–330). NJ: Erlbaum.

Clark, H. H., Schreuder, R., & Butterick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior, 22*, 245–258.

Colston, H. L. (2008). A new look at common ground: memory, egocentrism, and joint meaning. In K. Istvan & J. L. Mey (Eds.), *Intention, common ground and the egocentric speaker-hearer* (pp. 151–187). New York: Mouton de Gruyter.

Dunbar, R. I. M. (1996). *Grooming, gossip and the evolution of language*. London: Faber and Faber.

Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.

Enfield, N. J. (2008). Common ground as a resource for social affiliation. In K. Istvan & J. L. Mey (Eds.), *Intention, common ground and the egocentric speaker-hearer* (pp. 223–254). New York: Mouton de Gruyter.

Fillmore, C. J. (1982). Frame semantics. In P. Kiparsky & S. Yang (Eds.), *Linguistics in the morning calm* (pp. 111–138). Seoul: Hanshin.

Garfinkel, H. (1964). Studies of the routine rounds of everyday activities. *Social Problems, 11*, 225–250.

Gilbert, M. (1987). Modelling collective belief. *Synthese, 73*, 185–204.

Gilbert, M. (1989). *On social facts*. Princeton: Princeton University Press.

Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Grice, H.P. (1957). Meaning. *Philosophical Review* 66: 377–388. Reprinted in Grice, H.P. (1989). *Studies in the way of words*. Cambridge MA: Harvard University Press, pp. 213–223.

Grice, H. P. (1968). Utterer's meaning, sentence meaning, and word-meaning. *Foundations of language 4*, 225–242. Reprinted in Grice, H.P. (1989) *Studies in the way of words*. MA: Harvard University Press, pp. 117–137.

Grice, H.P. (1969). Utterer's meaning and intentions. *Philosophical Review 78*, 147–177. Reprinted in Grice, H.P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press, pp. 86–116.

Grice, H.P. (1981). Presupposition and conversational implicature. In C. Peter (ed.), *Radical pragmatics*. New York: Academic Press, pp. 183–198. Reprinted in Grice, H.P. (1989) *Studies in the way of words*. Cambridge: Harvard University Press, pp. 269–282.

Grice, H.P. (1982). Meaning revisited. In N.V. Smith (ed.) *Mutual knowledge*. New York: Academic Press, pp. 223–243. Reprinted in Grice, H.P. (1989) *Studies in the way of words*. Cambridge: Harvard University Press, pp. 283–303.

Grice, H. P. (1989). Indicative conditionals. In H. P. Grice (Ed.), *Studies in the way of words* (pp. 58–85). Cambridge: Harvard University Press.

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.

Holtgraves, T. M. (2002). *Language as social action: Social psychology and language use*. Mahwah NJ: Lawrence Erlbaum.

Horton, W. S. (2008). A memory-based approach to common ground and audience design. In K. Istvan & J. L. Mey (Eds.), *Intention, common ground, and the egocentric speaker-hearer* (pp. 189–222). New York: Mouton de Gruyter.

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processses in language production. *Discourse Processes, 40*, 1–35.

Karttunen, L., & Peters, F. S. (1979). Conventional implicature. In C.-K. Oh & D. A. Dinneen (Eds.), *Syntax and semantics 11: Presupposition* (pp. 1–56). New York: Academic Press.

Kecskes, I., & Zhang, F. (2009). Activating, seeking, and creating common ground: A socio-cognitive approach. *Pragmatics and Cognition, 17*, 331–355.

Kecskes, I., & Zhang, F. (2013). On the dynamic relations between common ground and presupposition. In A. Capone, F. L. Piparo, & M. Carapezza (Eds.), *Perspectives on linguistic pragmatics* (pp. 375–396). Cham: Springer.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics, 4*, 71–84.

Keysar, B., & Henly, A. S. (2002). Speakers' overestimation of their effectiveness. *Psychological Science, 13*, 207–212.

Labov, W. & Waletzky, J. J. (1967). Narrative analysis: Oral versions of personal experience. In H. June (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: American Ethnological Society.

Ladefoged, P. (1982). *A course in phonetics* (2nd ed.). New York: Harcourt Brace Jovanovich.

Lee, B. P. H. (2001). Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics, 33*, 21–44.

Lewis, D. (1969). *Convention*. Cambridge: Harvard University Press.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic, 8*, 339–359.

Mazzone, M. (2011). Schemata and associative processes in pragmatics. *Journal of Pragmatics, 43*, 2148–2159.

Minsky, M. (1977). Frame-system theory. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: readings in cognitive science* (pp. 355–376). Cambridge: Cambridge University Press.

Phillips, S. U. (1972). Participant structures and commmunicative competence: Warm Springs children in community and classroom. In C. B. Cazden, V. P. John, & D. Hymes (Eds.), *Functions of language in the classroom* (pp. 370–394). New York: Columbia Teachers Press.

Prince, E. (1978). A comparison of *wh*-clefts and *it*-clefts in discourse. *Language, 54*, 883–906.

Prince, E. (1981). Toward a taxonomy of given-new information. In C. Peter (Ed.), *Radical pragmatics* (pp. 223–256). New York: Academic Press.

Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language*. Chichester: Wiley.

Schank, R. (1984). *The cognitive computer*. Reading: Addison-Wesley.

Schank, R., & Abelson, R. C. (1977). *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. NJ: Lawrence Erlbaum.

Scheff, T. J. (1967). Toward a scoiological model of consensus. *American Sociological Review, 32*, 32–46.

Schiffer, S. R. (1972). *Meaning*. Oxford: Clarendon Press.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. London: Cambridge University Press.

Solheim, D., & Levin, R. (1989). The *Bloomsbury Review* interview. In J. Hepworth & G. McNamee (Eds.), *Resist much, obey little: some notes on Edward Abbey* (pp. 89–104). Tucson: Harbinger House.

Stalnaker, R. C. (1973). Presupposition. *Journal of Philosophical Logic, 2*, 77–96.

Stalnaker, R. C. (1974). Pragmatic presupposition. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and pilosophy*. New York: New York University Press.

Stalnaker, R. C. (1978). Assertion. In C. Peter (Ed.), *Syntax and semantics 9: Pragmatics* (pp. 315–332). New York: Academic Press.

Stalnaker, R. C. (2002). Common ground. *Linguistics and Philosophy, 25*, 701–721.

Twain, M., & Clemens, S. (1884). *The adventures of Huckleberry Finn: (Tom Sawyer's comrade): Scene, the Mississippi Valley: Time, forty to fifty years ago*. London: Chatto & Windus.

Layered Discourse Representation Theory

Bart Geurts and Emar Maier

Abstract Layered Discourse Representation Theory (LDRT) is a general framework for representing linguistic content. Different types of content (e.g. asserted, presupposed, or implicated information) are separated by putting them on different layers, all of which have a model-theoretic interpretation, although not all layers are interpreted uniformly. It is shown how LDRT solves so-called 'binding problems', which tend to arise whenever different kinds of content are separated too strictly. The power of the framework is further illustrated by showing how various kinds of contextual information may be accommodated.

1 Introduction

The information conveyed by any utterance is a mixed bag. Utterances carry content about the world as it is according to the speaker, but also about speakers' attitudes, the way they speak, what has been said before, and so on. There are many kinds of information that are conveyed by way of language, and differences in kind correlate with differences in status. Presupposed information exhibits a distinctive projection behaviour; conversational implicatures are cancellable in a way that asserted information is not; in French or German, a pronoun's grammatical gender may help to determine a referent, but is otherwise truth-conditionally inert; and so on.

B. Geurts (✉)
Department of Philosophy, University of Nijmegen, Nijmegen, Netherlands
e-mail: brtgrts@gmail.com

E. Maier
Department of Philosophy, University of Groningen, Groningen, Netherlands
e-mail: emar.maier@gmail.com

Interpreting utterances is as much a matter of integrating these various kinds of information as of keeping them apart. This much is uncontroversial. As far as we are aware, however, no attempts have been made thus far to devise a fully general framework within which processes of information integration can be modeled. There are partial theories, to be sure. For example, there are quite a few well-developed analyses of the interaction between presupposed and non-presupposed content. But to the best of our knowledge the problem of information integration as such has not been addressed before. So that is what this chapter is about: a general framework for representing and integrating all and sundry kinds of information that can be conveyed by linguistic means. This may seem like a grandiose project, and perhaps it is, but it is less ambitious than one might think. Our aim in this chapter is to develop a framework for *representing* different kinds of linguistic and para-linguistic information. How this information is processed is a different matter altogether, and not our main concern in the following.

2 Information Integration

In order to explain what we mean by information integration, we will discuss a few concrete cases. It will be seen that our examples are quite diverse, but this is to be expected in view of the broad aim of this chapter.

2.1 Presupposition

Our first example concerns the representation of presuppositions. In Discourse Representation Theory (DRT; Kamp 1981; Kamp and Reyle 1993), presuppositions are treated on a par with anaphoric expressions. Presuppositions prefer to link up to an antecedent, and if no suitable antecedent is available, they are interpreted by way of accommodation (van der Sandt 1992, Geurts 1999). Here is an example:

(1) Perhaps White met the Chinese Empress today.

The initial representation of this sentence is as follows:

(2) [y: White(y), ◇[x:Chinese-Empress(x), meet(y, x)]]

The underlining in this Discourse Representation Structure (DRS) reflects the fact that the definite NP 'the Chinese Empress' triggers the presupposition that there is a Chinese Empress. Assuming that this presupposition doesn't have a suitable antecedent, it may be construed by way of accommodation (provided the hearer is prepared to accept that there is a Chinese Empress), which means that the presupposition is added to the principal DRS, yielding the following representation:

(3) [x, y: Chinese-Empress(x), White(y), ◇[: meet(y, x)]]

This correctly captures what is expressed by (1), viz. that there is a Chinese Empress, and that White may have met her. Note that in the final representation of (1) the distinction between presupposed and asserted information is obliterated.

For many purposes this is fine, for we mainly need that distinction in order to account for the fact that presupposed material is processed in its own special way. But as it turns out, the distinction between presupposed and asserted information remains active after a sentence has been processed and the presupposition accommodated. To illustrate this, consider what would happen should another speaker object against (1) as follows:

(4) No, he had an encounter with the Japanese President.

Intuitively, this response only corrects what (1) asserts; the accommodated presupposition that there is a Chinese Empress remains unscathed. There are also ways to achieve the opposite, to deny the accommodated presupposition but leave the asserted content (Maier and der Sandt 2003; von Fintel 2004):

(5) Hey, wait a minute, China doesn't have an Empress!

It is not a good idea, apparently, to discard the division between presupposed and non-presupposed material once the mechanism of presupposition projection has performed its duty.

The moral of these observations is obvious and quite independent of the theory of presupposition we happen to prefer. It is simply that presuppositions will have to be separated from other types of information, because they have a special status: presuppositions are processed in their own way and once they have been accommodated they continue to be treated differently, as is shown by (5).[1]

2.2 Implicatures

What has been just said about presuppositions holds good for implicatures, too. By way of example, consider what is generally regarded a 'scalar implicature':

(6) The porridge is warm.

An utterance of this sentence presupposes that there is porridge, it asserts that the porridge is warm, and it implicates that the porridge is not hot; so the lexical meaning of 'warm' does not by itself rule out that the porridge is hot. According to this analysis, the information that the porridge is warm is of a different kind than

---

[1] Actually, there are two issues here. First, presuppositional material has to be separated from other types of content in order for the projection mechanism to perform its function. Secondly, once the presupposition has been processed, it must remain separated, as we have just argued. In the following, we confine our attention to the second issue. Whether or not presuppositions in preliminary DRSs must be interpreted, too, is a different matter, which we will not take a stance on here.

the information that the porridge is not hot, and one of the stock-in-trade arguments in favour of the distinction is that the implicature is cancellable in a way the assertion is not:

(7) a. The porridge is warm. As a matter of fact, it is hot.
    b. ?The porridge is warm. As a matter of fact, it is cold.

(7a) shows that the implicature is cancellable, and the oddness of (7b) suggests rather strongly that asserted information is more robust.

The upshot of these observations is analogous to that of the presuppositional case. We need to separate implicated information from other information conveyed by an utterance, and it will not do to discriminate between presupposition, assertion, and implicature only for the duration of sentence processing (as in Gazdar 1979, for example); for the subsequent discourse may need these distinctions, too.

2.3 Non-Literal Meaning

The heading of this rubric is somewhat tentative. What we have in mind are such phenomena as metaphor, metonymy, irony, and so on: non-truth-conditional content that is clearly part of the speaker's message, but may be at odds with its literal meaning. Even if such phenomena are to be treated in terms of conversational implicature, we prefer to distinguish them from run-of-the-mill cases of implicature, which merely add to the literal meaning of an utterance. Especially stark cases in point are irony and sarcasm. Suppose a connoisseur of modern art volunteers (8), pointing at what is obviously a fumbled attempt at self-expression:

(8) That is a beautiful painting.

Under the circumstances, this statement is probably intended to convey the opposite from what it literally says.

Another example to bring out the need for information segregation is sentence (9), as said by a father to his 15-year-old son:

(9) Someone used my after shave this morning.

The use of an indefinite would normally implicate that the subject is thought not to be present in the context of discourse, but in this particular case the utterance may be understood as implying that the addressee is the culprit, and if it is, the implicature is cancelled.

It is, mildly put, something of a mystery how such 'double meanings' are computed, and we don't have anything new to offer in this regard. However, we do have a proposal as to how different levels of meanings can be represented in such a way that some bits of information are shared (e.g., the reference of the pronoun is shared between the literal and the non-literal meaning of the sentence), while others are segregated.

2.4 The Former and the Latter

Consider the following example:

(10) If a beggar meets a bishop, then the latter will bless the former.

On the face of it, anaphoric devices like 'the former' or 'the latter' do not seem particularly troublesome. On reflection, however, they add an interesting wrinkle to the problems posed by definite NPs. As the wrinkle will appear, in some way or other, no matter what our theoretical predilections concerning definites are, we will follow our own. According to the DRT treatment of definite NPs that we favour, definites are presuppositional expressions, which is to say that they prefer to link up to a contextually given antecedent. In this respect, everything is fine in the present example, since 'the latter' as well as 'the former' have suitable antecedents: the former refers back to 'a bishop', the latter to 'a beggar'. However, problems begin to emerge when we ask ourselves how exactly these expressions manage to link up to their antecedents. To see this, consider how the story about 'the prelate' in (11) would go:

(11) If a beggar meets a bishop, then the prelate will bless him.

In this sentence, 'the prelate' establishes an anaphoric link in much the same way as 'the latter' does in (10). In this respect there is little difference between the two expressions. However, the descriptive content of 'the prelate' is very different from that of 'the latter'—so different, in fact, that some people would say the information contained in 'the prelate' is part of the sentence's truth-conditional content, whereas nobody would want to claim that the descriptive content of 'the latter' enters into the truth conditions of (10).

So the wrinkle is this. Since they appear to be just a special case of presuppositional (or anaphoric) expressions, we would like to analyse 'the former' and 'the latter' as being on a par with any other definite NP, except of course that they constrain the process of interpretation by referring not to *what* has been said but *how* it was said. Qua presuppositional expression, 'the latter' presents its referent as given in the same way 'the prelate' does; it is just that we have two rather different modes of givenness, so to speak. The problem is, therefore, how we can distinguish two very different kinds of information—about the discourse and about the world—and have a uniform account of definites at the same time.

Note, incidentally, that in some cases ordinary pronouns may be used in the same way as English 'the former/latter', so a strict distinction between 'referring' and 'formal' definites becomes even less desirable:

(12) Am Ende besteht ein wesenhafter Unterschied zwischen [dem Erfassen des Ganzen des Seienden an sich]$_i$ und [dem Sichbefinden inmitten des Seienden im Ganzen]$_j$. Jenes$_i$ ist grundsätzlich unmöglich. Dieses$_j$ geschieht ständig in unserem Dasein. (Martin Heidegger, *Was ist Metaphysik?*)

We will not endeavour to render this passage in colloquial English. Suffice it to say that it illustrates how in German the distal and proximal demonstrative pronouns 'jenes' and 'dieses' are used precisely as 'the former' and 'the latter' would be used in English. The following quote from Somerset Maugham shows that English pronouns have the same meta-linguistic use, too, although this may be a more isolated example:

(13)   For it was clear that the two were irreconcilable, the state and the individual conscious of himself. *That* uses the individual for its own ends, trampling upon him if he thwarts it, rewarding him with medals, pensions, honours, when he serves it faithfully; *this*, strong only in his independence, threads his way through the state, for convenience' sake, paying in money or service for certain benefits, but with no sense of obligation; and, indifferent to the rewards, asks only to be left alone. (W. Somerset Maugham, *Of Human Bondage*)

2.5 Grammatical Gender

Up to a point, grammatical-gender pronouns function not unlike 'the latter' and 'the former', as witness the following example from German:

(14)   Braun hat {einen Wagen/ein Auto} gekauft. {Er/Es} ist grün.
        Braun has bought {a car$_{neut}$/a car$_{masc}$}. {Pro$_{neut}$/Pro$_{masc}$} is green.

On pains of unintelligibility, the pronoun in the second sentence has to agree in gender with its antecedent in the first, and although the term 'grammatical gender' may suggest otherwise, this is not a grammatical phenomenon. Neither pronoun in (14) is bound syntactically: they are perfectly ordinary referential anaphors, whose duty it is to retrieve a discourse referent from the context. What makes these pronouns special is the requirement that, to a first approximation at least, the last mention of their referents should have employed an expression of the same gender.

That this is not quite right yet appears from the fact that gender pronouns may be used deictically, that is, without a linguistic antecedent. For instance, a Frenchman watching someone trying to get a table ('la table$_{fem}$') into his car might remark (Tasmowski-De Ryck and Verluyten 1982):

(15)   Tu n'arriverais jamais à {la/*le} faire entrer dans la voiture.
        You'll never manage to get {pro$_{fem}$/*pro$_{masc}$} into the car.

Here the pronoun has to agree in gender not with an earlier expression, for there was no previous mention of the table, but rather with the noun that would have been used by default to refer to the table. Observations like this highlight the fact that grammatical gender, too, depends on the non-linguistic context for its interpretation, although the information it carries is of a linguistic nature, and must therefore be represented on a different level.

2.6 Direct Reference

Due to the work of Kripke (1972) and Kaplan (1989) it has become widely accepted that certain types of singular terms, especially proper names and indexicals, 'refer directly'. This is not to deny that these terms have descriptive content. For it is obvious that, for example, 'I' carries the information that the speaker is referring to himself, and that the name 'Brown' refers to someone who is called 'Brown'. It is just that this sort of content is not truth-conditional content (Geurts 1997).

Direct reference poses a problem for theories of meaning that treat all descriptive content alike. Kripke and Kaplan have argued against such mono-semantic accounts, observing that a sentence like (16a) does not have the same truth-conditions as (16b); the first is a contingent truth, while (16b) is necessarily true.

(16) a. Brown is called 'Brown'.
     b. Brown is Brown.

The same point can be made with indexicals:

(17) a. I am the speaker.
     b. The speaker is the speaker.

Whereas the proposition expressed by (17a) might have been false, (17b) is, on one of its readings at least, necessarily true.

Kaplan's well-known analysis of direct reference involves dividing the Fregean notion of sense into two components, which he calls 'character' and 'content'. The character of an expression is its linguistic meaning, which in a given context determines the expression's truth-conditional content. The descriptive content of an indexical is unlike that of a definite description in that it remains at the level of linguistic meaning, and doesn't enter the truth-conditional level. We will see later on how this distinction can be captured in our representational framework.

3 Binding Problems

We have discussed a number of phenomena that illustrate the mundane truth that different kinds of linguistic and para-linguistic information need to be kept apart. But although we must separate between different kinds of content, the separation had better be not too strict. This is the lesson taught by a problem first noted by Karttunen and Peters (1979), which has come to be known as the 'binding problem' of presupposition projection (we will shortly see, however, that the problem is quite general). Karttunen and Peters' example is the following:

(18) ?Someone managed to succeed George V on the throne of England.

This sentence has a question mark because it is pragmatically infelicitous: it suggests that the person who succeeded George V found it difficult to do so, which

can hardly be the case (at least not in the sense intended here; George V's successor may have had problems adjusting to his new station, but he obtained it without effort). Apparently, the presupposition triggered by the verb 'manage' fails in this case. The problem is that many theories of presupposition (including Karttunen and Peters' own) cannot account for this kind of infelicity, because they strictly separate between asserted and presupposed information, as a consequence of which the content of (18) is predicted to have the following components:

(19) Assertion: Someone succeeded George V on the throne of England. Presupposition: It was difficult for someone to succeed George V on the throne of England.

Unfortunately, the presupposition triggered by 'manage', thus construed, comes out true: practically everybody would have had a hard time succeeding George V. This is a problem not only for Karttunen and Peters' own treatment of presupposition, but for theories of a younger vintage, too. The problem arises, obviously, because presupposed and asserted content are separated too strictly, and it is the opposite from the problem discussed above: the DRT treatment of presupposition does not run into the binding problem because it keeps presuppositions and assertions together, though for other reasons they should be differentiated more than they currently are, as we have seen in Sect. 2.1.

Although there has been much discussion of the binding problem in the literature (Krahmer 1998; Beaver 2001), it has rarely been noted that the problem is not confined to presuppositions. But as van der Sandt (1992) points out, binding problems are liable to crop up whenever a strict separation is made between different kinds of information with interdependencies between them. We should expect, therefore, that they also arise in connection with implicatures—and they do:

(20) Some years ago, a young Russian pianist recorded some of the Beethoven sonatas.

Applying the standard Gricean reasoning, we observe that a speaker who uttered (20) could just as easily have made a stronger statement:

(21) Some years ago, a young Russian pianist recorded all the Beethoven sonatas.

Why didn't the speaker utter (21) rather than (20)? Presumably, because he believes that (21) isn't true.[2] But if (21) isn't true, then the speaker is committing himself to the claim that no young Russian pianist ever recorded all the Beethoven sonatas—which in a normal run of events would not be implied by an utterance of (20).

---

[2] We're cutting a few corners here for dramatic effect. See Geurts (2010) for extensive discussion.

As discussed at length by Geurts (2006, 2010), the problem with this pseudo-implicature is caused by the assumption that pragmatic reasoning is conducted solely in terms of sentence-sized semantic units that are disconnected from each other; so the heart of the trouble is the same as in the case of Karttunen and Peters' binding problem. If instead of asking ourselves why the speaker didn't say (21) instead of (20), we would have asked why the speaker didn't say that *the pianist in question* recorded all the Beethoven sonatas, we would have obtained the implicature that, to the best of the speaker's knowledge, the pianist in question didn't record all the Beethoven sonatas—which is correct. But this requires that the implicature is not fully segregated from the assertion: they are about the same individual.

In this section and the last one we have discussed various phenomena illustrating that different kinds of information need to be kept apart, but in such a way that certain interdependencies between them are captured. In the remainder of this chapter we present a unified account that attempts to accomplish just this.

4 Layered DRT

The basic idea underlying Layered DRT (or LDRT for short) is straightforward enough. It is that a discourse representation should consist of more than one layer of information. All the information that is exchanged between speakers will go into the same representation, but within this representation we want to distinguish between information that is asserted, presupposed, implicated, and so on. So within a layered DRS (LDRS) there will be layers for assertions, presuppositions, implicatures, grammatical features of utterances, formal properties of the discourse, and so on. In many cases, information will reside on a single layer, but occasionally the same information can be on more than one layer. This holds, in particular, for discourse referents, which may be seen as inter-layer communication switches.

Formally, layers are implemented as sets of labels on discourse referents and conditions. Every layer has its own label, and as the same piece of content may be on several layers at once, discourse referents and conditions will be assigned sets of labels. In the following, we will show how to add layers to the standard DRT language (Kamp 1981; Kamp and Reyle 1993).

4.1 The LDRT Language

The vocabulary of LDRT extends the standard DRT language with a set of layer labels. We start, as usual, with sets of discourse referents, predicates, and logical constants. All conditions in an LDRS will bear zero or more labels; discourse referents will be labeled, too, but only when they are introduced, not when they

occur as arguments.[3] Taking as given inventories of discourse referents, predicates, logical constants, and layer labels, the following clauses simultaneously define the set of LDRSs, labeled discourse referents and labeled conditions:

(22) In the following clauses, L may be any set of layer labels:

     a. An LDRS φ is a pair $\langle U(\varphi), Con(\varphi)\rangle$, where $U(\varphi)$ is a set of labeled discourse referents and $Con(\varphi)$ is a set of labeled conditions.

     b. If u is a discourse referent, then u_L is an L-labeled discourse referent.

     c. If P is an n-place predicate and $u_1, \ldots, u_n$ are discourse referents, then $P_L(u_1, \ldots, u_n)$ is an L-labeled condition.

     d. If u and v are discourse referents, then $u =_L v$ is an L-labeled condition.

     e. If φ and ψ are LDRSs, then $\neg_L\varphi$, $\varphi \vee_L \psi$, and $\varphi \Rightarrow_L \psi$ are L-labeled conditions.

Furthermore, to be able to focus on an L-part of an LDRS, we define $U_L(\varphi)$ as the subset of discourse referents in $U(\varphi)$ whose label sets overlap with L, and similarly for $Con_L(\varphi)$:

(23) a. $U_L(\varphi) := \{u_K \in U(\varphi) \mid K \cap L \neq \emptyset\}$
     b. $Con_L(\varphi) := \{\psi_K \in Con(\varphi) \mid K \cap L \neq \emptyset\}$

Officially, LDRSs are set-theoretic constructs, but unofficially we will employ the following notation, which we find easier to read. First, instead of $\langle\{u_1, \ldots, u_m\}, \{\varphi_1, \ldots, \varphi_n\}\rangle$ we will write $[u_1 \ldots u_m : \varphi_1 \ldots \varphi_n]$. Secondly, if $\{\alpha_1, \ldots, \alpha_n\}$ is a non-empty set of layer labels we will write $\alpha_1 \ldots \alpha_n$; hence x_{abc}, $P_{abc}(x_1, \ldots, x_n)$, etc. Thirdly, if a discourse referent or condition resides on all layers, we will omit the label set; for example, if there are only three labels, x_{abc} will sometimes be shortened to x (note the difference between x and $x_\emptyset$).

To illustrate how the LDRS language may be used, the interpretation of example (6), repeated here as (24a), may be rendered in LDRT as (24b):

(24) a. The porridge is warm.
     b. $[x_p: porridge_p(x)\ warm_a(x)\ \neg_i[: hot_i(x)]]$

The LDRS in (24b) has three layers: a, p, and i, which contain asserted, presupposed, and implicated material, respectively. The intuitive interpretation of (24b) is that it is presupposed that there is some x that is porridge, that x is asserted to be warm, and that x is implicated not to be hot. Note that the asserted and implicated parts use a discourse referent that is presupposed: we simply cannot say what (24a) asserts or implicates without referring to whatever it is that is being presupposed

---

[3] Cf. Maier (2006, 2009) for a modification of the current version of LDRT, in which discourse referents are never labeled (except in preliminary DRS structures, where labels indicate layered resolution restrictions, which we will not discuss here). In the terms of this paper, Maier assumes that every discourse referent carries all available labels. Intuitively, this may seem to give rise to unwanted existence claims, but in fact these are quite harmless, as long as conditions are sensibly labeled.

by the subject term. This is to say that structures like (24b) cannot be interpreted as consisting of three fully independent layers of information: discourse referents serve to connect information on different layers.

4.2 Semantics of LDRT

The idea underlying our semantics for the LDRS language is simply that, instead of specifying what is the truth-conditional content of an LDRS φ, we have to define what is the truth-conditional content of a selection L of layers in φ. That is to say, if φ contains a condition ψ_K, where K is the layer set associated with ψ, ψ_K is to be ignored unless K and L overlap; i.e. unless $K \cap L \neq \emptyset$.

In the standard semantics of DRT, an embedding function f is said to be extended by another function g, with respect to a given DRS φ, iff $f \subseteq g$ and $dom(g) = dom(f) \cup U(\varphi)$. Here we extend f only with discourse referents with relevant labels:

(25) $f[^{\varphi}_{L}]g : = f \subseteq g$ and $dom(g) = dom(f) \cup U_L(\varphi)$

Given a world w, a label set L, and an embedding function f, the principal semantic object associated with an LDRS φ is $\|\varphi\|^f_{L,w}$, which, if defined, is the set of embedding functions g that extend f and make the L-part of φ true at world w. If $\|\varphi\|^f_{L,w}$ is a non-empty set, the L-part of φ is true at w; if $\|\varphi\|^f_{L,w} = \emptyset$, the L-part of φ is false at w; and otherwise $\|\varphi\|^f_{L,w}$ is undefined. If φ is a labeled condition, $\|\varphi\|^f_{L,w}$, if defined, is either $\top$ or $\bot$.

Let M $= \langle$D, W$\rangle$ be a model, where D is a domain of individuals and W is a set of interpretation functions ('worlds'); w $\in$ W; f is a partial function from the set of discourse referents into D; and L is a set of layer labels:

LDRSs: definedness and interpretation

Let φ be an LDRS. Then:

a. $\|\varphi\|^f_{L,w}$ is defined iff $\exists g$: $f[^{\varphi}_{L}]g$ and $\forall \psi \in Con_L(\varphi)$: $\|\psi\|^g_{L,w}$ is defined.

b. If defined, $\|\varphi\|^f_{L,w} = \{g \mid f[^{\varphi}_{L}]g$ and $\forall \psi \in Con_L(\varphi)$: $\|\psi\|^g_{L,w} = \top\}$.

Labeled conditions: definedness

c. $\|P_K(u_1, ..., u_n)\|^f_{L,w}$ is defined iff $\{u_1, ..., u_n\} \subseteq dom(f)$.

d. $\|u =_K v\|^f_{L,w}$ is defined iff $\{u, v\} \subseteq dom(f)$.

e. $\|\neg_K \varphi\|^f_{L,w}$ is defined iff $\|\varphi\|^f_{L,w}$ is defined.

f. $\|\varphi \vee_K \psi\|^f_{L,w}$ is defined iff $\|\varphi\|^f_{L,w}$ and $\|\psi\|^f_{L,w}$ are defined.

g. $\|\varphi \Rightarrow_K \psi\|^f_{L,w}$ is defined iff $\|\varphi\|^f_{L,w}$ and $\|\varphi \oplus \psi\|^f_{L,w}$ are defined, where $\varphi \oplus \psi = \langle U(\varphi) \cup U(\psi), Con(\varphi) \cup Con(\psi)\rangle$ (LDRS-merge).

Labeled conditions: interpretation

If φ_K is a labeled condition, $\|\varphi_K\|_{L,w}^{f} = \top$ iff $\|\varphi_K\|_{L,w}^{f}$ is defined and one of the following holds:

h. φ_K is of the form $P_K(u_1, \ldots, u_n)$ and $\langle f(u_1), \ldots, f(u_n) \rangle \in w(P)$.

i. φ_K is of the form $u =_K v$ and $f(u) = f(v)$.

j. φ_K is of the form $\neg_K \psi$ and $\|\psi\|_{L,w}^{f} = \varnothing$.

k. φ_K is of the form $\psi \vee_K \chi$ and $\|\psi\|_{L,w}^{f} \cup \|\chi\|_{L,w}^{f} \neq \varnothing$.

l. φ_K is of the form $\psi \Rightarrow_K \chi$ and $\forall g \in \|\psi\|_{L,w}^{f} : \|\chi\|_{L,w}^{g} \neq \varnothing$.

$\|\varphi_K\|_{L,w}^{f} = \bot$ iff $\|\varphi_K\|_{L,w}^{f}$ is defined and $\|\varphi_K\|_{L,w}^{f} \neq \top$.

If φ is an LDRS, $\|\varphi\|_{L,w}^{f}$ is the set of embedding ...

If φ is an LDRS, $\|\varphi\|_{L,w}^{f}$ is the set of embedding functions that extend f and that make the L-part of φ come out true in w. Shifting to a more general notion of content, the following clauses define the set of worlds in which φ's L-layers are true:

(27) $\|\varphi\|_{L}^{f} = \{w \mid \|\varphi\|_{L,w}^{f} \neq \emptyset\}$, if $\exists w : \|\varphi\|_{L,w}^{f}$ is defined; undefined otherwise.

(28) $\|\varphi\|_{L} = \|\varphi\|_{L}^{f_0}$, where f_0 is the empty function.

In LDRT, every choice of labels engenders its own sort of information. For example, $\| (24b) \|_{\{p\}}$ is the set of worlds that contain porridge; $\| (24b) \|_{\{p,a\}}$ is the set of worlds with warm porridge; and $\| (24b) \|_{\{p,a,i\}}$ is the set of worlds containing porridge that is warm but not hot. Note that, for example, $\| (24b) \|_{\{a\}}$ and $\| (24b) \|_{\{i\}}$ are undefined. This is because the a- and i-layers of (24b) use a discourse referent that is introduced in the p-layer. The undefinedness of $\| (24b) \|_{\{a\}}$ and $\| (24b) \|_{\{i\}}$, as opposed to the definedness of $\| (24b) \|_{\{p\}}$, is due to the fact that the assertion and implicature of (24b) depend on what the sentence presupposes, but not vice versa. What (24b) asserts or implicates can only be specified relative to a given value (or range of possible values) of the discourse referent x.

If we collect all labels into one set L, then $\| \cdot \|_{L}$ captures everything that is somehow expressed by a sentence or discourse. We have seen that such content aggregates need not be consistent:

(29)   a.   That is a beautiful painting. (= (8))

   b. $[x_k : dem_k(x) \text{ beautiful-painting}_a(x) \neg_i[: \text{beautiful-painting}_i(x)]]$

On the intended reading of (29a), the sentence is meant to convey that the object under discussion is in fact not a beautiful painting at all. Assuming for the nonce that this is an implicature, and that demonstratives are represented on a special k-layer (more about which in Sect. 5), we obtain a representation along the lines of (29b). The literal meaning of (29a) is $\| (29b) \|_{\{k,a\}}$; the implicit message is $\| (29b) \|_{\{k,i\}}$; and $\| (29b) \|_{\{k,a,i\}} = \emptyset$.

One application of the rich representations and flexible semantics of LDRT is in the analysis of denials. In Sect. 2.1, we saw how denials can target different layers, e.g. asserted or presupposed information, and they can also be directed at several layers at once (see Maier and der Sandt 2003 for discussion and an analysis of denial in LDRT). In the following, we will demonstrate the power of LDRT with two other applications, each of which will be seen to require some minor additions to the basic semantics presented above.

5 Indexical Content: The k-Layer

The LDRS-semantics given in (26) is uniform in the sense that, once a group of layers have been selected, all layers are treated alike. (That is, they are treated alike by the semantics. If layers were alike in every respect, there would be no point in having them in the first place.) In the remainder of this chapter we discuss two classes of phenomena which show that this is not quite right, and adjust our semantics accordingly.

Standard DRT has trouble with names, indexicals, and demonstratives because it has no way of separating descriptive content from contextual, 'reference fixing' content. In LDRT we can simply represent the two types of content on two different layers: 'a' is for asserted content, and 'k' for contextual, rigid content. We already used these layers in the previous section (cf. (29b)); a crucial example is the following, which we encountered before in Sect. 2.6:

(30) a. I am the speaker. (= (17a))
    b. $[x_k : speaker_k(x) \ speaker_a(x)]$

However, it is not enough just to put indexical content on a layer of its own: $\| (30b) \|_{\{k,a\}}$ merely says that there is a speaker, whereas it should say of the individual who *in fact* is doing the talking that he is the speaker. In order to account for this, we follow Kaplan by making the content of an LDRS dependent on the context in which it occurs. A context may be regarded as a small world in the sense that it determines a unique speaker, hearer, etc. In our LDRS-semantics worlds are identified with interpretation functions, so if c is a context, then c(speaker), c(hearer), c(now), etc., are singleton sets. There are various ways of enforcing this restriction, two of which will be demonstrated in the following.

Relative to a given context c, we define the indexical content of the *L*-part of an LDRS φ as follows:

(31) $\mathbb{I}_{L,c}(\varphi) = \|\varphi\|_L^{\iota}$, with ι being the unique element of $\| \varphi \|_{\{k\},c}^{f_0}$, if such exists; otherwise undefined.

If $\|\varphi\|_{\{k\},c}^{f_0}$ is not a singleton set, c fails to determine unique values for all discourse referents in the k-layer, and $\mathbb{I}_{L,c}(\varphi)$ is undefined. Otherwise $\mathbb{I}_{L,c}(\varphi) = \|\varphi\|_L^{\iota}$, where ι is the unique embedding function determined by c. For example, if White is the

speaker in context c, $\mathbb{I}_{\{k,a\},c}(30b) = \| (30b) \|_{\{k,a\}}^{\{\langle x, White\rangle\}}$ = the set of worlds in which White is the speaker.

This type of LDRT implementation of direct reference is further developed and defended by Maier (2009), where it is combined with a theory of layered presupposition resolution. Although it is an attractive way of thinking about rigidity in DRT, we would like to explore here also a slightly different way of implementing a Kaplan-style notion of content in LDRT, which incorporates context parameters in the definition of $\| \cdot \|$. The main difference is that we can then interpret embedded k-layers, which is a feature that we need not so much for Kaplanian rigidity, but for our treatment of formal content in Sect. 6. We have to clear several choice points if we take this line, but the simplest solution we can think of mainly affects the interpretation of atomic conditions, which now comes out as follows:

a. $\|P_K(u_1, ..., u_n)\|_{L,w}^{f,c}$ is defined iff $\{u_1, ..., u_n\} \subseteq dom(f)$ and if $k \in K \cap L$, then $|c(P)| = 1$.
b. $\|P_K(u_1, ..., u_n)\|_{L,w}^{f,c} = \top$ iff $\|P_K(u_1, ..., u_n)\|_{L,w}^{f,c}$ is defined and one of the following holds:
   $- k \notin K \cap L$ and $\langle f(u_1), ..., f(u_n)\rangle \in w(P)$;
   $- k \in K \cap L$ and $\langle f(u_1), ..., f(u_n)\rangle \in c(P)$.

In other words: If an atomic condition is not on the k-layer, its interpretation is as in (26). If it is on the k-layer, its semantic value is undefined if the current context c fails to assign a unique value to its predicate. Suppose again that c(speaker) = White; then $\| (30b) \|_{\{k,a\},w}^{f_0,c} = \emptyset$ if White is not the speaker at w, or else $\| (30b) \|_{\{k,a\},w}^{f_0,c} = \{g\}$, where $dom(g) = \{x\}$ and $g(x) = White$. Hence, $\| (30b) \|_{\{k,a\}}^{c} = \mathbb{I}_{\{k,a\},c}(30b)$ = the set of worlds in which White is the speaker.

So the two methods produce the same result in this case, as they do in many others, but they are not fully equivalent. First, and most importantly, with the second method, all discourse referents and conditions labeled 'k' are interpreted at the contextual index. The first method by contrast presupposes, in effect, that all k-material resides in the principal LDRS, or otherwise it will not be interpreted at c. Secondly, whereas the second method requires that each k-marked predicate be unique, the first method is less stringent in this respect, since it requires merely that, between them, the conditions in the k-layer determine unique values for all k-marked discourse refererents. The choice between these methods depends on considerations that go beyond the scope of this chapter. We should like to note, however, that on the whole the second method is more versatile, and brings out more clearly the relation between indexical content and what we call 'formal content', which is the topic of the next section.

6 Formal Content

We argued in Sect. 2.4 that we should aim for an analysis of expressions like 'the former' and 'the latter' which treats them as regular definite descriptions whose content is somewhat special. In this section we shall see that LDRT can provide us with such an analysis. As it turns out, 'the former', 'the latter', and their kin are context dependent in a way that resembles the context dependence of indexical expressions.

The LDRS in (33b) is a first stab at capturing the intuition that the expressions 'the former' and 'the latter' as used in (33a) are presuppositional devices whose content refers to formal properties of the previous discourse:

(33) a. As the beggar approached the bishop, the latter blessed the former.
b. $[x_p, y_p: beggar_p(x) \, bishop_p(y) \, approach_a(x,y) x \prec_{pf} y \, bless_a \, (y,x)]$

The intended interpretation of condition '$x \prec_{pf} y$' is something like: 'The last mention of x precedes the last mention of y.' (There may be other, and perhaps better, ways of rendering the meanings of 'the former' and 'the latter', but this one will do for our purposes.) This condition is part of the presupposition triggered by 'the former/latter', so it is on the p-layer, and it is also on the f-layer, because it refers to the form of the preceding discourse.

There is one problem with this proposal, for which we shall present a tentative solution. As it stands, our treatment of the interpretation of 'the former/latter' and related devices, such as grammatical gender, presupposes that the properties denoted by these expressions are properties of regular individuals: $'X \prec_{pf} Y'$ is true iff the last mention of the individual associated with x preceded the last mention of the individual associated with y. Example (10), repeated here as (34a), with its LDRS in (34b), demonstrates that this is not correct in general:

(34) a. If a beggar meets a bishop, then the latter will bless the former.
b. $[: [x_a, y_a : beggar_a(x) \, bishop_a(y) \, x \prec_{pf} y \, meet_a(x,y)] \Rightarrow [: bless_a(y,x)]]$

That is, for all beggar–bishop pairs in which the beggar is mentioned before the bishop, the second blesses the first. In order to heighten the dramatic impact of the example, we might suppose that the conditional is given a modal interpretation, and is construed as quantifying over worlds. But whatever the sentence quantifies over, its domain is not confined to states of affairs containing pairs of persons one of whom was mentioned before the other. The only mentioning event that is relevant is the actual utterance, which is part of the actual context.

If the predicate '$\prec$', as used in (34b), is not about beggars and bishops, be they possible or real, then what *is* it about? The answer, we would like to suggest, is 'discourse referents': the condition '$x \prec y$' states that the most recent use of x preceded the most recent use of y. Hence, a speaker who employs the 'former/latter' idiom in effect instructs the hearer to retrieve from the context a pair of

recently-used discourse referents. Note that this requires a notion of context that goes beyond the original Kaplanian context of utterance in containing not only individuals but also a structured representation of the actual utterance, as well as the discourse referents that that utterance has given rise to.[4]

In order to implement this idea, we propose to expand the interpretation of atomic conditions given in 4 along the following lines:

a. $\|P_K(u_1,...,u_n)\|_{L,w}^{f,c}$ is defined iff one of the following holds:
   - $k \in K$ and $\{u_1,...,u_n\} \subseteq dom(f)$ and $|c(P)| = 1$;
   - $f \in K$ and $|c(P)| = 1$;
   - $\{f,k\} \cap K = \varnothing$ and $\{u_1,...,u_n\} \subseteq dom(f)$;
b. $\|P_K(u_1,...,u_n)\|_{L,w}^{f,c} = \top$ iff $\|P_K(u_1,...,u_n)\|_{L,w}^{f,c}$ is defined and one of the following holds:
   - $k \in K \cap L$ and $\langle f(u_1),...,f(u_n) \rangle \in c(P)$;
   - $f \in K \cap L$ and $\langle u_1,...,u_n \rangle \in c(P)$;
   - $\{f, k\} \cap K \cap L = \varnothing$ and $\langle f(u_1),...,f(u_n) \rangle \in w(P)$.

According to this analysis, the interpretation of f-marked conditions depends on the context, and in this respect f-conditions and k-conditions are alike. But on the other hand, f-conditions are special in that they are about discourse referents, whilst all other conditions, k-conditions included, are about 'real' things in the world.

7 Conclusion

Our main objective in this chapter was to provide a general framework for representing and integrating all sorts of information that may be conveyed by linguistic means. Our proposal is LDRT. From a syntactical point of view, LDRT is perfectly straightforward. Label sets allow us to separate between different types of information, without severing binding relations. This device is uniform in the sense that, syntactically speaking, the *only* difference between one type of content and another consists in the labels they bear. But of course different types of content will differ in other ways as well—if they didn't, they wouldn't have to be distinguished in the first place. Such differences may be procedural; for example, certain types of content are cancellable whilst others are not. Other differences may be semantic; for example, indexical and formal content are context dependent in distinctive ways. But at the representational level, all kinds of content are equal.

---

[4] Incidentally, such a notion of context seems extremely useful for the study of signed languages, where discourse referents correspond to actual, visible points in the signing space, which signers point to and keep track of in a discourse. Cf. Schlenker (2010) for a discussion of DRT discourse referents in the analysis of sign language.

References

Beaver, D. (2001). *Presupposition and assertion in dynamic semantics*. Stanford: CSLI.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.

Geurts, B. (1997). Good news about the description theory of names. *Journal of semantics, 14*, 319–348.

Geurts, B. (1999). *Presuppositions and pronouns*. Oxford: Elsevier.

Geurts, B. (2006). Implicature as a discourse phenomenon. In E. Puig-Waldmüller (Ed.), *Proceedings of Sinn und Bedeutung 11*, (pp. 261–275). Barcelona: Universitat Pompeu Fabra.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, & M. Stokhof (Eds.), *Formal methods in the study of language*, (pp. 277–322). Amsterdam: Mathematical Centre Tracts 135.

Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Dordrecht: Kluwer.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press.

Karttunen, L., & Peters, S. (1979). Conventional implicature. In C. Oh & D. A. Dinneen (Eds.), *Syntax and semantics 11: Presupposition* (pp. 1–56). New York: Academic Press.

Krahmer, E. (1998). *Presupposition and anaphora*. Stanford: CSLI.

Kripke, S. (1972). Naming and necessity. In G. Harman & D. Davidson (Eds.), *Semantics of natural language* (pp. 253–355). Dordrecht: Reidel.

Maier, E. (2006). *Belief in context: Towards a unified semantics of de re and de se attitude reports*. Ph.D. thesis, Radboud University Nijmegen.

Maier, E. (2009). Proper names and indexicals trigger rigid presuppositions. *Journal of semantics, 26*, 253–315.

Maier, E. & van der Sandt, R. (2003). Denial and correction in layered DRT. In *Proceedings of DiaBruck 2003*. http://www.coli.uni-sb.de/diabruck/.

Schlenker, P. (2010). Donkey anaphora in sign language I: E-type vs. dynamic accounts. In M. Aloni (Ed.), *Logic, language and meaning: Proceedings of the Amsterdam Colloquium 2009*, (pp. 405–415).

Tasmowski-De Ryck, L., & Verluyten, P. (1982). Linguistic control of pronouns. *Journal of semantics, 1*, 323–346.

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of semantics, 9*, 333–377.

von Fintel, K. (2004). Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. In M. Reimer, & A. Bezuidenhout (Eds.), *Descriptions and beyond*, (pp. 269–296). Oxford University Press.

On the Conversational Basis of Some Presuppositions

Mandy Simons

Abstract This paper, originally published in 2001, deals with the question of the source of presuppositions, focusing on the question of whether presuppositions are conventional properties of linguistic expressions, or arise as inferences derivable from ordinary content in combination with some general conversational principles. I argue that at least some presuppositions should be analysed as conversational inferences, on the grounds that they show two of the hallmarks of such inferences: contextual defeasibility and nondetachability. I make this case for the presuppositions associated with change of state predicates and with factives. I argue further for the need for a general principle for deriving presuppositions as inferences by illustrating a variety of cases of presupposition-like inferences not clearly involving a lexical presupposition trigger. In the second half of the paper, I move towards the development of a general conversational account of the relevant presuppositions. Building on a brief comment in Stalnaker (1974), I develop the following pair of ideas: first, that an utterance embedding a proposition P may be seen as raising the question whether P; and second, that P may be related to a further proposition Q in such a way that it would make sense to raise the question whether P only if one already believed Q to be true. It is these required prior beliefs that constitute conversationally derived presuppositions. Although the account developed here is only a preliminary attempt, the relevance of contextually salient questions, or sets of alternatives, to an account of presupposition has been taken up in subsequent work, notably Abusch (2010) and Simons et al. (2010).

M. Simons (✉)
Carnegie Mellon University, Pittsburgh, USA
e-mail: simons@andrew.cmu.edu

1 Introduction

The current literature on presupposition focuses almost exclusively on the projection problem: the question of how and why the presuppositions of atomic clauses are projected to complex sentences which embed them. Very little attention has been paid to the question of how and why these presuppositions arise at all. As Kay (1992: 335) observes, "treatments of the presupposition inheritance problem almost never deal with the reasons that individual words and constructions give rise, in the first place, to the particular presuppositions that they do".[1] This is the question on which this paper will focus.

There are two kinds of answer that one might give to the question of how presuppositions arise. One type of answer is that presuppositions are conventional properties of lexical items, as in the conventional implicature view of Karttunen and Peters (1979). On this view, certain lexical items have, in addition to their truth conditional content, a special presuppositional content, which is carried through the compositional process to produce a propositional presupposition. Alternatively, one might say that part of the ordinary content is categorized, in the lexical entry of the expression, as having presuppositional status. Although the Karttunen and Peters model for treating presupposition has been rejected by most current researchers, our talk about presupposition seems at least implicitly to take their view of the sources of presuppositions for granted: we talk about the presuppositions of *know*, of *too*, and so on, as if assuming that the presuppositions are properties of these items.[2]

Presuppositions, however, might be thought to have a very different source. Presuppositions might be conversationally derived, that is, they might be inferences which are licensed by general conversational principles, in combination with the truth conditions of the presupposing utterance. Stalnaker, from whom we have inherited the currently standard view of presupposition, suggests repeatedly that at least some presuppositions have a conversational source. Indeed, he sees one of the primary advantages of the move from a semantic to a pragmatic account of presupposition as being the possibility of explaining "some of the (presupposition) facts in terms of general assumptions about rational strategy in situations where people exchange information or conduct argument" (1974: 205). However, Stalnaker never attempts to work out any general derivation for conversational presuppositions, nor any means for distinguishing conventionally generated presuppositions from those with a conversational source. Many other authors have

---

[1] Kay adds with regret that his own study "will unfortunately make no improvement on this practice".

[2] We also attribute presuppositions to specific constructions, in particular *it*-clefts and *wh*-clefts. But it is particularly problematic to think of these presuppositions as conventional, given the assumption that there is neither a "lexicon" of constructions, nor any construction-specific syntactic rules. If there is no rule for the formation of clefts, then there is nothing to which to attach the conventional rule that a cleft has a presupposition. (This observation is due to Sally McConnell-Ginet.)

likewise suggested that some or all presuppositions can be accounted for in terms of general conversational principles, or even reduced to conversational implicatures. The most robust formulations of this view are proposed by Atlas and Levinson 1981, Levinson 1983, Kadmon 2001 and Atlas 2005 (see also many earlier works by Atlas), but the idea is also explored in Kempson (1975), Wilson (1975), Boër and Lycan (1976), Karttunen and Peters (1979), Grice (1981) and Chierchia and McConnell-Ginet (1990), among others.[3]

My primary goal in this paper is to motivate the claim that at least some presuppositions have a conversational source, and to set out a partial proposal for the conversational derivation of these presuppositions. Thinking about the derivation of presuppositions is worthwhile, I think, because it gives us a somewhat different perspective on what presuppositions *are*, and raises some questions about standard assumptions.

2 Evidence for the Conversational Basis of Some Presuppositions

We begin with the evidence that at least some presuppositions do have a conversational source. In this section, I'll show that two classes of presuppositions—the presuppositions of change of state sentences and of factives—possess two properties that are typical of conversational implicature, the paradigm case of a conversationally derived inference. The two properties are contextual defeasibility and nondetachability. We'll take these two in turn.

2.1 Contextual Defeasibility

It is well known that many presuppositions can fail to project in certain sentential contexts. In addition, some—but crucially, not all—presuppositions may fail to project in what I will call *explicit ignorance contexts*: situations in which it is apparent to the addressee that the speaker is ignorant with respect to the proposition that would normally be presupposed. This type of case has already been noted in the literature. We begin with an example from Geurts (1994). Imagine a casual conversation taking place between two people who are meeting for the first time. One remarks to the other:

---

[3] (2011) When I originally wrote this paper, I was shockingly ignorant about the prior literature, particularly that from the 1970s and early 1980s, containing related arguments and views. It was for that reason that the original version contained no references to that literature, a serious omission which I am happy to remedy here.

(1) I notice that you keep chewing on your pencil. Have you recently stopped smoking?

In this situation, the addressee knows that the speaker is ignorant of her current or prior smoking habits, and in particular cannot be assuming that she (the addressee) was recently a smoker. In this situation, no implication would arise that the speaker believes the addressee to have been a smoker. The speaker is understood merely as asking whether the addressee has undergone the relevant change of state from being a smoker to not being one, equivalent to:

(2) Is it the case that you have habitually smoked in the recent past and that you recently ceased to do so?[4]

What is involved here is the choice of a non-presupposing interpretation of the question over a presupposing interpretation. While the presupposing interpretation is usually preferred, we apparently have a stronger preference not to attribute to the speaker an unfounded assumption. So the presuppositional interpretation is not chosen.

Example (3) is a variation on the same theme. (To make the example natural, assume that there is some special symptom displayed by a person who has stopped smoking that Jane does not display.)

(3) I have no idea whether Jane ever smoked, but she hasn't stopped smoking.[5]

Here, the speaker begins with an explicit assertion of her ignorance about the relevant proposition, and thus the *stop* clause does not give rise to the presuppositional inference that Jane is or was a smoker. The clause is understood simply as denying that she has undergone the change of state.

Not all expected presuppositions fail to arise in explicit ignorance contexts. In particular, presuppositions which have clearly identifiable lexical triggers such as *even*, *too* and *again* are not cancelable in this way. As an example, suppose I happen to meet a slight acquaintance at a video rental store, someone whose history of video rental I clearly know nothing about. I ask her:

---

[4] (2011) While one might be inclined to talk about suppression or cancelation of the presupposition in this case, note that the typically projective content does not just "go away"; rather, it makes only a local semantic contribution under the scope of the question operator. (See Tonhauser et al. 2011).

[5] Notice that this sentence does *not* have the usual form of a presupposition-canceling conjunction. Cancellation is expected when the first conjunct entails the presupposition of the second, which is not the case here. This type of sentence is, however, reminiscent of examples with *but* due to Liberman (1973), such as:

1. Perhaps John has no children, but perhaps his children are away on vacation.
Liberman offers his examples as counterexamples to the standard characterization of projection patterns. I think in fact that these examples involve cancellation in the face of explicit ignorance. The first clause indicates that the speaker is not sure whether John has children, and thus has the same effect as the first clause in my example.

(4) Are you renting "Manhattan" again?

Despite the explicit ignorance context, my addressee would be forced to conclude that I believe for some reason that she has rented "Manhattan" before. In particular, there is no way for her to understand me as asking:

(5) Is it the case the you have rented "Manhattan" before, and are going to do so again?

The presupposition generated by *again* is not canceled even when the speaker explicitly asserts her ignorance, leading to anomaly in cases like (6):

(6) # I don't know if Jane ever rented "Manhattan" before, but perhaps she's renting it again.

Contrast (6) with (7), which shows that the anomaly would disappear if a non-presuppositional reading were available for the second clause.

(7) I don't know if Jane ever rented "Manhattan" before, but perhaps she has and is renting it again.

Examples (1) and (3), which show contextual defeasibility of a presupposition, both involve the presupposition of the change of state verb *stop*. Indeed, all change of state predicates give rise to presuppositions of the same type, and all these presuppositions are contextually defeasible. The presuppositions of factives show the same susceptibility to contextual defeasibility, as noted in the literature. Let's set up another explicit ignorance context: Suppose we are at a restaurant, and notice a couple at another table engaged in a furious argument. We are speculating as to what has upset them. I say to you:

(8) Perhaps she just discovered that he's having an affair.

As you know that I couldn't possibly know whether he has in fact been having an affair, you will not take me to presuppose that he has, but merely to be making the relatively weak supposition that he has been having an affair and that she has discovered it.

Chierchia and McConnell-Ginet (1990) offer a similar example. They observe that sentence (9) below could well be uttered as part of a conversation between two people who know that Henry is searching for Jane, but who don't themselves know where Jane is:

(9) If Henry discovers that Jane is in New York, there'll be trouble.

In this situation, there is no presupposition or implication that Jane is in fact in NY.

We can summarize these observations as follows: Presupposition projection in the case of factives and of change of state predicates has the effect of committing the speaker to the relevant proposition. If it is clear from the context that the speaker has no such commitment, then the presupposition fails to project.

Not surprisingly, completely parallel phenomena are observed with conversational implicatures, including generalized conversational implicatures. Consider for example the ignorance implication usually triggered by the antecedent of a conditional such as:

(10) If Bill is coming to the party, then Jane won't come.

Normally, an utterance of (10) would allow the interpreter to infer that the speaker does not know whether or not Bill is coming. But in the following discourse, the inference does not arise:

(11) A: Is Jane coming to the party?
     B: Well, Bill is coming. And if Bill is coming to the party, then Jane won't come.

In this case, the use of the conditional does not imply speaker ignorance about the truth of the antecedent, because to derive this implication would be to draw a conclusion about the speaker's beliefs that is evidently false.

The point of all of these examples is that the presuppositions of change of state predicates and of factives display the same kind of contextual defeasibility as do (generalized) conversational implicatures. This is our first piece of evidence that at least some presuppositions have a conversational basis. (For a similar argument, see Levinson 1983.)

2.2 Nondetachability

Grice (1967, 1989) observes that conversational implicatures are by necessity *nondetachable* from the content of whatever utterance gives rise to them. What this means is that if utterance of a sentence with a particular content C generates an implicature in a given conversational context, then utterance of any other sentence with the same content will give rise to the same implicature.[6] For example, consider the short exchange in (12):

(12) Jane: Do you want to go out for a drink?
     Julia: I have to finish writing this paper.

In the context of Jane's question, Julia's utterance generates the implicature that she does not want to go out for a drink. However, the generation of the implicature is not dependent upon the form of Julia's utterance. Any other form which expresses more or less the same content would do just as well to produce the implicature. The responses in (13) are all possible candidates.

---

[6] An exception must be made here for Manner implicatures.

(13) a. ...I need to finish this paper.
    b. ...This paper needs to get finished tonight.
    c. ...I have to work on this paper.

Conversational implicatures are nondetachable because they are due to the expression of a particular content in a particular conversational context. They are—by definition—not *conventionally* associated with any expression, so naturally we do not expect that the expression of the same content in a different form will affect the implicature.

Some presuppositions turn out to have this same property of nondetachability. Consider another example with *stop*:

(14) Jane didn't stop laughing.

Utterance of this sentence normally gives rise to the presupposition that Jane had been laughing immediately prior to the reference time of the sentence. We standardly say that this is due to the presupposition associated with *stop*. But notice that if we replace *stop* with any of its synonyms, the presupposition remains:

(15) a. Jane didn't quit laughing.
    b. Jane didn't cease laughing.
    c. Jane did not discontinue her laughter.

The same is true of the presuppositions of all other change of state predicates. As further illustration, consider the synonymous sentences in (16). Utterance of any of these would normally give rise to the presupposition that Jane was in the house immediately prior to the reference time of the sentence.

(16) a. Jane didn't leave the house.
    b. Jane didn't quit the house.
    c. Jane didn't go out of the house.
    e. Jane didn't depart from the house.

The same point is illustrated with the factive *realize* and its synonyms in (17).

(17) Harry didn't realize/come to know/become aware that he was a fool.

These observations strongly suggest that the presuppositions of change of state predicates and of factives are nondetachable, that is, that they attach to the content expressed, and not to any lexical item. But presuppositions or implications cannot attach to content by convention. Thus, these presuppositions must have a conversational source.[7]

---

[7] (2011) A related point is that some investigations indicate that presuppositions are largely stable across languages. So, for example, in other languages which have a lexical item with the meaning *stop*, use of that item tends to give rise to the start-state implication. However, there has not really been enough systematic study of presuppositions across languages to make a definitive claim. For some preliminary work, see Levinson and Annamalai 1992 and Tonhauser et al. 2011.

Of course, one could deny that the data show nondetachability, and argue instead that in each of these cases, the presupposition that arises is due to a conventional property of one of the lexical items used. However, this immediately raises the question of why all items with the same truth conditional content as, say, *stop* should also have the same presuppositional content. After all, it could be quite useful for language users to have in the language two items which are truth conditionally equivalent but presuppositionally distinct. In fact, the observation that some set of truth conditionally equivalent forms all give rise to the same presupposition is strong motivation to seek a principled connection between the truth conditional content and the observed presupposition.

2.3 More on Nondetachability

I must confess that the nondetachability argument is something of a double edged sword for me. The problem is that the very same arguments appear to apply to presuppositions which I would rather assume *are* in fact conventionally attached to particular lexical items, namely, the presuppositions triggered by *even*, *too*, *again* and the like. For example, all of the sentences in (18) give rise to the presupposition that some (relevant) person other than Harriet is going to the conference. (In all cases, assume focus on the subject NP.)

(18) a. Harriet might go to the conference too.
     b. Harriet might go to the conference as well.
     c. Harriet might also go to the conference.

The same point can be made about *again* and its synonyms; the point is harder to make with *even*, simply because it doesn't have any obvious synonyms in English.

There is, however, a difference between the cases of *again* and *too*, and the change of state and factive cases discussed above. It is possible to express the non-presuppositional content of the sentences in (18) independently of the presupposition, simply by omitting the presupposition trigger. In the case of the change of state predicates and the factives, there is no way to separate these two types of content. On one way of viewing things, this is to say that the presupposition in (18) a–c is *not* nondetachable: the truth conditional content of these sentences is just that Harriet might go to the conference, and this content can be expressed without giving rise to the presupposition. Hence, the presupposition is not attached to the truth conditional content. The danger of this argument, as Kent Bach has recently pointed out in another context (Bach 1999), is that it can become question-begging. It requires the prior assumption that whatever content is contributed by *too, as well* and *also* is not part of the truth conditional content of the sentence.

While the status of these presuppositions thus remains unclear, I can safely set the issue aside for now. My goal here is to demonstrate that at least some

presuppositions have the properties of conversationally generated inferences. The presuppositions of change of state predicates and of factives are robustly nonde-tachable, as well as being contextually defeasible. These observations strongly suggest that at least these presuppositions have a conversational basis.[8]

3 More Evidence: Projection Without Presuppositions

I turn now to a second kind of evidence that some general strategy for generating presuppositions is needed. This evidence involves cases where we get something that looks like a presuppositional inference in the absence of anything that we would normally think of as a presupposition trigger. My first set of examples all involve voting for Nader. Let's begin with the following exchange:

(19) Jane: George voted for Nader.
   Julia: No he didn't.

Note that Julia's response would normally be taken as a denial of a vote for Nader, not as a denial that George voted at all. The same would generally be true of an utterance of the full negation:

(20) George didn't vote for Nader.

This implication is easily cancelable, but the cancellation has the "flavor" of presupposition cancellation:

(21) George didn't vote for Nader. In fact, he didn't vote at all.

I must immediately acknowledge that the facts are changed by the introduction of focal stress. Focal stress on *George*, for example, gives rise to a presupposition or implication that someone other than George voted for Nader, and also seems to reduce the strength of the implication that George voted. Focal stress in other positions similarly changes the presuppositions, in ways that have been well documented in the focus literature. What is crucial for my current purposes is that there is no stress pattern naturally interpreted as a denial that George voted.

In examples (19) and (20), the clause *George voted for Nader* is embedded under negation. We have observed that one entailment of the clause "projects" over the negation, surviving as a non-entailed implication. The same "projection effects" occur in questions, under epistemic modals, and in conditionals, just as with standard presuppositions. Thus, utterance of any of the sentences in (22) would normally give rise to the implication that the speaker believes that George voted:

---

[8] These points echo the arguments given in Atlas and Levinson 1981, who made the same observations about defeasibility and nondetachability. They do not, though, make the distinction I have indicated here between different kinds of presupposition triggers.

(22) a. Did George vote for Nader?
    b. Perhaps George voted for Nader.
    c. If George voted for Nader, he's more politically motivated than he used to be.

One way to characterize these observations would be to say that the sentence *George voted for Nader* presupposes that George voted. But to what would we attach this presupposition? It cannot be attached to *vote* alone, for the "presupposition" disappears when the modifier is removed. Sentence (23) obviously does not presuppose that George voted.

(23) George didn't vote.

On the other hand, there is no mechanism to attach the presupposition to the complex expression *vote for Nader*, for this expression does not have a lexical entry and is not formed by a construction-specific rule (cf. note 2). Moreover, there is no intuition that the prior assumption that George voted is required in order to make sense of the sentences in (20) and (22). We get little explanatory benefit from saying that *George voted for Nader* presupposes that George voted. What is wanted is some principled explanation for the inferences observed.[9]

Once we find one example of this kind of projection without presupposition, the cases multiply rapidly. There are two different types of example, which I designate *lexical* and *modificational*. Sentences with the verb *win* are an instance of a lexical case. Consider sentence (24). This entails that Smith will participate in the race:

(24) Smith will win this race.

The sentences in (25) do not share this entailment, but a speaker who uttered either of these sentences would nonetheless normally be taken to believe that Smith will participate in the race. In other words, the entailment that Smith will participate in the race "projects" over the negation.

(25) a. Smith won't win this race.
    b. Will Smith win this race?

The adjectives *late* and *early* behave in the same way. To say either (26) a. or b. is to imply that Jones has an obligation to be somewhere at a particular time, a proposition which would be entailed by the corresponding affirmative but is not entailed by the negations:

(26) a. Jones isn't late.
    b. Jones isn't early.

---

[9] (2011) In light of the observations here about the effects of intonation, I would now be inclined to think that this class of cases is to be explained as an effect of semantic focus. Abrusán (2012) gives convincing arguments to this effect with respect to parallel effects in Hungarian. However, the data discussed in the remainder of this section do not involve focus effects.

Many morphologically complex verbs show the same kind of projection effects.[10] Thus, utterance of any of the following would imply that the soup had been heated:

(27) I didn't reheat/overheat/underheat the soup.

The same holds for any sentence whose main verb is formed with one of these prefixes.

Admittedly, these are cases where we could in principle attach a conventional presupposition to a lexical form or bound morpheme. Indeed, these cases are reminiscent of examples familiar from the earlier literature on presupposition, which attributed presuppositions to lexical items such as *bachelor*, *spinster* and even *boy*, on the basis of the following patterns of defeasible implication:

(28) Smith isn't a bachelor. *Defeasibly implies*: Smith is an adult human male.
(29) My cousin isn't a boy anymore. *Defeasibly implies*: My cousin is male.

However, although we could attribute these presuppositional implications to the lexical meaning of the items involved, I find that there is something unsatisfying about saying, for example, that the observed implications of the sentences in (25) are due to a presuppositional specification associated with the verb *win*. It seems much more plausible that some general principle is involved.

The second class of cases, from which the *vote for Nader* example is drawn, involve adverbial or adjectival modification. The robustness of the implication in these cases varies, and there may be some disagreement about particular cases. Here are some representative examples:

(30) George isn't arriving on Wednesday.
    Defeasibly implies: George is arriving some time.

(31) George didn't leave late.
    Defeasibly implies: George left.

(32) I didn't wash the windows with soap.
    Defeasibly implies: I washed the windows.

(33) The baby didn't cry loudly.
    Defeasibly implies: The baby cried.

The pattern that is emerging here is the following: Given utterance of a sentence of the form O[p], where O is an entailment-canceling operator and where p entails $q_1 \ldots q_n$, there is a tendency for the utterance to be interpreted as if some subset of $q_1 \ldots q_n$ were outside the scope of O.[11]

---

[10] This observation is due to Tom Werner.

[11] In addition, whatever is "scoped out" tends to be interpreted as "backgrounded," as is usual with presupposition. I will have nothing further to say about this here.

I frame this generalization somewhat tentatively. The "projection" interpretations are often preferred, but are not required: hence, the reference to a *tendency* towards this interpretation. Moreover, not all entailments "project". Among the non-projecting entailments are those which can be identified by replacing the content of an argument position with a variable and existentially closing the result. Thus, an utterance of (34) a. would not normally be interpreted as in (34) b. or c., nor (35) a. as in (35) b.[12]

(34) a. I didn't wash the windows.
　　 b. I washed something, not the windows.
　　 c. Someone washed the windows, not me.

(35) a. I didn't sleep.
　　 b. Someone slept, but not me.

I do not yet know how projecting entailments should be correctly distinguished from non-projecting ones, and indeed there is a great deal about these cases that I do not yet understand. However, the bottom line is this: In a variety of cases, we find projection behavior where there is no obvious candidate to which to attach a presupposition. And this provides further evidence that some general, conversational principle is responsible for producing projection behavior.[13]

4 En Route to a General Principle

4.1 An Interpretation Principle Based on Logical Ordering

I have now given several arguments that a general principle for generating presuppositional inferences is needed. But what kind of conversational principle would do the job? I cannot yet give a complete answer to this question, but I will here suggest a route towards one. This is a route which I began to pursue only to discover that Stalnaker had already been there before me, albeit just a short way. In

---

[12] These readings become available with the right placement of focal stress.

[13] Non-restrictive relatives are another source of projecting entailments, as has been pointed out in the literature. (See Wilson and Sperber 1979 and Chierchia and McConnell-Ginet 1990.) Utterance of either (i) or (ii) commits the speaker to the claim that Jane was hiding in the closet.

(i) I didn't see Jane, who was hiding in the closet.
(ii) Did you see Jane, who was hiding in the closet?

Similarly, in sentences with secondary predication, the primary predication projects. Thus, utterance of either (iii) or (iv) implies that the speaker believes that Cleo came home.

(iii) Cleo didn't come home drunk.
(iv) Did Cleo come home drunk?

speculating as to possible conversational sources of presuppositional constraints, Stalnaker (1974, p. 205) suggests that:

> The propositions that P and that Q may be related to each other, and to common beliefs and intentions, in such a way that it is hard to think of a reason that anyone would raise the question whether P, or care about its answer, unless he already believed that Q.[14]

The idea is very intuitive. If I say, *Jane didn't stop smoking* I am showing an interest in the question of whether or not Jane stopped smoking; and why should I care about that question unless I already think that Jane was a smoker? Similarly, I would normally have no reason to wonder who George voted for unless I was already reasonably sure that he voted. The question is whether we can turn this seed of an idea into a robust explanation for the full variety of cases.

There are two issues that must be addressed in order to spell out Stalnaker's idea. The first of these is the issue of when an utterance counts as raising a particular question or as showing interest in the answer to it. Obviously, explicitly asking a question counts as raising it. But offering an answer to a question is also a way to raise it, as suggested above. To say either *Jane stopped smoking* or *Jane didn't stop smoking* is to offer an answer to the question *Did Jane stop smoking?* To offer an answer to a question is to indicate acceptance of that question as a topic of interest, and thereby, in some sense, to raise that question.[15] Note that because a sentence and its negation both serve to raise the same question, it is expected that they would share any presuppositions generated by virtue of question-raising. It is also expected that an explicit utterance of the associated question would give rise to the very same presuppositions. This observation points towards a solution to the projection problem.

The other operators over which presuppositions project are epistemic modals and conditionals. We must therefore ask whether sentences in which a proposition P is embedded under either of these operators can serve to raise the question whether P. It seems that they can. Consider a modal sentence such as:

(36) Perhaps Jane stopped smoking.

Utterance of this sentence indicates a willingness to discuss the question of whether or not Jane stopped smoking, and certainly indicates an interest in the answer to this question. The modal utterance could indeed be seen as opening an inquiry into the question.

Concerning the antecedents of conditionals, Groenendijk and Stokhof (1984) observe that a conditional can be used to raise the question which would be

---

[14] It is worth noting that Stalnaker himself here invokes speaker *beliefs*, not speaker *presuppositions*, as a condition on appropriate utterance.

[15] This is an idea which originates in the work of Carlson (1983), who suggested that sentence topics be treated as questions which are answered by utterances of that sentence. The idea has since been pursued by a variety of researchers. See, among others, Ginzburg (1994), van Kuppevelt (1995), and Roberts (1996); see also Simons (2000).

answered by an assertion of the antecedent. Thus, suppose that I ask whether Jane is likely to get sick, and you reply:

(37) If she doesn't stop smoking, she will get lung cancer.

Although (37) does not provide a direct answer to my question, it suggests a strategy for getting closer to an answer, namely, attempting to answer the question of whether Jane will stop smoking. The conditional thus serves to raise this question. Conditionals also raise the question of the conditions under which the consequent will hold. Thus, if I say:

(38) If Jane isn't warned of the dangers, she won't stop smoking

I appear to have an interest in whether or not Jane will stop smoking. The utterance could naturally be followed up with a discussion of what would happen if Jane didn't stop smoking.

It is perhaps unnecessary to establish that each of these individual cases can be seen as raising a question. Searle (1969, p.124) suggests that predication itself involves question-raising: to predicate a property P of an object o is to raise the question of whether or not P is true of o. Searle argues that this is true whatever sentence type the predication is part of, and whatever the speech act function of the utterance as a whole. He says, "the man who asserts that Socrates is wise, the man who asks whether he is wise and the man who requests him to be wise may be said to raise the question of his being wise (of whether "wise" is—or in the case of requests will be—true of him)." On this view, then, all of the cases discussed involve raising the question of whether or not Jane will stop smoking.

The conclusion so far is that it is plausible to think of utterances embedding a proposition P as raising the question of whether P is true. This addresses the first issue raised by Stalnaker's suggestion. We turn now to the second, and more difficult, issue. Recall that the idea we are pursuing is that two propositions P and Q might be related in such a way that it would make sense to raise the question whether P only if one already believed that Q was true. The question is this: Just what relation between propositions would impose such a constraint?

As a first stab, it seems likely that the relevant relation is (non-mutual) entailment. After all, in all of the cases that we have considered, the proposition that displays projection behavior is an entailment of the embedded clause. The significance of entailment seems straightforward: if P entails Q, then Q is necessary for the truth of P. So it would make sense to establish the truth of Q before wondering about P. This suggests the following interpretation principle:

Interpretation Principle (tentative)
Suppose that P entails but is not entailed by Q. A speaker who raises the question whether P indicates a belief that Q is true.

This principle is going to be too strong, but it gets some nice initial results. Let's look at these nice results, and then turn to the problems.

4.2 Applications

Let's begin with example (39):

(39) Jane doesn't know that George dislikes Cleo.

This sentence offers a direct answer to the question *Does Jane know that George dislikes Cleo?*, and so can be said to raise this question. The questioned proposition is the proposition that Jane knows that George dislikes Cleo. Some of the entailments of this proposition are listed in (40).

(40) a. George dislikes Cleo.
     b. George exists.
     c. Jane exists.
     d. Cleo exists.
     e. Jane is sentient.
     f. George is sentient.

All of these propositions escape the negation in (40), that is, utterance of (40) appears to presuppose these propositions.

There are some less desirable consequences. In addition to the propositions just considered, sentence (40) entails all tautologies, as does every other sentence. Hence, it follows from the Interpretation Principle that utterance of any sentence will give rise to the implication that the speaker believes all tautologies. I will set this point aside as a technicality. In fact, other treatments of presupposition are dogged by the problem of vacuous presuppositions. For example, context change treatments of presupposition based on Heim (1983) define the presuppositions of a sentence S as those propositions entailed by all contexts which admit S. But every consistent context entails all tautologies, and hence on this account too, every sentence presupposes all tautologies.

Let's work through one more well-behaved example:

(41) Jane didn't vote for Nader.

Here, the question raised is:

(42) Did Jane vote for Nader?

The questioned proposition is:

(43) Jane voted for Nader.

This proposition has the entailments listed in (44):

(44) a. Jane exists.
     b. Nader exists.

    c. Jane voted.
    d. Nader was on the ballot.[16]

Again, all of these survive the negation in (41), and emerge as presuppositional implications of its utterance.

4.3 Problems

As noted at the outset, the Interpretation Principle is too strong, predicting presuppositional implications which do not occur. We have already seen that V + complement structures do not show projection behavior, but the Interpretation Principle predicts that they will. Thus, for example, an utterance of the sentence *Jane didn't eat the sandwich* is incorrectly predicted to imply that the speaker believes that Jane ate something. These entailments, for a reason I do not yet understand, have a different status from those which arise by virtue of adverbial modification.

A different type of problem arises with the change of state verbs on which I have focused attention here.[17] Consider example (45):

(45) Jane didn't stop laughing.

Sentence (45) raises the question in (46) a., in which the proposition in (46) b. is questioned.

(46) a. Did Jane stop laughing?
    b. Jane stopped laughing.

According to the Interpretation Principle, a speaker who utters (45) should imply that she believes the propositions which are entailed by, but do not entail, (46)b. But these entailments include both of the propositions in (47).

(47) a. Jane was laughing immediately prior to the reference time.
    b. Jane was not laughing immediately after the reference time.

Now, the speaker could not coherently assume both of these propositions, given her assertion: she must be taken to assume at most one of these, and to be *denying* the other. The question is, on what basis does the hearer determine which is which? And in particular, why is it that in almost every case the speaker will be understood to assume that Jane was laughing prior to the reference time, and to be denying that the laughter ceased?

---

[16] This is not a logical entailment, but a contextual entailment, in the sense of Chierchia and McConnell-Ginet (1990). It follows given what we know about the world and about how voting works.

[17] This problem was pointed out to me by Roger Schwarzschild.

One idea which seems promising is that the two entailments in (47) do not have the same status. Specifically, (47)a. is a *precondition* of the change of state described by *stop laughing*. The notion of precondition I intend to invoke here is an ontological one. I conceive of preconditions as holding of events: a precondition on an event E is a condition which must be satisfied by the world in order for the event E to take place. Any event of change of state will have as a precondition that the changing entity be in the appropriate start state for that change of state.[18] The observation that (47)a. is the usual presupposition of (45) suggests that raising the question of whether a particular change of state took place (or will take place) gives rise to the implication that the speaker believes the relevant start held (or will hold).

While this idea is suggestive, it raises further questions. Why should the precondition have a privileged status relative to other entailments? Can the notion of an ontological precondition be defined for event types other than change of state? Is it the case that all propositions which show projection behavior are characterizable as preconditions? These questions currently remain unanswered.

4.4 Section Summary

In this section, I have tried to cash out an idea of Stalnaker's to account for the "projection" of entailments over entailment-canceling operators in a range of cases. The idea, as I have interpreted it, is that an utterance embedding a proposition P may be understood as raising the question whether P; and P may be related to a further proposition Q in such a way that it would make sense to raise the question whether P only if Q were already believed to be true. Hence, utterance of the embedding sentence would give rise to the implication that the speaker believes Q.[19]

In Sect. 3.1., I argued that the relevant range of utterance types indeed can be seen as raising questions. In Sect. 3.2., I suggested that the relevant relation between propositions is simply entailment. The discussion in Sect. 3.3. makes clear that this suggestion needs further refinement. However, the idea that the projection facts may be amenable to a treatment in terms of logical ordering remains attractive.

---

[18] This notion of preconditions on events is to be distinguished from preconditions on the truth and falsity of sentences. I claim, for example, that Jane's laughing immediately before t is a precondition on the event of her stopping laughing. However, this claim does not entail any commitments as to whether the sentence *Jane didn't stop laughing* (at t) is true, false, or truth valueless in a situation in which the precondition is not met.

[19] (2011) For a different way of utilizing implicit questions in an account of projection, see Simons et al. (2010)

5 Conclusion

What we have established in this discussion is that at least some presuppositions must have a conversational basis, that is, they must be inferences derivable from some general conversational principle or principles. I would like to conclude by considering what consequences this conclusion has for the general treatment of presupposition.

On the view of presupposition now most standard in the literature, presuppositions are thought to be propositions which must be entailed by the presumed common ground of the discourse participants. However, if at least some presuppositions are derived by the kind of mechanism which gives rise to other conversational inferences (i.e. conversational implicatures), then it is more appropriate to view them as propositions which the addressee can infer the speaker to believe on the basis of what the speaker has said, plus the assumption that the speaker is behaving cooperatively. On this picture, the derivation of a presupposition may require speaker and addressee to share certain assumptions (e.g. that the speaker is behaving cooperatively), but the presuppositions themselves are neither required nor expected to be entailed by the common ground.

Moreover, on this picture, presuppositions are not attached to atomic clauses, but are inferences derivable from the utterance as a whole, given the conversational situation. This raises a question about algorithmic treatments of presupposition projection, which are predicated on the assumption that presuppositions are locally generated.

This discussion also suggests the possibility that presuppositional phenomena are not homogeneous. If the hypothesis presented here is correct, then the presuppositions of, say, change of state sentences are derived very differently from the presuppositions generated by *even*, *too*, and *again*, which plausibly have a conventional source. On the other hand, this picture of (some) presuppositions as conversationally generated inferences suggests the possibility of unifying the treatment of these presuppositional data with that of other types of "preferred readings," such as those illustrated in Sect. 2.

Current treatments of presupposition in the dynamic semantics literature have a great deal of appeal. They are formally rigorous, make clear predictions, and are very successful with respect to the range of data which they cover. Unfortunately, these accounts disassociate the question of presupposition projection from the prior question of the sources of presuppositions. The investigation of this question leads us down the notoriously swampy path of conversational inference and interpretative strategies. However, I think that we will not really understand presupposition until we see where this path leads.

References

Abusch, D. (2010). Presupposition Triggering from Alternatives. *Journal of Semantics, 27*(1), 37–80.

Abrusán, M. (2012). A note on Quasi-presuppositions and focus. *Journal of Semantics*. doi:10.1093/jos/ffs010.

Atlas, J. D. (2005). *Logic, meaning and conversation*. New York: Oxford University Press.

Atlas, J. D., & Levinson, S. (1981). It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In P. Cole (Ed.), *Radical pragmatics* (pp. 1–62). New York: Academic Press.

Bach, K. (1999). The Myth of conventional implicature. *Linguistics and Philosophy, 22*(4), 327–366.

Boër, S. E., & W. G. Lycan. (1976). *The Myth of semantic presupposition*. Ohio State working papers in Linguistics (No. 21). Columbus: Department of Linguistics, Ohio State University.

Carlson, Lauri. (1983). *Dialogue games: An approach to discourse analysis*. Boston: D. Reidel Publishing Company.

Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and grammar*. Cambridge, MA: MIT Press.

Geurts, B. (1994). Presupposing. PhD. *Dissertation*, Universität Osnabrück.

Ginzburg, J. (1994). An update semantics for dialogue. In H. Bunt (Ed.), *Proceedings of the 1st International Workshop on Computational Semantics*. Tilburg: ITK, Tilburg University.

Grice, H. P. (1967). *Logic and conversation*. Harvard: the William James Lectures. [Published in Grice (1989)].

Grice, H. P. (1981). Presupposition and conversational implicature. In P. Cole (Ed.), *Radical pragmatics* (pp. 183–197). New York: Academic Press.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Groenendijk, J., & Martin S. (1984). Studies on the semantics of questions and the pragmatics of answers. PhD. *Dissertation*, University of Amsterdam.

Heim, I. (1983). On the projection problem for presuppositions. *Proceedings of the West Coast Conference on Linguistics, 2*, 114–126.

Kadmon, N. (2001). *Formal pragmatics*. Maldan, MA: Blackwell.

Karttunen, L., & Stanley P. (1979). Conventional implicature. In C.-K. Oh & D. A. Dinneen (Eds.), *Syntax and semantics*: Presupposition (Vol. 11, pp. 1–56). New York: Academic Press.

Kay, P. (1992). The inheritance of presuppositions. *Linguistics and Philosophy, 15*, 333–379.

Kempson, R. (1975). *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics, 31*, 109–147.

Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S., & Annamalai, E. (1992). Why presuppositions aren't conventional. In R. N. Srivastava et al. (Ed.), *Language and text: Studies in honor of Ashok R. Kelkar*. Delhi: Kalinga Publications.

Liberman, M. (1973). Alternatives. *Proceedings of the Chicago Linguistics Society, 9*, 346–355.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics* 49. (pp. 91–136). Columbus: Department of Linguistics, Ohio State University. Reprinted in Semantics and Pragmatics 5, Article 6:1–69. doi: http://dx.doi.org/10.3765/sp.5.6

Searle, J. R. (1969). *Speech acts*. New York: Cambridge University Press.

Simons, M. (2000). *Issues in the semantics and pragmatics of disjunction*. New York: Garland.

Simons, M., Tonhauser, J., Beaver, D., & Roberts C. (2010). What Projects and Why. *Proceedings of Semantics and Linguistic Theory (SALT), 20*, 309–327.

Stalnaker, R. (1974). Pragmatic presuppositions. In Milton K. Munitz & Peter K. Unger (Eds.), *Semantics and philosophy*. New York: New York University Press.

Tonhauser, J., Beaver D., Roberts, C., & Simons M. (2011). *Toward a taxonomy of projective meaning*. Talk given at Annual Meeting of the Linguistic Society of America, Pittsburgh, PA.

Wilson, D. (1975). *Presupposition and non-truth conditional semantics*. New York: Academic Press.

Wilson, D., & Sperber, D. (1979). Ordered entailments: An alternative to presuppositional theories. In C. K. Oh & D. Dinneen (Eds.), *Syntax and semantics* (Vol. 11, pp. 299–323). presupposition New York: Academic Press.

The Salience Theory of Definiteness

Klaus von Heusinger

Abstract The salience theory of definiteness combines the best features of the uniqueness theory and the familiarity theory to a novel concept of definiteness. A definite expression refers to the most salient element of a given set. Thus, this theory does not suffer from the notoriously problematic uniqueness condition nor from the often too globally interpreted familiarity condition. The paper provides the theoretical and empirical foundations for the salience theory of definiteness and illustrates its range by successfully analyzing different uses of definite noun phrases.

1 Introduction

The concept of definiteness in natural language is of special interest because it seems to be pragmatic in nature but it has semantic impact. The analysis of definite expressions exhibits some aspects of the fuzzy borderline between semantics and pragmatics and the interaction between the two areas. In this paper, I will examine four semantic theories about definiteness with particular view on English. I conclude that the pragmatic concept of "salience" is the underlying principle for definiteness. However, no theory has given a formal account of this pragmatic principle. I show that choice functions provide the adequate means to reconstruct salience in a formal theory. They are functions that assign to each non-empty set one of its elements. In this formal approach the pragmatic principle of salience gets its semantic reconstruction, which yields a unified account of the semantics of definite noun phrases and pronouns.

K. von Heusinger (✉)
Universität zu Köln, Cologne, Germany
e-mail: Klaus.vonHeusinger@uni-koeln.de

The paper is organized in the following way: In the second section I introduce five different groups of definite expressions, namely proper names, definite NPs, demonstratives, personal pronouns, and possessive constructions. In the third section, I focus on definite NPs as the most complex kind of definite expressions and discuss the relevant contexts where they are used: the anaphoric linkage, the relational dependency, the situational salience, and the unique case.

In the fourth section, I shortly sketch three semantic theories of definiteness. Each of the theories focuses on one of the typical contexts of definite expressions: Russell's Theory of Descriptions focuses on uniques, Kamp and Heim's familiarity theory takes the anaphoric use as fundamental, and Löbner's relational approach bases definiteness on relational dependencies. However, none of these three theories gives a complete picture of all uses of definite NPs. Therefore, the more general salience approach is presented in the fourth section. In this approach, the context crucially contributes to the interpretation of the definite NP by forming a salience hierarchy among the potential referents. It is assumed that each context can be associated with an ordering among the elements of subsets of the domain of discourse. This ordering reconstructs the intuitive idea of a salience hierarchy. The three historical sources of this salience theory are outlined: Lewis' semantic criticism of Russell, the linguistic conception of the Prague School and the investigation of AI researchers. However, there has not been any attempt to formalize the principle of salience.

In the sixth section, I give a formal representation of the concept of salience by means of context dependent choice functions, which pick out from a set one of its elements or a "representative". Due to this formal account of the pragmatic principle of salience it becomes possible to reconstruct definiteness in the logical representation of natural expressions. It will be shown that the developed formalism can uniformly describe all four different uses of definite NPs.

2 Definite Expressions

In a pretheoretical definition, a definite singular expression unambiguously denotes or refers to one object, i.e. the object can be identified as the only one that is denoted by the expression. The fixed reference of a definite expression depends on different grounds: it can be determined by lexical material, by semantic rules or by pragmatic strategies. Traditionally, proper names, definite NPs, demonstratives, personal pronouns and possessive constructions are regarded as definite. In this section I will give a short overview of these types on the example of English expressions and discuss some of their properties. I confine the presentation to expressions referring to singular countable objects.

2.1 Proper Names

A proper name is a prototypical definite expression. It refers to exactly one individual, namely the bearer of the name. The reference is purely conventional since no internal part of the expression points or gives any relation to its bearer. Despite their treatment as constants in formal semantics, proper names are highly context dependent as the list (1) shows. There are many Baraks and Angelas and there is even more than one Barak Obama and one Angela Merkel. However, these problems of proper names should not concern us here too much.

(1a) *Barak*
(1b) *Angela*
(1c) *Barak Obama*
(1d) *Angela Merkel*

2.2 Definite NPs

Definite NPs (here short for "definite descriptions") as in (2) refer to their objects not by convention but due to their descriptive content. Since there is only one person who has been the first man on the moon the definite NP *the first man on the moon* refers to exactly that man. This behavior of definite NPs caused their use in mathematics and epistemology for definitions. In the context of a definition, a definite NP (or definite description, as this term is more common in this literature) refers to the unique object that satisfies the descriptive material. Russell's Theory of Descriptions is based on such cases and, therefore, entails the uniqueness condition for definite descriptions. However, in normal natural language discourse we find definite NPs whose descriptive material can be satisfied by more than one individual, like *the sun, the university, the table* etc. Such NPs are sometimes called incomplete definite descriptions. They refer uniquely to one object due to their descriptive material and further information, like our shared background knowledge about the astronomical system of the earth, or contextual information about the place and time of utterance.

(2a) *the first man on the moon*
(2b) *the sun*
(2c) *the university*
(2d) *the table*

2.3 Demonstratives

Demonstrative expressions include demonstrative pronouns or demonstrative NPs. Demonstrative pronouns like *this* or *that* refer to an object only if the linguistic

utterance is accompanied by a non-linguistic demonstration or ostension. They form a borderline case of the semantic-pragmatic interface since they do not determine the referent of the expression by themselves but rather indicate that an additional demonstration is to be undertaken. Like deictic expressions (*here, now* etc.) demonstrative pronouns have a very impoverished lexical content. They express the *here-there*-distinction in English and can indicate gender, case and number in other languages. Demonstrative NPs like *this man, that book* etc. consist of a demonstrative and a descriptive part. Thus, they identify their referent by combining a demonstrative action with descriptive information about the referred object.

(3a) *This* is my teacher.
(3b) I take *that.*
(3c) *This man* is very late.
(3d) I bought *that book*.

It is noteworthy, that in all Indo-European languages that have a definite article the form of the article has developed out of the demonstrative pronoun. We come back to this point later.

2.4 Personal Pronouns

The use of personal pronouns is traditionally analyzed either as deictic or as anaphoric. In absence of any linguistic context, the pronoun *he* in (4a) most likely refers to an object that must be in some way prominent in the context or "easy to access". This deictic interpretation of the pronoun is licensed if the pronoun is accompanied by a demonstration or if the non-linguist context contains some prominent or salient object. Background knowledge may play an important role, too. A pronoun is interpreted anaphorically, if it refers to an object that has been already introduced into the discourse, as in (4b). The analysis of pronouns is crucial for any theory of reference. Therefore, examples similar to (4c) and (4d) have been discussed since classical times illustrating the interaction with other expressions and constructions, like conditionals.

(4a) *He* will be late again.
(4b) A man walks. *He* whistles.
(4c) If a man is in Athens *he* is not in Rhodes.
(4d) If a man has a donkey *he* beats it.

2.5 Possessives

Possessive constructions like *John's car* consist of a common noun or head noun (*car*) that is preceded by a definite expression or a modifier, like a pronoun, a

proper name or a definite description, but not by a demonstrative pronoun. Both expressions are conjoined by the possessive "s" which indicates the definiteness of the whole expression. Personal pronouns and the possessive "s" merge to possessive pronouns as in (5a). The possessive expression denotes exactly the object that fulfills the property that is expressed by the common noun (cf. *car*) and that further stands in a certain relation to the object that is denoted by the modifier (cf. *John*). This relation can be determined by the lexical material of the head noun if it is a functional concept, like *father*. Since for each person there exists exactly one father, an expression of the kind *X's father* denotes always one person. If the head noun does not denote a functional concept, but rather a sortal one as in (5b) the relation is usually the possessor relation. *John's car* is that object that is a car and has a certain relation to John, which is probably the car that John owns. Possessive constructions of this kind should not be mixed with constructions of the kind *the car of John*, because the definiteness in the latter case comes from the definite article and not from the possessive relation.

(5a) *his claim*
(5b) *John's car*
(5c) *Lisa's father*
(5d) *the man's bag*

In the following I will concentrate on the use of definite NPs in natural language since they form the most complex group of definite expressions. Definite NPs need for their reference not only descriptive content but also contextual information of a different kind. This combination of descriptive content and contextual information makes their analysis not only difficult and controversial but also a very challenging enterprise for semantic analysis. Definite NPs exhibit an interaction between the different mechanisms and, hence, call for general principles explaining the way they are linked with their referents.

3 The Uses of Definite NPs

There are several different uses of definite NPs and even a more subtle categorization of these uses. We will start with the overview that was presented by Christophersen (1939). His work on articles is very prominent for two reasons. Firstly, he not only summarizes the descriptive state of art, but also tries to give a more abstract categorization of definiteness. And secondly, he was one of the first who reacted to Russell's Theory of Descriptions. It is interesting to note that all approaches except Russell's theory refer to Christophersen's work as precursor of their ideas.

Christophersen (1939, 29) distinguishes between *the explicit contextual, the implicit contextual* and *the situational basis* use for definite NPs. According to the contemporary terminology in the literature we will call these three main groups

anaphoric, relational and *situational* use, respectively. I discuss a fourth group of unique uses though Christophersen does not recognize it as a proper use, but rather subsumes it under the three other uses (see Hawkins 1978 and Lyons 1999 for further types of uses).

3.1 Anaphoric Linkage

In the anaphoric use (Christophersen's explicit contextual), the definite NP refers to an object that is explicitly introduced by the linguistic context. Thus, definiteness is based on the principle of coreference.

(6) Once upon a time, there was a king, … and *the king* …

The object is introduced by the indefinite expression *a king* and then the reference is picked up by the definite NP *the king*. It could be picked up by the pronoun *he* or by the demonstrative NP *that king,* as well. However, there are differences in application of anaphoric pronouns, demonstratives or definite NPs. One principle concerns the distance between the antecedent and the anaphoric expressions: The further the distance between the first mention and the resumption, the more likely it is to use the definite NP.

3.2 Relational Dependency

In the relational (associative, implicit contextual) use, the definite NP refers to an object due to another already mentioned object in the discourse. However, it does not refer to the same object like in the anaphoric linkage discussed in the last subsection. The definite NP *the author* receives its referent not by coreference with an antecedent expression, but rather by a significant association relation to the antecedent *a book*.

(7a) I read a book. I cannot remember *the author*.
(7b) I bought a new car. I had to change *the motor*.
(7c) I bought a new car. ?I had to change *the wheel*.

The definite NP *the author* does not pick up the referent of another expression, but it refers to an object that is unequivocally linked to a just mentioned object. This is possible due to the relational (or functional) nature of the expression. An author has to be an author of something, probably a book. The definite NP expresses two things: its descriptive material delimits the class of potential referents and then establishes a relation to a mentioned object in discourse. In the example (7a) this is done by the common noun *author*, which expresses the

relational concept between a person and a written text such that the person has produced the text.

The link between a definite NP and an expression it is related to must be in some way unique. Since nothing else than the relation is expressed, the relation itself must unequivocally determine exactly one object. Otherwise one has to use the indefinite article. The sentence *I bought a new car. I had to change the wheel* is awkward without any further context. Therefore, functional expressions like *the father* are preferred to relational expressions like *the wheel*. One can think of such relational definite NPs as abbreviated possessive constructions. *The author* stands for *the author of the book* or *its author* etc. It seems that the definite article stands for the possessive construction discussed in Sect. 2.5 and could be easily replaced by the appropriate possessive pronoun. However, the definite NP cannot be substituted by a pronoun or by a demonstrative expression as illustrated in (7d).

(7d) I read a book. ?I cannot remember *this author/him.*

The relational concept of an definite NP must be lexically determined, whereas possessive construction can be used in a wider range of contexts. The relational property need not be lexically expressed, but can also be given by the context.

3.3 Situational Salience

Definite NPs that are neither relational nor just mentioned can be used if the situation or the non-linguistic context delivers additional information to single out the referent.

(8a) *The island* is beautiful.
(8b) *The sun* shines.
(8c) *The talk* will start soon.
(8d) *The train* left two minutes ago.

The isolated sentences in (8) can only be uttered felicitously if the non-linguistic context specifies which object is uniquely meant. This non-linguistic context can consist in the shared background knowledge or in the actual circumstances. The latter should be the case when uttering (8a). If we stand at the University of Konstanz and look around the lake uttering (8a) we mean the only visible island, namely the Mainau. This use is sometimes called deictic or demonstrative and has a special relation to demonstrative NPs (cf. 2.3), as the definite article can be substituted by the demonstrative pronoun *this* or *that*. It is interesting to note that in all Indo-European languages the definite article is derived from the demonstrative pronoun. Therefore, Lyons (1977, II, 671ff.) assumes that every definite NP contains a deictic element. This idea will be formalized in Sect. 6. However, there are some cases in which we cannot replace the definite article by the demonstrative: The definite NP *the sun* in (8b) refers

uniquely due to our background knowledge that there is only one sun (in the relevant circumstances). In this case we cannot replace the definite article with the demonstrative pronoun.

3.4 Uniques

Despite the fact that uniques do not form an independent class of definite NPs in Christophersen's classification they should be discussed here. Uniques are nouns whose lexical content is such that only one object can fit it. Thus, we find such nouns in the latter two groups of definite NPs discussed above: A unique can consist in a noun that expresses a functional concept, i.e. a concept that gives exactly one value for each argument. It can also consist in a complex nominal expression that due to its meaning refers only to one object (in the relevant context) like *the first man on the moon*. *The sun* refers uniquely because there is only one sun in our solar system. Or one can argue that *the sun* stands for the relational concept of *sun of something* and given the case that all of us live on the same planet, the sun of this planet refers to the only sun we have. Finally, the definiteness could be reduced to the principle of salience as well: we refer to the sun with "the sun", because it is the most salient sun. Uniques are used for definitions and have got, therefore, a special place in logic and epistemology. Certainly, in formal semantics their role is overestimated because they can be captured by the other classes. In the remainder, we will disregard uniques as an independent class and consider only the other three classes.

These uses of definite NPs are not independent of each other and sometimes it is hard to classify a particular use. They often overlap and a definite NP refers uniquely because there are linguistic and non-linguistic pieces of information given in distinct ways. The question that arises is whether there is one basic use or function of the definite NP and how we can describe it. In the next section we will see that different approaches take different uses as primary and try to define the other uses in terms of the chosen one.

4 Three Theories of Definiteness

We have mentioned above that definiteness is a pragmatic principle that has a semantic impact. An analysis of definite expressions is a central task for every semantic theory. In this section, I will characterize three alternative theories of definiteness: Russell's classical Theory of Descriptions, Heim and Kamp's Familiarity theory and Löbner's relational approach to definite expressions. In Sect. 5, I introduce the salience approach which is based on the situational salience of the referred object. Though the theories are confronted with the multiple uses of definite NPs discussed in the last section, they assume that there is only one

underlying meaning of the definite NP that can be found in all of its uses. However, each of the theories chooses a different use of definite NPs as the primary one and gives an adequate analysis of this use. The analysis is then extended to the other uses. Further arguments for each of the discussed theories are gained if other definite expressions, as discussed in Sect. 2, can be described in the same format or according to the same principles. The first three theories mentioned are successful in their primary area, but they cannot convincingly describe other uses of definite NPs. Therefore, a more general approach will become necessary.

The Russellian Theory of Descriptions is the clearest and the best understood approach. It gives a clear formal representation of definite and indefinite NPs as quantifier phrases. In this way certain ontological and epistemological problems with non-existent objects are solved. The definite article expresses the uniqueness condition, either as an assertion (Russell 1905) or as a presupposition (Frege 1892; Strawson 1950, and most contemporary theories, as presented in Abbott 2004; Ludlow 2007 or Heim 2011). Definite NPs are represented as quantifier phrases, and typical ambiguities can be explained in terms of quantifier interaction and scope. The problematic uniqueness condition is amended by a rule of domain restriction, which is also necessary for the interpretation of other quantifiers. Still, this approach is conceptually and technically grounded on the unique use of definite NPs, and not easily transferable to other uses of definite NPs. The uniqueness condition poses a general problem, and several arguments show that definite and indefinite NPs are not quantifier phrases but terms (e.g. Löbner 1985; Egli 1991; von Heusinger 1997; but see for arguments in favor of the quantifier view Abbott 2004; Ludlow 2007; Heim 2011).

All of the three following theories, namely Heim and Kamp's familiarity theory, Löbner's functional approach and the salience approach, introduced in the next section, can be understood as a reaction to the very strong Russellian assumptions. It is noteworthy that all of them refer in one way or another to Christophersen's original work and claim that they spelled out his original ideas. Heim and Kamp's approach focuses on the anaphoric use of definites in a discourse. This view gave rise to the new generation of dynamic semantic theories, which do not analyze isolated sentences, but an entire discourse. Heim and Kamp's familiarity theory claims that there is a uniform representation of definite and indefinite NPs as open sentences with free variables at the additional level of discourse representation. The indefinite article indicates that a new variable has to be introduced whereas the definite article expresses that the open sentence has to be linked to an already introduced variable, i.e. to a familiar variable. Thus, Heim and Kamp claim to have adapted Christophersen's familiarity on the level of discourse representation. Anaphoric pronouns can be described by means of the same formalism and for deictic expression the formalism can be extended in an acceptable way. In this view, definite and indefinite NPs are not represented as quantifier phrases but as singular terms, which nevertheless can be bound by higher operators.

Löbner's relational approach occupies a position between the two former theories. On the one hand he focuses on the relational use of definites like Russell (i.e.

narrow scope definites) and rejects Heim and Kamp's approach that concentrates on the anaphoric use. On the other hand he refuses all three of Russell's claims, namely (1) that definite NPs are quantifier phrases, (2) that there is a uniform semantics of definite and indefinite NPs, and (3) that uniqueness is a property of the descriptive material of a definite NP. He rather takes definite NPs as terms like proper names, whereas indefinites are quantifier phrases. Instead of Russell's uniqueness condition he uses Christophersen's view according to which definite NPs refer unambiguously. This fits well into the formal representation of definites as terms since a term refers uniquely *per definitionem.* In contrast to Heim and Kamp's approach, definites do not express a global definiteness (wide scope) but rather a local definite relation. A global relation can be constructed by chains.

4.1 Russell's Theory of Descriptions

Russell takes the uniques as the prototype of definite NPs or definite description. His uniques are generally functional concepts, like *the center of the solar system* or *the father of Bertrand Russell*. He does account for context dependencies, which do not play any role in mathematics and logic. Furthermore, context has no place of its own at the formal level of analysis in his conception of a language as formal system. However, Russell's Theory of Descriptions is a very common view among formal semanticists since it is a well developed theory, which fulfills logical, ontological and epistemological standards. Russell (1905) represents the definite article with the "iota operator" as in (9a), which is contextually defined as a complex quantifier phrase consisting in the uniqueness condition, the existential condition and the matrix predication, as spelled out in (9b). The iota operator can represent complex possessive constructions, like in (10):

(9)   The father of Bertrand Russell was English.
(9a) English(ιx Father_of(b, x))
(9b) $\exists$x [Father_of(b, x) & $\forall$y [Father_of(b, y) $\rightarrow$ x = y] & English(x)]
(10) Bill's father's dog's basket = the basket of the dog of the father of Bill
(10a) ιx [Bx(ιy (Dy ιz(Fzb)))]

Neale (1990) gives an excellent defense of the Russellian approach and extends it to more sophisticated problems. Especially, he successfully exploits the Russellian iota terms for describing functional dependencies as in (11). He further integrates the treatment of so called "E-type pronoun", i.e. cross-sentential pronouns, into this formalism by using complex iota terms like in (12b):

(11) Every man loves the woman that raised him.
(11a) $\forall$x Mx $\rightarrow$ Lx(ιy)(Wy & Ryx))
(11b) $\forall$x Mx $\rightarrow$ $\exists$y [(Wy & Ryx) & $\forall$z [(Wz & Rzx) $\rightarrow$ z = y] & Lxy]
(12) A man walks. He whistles.

(12a) A man walks. The man who walks whistles.
(12b) $\exists x \; [Mx \; \& \; Wx] \; \& \; Wh(\iota x \; [Mx \; \& \; Wx])$

However, there seem to be some problems with Russell's theory that concern the uniqueness condition: it is too strong for natural language descriptions. And even if we assume domain restrictions, as for other quantifiers, it is still an open question, whether we can restrict the relevant domains such that the definite NP always corresponds to exactly one referent that fits its descriptive content. An additional problem is that the difference between the definite and the indefinite article lies only in this problematic uniqueness condition. Finally, in this analysis definite NPs do not belong to the class of referring terms like proper names and pronouns, but to the class of denoting phrases like quantifiers.

4.2 Heim and Kamp's Familiarity Theory

With the beginnings of the eighties a new generation of semantic theories was developed (Kamp 1981; Heim 1982) that uses an additional level of representation. This representational level was motivated by linguistic investigation into anaphora (Karttunen 1976), by research of artificial intelligence into the representation of discourse (e.g. Webber 1983) and by philosophical investigations (cf. Stalnaker 1978). There are two main aims of these new theories. One is to extend the semantic representation from the sentence to discourse phenomena. The discourse representation level should model not only the meaning of a sentence, but also the information of a whole discourse. The second aim is to represent definite and indefinite NPs in a uniform way as discourse referents that 'live' on the discourse representation level, but not necessarily in the real world. Hence, ontological problems with non-existent objects can be solved by describing them as discourse referents with a short 'livespan'. The emphasis of these theories lies in the investigation of discourse anaphora that carry on certain information from one sentence to the following sentences. This is also the beginning of a dynamic view of meaning.

Thus, the most prominent discourse phenomenon that is treated in this approach is the anaphoric linkage between sentences. The core meaning of definite NPs is seen in the anaphoric use. An indefinite NP introduces a new discourse referent into the discourse representation, whereas a definite NP is anaphorically linked to an already introduced or 'familiar' discourse referent. This view on definiteness is traced back to Christophersen (1939) and his familiarity theory, which says that an indefinite NP introduces a new referent and a definite NP refers to an old or familiar referent. However, Heim and Kamp transfer this principle to the level of discourse representation to avoid ontological problems. The indefinite NP *a man* in (13) introduces a new discourse referent d_1 in (13a). The definite pronoun *he* in the second sentence of (13) introduces the discourse referent d_2 which is identified with the first one in (13b) expressing the anaphoric relation in (13). Discourse

referents can also be bound by other operators like conditionals in (14) and (15) which are interpreted as universal adverbial quantifiers (cf. Lewis 1975):

(13) A man walks. He whistles.
(13a) $\{d_1 \mid M(d_1) \ \& \ W(d_1)\}$
(13b) $\{d_1, d_2 \mid M(d_1) \ \& \ W(d_1) \ \& \ d_1 = d_2 \ \& \ Wh(d_2)\}$
(14) If a man is in Athens he is not in Rhodes.
(14a) $\forall \{d_1 \mid M(d_1) \ \& \ A(d_1)\} \ \{d_1 \mid \neg Rh(d_1)\}$
(15) If a man has a donkey he beats it.
(15a) $\forall \{d_1, d_2 \mid M(d_1) \ \& \ D(d_2) \ \& \ O(d_1, d_2)\} \ \{d_1, d_2 \mid B(d_1, d_2)\}$

In this analysis indefinite NPs are not scope-bearer by themselves, but get scope assigned by some other operator, such as the conditional in (14) or a text-level existential operator as in (13). Definite NPs get wide scope, i.e. at least the scope over the sentence they are constituents of. This mechanism explains the anaphoric use of definite NPs. It shows how the information that is needed for establishing anaphoric linkages is carried on in discourse.

However, such theories face problems with the other uses of definite NPs. The situational use is explained by the assumption that in such cases non-linguistic information may introduce discourse referents to which definite NPs can be linked. The sentences listed in (8) can only be uttered if the non-linguistic context delivers an object that introduces a discourse referent. This mechanism allows for an analysis of both deictic pronouns and definite NPs by creating one domain for linguistic and non-linguistic information. However, there may be a problem of delimiting the non-linguistic information that is needed for the semantic analysis.

Relational definite NPs cause a different problem. They can be bound by a higher operator in the same sentence (cf. Heim 1982, 245ff).

(16) Every man saw the dog that barked at him.

This problem is generally solved by introducing a new kind of rule, namely accommodation according to Lewis (1979). An accommodation is possible if the sentence cannot be interpreted felicitously. This may be the case if one processes a sentence and comes across a definite NP without an antecedent. Then the alternative consists in rejecting the whole sentence or in accommodating it. If one has good reasons to think that the given sentence is felicitous, one has to apply accommodation. The accommodation rule says that one can add a new property that stands for a functional concept whose argument must already be given. In the following sentence, the definite NP *the dog that barked at him* introduces a new complex $D(x) \ \& \ B(x, y)$ for the functional concept *dog that barked at y* and the argument y refers to the discourse referent d_1 that is introduced by the NP *a man*.

(17) A man saw the dog that barked at him.
(17b) $\{d_1, d_2 \ [D(x) \ \& \ B(x, y)] \ M(d_1) \ \& \ d_2 = \iota x \ [D(x) \ \& \ Bark(x, d_1)] \ \& \ S(d_1, d_2)\}$

To sum up, we have seen that the representational approach with the familiarity principle explains the anaphoric use of definites in an elegant way. However, for

the situational and relational use, some modifications are necessary. The situational use is explained by stipulating that non-linguistic context can establish discourse referents as well. In this way, the deictic use of definites in general (i.e. deictic NPs and deictic pronouns) gets a uniform analysis together with the anaphoric use (of NPs and pronouns). The most obvious problem with this stipulation is that it is difficult to delimit the non-linguistic information that is necessary. The relational use of definites is explained by accommodation, i.e. a pragmatically determined repair of semantic procedures. If the semantic analysis does not find an antecedent for a definite expression, one may introduce the relational concept such that one argument is filled by an antecedent expression. This move to save the theory is not unproblematic since the restriction of this very powerful rule is not obvious. And if one needs such powerful mechanism, the question arises whether this mechanism is only a repair mechanism or whether it represents the real character of the definite NP. This position is discussed in the next section.

4.3 Löbner's Relational Approach

Löbner (1985) takes the complementary position to Heim and Kamp, namely that the prototypical use of definite NPs is not the anaphoric but the relational or functional use. However, he differs also dramatically form the Russellian approach. According to Löbner the definite article has no lexical meaning, but just indicates the way the reference is established, namely that the expression refers non-ambiguously.[1] This function was already defined by Christophersen. "I agree with Christophersen that the crucial feature of definiteness is non-ambiguity of reference" (Löbner 1985, 291).[2] It means that a definite NP cannot be represented by a quantifier phrase, but must be reconstructed by a term, like proper names and pronouns. The Russellian case, where the definite NP refers due to its descriptive material that uniquely denotes an object, comes out as a special case of unambiguous reference.

Löbner (1985, 299) distinguishes between semantic and pragmatic definites. "Semantic definites refer unambiguously due to general constraints; Pragmatic definites depend on the particular situation for unambiguous reference." Thus, he merges the anaphoric use and the situational (or deictic) use into one class, which he coins pragmatic definite. The relational use becomes the semantic definites and the paradigm of definite NPs. "An NP is semantic definite if it represents a functional concept, independently of the particular situation referred to" (Löbner 1985, 299). An expression is inherently functional if it needs a further argument to refer to an object. This argument can be implicitly expressed by the situation like

---

[1] Löbner (1985, n8) notes that the German word *eindeutig* expresses this very accurately.

[2] He further rejects the claim of Heim and Kamp to have reformulated Christophersen's familiarity theory, but argues that they have deviated from the original idea.

weather, prime minister, post office etc. and like proper names. This is what we have called the larger situational use of the definite article. The argument can also be explicitly expressed by an overt object argument like *father of_, capital of_,* The argument slot need not be filled by another definite expression. It can also be filled by an indefinite or quantificational expression:

(18) The mayor of a small town in Wales.
(19) Every man loves his wife.

Examples like these suggest that the definiteness has not to be considered as a property of (global) reference (cf. Lyons 1977) but as a local property of the link between the head and its argument. (18) means that there is a definite relation from the town (whatever it is) to its mayor. Löbner confirms this view of definiteness by the following class of examples, which he calls configurational use.

(20) He was the son of a poor farmer.
(21) He put his hand on her knee.

Again, the definiteness expresses a local determined relation between two arguments. It expresses neither a global definite reference nor any uniqueness condition of the definite term.

Pragmatic definites consist in anaphoric and deictic uses of definites. Löbner explains their use in terms of functional concepts. A pragmatic definite is a function from an established situation to an (unique) object. He develops some kind of discourse network to show that definite relations exist in local relation. However, Löbner does not give any formal definition of what a discourse consists of and which parts influence the definite NPs. Since he focuses on the local effect of definiteness he cannot account for the discourse phenomena of definite NPs. Therefore, he regards anaphora only as an epiphenomena and not as the central use of definite NPs.

5 The Salience Theory of Definiteness

Neither Russell's Theory of Description, nor Heim and Kamp's discourse representation or Löbner's relational view can analyze all uses of definite NPs. Therefore, a more general approach is necessary, which takes the situational use as the central one of definite NPs. The salience approach essentially incorporates contextual information into the representation of definite expression. The contribution of the context to the interpretation of the definite NP consists in a salience hierarchy. It is assumed that each context can be associated with an ordering among the elements of subsets of the domain of discourse. The definite NP *the F* denotes the most salient F according to the situation *i*. This representation completes the ideas of discourse representation theories by producing a more comprehensive picture: a definite NP is not only linked to an already introduced

discourse referent, it is rather linked to the most salient discourse referent of the same kind so far.

The salience theory of definiteness has three historical sources: first, Lewis (1979) criticizes Russell's Theory of Descriptions and sketches an alternative theory using a salience ranking instead of Russell's uniqueness condition. Second, the investigation of the Prague School (cf. Sgall et al. 1973; Hajicová et al. 1995) developed an information structure of a sentence the pragmatic background of which is a hierarchy of "activated" referents. Third, research in artificial intelligence showed that discourse models need a structure or hierarchy of referents that is very similar to Lewis' concept of salience (cf. Grosz et al. 1995).

5.1 Lewis' Theory of Salience

Lewis (1970, 63) develops the concept of salience in the philosophical and linguistic discussion of the Russellian Theory of Descriptions:

> Second, consider the sentence 'The door is open'. This does not mean that the one and only door that now exists is open; nor does it mean that the one and only door near the place of utterance, or pointed at, or mentioned in previous discourse, is open. Rather it means that the one and only door among the objects that are somehow prominent on the occasion is open. An object may be prominent because it is nearby, or pointed at, or mentioned; but none of these is a necessary condition of contextual prominence. So perhaps we need a *prominent-objects coordinate*, a new contextual coordinate independent of the other. It will be determined, on a given occasion of utterance of a sentence, by mental factors such as the speaker's expectation regarding the things he is likely to bring to the attention of his audience.

Lewis (1979, 178) rejects Russell's uniqueness condition for definites or any modified version of it: "It is not true that a definite description 'the F' denotes x if and only if x is the one and only F in existence. Neither is it true that 'the F' denotes x if and only if x is the one and only F in some contextually determined domain of discourse." He considers the following examples, in which two individuals are introduced by the same definite NP (in the non-generic reading):

(22) The pig is grunting, but the pig with floppy ears is not grunting.
(23) The dog got in a fight with another dog.

In both examples two individuals with the same property are introduced into the discourse. However, the definite NP should unambiguously refer to one object. Note that no functional concept plays a role, since *pig* and *dog* are sortal concepts (except one would claim a functional concept from situations into objects of the mentioned kind). An anaphoric link to another expression seems not to be relevant here. Thus, the definite NP must refer uniquely according to another and more general principle. Lewis (1979, 178) names this principle *salience*:

> The proper treatment of description must be more like this: 'the F' denotes x if and only if x is the most salient F in the domain of discourse, according to some contextually determined salience ranking.

However, there has been no attempt to formalize this concept in order to integrate it into formal semantics.[3]

5.2 The Praguian School

The Prague School has developed a dynamic view of the information expressed in a sentence. In this approach, the "stock of shared knowledge" (Sgall et al. 1973, 70) constitutes the common background of the speaker and the hearer. It is the set of potential referents for definite expressions. This set is further divided into background and foreground information, which depends on encyclopedic knowledge, context information and thematic structure of the sentence. Besides this dichotomy, there is a further structure which are described in the following way (Sgall et al. 1973, 70f.):

> There is no clear-cut dichotomy in the stock of shared knowledge, and it would be, probably, more adequate to work here with a kind of ordering than with two subclasses. Let us remark that the mentioning of an element of the stock of shared knowledge brings this element into the foreground of the stock, and, in some respects, it is possible to conceive the last mentioned element to be more foregrounded that the elements mentioned before, the foregrounding of which already shades away step by step, if it is not supported by some specific moments due to the given situation.

In the extended system of Sgall et al. (1986, 54f.), different ways of shifts in a discourse model ("hearer's image of the world") are assumed. One of this shift is described in terms of a salience hierarchy:

> not the repertoire [of objects, relations etc., K.v.H.] itself is changed, but a certain relationship between its elements, namely their salience, foregrounding, or relative **activation** (in the sense of being immediately 'given', i.e. easily accessible in memory).

Hajicová et al. (1995, 14ff.) show how the position of an element in a sentence may effect its force to shift the salience: "(…) the activation of an item in SSK [=stock of shared knowledge, K.v.H.], if conceived as its attractiveness towards pronominal anaphora, seems to depend on in which position the item has been mentioned for the last time and on how many utterances have passed since that time point." They show that the choice of different pronouns (weak or strong) in

---

[3] Heim (1982, 21f.) additionally shows that the pragmatic concept of salience is too coarse-grained (the argument is due to Barbara Partee). In examples (1) and (2), the salience of the lost marble is raised by the preceding sentence. However, only in (1) the anaphoric linkage is possible. It seems that the structure of the expression plays an important role:
(1) I dropped ten marbles and found all of them, except for one. It is probably under the sofa.
(2) I dropped ten marbles and found only nine of them. # It is probably under the sofa.

Czech depends on this hierarchy of salience in the stock of shared knowledge. This view differs from Lewis' concept in that salience is regarded as a property of the cognitive discourse model, rather than as a property of the discourse as such. Furthermore, it concentrates on the use of pronouns rather than on the analysis of definite NPs.

5.3 The AI Approach

Computational analyses of discourse assume additional structures for discourse models in form of a hierarchy. Such analyses treat a referential process on par with the representation of the discourse in structured models. Sidner (1983) develops a system in which a focus-algorithm administrates the activation and focusing of potential referents such that anaphoric expressions can be linked to a focused expression. According to Grosz and Sidner (1985, 3), a general discourse model consists of three components: "a linguistic structure, an intentional structure, and an attentional state." The third component encodes the dynamic hierarchy between the different discourse objects. Grosz and Sidner (1985, 9) define them in the following way:

> The third component of discourse structure, the attentional state, is an abstraction of the participants' focus of attention as their discourse unfolds. The attentional state is a property of discourse, not of discourse participants. It is inherently dynamic, recording the objects, properties, and relations that are salient at each point in the discourse.

In contrast to the Praguian approach, this structure does not depend on the hearer or speaker, but it is a property of the context (like in Lewis' view). Webber (1983, 335) distinguishes between the act of reference by the speaker, and the referential behavior of expression in a certain discourse:

> That is, "referring" is what people do with language. Evoking and accessing discourse entities are what texts/discourses do. A discourse entity inhabits a speaker's discourse model and represents something the speaker has referred to. A speaker *refers* to something by utterances that either *evoke* (if first reference) or *access* (if subsequent reference) its corresponding discourse entity.

Grosz et al. (1995, 205) use the term "centering" instead of "focusing" or "evoking". They distinguish between "forward looking centering", which raises certain entities to salience, and "backward looking centering", which links anaphoric expression to such salient entities. The elements of the set of forward looking centers "are partially ordered to reflect the relative prominence" (209). They discuss a number of factors that may affect the ordering on these elements. However, they do not give a formal account of this that could be integrated into a formal approach to sentence and discourse meaning.

5.4 Salience and Discourse

According to Lewis (1979), a definite NP refers to the most salient object in the discourse that fits the descriptive content. And he notes further that the salience ranking depends on the context, i.e. it is not global in the sense that each expression gets its referent for global constraints nor it is local in the sense of Löbner, since once established it can keep its ranking during the whole discourse if there is no other salience changing expression. This property of changing the salience may be exemplified by the following example given by Lewis (1979, 179):

> Imagine yourself with me as I write these words. In the room is a cat, Bruce, who has been making himself very salient by dashing madly about. He is the only cat in the room, or in sight, or in earshot. I start to speak to you:

(24) The cat is in the carton. The cat will never meet our other cat, because our other cat lives in New Zealand. Our New Zealand cat lives with the Cresswells. And there he'll stay, because Miriam would be sad if the cat went away.

In terms of discourse representation theory, where the salience shifting potential cannot be encoded, the representation would look as follows: The first sentence in (24) introduces a discourse referent, that must be linked to an already introduced one. The second sentence refers to this referent by the expression *the cat* and introduces a new discourse referent with the same property of being a cat and the further relation that belongs to the speaker (and the presupposition that the first cat belongs to the speaker, as well). The third sentence refers to the second introduced cat by the expression *our New Zealand cat*. And the fourth sentence is anaphoric linked to that cat by the expression *he* and *the cat*. However, in a discourse representation there would be no difference in the accessibility of the discourse referents. Therefore, the theory must rely on further information.

However, if we modify the theory and let the indefinite NP not introduce a discourse referent but let it give the highest salience ranking to an individual that fits the description, a definite NP would then refer to the object that fits the description and that has the highest salience rank.[4] The first sentence introduces a new cat, lets say Bruce, into the discourse and raises him to the most salient cat, such that the definite NP *the cat* in the next two sentence can refer to this salient cat Bruce. The third sentence refers to this cat and introduces a second cat Albert, that gets a lower rank. Therefore, in the following two sentences we have to refer to Albert by an unambiguous description (*our other cat* and *our New Zealand cat*). Since in sentences (iv) and (v) we talk only about Albert, he gains it the first rank

---

[4] "Thus although indefinite descriptions – that is, idioms of existential quantification – are not themselves referring expressions, they may raise the salience of particular individuals in such a way as to pace the way for referring expressions that follow" (Lewis 1979, 180).

of the salience hierarchy such that in the last sentences we can refer to Albert by the pronoun *he* and by the definite NP *the cat*.

(25) Discourse   Ranking
    (i)   In the room is a cat   Bruce
    (ii)   The cat is in the carton.   Bruce
    (iii) The cat will never meet our other cat,   Bruce > Albert
    (iv) Because our other cat lives in New Zealand.   Albert, Bruce
    (v)   Our New Zealand cat lives with the Cresswells. Albert, Bruce
    (vi) And there he'll stay, because Miriam would   Albert > Bruce
    (vii) Be sad if the cat went away.   Albert > Bruce

This salience ranking can be represented in the following schema, which goes back to Hajicová (1993, 77). The mechanism used there is more fine-grained since it also considers the topic-focus structure of the text. This is especially important for the resolution of anaphoric pronouns. However, it seems that it is less important for anaphoric definite NPs. Moreover, it is not clear how the Praguian approach integrates the descriptive material of the NPs in questions. This becomes relevant in cases where we have more than one individual of the same kind, like in sentence (iii). The anaphoric reference in sentence (iv) is possible because the definite NP contains the description "other cat", which identifies only one cat.

(25a) *Schematic representation of the salience ranking*
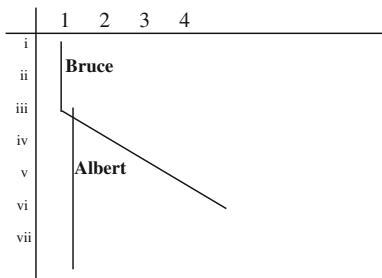
With the illustration of this small discourse the anaphoric use of definite descriptions is explained in terms of salience. That means that the anaphoric use can be seen as a specialized form of deictic use. In this way a uniform conception of definite NPs and deictic and anaphoric pronouns is possible.

6 Salience and Choice Functions

The concept of salience was never formally reconstructed although it was often regarded as an essential part for fixing the referent of definite expressions. In this section I develop a formal reconstruction of salience by means of context dependent choice functions. A choice function f is defined as the operation of assigning to a non-empty set one of its elements (It is not defined for empty sets).

Recently, choice functions are used to represent wide scope indefinites (Reinhart 1997; Winter 1997; Kratzer 1998). This type of choice functions are *local* choice functions (see below), while I use *global* choice functions in the sense of Egli (1991) and Egli and von Heusinger (1995). A global choice function depends on the shared knowledge between speaker and hearer or the common ground. A choice function selects the first element of an ordered set. Different choice functions can select different elements from one and the same set, i.e. the ordering of the elements in the set may differ. Peregrin and von Heusinger (2004) and von Heusinger (2004) combine the choice function approach with a dynamic logic. I try to keep the choice function mechanism as informal as possible.

Let us consider a situation where we have three cats *Albert, Bobby* and *Casimir* and three owners of the cats, *Ann, Beatrice* and *Carola*, respectively. The definite NP *the cat* is represented as the context dependent choice function applied to the set of cats $f_i(cat)$, which refers to the most salient cat in the context i. The different situations and accordingly the choice functions vary in the cat that is the most salient cat of the set of the three cats. We can define three choice functions (I use bold letters for indicating the objects of the model: **bobby** is the object we refer to by the name "Bobby"):

(27a) f_{ann} assigns to the set {**albert**, **bobby**, **casimir**} the cat **albert**
(27b) $f_{beatrice}$ assigns to the set {**albert**, **bobby**, **casimir**} the cat **bobby**
(27c) f_{carola} assigns to the set {**albert**, **bobby**, **casimir**} the cat **casimir**

Given this model with the defined choice functions, we can represent sentence (28) by the logical form (28a). The context index is informally integrated into the logical form in (28b). The interpretation (28c) of this representation proceeds according to compositional rules: The sentence is true if the extension of the definite NP *the cat* lies inside the extension of the predicate *very intelligent*. In order to fix the extension of the definite NP, the choice function $f_{beatrice}$ is applied to the set of cats yielding the individual **bobby** as value:

(28) The cat is very intelligent uttered by Beatrice
(28a) Very_Intellegent(f_i(cat) uttered by Beatrice
(28b) Very_Intelligent($f_{beatrice}$(cat)

A sentence with two individuals of the same characterization can be analyzed like (29). The two individuals are described by choice functions applied to sets of dogs. Additionally, the second mentioned dog is represented by the choice function applied to the set of dogs that does not contain the most salient dog, i.e. the functions picks out the second most salient dog: $f_i\,(\lambda y \mid y$ *is a Dog* $\&\ y \neq f_i(Dog)]$, which indicates that the referred object is not identical with the first chosen dog. i.e. it is the second most salient dog (cf. Egli and von Heusinger 1995, von Heusinger 1997):

(29) The dog got in a fight with another dog.
(29a) Got_a_Fight(f_i(Dog), f_i($\lambda y \mid y$ is a Dog $\&\ y \neq f_i$(Dog)]

6.1 The Situational Use

In the following subsections we will apply this formal reconstruction of salience, and hence definiteness, to the different uses of the definite NP, which were already discussed in Sect. 3. In the last example we saw how the situational context determines the choice of the object. Definite descriptions of the following kind crucially depend on context information. We will encode this information into the context index:

(30a) the sun $f_i(Sun)$
(30b) the university $f_i(University)$
(30c) the republic $f_i(Republic)$
(30d) the table $f_i(Table)$

We can now insert an argument in the situational index and fix the choice function. For example, if we are here in Cologne and speak of the republic we can fill the index slot with *cologne* and get the following expression:

(30e) $f_{cologne}(Republic)$

This term denotes that object that is a republic and that is first selected by a choice function, called *cologne*. Of course, we would define this choice function in such a way that it picks up first the German Federal Republic. This formalism implies that definite NPs contain an indexical element (see Wettstein 1981).

6.2 The Anaphoric Use

The representation of definite NPs as context dependent choice functions is a very general analysis and can be adapted to more specific uses. In the case of the anaphoric use the situation index has to be made exclusively dependent on the linguistic information of the discourse. We assume that the linguistic context in the discourse can raise the salience of an object by different means. One very obvious means is to refer to this object by a definite or an indefinite NP. The indefinite NP is used when the object has not yet been mentioned and the definite NP is used if the object was mentioned before. However, both make the object salient as the example (24) with the cats showed. In order to represent indefinite NPs we use *local* choice functions, i.e. choice functions that are different from the global choice function that is used for interpreting the definite NPs. A local choice function (see Reinhart 1997; Winter 1997; Kratzer 1998; von Heusinger 2002) is a newly introduced choice function, either bound by a local salient agent or existentially bound at some structural configuration, but not higher than the text level. We index such local choice functions by *x, y, z*.

The anaphoric linkage can be decomposed into the salience change potential of
an expression and the contextually dependent interpretation of another expression
as illustrated in example (31). In (31) the indefinite NP *a man* in the first sentence
introduces an arbitrary object d, which then becomes the most salient object of the
set of men. Thus, the indefinite not only updates the set of referents but also
updates the salience structure of the set of men (this feature distinguishes the
salience theory of definiteness from the familiarity theory). Therefore, the definite
NP, which refers to the most salient man, denotes the same object d as the
indefinite. In the representation, we assume that the indefinite NP changes the
given context i to the context j. The difference between the two context indices
reduces to the difference of choice function assignments. The assignment of the
updated global choice function f_j is equal to that of the initial global choice
function f_i except for the value of the set of men, which is d. This individual has
been introduced by the indefinite NP *a man* (for a more detailed formalism, see
Peregrin and von Heusinger 2004, von Heusinger 2004).

(31) A man comes. The man smokes.
(31a) Comes(f_x(Man)) & Smokes(f_j(Man)) with f_x(Man) = d
(31b) $f_j = f_i$ « $[[\text{Man}]]^{M,g}/d$ » with f_j(Man) = d

We generally indicate the update of a choice function by a set **s** and its new
assignment **a** inside double angle brackets: $f_j = f_i$ « s/a » : f_j is equal to f_i except
for the assignment to the set **s**, which is **a**.

We can account for anaphoric pronouns in the same way. They are represented
as very general choice function terms: f_i (λx [$x = x$]). The property [$x = x$]
denotes the individual domain D. Such a choice function term picks up the most
salient object in discourse, which is in sentence (32) identical with the most salient
man: $f_i(\lambda x$ [$x = x$]) = f_j(Man). In order to license the link between the indefinite
NP *a man* and the anaphoric pronoun, we must modify the salience change
potential of NPs. It does not only change the assignment for the set of men, but
also for certain supersets, e.g. the set of all (male) objects (in the following we
disregard gender differences):

(32) A man comes. He smokes.
(32a) Comes(f_x(Man)) & Smokes(f_j(λx [$x = x$]))
(32b) $f_j = f_i$ «$[[\text{Man}]]^{M,g}/d$, D/d»

We have now created the adequate means to describe even longer discourse
fragments like (25), which is repeated as (33). We assume that each sentence has
its own contextual index, i.e. is interpreted according to an optionally updated
global choice function. The relation between the different choice functions is
indicated by the equations. Generally, the choice functions are identical except for
the assignment of the sets that are denoted by the properties in the NPs and the
domain D of individuals, i.e. they are updates of the preceding choice functions in
respect to the used NPs. In (33i), the indefinite NP *a cat* refers to Bruce and
changes the choice function f_i to the choice function f_1. f_1 is equal to f_i except that

is assigns **bruce** to the set of cats and to *D*. Therefore, the definite NP *the cat* refers to **bruce**, too. Since **bruce** is already the most salient cat, sentence (33ii) does not change the actual salience hierarchy and its formal counterpart, the choice function f_2. Sentence (33iii) changes the assignment to the set of *other cats* to **albert**, and the next two sentences change the assignments to the set of cats and the universal set to **albert**, too. The definite expressions *he* in (33vi) and *the cat* in (33vii) refer to this very cat **albert**:

(33i) In the room is a cat

$$\text{In\_the\_Room}(f_1(\text{Cat})) \qquad\qquad f_1 = f_i \text{«}[[\text{Cat}]]^{M,g}/\textbf{bruce, D}/\textbf{bruce»}$$

(ii) The cat is in the carton.

$$\text{In\_Carton}(f_2(\text{Cat})) \qquad\qquad\qquad\qquad\qquad f_2 = f_1$$

(iii) The cat will never meet our other cat,

$$\text{Never\_Meet}(f_3(\text{Cat}(x)), f_3(\lambda y\ [\text{Cat}(y)\ \&\ y \neq f_3(\text{Cat})]))$$
$$f_3 = f_2\text{«}[[\text{other cat}[[M,g/\textbf{albert»}$$

(iv) because our other cat lives in New Zealand.

$$\text{Lives\_in\_New\_Zealand}(f(\lambda y\ [\text{Cat}(y)\ \&\ y \neq f_4(\text{Cat})])) \qquad f_4 = f_3$$

(v) Our New Zealand cat lives with the Cresswells.

$$\text{Lives\_with\_Cresswells}(f_5(\text{Cat \& In\_New\_Zealand}))$$
$$f_5 = f_4 \text{«}[[\text{New Zealand cat}]]^{M,g}/\textbf{albert, D}/\textbf{albert»}$$

(vi) And there he'll stay,

$$\text{Stay}(f_6([x = x])) \qquad\qquad\qquad\qquad\qquad f_6 = f_5$$

(vii) because Miriam would be sad if the cat went away.

$$\text{Miriam\_Would\_Sad\_If\_Went\_Away}(f_7(\text{Cat})) \qquad\qquad f_7 = f_6$$

6.3 *The Relational Use of Definite NPs*

A choice function term can express complex dependencies by embedding, i.e. if a term is dependent on other terms this can be expressed by a parameter inside the term. Definites without further modifications have wide scope since they are dependent on the situation whose scope is certainly wider than the sentence in which the definite NP stands. The definite NP in (34) has wider scope than the quantifier expression *every man*. However, if we add the relative clause *that barked at him* the definite NP is narrow scoped, since the universal quantifier binds

a variable inside the term. The denotation of the set depends on the choice of the variable of the universal quantifier.

(34) Every man saw the dog.
(34a) $\forall x \, (Man(x) \rightarrow Saw(x, f_i(\lambda y \, Dog(y)))$
(35) Every man saw the dog that barked at him.
(35a) $\forall x \, (Man(x) \rightarrow Saw(x, f_i(\lambda y \, [Dog(y) \, \& \, barked\_at(x, y)]))$

7 Summary

The different uses of definite NPs can be best reconstructed with context dependent choice function terms. This representation focuses on the situational use of definite NPs and extends its analysis to the anaphoric and relational uses as well. Choice functions allow capturing the uniqueness condition of classical theories in a very elegant way: They select exactly one element of a set, but the set itself need not be unique. They also capture one of the main insights of the familiarity theory: Indefinite and definite NPs are updates on the context, here on the salience structure of the discourse. The salience theory of definiteness also mirrors the diachronic development of definite articles from demonstratives and other indexical items. While demonstratives clearly need additional information such as an ostension, the definite article expresses a contextually given salience ordering. The salience theory of definiteness also allows for a unified account of definite and indefinite NPs in terms of global versus local choice functions (see Chierchia 2005). Thus, it raises many new and challenging questions to our semantic interpretation of noun phrases in general.

References

Abbott, B. (2004). Definiteness and indefiniteness. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 122–149). Oxford: Blackwell.

Chierchia, Gennaro. 2005. Definites, locality, and intentional identity. In G. N. Carlson & F. J. Pelletier (Eds.) *Reference and quantification. The Partee effect* (pp. 143–177). Stanford: CSLI Publications.

Christophersen, Paul.1939. The Articles. A Study of their Theory and Use in English. Copenhagen: Munksgaard.

Egli, Urs. (1991). (In)definite Nominalphrase und Typentheorie. In *Zwei Aufsätze zur definiten Kennzeichnung*, eds. U. Egli and K. von Heusinger. Arbeitspapier 27. Fachgruppe Sprachwissenschaft, Universität Konstanz.

Egli, U., & von Heusinger, K. (1995). The epsilon operator and E-type pronouns. In U. Egli et al. (Eds.) *Lexical knowledge in the organization of language* (pp. 121–141). Amsterdam: Benjamins.

Frege, G. (1892). Über Sinn und Bedeutung. Zeitschrift für Philosophie und philosophische Kritik *100*, 25–50. English translation in: Translations from the Philosophical Writings of Gottlob Frege. P. Geach, & M. Black (Eds.) 1960. (pp. 56–78). Oxford: Blackwell.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics, 21*, 203–225.

Grosz, B., & Sidner, C. (1985). *The structure of discourse structure*. CSLI-Report 85-39. Ventura Hall: CSLI.

Hajicová, E., Hoskovec, T., & Sgall, P. (1995). Discourse modelling based on hierarchy of salience. *The Prague Bulletin of Mathematical Linguistics, 64*, 4–24.

Hajicová, E. (1993). *Issues of sentence structure and discourse patterns*. Prague: Charles University.

Hawkins, J. (1978). *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croom Helm.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. PhD Dissertation. University of Massachusetts, Amherst. Ann Arbor: University Microfilms.

Heim, I. (2011). Definiteness and indefiniteness. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.) *Semantics. An international handbook of natural language meaning,* vol. 2, (pp. 996–1025). Berlin: de Gruyter,

von Heusinger, K. (1997). *Salienz und Referenz. Der Epsilonoperator in der Semantik der Nominalphrase und anaphorischer Pronomen*. Studia Grammatika 43. Berlin: Akademie Verlag.

von Heusinger, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics, 19*, 245–274.

von Heusinger, K. (2004). Choice functions and the anaphoric semantics of definite NPs. *Research on Language and Computation, 2*, 309–329.

Kamp, H. (1984)[1981]. A theory of truth and semantic interpretation. In J. Groenendijk, T. M. V. Janssen, & M. Stokhof (Eds.) *Truth, interpretation and information.* (pp. 1–41). Dordrecht: Foris.

Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground* (pp. 363–385). New York: Academic Press.

Kratzer, A. (1998). Scope or pseudoscope? Are there wide-scope indefinites? In S. Rothstein (Ed.), *Events and grammar* (pp. 163–196). Dordrecht: Kluwer.

Lewis, D. (1970). General semantics. *Synthese, 22*, 18–67.

Lewis, D. (1975). Adverbs of quantification. In E. L. Keenan (Ed.) *Formal semantics of natural language* (pp. 3–15). Cambridge: Cambridge University Press.

Lewis, D. (1979). Scorekeeping in a language game. In R. Bäuerle, U. Egli, & A. von Stechow (Eds.) *Semantics from different points of view* (pp. 172–187). Berlin, Heidelberg, New York: Springer.

Löbner, S. (1985). Definites. *Journal of Semantics, 4*, 279–326.

Ludlow, P. (2007). Descriptions. In E.N. Zalta (Ed.) *The Stanford encyclopedia of philosophy* http://plato.stanford.edu/entries/descriptions. Accessed 15 Jan 2012.

Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.

Lyons, J. (1977). *Semantics*. (Vols. 2) Cambridge: Cambridge University Press.

Neale, S. (1990). *Descriptions*. Cambridge/Mass: MIT Press.

Peregrin, Jaroslav andKlaus von Heusinger. (2004) [1996]. Dynamic semantics with choice functions. In H. Kamp and B. Partee (Eds.) *Proceedings of the Workshop "Context Dependence in the Analysis of Linguistic Meaning".* (pp. 255–274). Amsterdam: Elsevier.

Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy, 20*, 335–397.

Russell, B. (1905). On denoting. *Mind, 14*, 479–493.

Sgall, P., Hajicová, E., & Benesová, E. (1973). *Topic, focus and generative semantics*. Kronberg/Taunus: Scriptor.

Sgall, P., Hajicová, E., & Panevová, J. (1986). *Meaning of the sentence in its semantic and pragmatic aspects*. J. Mey. (Ed.) Dordrecht: Reidel.

Sidner, C.L. (1983). Focusing in the comprehension of definite anaphora. In M. Brady, & R. Berwick (Eds.) *Computational models of discourse* (pp. 265–328). Cambridge/Mass: MIT Press.

Stalnaker, R. (1978). Assertion. In P. Cole (Ed.) *Syntax and semantics 9: Pragmatics* (pp. 315–332). New York: Academic Press.

Strawson, P. (1950). On referring. *Mind, 59*, 320–344.

Webber, B. (1983). So what can we talk about now? In M. Brady, & R. Berwick (Eds.) *Computational models of discourse* (pp. 331–370). Cambridge/Mass: MIT Press.

Wettstein, H. (1981). Demonstrative reference and definite descriptions. *Philosophical Studies, 40*, 241–257.

Winter, Y. (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy, 20*, 399–467.

On the Dynamic Relations Between Common Ground and Presupposition

Istvan Kecskes and Fenghui Zhang

Abstract The common ground theory of presupposition has been dominant since the seventies (Stalnaker 1974, 1978, 2002). This theory has resulted from a view of communication as transfer between minds. In this view interlocutors presume that speakers speak cooperatively, they infer that they have intentions and beliefs that are necessary to make sense of their speech acts, and treat such entities as pre-existing psychological ones that are later somehow formulated in language. Common ground is considered as a distributed form of mental representation and adopted as a basis on which successful communication is warranted (Arnseth and Solheim 2002; Kecskes and Zhang 2009). However, the theory has not gone without objection and criticism (e.g. Abbott 2008; Beaver and Zeevat 2004; von Fintel 2001, 2006; Simons 2003) because it is based on "an oversimplified picture of conversation" (Abbott 2008), and as a consequence the relationship between common ground and presupposition has also been oversimplified. In this approach presupposition is often considered as a conventional or conversational constraint of common ground, or requirement on common ground that must be satisfied in order to make an appropriate utterance. The problem of accommodation is a critical issue that has been raised against this view, and caused great challenge to the theory by stimulating diverse alternatives. The goal of this paper is to redefine the relationship between common ground and presupposition within the confines of the socio-cognitive approach (SCA). SCA (Kecskes 2008; Kecskes and Zhang 2009; Kecskes 2010a, b) adopted in this paper offers an alternative view on communication, which claims that communication is not an ideal transfer of information, and cooperation and egocentrism (Barr and Keysar 2005; Colston 2005; Keysar 2007), are both present in the process of communication to a varying extent. The SCA emphasizes the dynamics of common ground creation and

I. Kecskes (✉)
University at Albany, Albany, NY 12222, USA
e-mail: ikecskes@albany.edu

F. Zhang
East China University, Shanghai, China

updating in the actual process of interaction, in which interlocutors are considered as "complete" individuals with different possible cognitive status being less or more cooperative at different stages of the communicative process. Presupposition is a proposal of common ground, and there is a vibrant interaction between the two. They enjoy a cross relation in terms of content and manners in which they are formed, and their dynamism is inherently related and explanatory to each other. This claim has important implications to the solution to presupposition accommodation. After the introduction Sect. 2 describes the socio-cognitive approach. Section 3 reviews the assumed common ground, and Sect. 4 introduces the speaker-assigned presupposition. Section 5 discusses the dynamism of presuppositions and common ground, and claims that their dynamic observations are coherent and explanatory to each other. Section 6 readdresses the accommodation problem with redefinition of the relations.

Keywords Presupposition · Common ground · Socio-cognitive · Egocentrism · Cooperation

1 Introduction

The common ground theory of presupposition started to be formulated in the seventies. It was first proposed by Stalnaker (1974, 1978) and adopted by several formal pragmatists (e.g. Heim 1983; Beaver 1997; von Fintel 2006). In this theory propositions that a sentence presupposes are just those that must be entailed by the common ground of any context that is to admit that sentence. This notion of presupposition relies on a concept of "common ground" according to which the common ground of a context of utterance is the conjunction of all those propositions that interlocutors take for granted in that context either because they are permanently shared beliefs in their community, or because they have been established in the course of the preceding conversation (cf. Heim 1990).

Stalnaker proposed a version of information-gathering discourse, in which assertion is intended to update common ground, and presupposition is intended to shape or narrow down common ground. Presupposition is interpreted as the speaker's belief in the common ground status of the proposition: "A proposition *P* is a pragmatic presupposition of a speaker in a given context just in case the speaker assumes or believes that *P*, assumes or believes that his addressee assumes or believes that *P*, and assumes or believes that his addressee recognizes that he is making these assumptions, or has these beliefs" (Stalnaker 1974: 573). The formula below describes the speaker's cognitive state in presupposing:

(1) $K_1p, K_1K_2p, K_1K_2K_1p, \ldots$

In this formula number 1 denotes the speaker, number 2 denotes the hearer or the addressee, and the letter K denotes the state of having belief or assumption in a weak sense or knowledge in a strong sense.

Formal pragmatists (cf. Beaver and Zeevat 2004; Simons 2003; von Fintel 2001) adopted Stalnaker's common ground theory and adapted it to their own needs. They have aimed at formal description of presupposition projection which takes place in an unfolding context. The sentence requires that the presupposed proposition be taken for granted and not subject to (further) discussion and the utterance of it requires that the speaker assume that its requirements are satisfied.

The main problem with this approach is that it takes into account only the collective core part of common ground that is encapsulated in the utterance and means relatively the same for all speakers of that language community. At the same time this approach ignores the privatized knowledge and beliefs of interlocutors. It does not actually consider what is on the speaker's mind in that particular situational context. Instead, it formalizes what's going on with the sentence and its linguistic context. The formula below describes the minimal requirements of presupposition on the interlocutors' cognitive states:

(2) K_1p, K_2p

It might be difficult to clarify what exactly "taken-for-grantedness" means, but the formula may suffice to indicate that the proposition is noncontroversial and requires no further discussion. Researchers studying pragmatic presuppositions from socio-cultural or other perspectives (e.g. Soames 1982; Yule 1996) have shown concerns about the interlocutors' cognitive states.

However, in whatever directions the common ground theory of pragmatic presupposition has been developed, it still has drawn criticism. The critical issue is accommodation. A special type of informative presupposition raises this issue which has become a problem that no one could get by without handling it first. The following conversation shows this problem:

(3) Bob (who met Alice for the first time): Are you going to lunch?
    Alice: No, I've got to pick up my sister.

This is a problematic case of presupposition concerning its common ground status. We can see this problem more clearly by the formula below:

(4) K_1p, $\sim K_2p$, $K_1 \sim K_2p$, …

Obviously this description of cognitive states does not go with either (1) or (2) Because this special case of presupposition violates the common ground theory, many researchers adopted the "common ground + accommodation" view (e.g. Karttunen 1974; Lewis 1979; Stalnaker 1974, 2002, 2008; Soames 1982; von Fintel 2000, 2006). Although in different forms of interpretation, this view generally follows the rule depicted by Lewis (1979: 417): "If at time t something is said that requires presupposition P 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".

Although several attempts have been made to explain accommodation, none has managed to clarify properly why the speaker presupposes as such when s/he is fully aware of the presupposed not being part of common ground knowledge, and

how accommodation is achieved in the time gap after the utterance event has taken place and before it has been accepted or rejected (Zhang 2009).

This case of informative presupposition and the approach of accommodation have called the common ground theory of presupposition into question. Several researchers have proposed alternative ways which greatly challenged the extant theory. Burton-Roberts (1989) regarded accommodation as a fatal problem to the theory and insisted on a semantic version of presupposition instead. Abbott (2000, 2006, 2008) approached the issue from the angle of information structure and proposed to give up the theory altogether. Simons (2001, 2003, 2004) argued that accommodation, as construed on the common ground view, requires one to posit a conventional constraint on sentences, and therefore she was forced to abandon the common ground view of presupposition and investigated conversational constraints on utterances by analyzing the assumptions of relevance and cooperativeness.

From the above we can see that not only is the "common ground + accommodation" view problematic but objections to and criticisms of it also vary greatly. What makes things more complicated is that Stalnaker (2002, 2008) appears to have obscured the dividing line between common ground and presupposition, and occasionally tends to use them as alternative terms. This may cause even more confusion to the problems at issue. It is therefore crucial how one views the relationship between common ground and presupposition. The socio-cognitive approach with its notions of assumed common ground and speaker-assigned presupposition gives us the chance to reexamine the dynamism of these two entities and redefine their relationship.

2 The Socio-cognitive View of Communication

We think that the main problem with the common ground theory of presupposition is that it considers common ground an a priori concept, partly ignores its dynamism and relies on the assumption that cooperation is always present to the same extent in the communicative process. The Alice's sister (3) issue may be solved within the confines of the socio-cognitive approach (SCA) to communication proposed by Kecskes (2008, 2010b) and (Kecskes and Zhang 2009). This approach is based on two claims. First, speaker and hearer are equal participants of the communicative process. They both produce and comprehend relying on their most accessible and salient knowledge expressed in their private contexts in production and comprehension. Consequently, only a holistic interpretation of utterance from both the perspective of the speaker and the perspective of the hearer can give us an adequate account of language communication. Interlocutors should be considered as "complete" individuals with different possible cognitive status, with possible different interpretation of the same core common ground information, which has a profound effect on what the same linguistic structure may mean for any of them. Second, communication is a dynamic process in which individuals are not only constrained by societal conditions but they also shape them at the same time. As a

consequence, communication is characterized by the interplay of two traits that are inseparable, mutually supportive and interactive:

Individual trait:	Social trait:
Prior experience	Actual situational experience
Salience	Relevance
Egocentrism	Cooperation
Attention	Intention

In the socio-cognitive approach interlocutors are considered as social beings searching for meaning with individual minds embedded in a socio-cultural collectivity. Individual traits (prior experience - → salience - → egocentrism - → attention) interact with societal traits (actual situational experience - → relevance - → cooperation - → intention). Each trait is the consequence of the other. Prior experience results in salience which leads to egocentrism that drives attention. Intention is a cooperation-directed practice that is governed by relevance which (partly) depends on actual situational experience. Kecskes (2010b; 2012) argued that SCA integrates the pragmatic view of cooperation and the cognitive view of egocentrism and emphasizes that both cooperation and egocentrism are manifested in all phases of communication, albeit to varying extents. *Communication is the result of interplay of intention and attention motivated by socio-cultural background* that is privatized/subjectivized by the individuals in their linguistic behavior. The background is composed of knowledge of interlocutors deriving from their private prior experience and current situational experience that are both socio-cultural in nature (Kecskes 2008).

3 Assumed Common Ground

3.1 Core Common Ground and Emergent Common Ground

Kecskes and Zhang (2009) postulated that there are two sides of assumed common ground: *core common ground* and *emergent common ground*. *Core common ground* refers to the relatively static, generalized, common knowledge that belongs to a certain speech community as a result of prior interaction and experience, whereas *emergent common ground* refers to the relatively dynamic, actualized and particularized knowledge co-constructed in the course of communication that belongs to and is privatized by the individual(s). The former is a repertoire of knowledge that can be assumed to be shared among individuals of a speech community independent of the situational circumstances, such as when and where the conversation occurs, between whom it occurs, etc. In contrast, the actual

contextual part (emergent common ground) is knowledge that is aroused, co-constructed and/or involved as shared enterprises in the particular situational context that pertains to the interlocutors exclusively. This contingent circumstance draws attention of the interlocutors to the same entities or states and, with the formation of particular intentions therein, activates some of their prior individual experiences that join in this intention-directed action.

When critiquing the common ground view of presupposition Abbott (2000) underlined that the driving idea behind this theory is that presuppositions are identified with "old" information, or information that the speaker is treating as "old." This is not the case in the socio-cognitive approach in which common ground is perceived as an effort to converge the mental representation of shared knowledge present as memory ("old" information) that we can activate, shared knowledge that we can seek, and rapport as well as knowledge that we can create in the communicative process ("new" information). According to this approach common ground is present throughout the whole communicative process. The core and actual (emergent) components join in the construction of common ground in all stages and motivate the interplay of intention and attention in this process, although they may contribute to the construction process in different ways, to different extent, and in different phases of the communicative process.

3.2 Dynamism of Common Ground

Common ground is an assumption that we make in the course of actual communication. Both core common ground and emergent common ground are integrated parts of this assumed common ground. Core common ground is a general assumption in two ways. First, although core common ground is relatively static and shared among people, it can *change diachronically*. During a certain period, say a couple of years, we may safely assume that interlocutors have access to relatively similar common knowledge because components of core common ground won't change dramatically. However, in the long run it definitely will change; people's social life, both material and spiritual, will experience some changes over a long period of time, and as a consequence their core common ground will also be changed. For instance:

(5) Jill: I need some money.
    Jack: There is an ATM over there.

It is part of core common ground what "ATM" refers to. However, thirty years ago that conversation would not have made much sense since "ATM" did not exist as a part of core common ground.

Second, core common ground may also vary among different groups of individuals within a speech community. Type of shared knowledge may be determined by different factors such as geography, life style, and educational, financial and

racial factors. This fact may restrain the accessibility of certain core common ground to particular groups only within that speech community.

Emergent common ground is assumptive in that it is contingent on the actual situation, which reflects a *synchronic* change between common grounds in different situations. However, emergent common ground is not only new shared knowledge created in the course of communication but also the use and modification of shared prior knowledge or experience. There is a dialectical relationship between core common ground and emergent common ground. The core part may affect the formation of the emergent part in that it partly restricts the way the latter occurs. In many cases the emergent part may partly originate in instances of information that are predictable in the core part. On the other hand, the emergent part may contribute to the core part in that the contingent emergent part in a frequent ritual occurrence potentially becomes public disposition that belongs to the core part. In other words, they are different components of assumed common ground, which are interconnected and inseparable.

The dialectical relationship between the two sides of common ground (core and emergent) can be illustrated by the following conversation.

(6) Jill: I met someone today.
    Jane: Good for you.
    Jill: He is a police officer.
    Jane: Are you in trouble?
    Jill: Oh, no…

Jill met someone who was a policeman. Conforming with our society's collective salience, the concept of 'policeman' is identified with some kind of trouble. This knowledge is part of core common ground. However, this understanding of the concept is privatized in Jill's case and acquires a positive overtone, as the result of her positive (maybe even romantic) encounter with the policeman. Jane did not have this experience, so she processed the word in accordance with core common ground. What the speaker meant differed from what the hearer inferred from the same utterance. Emergent common ground was created as a modification of core common ground as required by the given situation. In brief, both shared sense and current sense can vary from case to case according to the identification of relations or roles of interlocutors, their memory of prior experiences, and their cognitive perception of the actual situational context available to them.

In the socio-cognitive view *assumed common ground works as a background on which the interplay of intention and attention occurs and communication takes place*. There are three different ways intention and attention affect the construction of common ground in the process of communication (Kecskes and Zhang 2009). One is that the interlocutors *activate* mental representations of shared information that they already have. A second way of constructing common ground is that interlocutors *seek* information that potentially facilitates communication as mutual knowledge. Before the speaker makes the seeking effort, the piece of information is not salient in the hearer as background underlying the upcoming conversation.

The third contribution to common ground is when the speaker brings in her private knowledge and makes it a part of common ground. The speaker has some private information that she knows is non-accessible to the hearer, and she adopts it as common ground in the belief that it facilitates the conversation and that the hearer will accept it willingly. Example # 6 demonstrates this case.

4 The Speaker-assigned Presupposition

In this section we intend to relate presupposition to the understanding of common ground in the socio-cognitive approach. The most commonly accepted view is that presupposition is taken for granted in the sense that its assumed truth is a pre-condition for felicitous utterance of the sentence and places a kind of constraint on discourse contexts that admit the sentence for interpretation (Chierchia and McConnell-Ginet 1990: 283). For a sentence to be appropriate in a given context its pragmatic presuppositions should already be part of the conversational background or common ground, though it may also be easy for the hearer to *accommodate* them.

According to the SCA this is a restricted view on common ground that relies mainly on what we called "core common ground" that the interlocutors take for granted in that context either because they are permanently shared beliefs in their community. *In the SCA presupposition is always related to the speaker's and hearer's state of mind that works with both prior and current experience.* Presupposition is a joint business, in which the speaker and the hearer play different roles. An analogy to an oral contract can be made. The first party (the speaker) draws a draft of this contract by uttering a sentence, and once it is agreed on by the second party (the hearer) giving a positive response to it, the contract becomes valid between them. Similarly, presupposing by the speaker is to common ground as what draft is to a contract. The speaker proposes a background of the conversation in his/her presupposition, and this action will receive a response from the hearer, who either agrees with it and a mutual background is formed, or has no idea about it and (or) feels doubt about it, and then the common ground is at stake. This analogy illustrates the dynamic nature of presupposition. Aiming to shape out a systematic vision of dynamic presupposition, Zhang (2009) proposed the definition of speaker-assigned presupposition, which can be formalized as follows:

The speaker presupposes that p in her/his utterance, iff:

1. s/he proposes that p be common ground;
2. s/he assigns propositional attitudes and communicative interests about p to the proposal; and
3. s/he observes truthfulness principles and intention principles in the proposal.

This definition emphasizes the dynamic nature of presupposition, and reveals the effect of attention (substantiated by the speaker's propositional attitudes) and intention (by form of communicative interests) on the formation of presupposition.

The three points of the definition offer us answers to the basic questions concerning the nature of presupposition: what is presupposition, where does it come from, and how is it made. According to the SCA it is not the truth value of the proposition or its common ground status as triggered by linguistic expressions or means that must be satisfied as a precondition for the utterance of the sentence; instead, it is the propositional attitudes and the communicative interests concerned as assigned by the speaker that make the utterance appropriate and comprehensible.

4.1 Categorization

The formal analysis of the speaker's knowledge set (or belief set)[1] enables us to investigate specific ways the speaker relies on her propositional attitudes and communicative interests and makes his/her proposal through presupposition. Presupposition can be categorized into three groups according to values of the speaker's belief set: truthful presuppositions, assumptive presuppositions, and fake presuppositions.[2] The table below is a summary.

As we can see from the Ep value in the table, the speaker commits herself/himself to the truth value of the proposition for the group of truthful presuppositions, suspends it for assumptive ones, and forges a fake value for fake ones. This indicates that the speaker doesn't necessarily commit himself/herself to the truth value of presupposition. In addition, Cp-1 and Cp-2 indicate that the speaker assumes the truth value of presupposition to be shared knowledge for canonical presupposition, but not for the rest groups or subtypes. That is to say, the speaker doesn't always hold a common ground belief of presupposition.

We have no space here to explain each sub-categories so we will give only one example for each category.

The most often addressed case is *informative presupposition*. It deviates from the speaker's common ground belief and reveals her communicative interests. Previous research, however, did not establish a systematic view of it. By defining presupposition as the speaker's assignment of both propositional attitudes and communicative interests, we are able to examine informative presupposition in a more reliable way. Let us turn to Alice's case.

---

[1] According to epistemic logic, we assume that the agent's knowledge that p is a strong version of her belief that p. As presupposition usually reflects the speaker's beliefs about p and its common ground status, we use the logic operator K to denote the agent's strong belief; i.e., if the speaker who utters "John regrets beating his dog" presupposes that p ("John has beaten his dog before") the way she believes that p is common ground, then the formal analysis of the speaker's knowledge set about p is: K_1p, K_1K_2p, $K_1K_2K_1p$, ….

[2] The first two groups are categorized in Zhang (2009), and the third group will be elaborated in an upcoming paper.

(7) (adopted from 3) Bob (who met Alice for the first time): Are you going to
     lunch?
     Alice: No, I've got to pick up *my sister*
     ≫ p: Alice has a sister

The presupposition is proposed when Alice is aware of Bob's ignorance of it, as
she intends to speak in an economic way.

(8) A daughter to her dad who has no idea about her engagement: - Oh Dad, I
     forgot to tell you that *my fiancé* and I are moving to Seattle next week.
     (Simons 2004: 14)
     ≫ p: The speaker is engaged

In this example the speaker intends to achieve certain rhetoric effects via
indirect conveyance of new information that p.[3]

Assumptive presuppositions
Subtype: Partial presupposition

The speaker does not have common ground beliefs about p; however, s/he
believes that p will be common ground if the hearer contributes that p.

(9) Mary does not know if Jill is married or not. However, she wants to find it out
     indirectly.
     Mary: Why do you want to take a bus to the meeting? Can't your husband
     drive you?
     Jill: No, he cannot. He is too busy.
     ≫ p: Jill has a husband (Jill is married)

Fake prepositions
Deceptive presuppositions:

The speaker does not have common ground beliefs about p, the utterance does
not have truth-value. Some presupposition acts are designed only to facilitate
certain communicative interests and the truth-conditional commitment is falsely
made. For example:

(10) Washington to the neighbor who has stolen his horse (the fact is that the
     horse is not blind): *Which of the horse's eyes is blind*?
     ≫ p: One of the horse's eyes is blind

When interrogating the neighbor who had stolen his horse, Washington
deliberately offers a false presupposition, for which he made a false commitment

---

[3] However, at present there is no reliable criterion based on which we can distinguish the two
effects. .

Fig. 1 Mechanism for
presupposition processing
Agent1 Agent2

Agent1   Agent2
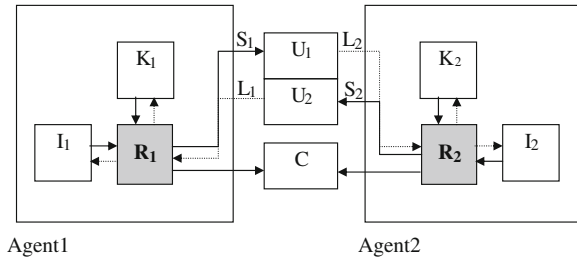
for its truth value so as to mislead the neighbor. In such presuppositions the speaker violates, rather than just exploits the common ground status in order to realize certain communicative effects.

In brief, the categorization of speaker-assigned presupposition describes the dynamism of various presuppositional phenomena in an exhaustive and systematic way. Presupposition is a proposal of common ground, and during the process of communication the speaker's attentional resources and intentional state affect the ways presupposition is proposed and common ground is updated. The mechanism below formalizes the process.

4.2 The Mechanism

We propose the following mechanism to regulate the dynamic processing of presupposition by the interlocutors. Figure 1 describes that two agents (Agent1 and Agent2), with their cognitive states (consisting of knowledge set K and intention set I), make speech acts (S) and listen to (L) the utterance set (U_1 and U_2 by two agents) by observing the rules set (R). The presupposition act as component of the speech act may result in the update of the common ground set (C) of the utterance. In this mechanism the presupposed proposition p, which is proposed to be common ground, may be added to the common ground set (C) but not necessarily so.[4]

There are altogether the following components involved in the mechanism: two agents (Agent1 and Agent2), four elements that interact with each other (K, I, U, C), two components of speech act (S, L), and one set that regulates all connections (R). So we get the Cartesian products below:

---

[4] Here the 'update of' or 'adding to' the common ground should be perceived from a technical perspective; i.e., in formal computation of information conveyed in a conversation, there is an update effect of presupposition on the common ground set, as dynamic semantics has claimed (cf. Heim 1983). It is different from psychological concerns; it does not mean every presupposed proposition is proposed to update the common ground that the interlocutors have already had in mind at the utterance time, as we understand that the speaker does hold common ground belief for some cases of presupposition. .

$R = \{R_K, R_I, R_U, R_C\}$
$R_K: C \times U \times I \Rightarrow K$
$R_I: C \times U \times K \Rightarrow I$
$R_U: C \times K \times I \Rightarrow U$
$R_C: U \times K \times I \Rightarrow C$

These products reveal that each of the four elements (K, I, U, C) are functions of the other three by the relation R. Each time an utterance is made (update of U), all the other three elements (K, I, C) will be updated accordingly. The question is about the nature of the information encoded in the presupposed proposition of U and the way the relation R is specified. *Does the proposition refer to/activate/seek prior knowledge and/or information, or does it create new knowledge and/or information that can be considered emergent common ground?* As the categorization shows, presupposition may contribute to common ground of the conversation in many different ways, but sometimes it does not contribute to it at all. The mechanism connecting the four elements and regulating the update of common ground information by presupposition is based on truthfulness principles and intention principles formulated by Zhang (2009). By observing the principles or 'exploiting' them, the speaker shapes presupposition in different ways, and the update of common ground also varies accordingly.

4.3 Dynamism of Presupposition

The dynamic nature of presupposition has been explored in the definition of speaker-assigned presupposition and the subsequent categorization and processing mechanism. In the SCA dynamism reveals itself in two dimensions, namely *synchronic variation and diachronic change.* Synchronic dynamism is reflected by the variety of participants' common ground belief at the utterance time, whereas diachronic dynamism by the change of such belief during the course of communication.

Synchronic dynamism means that the speaker's common ground beliefs about p and p's common ground status at the utterance time vary in different types of presuppositions. The speaker's belief in common ground should be:

$K_1 = \{p: K_1p, K_1K_2p, K_1K_2K_1p\}$

As we can see in Table 1, the speaker holds different beliefs about p in different types of presuppositions, each related to common ground in different ways. The speaker's belief about p varies from case to case, and many do not go with the expected common ground status. Informative presupposition, for instance, is a case where the speaker deviates from his/her common ground belief as we illustrated in example (3). Also see the example below:

(11) Joe: Look at that poster.
     Bill: Which one?
     Joe: *The green one about car insurance*. Just over there, on the wall.

Table 1 Categorization of presuppositions (ps)[10]

Group	Subtype	Ep	Cp-1	Cp-2
Truthful ps	Canonical	K_1p	K_1K_2p	$K_1K_2K_1p$
	Informative	K_1p	$K_1 \sim K_2p$; $K_1{\wedge}K_2p$; $\sim K_1?K_2p$	$K_1K_2^+K_1p$
Assumptive ps	Partial	$\sim K_1p$	$K_1K_2p^!$	$K_1 \sim K_2 \sim K_1p$; $K_1K_2 \sim K_1p$
	Temporary	$\sim K_1p$	$K_1 \sim K_2p$	$K_1K_2 \sim K_1p$
Fake ps	Figurative	$K_1\neg p$	$K_1K_2\neg p$	$K_1K_2K_1\neg p$
	Deceptive	$K_1\neg p$	$K_1 \sim K_2p$	$K_1 \sim K_2K_1\neg p$

> Bill: Wow, it's huge!
> Joe: Soon I will need to renew *my insurance*.

In this conversation, Joe makes the two presuppositions of insurance in different ways. These are different cases of informative type (see Table 1 for formal analysis). Whereas the first one is seeking shared information in that current situation which may add to current sense of emergent common ground, the latter also contributes to current sense, but by way of creating it with the speaker's individual knowledge.

Diachronic dynamism means that p's common ground status changes at different times of the conversation. After the presupposition is made, it goes through a process in which the participants' knowledge adapts to each other's, and p's common ground status may experience some change. The proposition p that is not common ground at the utterance time may be added to common ground set later. The addition of p to C is a joint effort of the interlocutors. A proposition p that is eligible for common ground should be:

$C = \{p: K_1p, K_2p, K_1K_2p, K_2K_1p, K_1K_2K_1p, K_2K_1K_2p, \ldots\}$

In example (11) after the hearer *accommodates* his knowledge the propositions will then be added to the common ground of the actual conversation.

---

[10] The formal system goes in this way: "1" denotes the speaker. "Ep" ("everyone knows that p") here denotes the status of p in the speaker's knowledge; "K_1p" denotes that the speaker knows that p, "$\sim K_1p$" the speaker does not know that p (which entails that the speaker believes that $\diamond p$), and "$K_1\neg p$" the speaker knows that p is false.

"Cp" ("it is common knowledge among the agents that p") here denotes the speaker's belief about status of p in the agent's knowledge, which consists of two parts (Cp-1 and Cp-2). In Cp-1: "K_1K_2p" denotes that the speaker believes that the hearer knows that p, "$K_1 \sim K_2p$" denotes that the speaker believes that the hearer does not know that p, "$K_1 \wedge K_2p$" denotes that the speaker believes that the hearer has a false belief about p, "$\sim K_1?K_2p$" denotes that the speaker does not know whether the hearer knows that p, "$K_1K_2p^!$" denotes that the speaker believes that the hearer knows that p or not p, "$K_1K_2\neg p$" denotes that the speaker believes that the hearer knows that p is false. In Cp-2: "$K_1K_2K_1p$" denotes that the speaker believes that the hearer believes that the speaker knows that p, "$K_1K_2^+K_1p$" denotes that the speaker believes that the hearer will believe (upon hearing the utterance) that the speaker knows that p, "$K_1K_2K_1\neg p$" denotes that the speaker believes that the hearer believes that the speaker knows that p is false, and so forth

Also see example (12) illustrating a partial presupposition:

(12) Sally: Sorry, I must go now.
     Mary: So you are going to meet *your boy-friend* tonight?
     Sally: Yes, he will come to pick me up.

In this example interlocutors make joint effort to add partial *p* to common ground. Sally's contribution to common ground can be dubbed as *compensation*. Both accommodation and compensation are instances of dynamic change of common ground belief (and status) of *p* in an unfolding conversation.

The projection problem also can be perceived as disclosure of diachronic dynamism. Let us explain this by the example below.

(13) Student: The prime minister of Tahiti is a woman.
     Teacher: Tahiti's prime minister is not a woman because there is no prime minister in Tahiti.[5]

     # ≫ There exists a prime minister of Tahiti.

The speaker's (teacher) belief is that not p remains unaltered, but the hearer (student) will experience an update of her belief, and this is done diachronically: she may not change her belief that p when hearing the first clause of the utterance but change it soon when hearing the second clause. The utterance can be interpreted in two steps: the first is that the speaker claims that the student's assertion does not hold, and the second is that she justifies her claim by arguing that the presupposition (precondition) of the false assertion does not hold. It is after the second step is made that the hearer detects the negation of presupposition and updates her knowledge accordingly.

5 The Dialectic Relation Between Common Ground and Presupposition

In this section we will analyze the dialectic relations between the two phenomena and argue that they enjoy a cross relation in respect of their content and the particular ways they are formed.

Dynamism of common ground in terms of its components (core common ground and emergent common ground) explains why dynamic processes of presupposition formation should occur. As there is no perfect match between the

---

[5] We hold that the speaker wouldn't make the presupposition out of the blue (see Kecskes and Zhang 2009). S/he makes it to achieve coherence in the conversation; that is, it is her/his priority to deny what is asserted by the student. In other cases when such coherence is not needed, the speaker would assert "There is no prime minister of Tahiti" directly without presupposing as such.

interlocutors' common ground in a particular situation, for the benefit of smooth communication a variety of methods to build up and develop common ground becomes necessary. Presupposition is to propose common ground based on which the present utterance can be made and comprehended. Different types of presupposition are required by different mental representations of shared knowledge the speaker has obtained, and this variety also enables her/him to aim for particular communicative interests.

Presupposition is among the various ways common ground can be formed and updated. Assertion and implicature may also contribute to common ground in different ways. All these sources converge to build up common ground and facilitate smooth communication. *In the socio-cognitive approach we proposed it is not the quantity but the quality of common ground that counts*. Efficiency of common ground constructions depends on their attention-raising quality that must be adjusted to the actual situational context. Hearers sometimes may ignore common ground that is activated by a presupposition, or they may also miss the information updated by an assertion and so forth. This may occur because of lack of attention, or there exist other cognitive obstacles, such as amnesia or other mental disorders. In such cases more strenuous efforts are called for to achieve common ground for the participants. The following dialogue[6] demonstrates this point.

(14) Mother: Josh, *your grandma' called* (Assertion 1)
       Josh (working on the computer): What?
       Mother: *She is having a birthday party on Sunday* (Assertion 2)
       Josh: Who is having a party?
       Mother: You never listen.. *Your grandma' does* (Assertion 3)

Josh's mother made three assertions, among which the first two were not easily received by Josh or added to common ground of the conversation because of attention and quality problems. The information "grandma's called' in Assertion 1 did not get common ground status because Josh did not pay close attention to what was said, and as a remedy it was restated by Assertion 3 in an explicit way.

On the other hand, dynamism of presupposition, revealed in its variety of types and change in the communication, affects the dynamic process of common ground construction. As we have explained earlier, different types of presupposition may contribute to common ground in different ways. They enjoy different belief representations at the utterance time and add to common ground after the joint effort of the interlocutors. Not all presuppositions will invariably contribute to common ground successfully. Sometimes they fail. The failure may be caused by

---

[6] This can be taken as counterevidence to Stalnaker's (2002) claim that what is asserted is no longer new information and must be treated as common ground. In the grandma case, the proposition p is not new information to the participants, but it's not part of common ground either, as it is not mutually activated as relevant information in this conversation. Also see Kecskes and Zhang (2009: 351).

undesirable complexity of the interlocutors' cognitive state. They may have a store of different set of core common ground knowledge because of age as illustrated by (15), or they lack facilities to achieve common ground status as illustrated by (14) or the same emergent one as in (16). For different components of knowledge or belief about/in a proposition their sharedness varies according to situational factors, such as absent-mindedness in (14), cultural gap in (15), and loss of deictic tracking in (16). When these occur, the presupposed part has to be readdressed so that it can be added to common ground. In (14) 'grandma's called' is asserted, in (15) 'ATM' is brought under discussion, and in (16 below) no remedy can be made, as 'blonde hair' is not longer within sight.

(15) Jiang: I need some money.
     Jack: There is an ATM over there.
     Jiang: I beg your pardon? What is that, uh, 'ATM'?

(16) Bob: Look at that girl. Her blonde hair looks so nice.
     Tom: Where is she?
     Bob: Oh forget it, she's gone.

Still there are cases when it is the speaker's intention not to take presupposed propositions as common ground. S/he 'exploits' their common ground status. Temporary presuppositions and fake ones are such cases. These presuppositions are designed for the speaker's intention to achieve certain communicative interests and as a consequence their truth conditions are rated lower in value. The exploited presuppositions are distinct from common ground, or they are 'contaminated' by communicative intentions. They don't contribute to common ground in a traditional way, but enjoy some similar temperaments of implicature with intentions joined in.[7] Some of them may be added to common ground later after the truth values are mutually recognized, as (12) illustrates, and others may not, as (17) goes (see also 10).

(17) Washington to the neighbor who has stolen his horse (the fact is that the horse is not blind): *Which of the horse's eyes is blind*?
     ≫ p: One of the horse's eyes is blind

From the above analysis, we can see that there is a vibrant interaction between presupposition and common ground. Their dynamism is inherently related and explanatory to each other. Both of them can be explained within the confines of the socio-cognitive approach that we have proposed. The individual factors of attention and societal factors of intention affect the dynamic processes in which

---

[7] Simons (2001, 2004) observed this similarity and attempted to search for the same theoretical frames to interpret presupposition and implicature. However, this effort was not quite successful as she blurred their distinctions and did not offer an adequate approach to identify and explain different phenomena of presupposition (Zhang 2008).

presupposition is proposed and common ground is formed. As a result, they enjoy a cross relation in terms of content and manners in which they are formed. In terms of content, presupposition partly contributes to the formation of common ground, as it is just one source of common ground. It is added to common ground selectively, and there are cases when it fails.

Presupposition and common ground also overlap in terms of manners they are formed. Different types of presupposition serve to activate, seek shared knowledge present as memory in both/all interlocutors, and create new knowledge to enrich their common ground. While canonical presupposition is a process in which old information is activated, informative one is a typical example when common ground is actually created on the spot. Common ground can be sought, and created by other forms, such as Assertion 3 in (14). The ways through which common ground can derive from assertion and implicature are still generally ignored.

6 The Accommodation Problem Revisited

As said above the accommodation problem has brought great challenge to the common ground theory of presupposition. Not only supporters of this theory vary in their arguments, but its opponents hold different views as well. Zhang (2009) pointed out that the attempts to justify accommodation did not clarify properly why the speaker presupposes as such when s/he is fully aware of the presupposed not being part of common ground knowledge, and how accommodation is achieved in the time gap after the utterance event has taken place, and before it has been accepted or rejected.

The main problem with those attempts is that they take into account only the collective core part of common ground that is encoded in the utterance, and means relatively the same for all speakers of that language community. In other words, a relatively static view of presupposition and common ground and their relations has hampered the development of a proper theory. Accommodation is an inherent problem for the common ground theory which might not be solved properly within its confines.

SCA offers an alternative approach to this problem. We argue that accommodation is normal occurrence, not a problem or exception to presupposition theory. The case of informative presupposition is one form of speaker-assigned presuppositions, and its accommodation process can be well explained by the dynamism in two dimensions (synchronic and diachronic). As shown in the categorization (4.1), informative presupposition is one type of truthful presuppositions to which the speaker assigns a truth value commitment, but differs from canonical one in that its common-groundness is an outcome of co-constructing by both/all participants. The speaker creates new information to become part of common ground, and only after the hearer accommodates her belief to it can this common ground be achieved.

The two unsolved issues concerning accommodation can be explained through the analysis of the dynamic and dialectical relationship between presupposition

and common ground. Why should (or can) the speaker presuppose as such when s/he lacks common ground belief? This is answered by the *synchronic dynamism*. As the speaker holds different propositional attitudes and targets for different communicative interests, presuppositions she makes are also of different types. Informative presupposition is proposed for economic effect, saving energy by packing new information in the form of presupposition and making space for more new information coming up in the form of assertion. This may also achieve an effect of coherence, as the information most closely related to communicative intentions is asserted and made focus of the participants' attention. Then the second issue is: how is accommodation achieved? Accommodation is the process through which presupposition gets added to common ground; the participants' common ground beliefs for the proposition are co-constructed. There is a problem, however, about how accommodation is fulfilled. From the formal analysis below, it is unknown why the speaker should assume that the hearer will accommodate $(K_1K_2^+p?)$, and why the hearer should accommodate as expected by the speaker $(K_2^+p?)$ and even believe that the latter should believe so $(K_2K_1K_2^+p?)$.

(18) a. at t (utterance time): K_1p, $\sim K_2p$, $K_1 \sim K_2p$, $K_1K_2^+p$;
   b. at t + 1 (after t): K_2^+p, $K_2K_1K_2^+p$;
   c. at t + 2 (after added to CG): K_1p, K_2p, K_1K_2p, K_2K_1p, $K_1K_2K_1p$, $K_2K_1K_2p$, ...

According to *diachronic dynamism* of speaker-assigned presupposition, accommodation is one of the dynamic processes through which presupposition gets added to common ground. The addition of p to common ground is a joint effort of the interlocutors, and in this joint effort of accommodation, they abide by related truthfulness principles.[8] *Truthfulness principles 1* (for truthful presuppositions): The speaker's principle: presuppose that p iff K_1p; The hearer's principle: accept that p unless $K_2\neg p$. For more details, please refer to Zhang (2008, 2009). The formal analysis below describes the process.

(19) a. at t_0 (before t): K_1p, $\sim K_2p$, $K_1 \sim K_2p$;
   b. at t: $\sim K_2K_1p$, $K_1K_2^+K_1p$, $K_1K_2^+p$;
   c. at t + 1 (after t): $K_2^+k_1p$, K_2^+p, $K_2K_1K_2^+p$;
   d. at t + 2 (after added to common ground): K_1p, K_2p, K_1K_2p, K_2K_1p, $K_1K_2K_1p$, $K_2K_1K_2p$, ...

At time t, the speaker infers according to truthfulness principles that the hearer will infer her truth value commitment $(K_1K_2^+K_1p)$ and therefore assumes that the latter will accommodate $(K_1K_2^+p)$. At time t + 1, the hearer, also based on truthfulness principles, infers that the speaker knows p $(K_2^+k_1p)$, accommodates her belief (K_2^+p), and also infer that the speaker believes so $(K_2K_1K_2^+p?)$. In this

---

[8] Please refer to Sect. 4. The related principles for truthful presuppositions are:

process the interlocutors 'take for granted' that each part should abide by the truthfulness principles and therefore truth values are derived.

In addition, informative presupposition and its accommodation are indispensable processes for the formation of common ground. SCA distinguishes core common ground from emergent common ground. In the course of communication there is always a chance (or necessity) to bring in some new information as emergent common ground, as the current situation adapts to changes of various elements and the same with the interlocutors' perceptions of them. Informative presupposition plays an important role in creating emergent common ground. It is a reliable source for common ground, as the speaker assigns a true proposition to it and the hearer will safely adopt it and update their common ground information. The exception, which is rare, is that the speaker's knowledge turns out to be false.

7 Conclusion

In this paper we examined the dynamism of presupposition-common ground relation within the confines of SCA and redefined deictic relations between them. The SCA emphasizes the dynamics of common ground creation and updating in the actual process of interaction, in which interlocutors are considered as "complete" individuals with different possible cognitive status, being less or more cooperative at different stages of the communicative process. Presupposition is a proposal of common ground, and there is a vibrant interaction between the two. They enjoy a cross relation in terms of content and manners in which they are formed, and their dynamism is inherently related and explanatory to each other.

Presupposition and common ground share similarities and differ from each other. When we agree on their close correlations, we should also pay respect for their separateness. While presupposition is an important subject of linguistic pragmatics which relies on formal analysis, common ground embraces a broader area with rich resources of information that do not always derive from verbal cues. Presupposition is significant in its linguistic neatness,[9] truth-conditional concern and recent conversational incorporation. In contrast, common ground outstands for its psychological, philosophical reflection and recent experimental observation. SCA offers an alternative to explain their complex relationship.

---

[9] It is not as neat as expected, though.

References

Abbott, B. (2000). Presuppositions as nonassertions. *Journal of Pragmatics, 32*, 1419–1437.

Abbott, B. (2006). Unaccommodating presuppositions: A NeoGricean view. http://www. pragmatics.osu.edu/links/events/BarbaraAbbott_Unaccompresupps.pdf (31 August 2008). Draft paper for Workshop on Presupposition Accommodation at The Ohio State University.

Abbott, B. (2008). Presupposition and common ground. *Linguistics and Philosophy, 21*, 523–538.

Arnseth, H. C. & Solheim, I. (2002). Making sense of shared knowledge. In G. Stahl (Ed.), *Proceedings of CSCL 2002, Computer Support for Collaborative Learning: Foundations for a CSCL Community*, (pp.102–110), Boulder, Colorado, January 2002.

Barr, D. J. & B. Keysar. (2005). Making sense of how we make sense: The Paradox of egocentrism in language use. In Colston, Herbert, L. & Albert N. Katz. (eds.), *Figurative language comprehension* (pp. 21–43). Mahwah, N.J: Lawrence Erlbaum.

Beaver, D. I. (1997). Presupposition. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 939–1008). Amsterdam: Elsevier.

Beaver, D & Zeevat, H. (2004). Accommodation. MS.

Burton-Roberts, N. (1989). *The limits to debate: A revised theory of semantic presupposition*. Cambridge: Cambridge University Press.

Chierchia, G., & McConnell-Ginnet, S. (1990). *Meaning and Grammar*. Cambridge, Mass.: The MIT Press.

Colston, Herbert L. (2005). On sociocultural and nonliteral: A synopsis and a prophesy. In H. Colston & A. N. Katz (Eds.), *Figurative language comprehension: Social and cultural influences* (pp. 1–20). Hillsdale, N.J.: Lawrence Erlbaum.

Heim, I. (1983). On the projection problem for presuppositions. *West Coast Conference on Formal Linguistics*: 114–125. Republished in 1991 Davis, S (ed.), *Pragmatics: A Reader*. New York: Oxford University Press. 397–405.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy, 13*(2), 137–177.

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics, 1*, 181–193.

Kecskes, I. (2008). Dueling contexts: A dynamic model of meaning. *Journal of Pragmatics, 40*, 385–406.

Kecskes, I. (2010a). Situation-bound utterances as pragmatic acts. *Journal of Pragmatics, 42*(11), 2889–2897.

Kecskes, I. (2010b). The paradox of communication: A socio-cognitive approach. *Pragmatics and Society, 1*(1), 50–73.

Kecskes, I. (2012). Is there anyone out there who really is interested in the speaker? *Language and Dialogue, 2*(2), 285–299.

Kecskes, I., & Zhang, F. (2009). Activating, seeking and creating common ground: A socio-cognitive approach. *Pragmatics and Cognition, 17*(2), 331–355.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics, 4*(1), 71–84.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8: 339–359. Republished in 1991 Davis, S (ed.), *Pragmatics: A Reader*. New York: Oxford University Press. 416–427.

Simons, M. (2001). *On the conversational basis of some presuppositions*. Proceedings of Semantics and Linguistic Theory 11.

Simons, M. (2003). Presupposition and accommodation: Understanding the Stalnakerian picture. *Philosophical Studies, 112*, 251–278.

Simons, M. (2004). Presupposition and relevance. In Z. G. Szabo (ed.), *Semantics vs. Pragmatics*. New York: Oxford University Press.

Soames, S. (1982). How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry, 13*, 483–545.

Stalnaker, R. C. (1974). Pragmatic presuppositions. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and Philosophy* (pp. 197–214). New York: New York University Press.

Stalnaker, R. C. (1978). Assertion. *Syntax and Semantics, 9*, 315–332.

Stalnaker, R. C. (2002). Common ground. *Linguistics and Philosophy, 25*(5–6), 701–721.

Stalnaker, R. C. (2008). A response to Abbott on presupposition and common ground. *Linguistics and Philosophy, 31*, 539–544.

Von Fintel, K. (2000). What is presupposition accommodation? http://web.mit.edu/fintel/www/accomm.pdf. Ms., MIT.

Von Fintel, K. (2001). Counterfactuals in a dynamic context. In M. Kenstowicz (Ed.), *Ken hale: A life in language*. Cambridge, Mass: MIT Press.

Von Fintel, K. (2006). What is Presupposition Accommodation, Again? Draft paper for Workshop on Presupposition Accommodation at The Ohio State University.

Yule, G. (1996). *Pragmatics*. Oxford: Oxford University Press.

Zhang, F. (2008). Conversational constraint of truthfulness on presuppositions. *Intercultural Pragmatics, 5*(3), 367–388.

Zhang, F. (2009). Speaker-assigned presupposition: A cognitive-pragmatic approach. *Proceedings for The 2nd International Conference on English, Discourse and Intercultural Communication*. Macao: University of Macao.

What can Pragmaticists Learn from Studying Artificial Languages?

Alan Reed Libert

Abstract Among the hundreds of artificial languages put forth as possible international auxiliary languages, relatively few (e.g. Esperanto, Interlingua) have seen a substantial amount of actual use. Given this, one might think that the study of such languages might have little to offer pragmaticists, and indeed there has been very little pragmatic work on them. However, I would argue that the pragmatic investigation of artificial languages can provide useful insights and information. Most designers of artificial languages are not professional linguists. Although they usually say little or nothing about the pragmatics of their languages, what they do say can reveal popular ideas about pragmatics, which may otherwise be difficult to discover. I shall present and discuss relevant remarks by some artificial language designers. I shall also look at several pragmatic features of artificial languages. Although the amount of textual material available in most artificial languages is limited, what exists can be subjected to pragmatic analysis. Perhaps most intriguing are the a priori artificial languages (e.g. aUI), attempts to build a language without borrowing anything from natural languages, as, on the surface, these languages can appear quite odd. I shall present some texts from several artificial languages with a view to seeing whether even apparently exotic artificial languages have the same pragmatic properties as natural languages. Such work can be seen as contributing to the study of cross-linguistic pragmatics.

1 Introduction

Given that the vast majority of artificial languages (henceforth ALs) have seen very little, if any, use, one might have serious doubts about whether there could be much of a pragmatic study of them. Indeed, there has been very little pragmatic

A. R. Libert (✉)
University of Newcastle, Newcastle, Australia
e-mail: Alan.Libert@newcastle.edu.au

research on ALs, which have generally received little attention from theoretical linguists (who might feel that they are not serious objects of study, a matter not helped by the perhaps biased work of some Esperantists and by the existence of "languages" such as Klingon), but some linguists may believe that there would be even weaker grounds for pragmatic analysis of ALs than for e.g. morphological or syntactic analysis, since most ALs have not been used much, if at all.[1]

However, many works presenting an AL do provide (sometimes extensive) textual material, which one can analyze pragmatically (as well as syntactically). Since such texts were largely intended as instructional material, one might say that they do not represent a typical use of language, but they do represent a kind of use (and one meant to mirror more typical uses). It could be claimed that even example sentences and reading exercises constitute a certain kind of language use. One of the few works treating artificial language pragmatics is Huang (2002), but, in spite of the fact that "Artificial Language" is contained in its title, the discussion of ALs makes up only a small part of it. More space is given to ALs by Traunmüller (1991/1996); however, some of the phenomena which he deals with (e.g. number marking in NPs) are not those usually dealt with by pragmatics.

If we now look at particular views on pragmatics and ALs, Galdia (2009) would probably deny the possibility of a serious pragmatic study of ALs, as he says (p. 334):

> No natural language has been meaningfully characterized without the pragmatic dimension nor can it be properly spoken without it. The lack of knowledge about this dimension makes a full characterization (and also the full command) of ancient languages like Latin or Greek impossible. This theoretical problem is even more manifest in the case of artificial languages such as Esperanto which cannot be fully determined in terms of pragmatics simply because they are artificial. An artificial pragmatics, in turn, would not make much sense.

One may compare this opinion with the remarks of van Cranenburgh et al. (2010) about Esperanto: "Although it was designed as an easy-to-learn language, with regular and transparent syntax and morphology, its semantic and pragmatic components have evolved naturally" (p. 2); "Esperanto has a regular and transparent morphology while featuring rich semantics and pragmatics" (p. 7). Similarly Dellert (2008: 2) states:

> since the language has developed into a full replacement for natural languages in all situations, all aspects of semantics and pragmatics that NLP [natural language processing] wants to address are present in Esperanto as much as in any natural language

About ALs more generally Gobbo (2008: 39) says:

> From the point of view of theoretical linguistics, planned languages are fully human languages, being non-natural without necessarily being unnatural, since they are acquired

---

[1] I use the following abbreviations: lit.—literally, sec.—section, tr.—translation. Translations of quotations from languages other than English are mine, while translations of texts and examples are those of the source, unless otherwise indicated. In some quotations and texts I have modified punctuation and/or formatting in minor ways.

or acquirable as a normal part of the process of maturation and socialization (Lyons 2006). Consequently, they will be scrutinized by linguistic level, from language core to language use, that is to say phonetics, morphology, syntax, semantics, pragmatics.

One might think that if one were going to pragmatically analyze an AL, Esperanto would clearly be the best choice, since it has seen far more use than any other ALs (and in fact has some first language speakers), and second choices would be ALs which have been used to a relatively high degree (for ALs), such as Ido and Interlingua, while it would be unproductive, if not silly, to try to treat ALs which have been used very little or not at all (i.e. the vast majority of them).

However, although I shall discuss Esperanto and Ido, I take a different position. ALs are commonly classified on the basis of how much material, if any, they take from natural languages. Those which, like Esperanto, are largely or entirely based on one or more natural languages are known as *a posteriori* ALs, and it is these which have been most successful and popular. Those which represent attempts to build a language from scratch, i.e. for which little or nothing is (consciously or intentionally) taken from natural languages are called *a priori* ALs. This is a spectrum rather than a strict dichotomy: most a priori languages have borrowed at least a small amount from natural languages, and many languages have substantial amounts of both original and a posteriori material; such languages are referred to as *mixed* ALs.

In the early days of AL language construction, i.e. in the 17th and 18th centuries, most ALs were designed along a priori lines, though later the mixed and a posteriori methods of language creation became much more popular. Nevertheless even in the 20th century some a priori languages, such as Ro, Suma, and aUI, were constructed. No a priori language has come close to the level of use of Esperanto or Volapük, the most successful mixed language.

However, it might be argued that if one is interested in the pragmatic possibilities of human language, the languages to look at would be the a priori languages, for these would be good testing grounds for putative pragmatic universals. If an AL is meant to be a completely original creation, owing nothing to natural languages, and if even in it such universals hold, then we may have an idea of the limits of pragmatic variation.[2]

Another general area where we may be able to learn something from ALs, of whatever sort, is laymen's ideas about pragmatics. The vast majority of AL designers or describers are not well-versed in linguistics, and what they say (or perhaps more importantly, do not say) about pragmatic features of the language can tell us what the average person thinks and knows about pragmatics.

It may be significant for our purposes that ALs are usually second languages, and so there may be pragmatic (and other) influence from a first language. Given

---

[2] In general I am only interested in ALs that seem reasonably serious in purpose, and thus I do not deal with ALs created in connection with a work of fiction or "artistic" or "personal" ALs. Most of the more serious ALs are meant to be auxiliary languages for international use. I also do not treat computer languages such as BASIC.

this, one might ask which texts in an AL would be suitable for pragmatic analysis; that is, presumably only texts created by those with a certain degree of proficiency would be of interest (unless we were specifically interested in early stages of the learning process)—what is that level and how can we determine it, or can we? (The same issues come up if we are analyzing texts at some other level, e.g. for syntax, but there may be a need for a larger amount of textual material if one is looking at pragmatics). Most ALs do not have official (or any) courses which could give some certification or evidence of having reached some level of skill in the language. It would appear reasonable to use texts written by the AL designer (and sometimes the only texts in an AL are by its designer), although sometimes even they make errors in their languages. It would also appear reasonable to use texts in material which in some way is (explicitly or implicitly) approved of by the designer (e.g. in webpages which he gives links for in his website) or by some official organization or webpage for the AL.

I shall deal with the following areas: politeness, different styles and levels of language (including formal language), non-descriptive meaning, illocutionary force, and non-literal language. I shall thus not be discussing presuppositions, because I have found neither many relevant examples nor any significant discussion of them by AL designers.

2 Politeness

Politeness may be the most obvious pragmatic feature of language and the one which most often comes up in presentations of ALs (compare e.g. conventional implicature, which most AL designers, like most speakers of any language, are probably not consciously aware of).

2.1 Pronouns

ALs, like natural languages, differ in whether they have polite pronouns. Several AL designers explicitly state that their language lacks them. For example, Russell (1966: 5) says of Suma, "There is no polite form of the pronoun 'you'".

Bollack, the designer of the Blue Language, believes that an AL, or at least his AL, must have 2nd person polite pronouns (the Blue Language has not only a singular, but also a plural, polite form); he states (1900: 19), "This creation is absolutely necessary, as the «civilization», when to address somebody, imposes on us a *familiar* and a *respectful* form".

Not all presentations of ALs with polite pronouns give instructions on how or when to use them (although the same is true of many pedagogical grammars of natural languages). For example, Talmey (1925: 7) gives *vu* as the "singular, polite" 2nd person pronoun of Arulo (cf. *tu* 'thou, thee, you' and *vi* 'you (plural)'),

but says nothing about situations in which it should appear. A lack of instructions could be seen as unfortunate, since it is not always obvious in which contexts polite pronouns are appropriate, especially for native speakers of languages such as English which lack them. Since the use of polite pronouns differs from language to language, even native speakers of languages with such pronouns may be uncertain about this feature of an AL.

De Beaufront (1925/2005) is a source which does give instructions about them; he says (p. 25) concerning Ido (which has three 2nd person pronouns, two singular ones, *tu* and *vu*, and the plural form *vi*), "Por la duesma persono singulara existas anke formo familiara: *tu*, quan on darfas uzar nur ad amiki tre intima, a frati o parenti kun qui on uzas, en sua linguo matrala, formo familiara korespondanta" ('For the second person singular there also exists a familiar form, *tu*, which one may use only with very intimate friends, with siblings or relatives with whom one uses, in his mother language, a corresponding familiar form').

ApGawain et al. (2008: 37) have more complete instructions on the use of *tu* in Ido:

1. "tu" refers to one person only. It shows affection towards the person addressed, and is therefore only to be used in special circumstances: (a) within the family, (b) between close friends, (c) when addressing small children, (d) perhaps when addressing an animal or pet.
2. "vu" also refers to one person only. It is the usual word for "you".
3. "vi" refers to more than one person, and is the plural of both "tu" and "vu".

Concerning Sambahsa Simon (2010: sec. 3) says:

"Yu" can apply to a group of persons or to a single person as a sign of formal respect. "Tu" (cf. archaic English "thou") is used only to address close relatives or friends and children, as, for example, in French, Russian or Farsi.

Parke (2008: 9) gives a rather detailed description of the use of the familiar and polite pronouns of Frenkisch:

Unlike in English, there are both plural and singular forms for the second person pronoun, *ji* and *dou*. *Ji* should always be used when addressing more than one person. When addressing one person in a formal context, *ji* should also be used. *Dou* is a more familiar, intimate and informal pronoun than *ji*. *Dou* should be used mostly for addressing friends, lovers, family members and young children. Its use may be extended to colleagues and fellow students but then care should be taken to use it only informal situations. Used in the wrong situation, *dou* can be seen by some to be rude or excessively intimate—the verbal equivalent of invading someone's personal space.

Esata has both a 2nd person singular pronoun, *yu*, and a 2nd person plural and formal pronoun, *yi*, but it also has a sort of honorific affix for pronouns; Bothi (2006: sec. 6) says, "Deference and respect can be indicated by prefixing (stronger) or appending the pronoun modifier *ji*: e.g. *yuji*—your honor, *jiyu*—your excellency". It is interesting that the level of respect it marks differs depending on its position.

In Latinulus various pronouns can replace other pronouns; Martellotta (1919: 103) gives the following instructions on pronoun usage:

> Il pronome di prima persona plurale viene adoperato in vece del singolare dai sovrani, dai magistrati e dagli scrittori in senso maiestativo.
> Il pronome *tu* si usa parlando fra colleghi, o con inferiori di dignità, o ad animali, o ad esseri sacri.
> Il pronome di seconda persona plurale si usa sovente invece del singolare quando si parla o si scrive ad una sola persona.
> Il pronome di terza persona singolare e plurale *leis* e *leise* si usa invece del pronome di seconda quando si parla o si scrive ad una persona di riguardo.
> ('The pronoun of the first person plural is used instead of the singular by sovereigns, by magistrates, and by writers in a majestic sense.
> The pronoun *tu* ['you'] is used when speaking among colleagues, or with those lower in rank, or to animals, or to sacred beings.
> The pronoun of the second person plural is often used instead of the singular when one is speaking or writing to a single person.
> The pronouns of the third person singular and plural *leis* ['she'] and *leise* ['they' (feminine)] are used instead of the pronouns of the second person when one is speaking or writing to a person of regard.')

Esperanto has a 2nd person pronoun in addition to *vi*, but it is very rare and its function is not clear; Wennergren (2005: 102) states:

> *Ci* estas unu-nombra alparola pronomo… *Ci* kaj *cia* ekzistas nur teorie, kaj estas preskaŭ neniam praktike uzataj. Eblus imagi *ci* kiel pure unu-nombran *vi*, aŭ kiel intiman familiaran (unu-nombran) *vi*, aŭ eĉ kiel insultan (unu-nombran) *vi*. Sed estas fakte tute neeble diri, kian nuancon ĝi montras, ĉar ĝi apenaŭ estas uzata
> ('*Ci* is a singular addressing pronoun… *Ci* and *cia* exist only in theory and are almost never used in practice. One could imagine *ci* as purely a singular of *vi*, or as an intimate familiar (singular) *vi*, or even as an insulting (singular) *vi*. But it is in fact completely impossible to say what nuance it indicates, because it is hardly ever used')

This raises the important point that one cannot know all of the details of a word which does not occur, or, to see it another way, it is only through use that a word acquires all of its meaning. (This recalls the argument (which I do not fully agree with) that there cannot be a pragmatics of an artificial language, or at least of most artificial languages, as they saw little or no use.)

However, Kellerman (1910: 15) gives a different impression of *ci*:

> There is another pronoun [besides *vi*] **ci**, *thou*, for the second person singular, used in solemn style, as in the Bible, in poetry, and also for intimate or familiar address when desired, like German *du*, French *tu*, etc.

Gledhill (2000:103) says the following about this word:

> The pronoun *ci* was originally devised by Zamenhof to represent the archaic informal or Shakespearean 'thou' in translations and is used by some Esperantists as a familiar 'you' because of the influence of European languages (French *tu,* German *du,* Russian *ti* etc.). Zamenhof discouraged *ci* with the justification that different languages had different

conventions for its use, and there may also have been an ideological motive to establish parity for all language speakers. There are no examples of *ci* in the corpus,[3] although its use is attested among some reform-minded Esperantists. Conversely, a polite form *Vi* (i.e. capitalized *vi*) was widely used as a polite or plural second person form in early writing, and there are a number of instances in the literary sections of our corpus, including the early writings of Zamenhof.

There are some interesting points here. The fact that some natural European languages have played a role in the occurrence of *ci* indicates that at least in this area of pragmatics ALs can be affected by natural languages. (One might have assumed that this could happen, but one also might want evidence for it, as we now have.) Also intriguing is the connection between ideology (in this case egalitarianism) and a view about a pragmatic feature, the familiar/polite distinction in pronouns. It would appear that there is some sort of impulse, at least on the part of some language speakers, to have a polite pronoun, since there has been both the aforementioned effect from European languages with the respect to *ci* and the creation of the other polite pronoun *Vi*.

The situation for Volapük's polite pronoun *ons* is somewhat similar to that of Esperanto *ci*; Linderfelt (1888: 14) states:

> The English "you" having usurped the functions of both "thou" and "ye", observe that in Volapük *ons* represents a polite "you", whether addressed to one or more persons, *ol* the "you" of familiar intercourse to one person and *ols* to two or more persons. The *ons* is, however, of doubtful value and might as well be dropped, the English usage having amply demonstrated that one form of address may be employed to everybody, without giving offense to anyone, though keeping a sharp distinction between the singular *ol* and and the plural *ols* is indispensable to clearness. At the Volapük conference in Münich, it was decided to discourage the use of the form in *ons*.

There are thus different views about whether ALs should have polite pronouns.

2.2 Forms of Address

As with polite pronouns, sources on ALs sometimes do not give instructions on when to use forms of address. Again de Beaufront, on Ido (1925/2005), does give instructions; on p. 184 he deals with salutations in letters. He begins by noting the need for establishing standard practices:

> Ta formuli esas afero di nacionala kustumo e stilo, e la simpla traduko di tala nacionala formuli genitus ne nur senfina diverseso, ma frazi stranja, nekomprenebla o miskomprenebla. Semblis do necesa fixigar, per konveniono, to quo devas konsideresar kom polita formuli.
>   ('These formulae are a matter of national custom and style, and the simple translation of such national formulae would generate not only endless diversity, but sentences which were strange, incomprehensible, or liable to be misunderstood. It therefore seemed necessary to fix, by convention, that which should be considered as polite formulae.')

---

[3] Gledhill's grammar of Esperanto is "corpus based", as it says in the title.

He then gives details:

> Ye la komenco di letro, ni uzez nur *Sioro*, e se la korespondanto havas ula titulo, funciono
> o profesiono, qua konsideresas en nia korespondado, ni skribez: *Sioro Prezidero*, *Sioro
> Profesoro*, e. c. Por iti, qui havas funciono o situeso, por qua la simpla *Sioro* ne semblas
> suficanta, ni havas la vorto *Sinioro*: *Sinioro Episkopo*, *Sinioro Ministro*.
>
> Ye la fino di letro, ni generale uzez: *Kun sincera saluto*. Por siniori e la personi, quin ni
> qualifikas «sinioro» ni dicez: *Kun respektoza saluto*.
>
> ('At the beginning of a letter, let us use only *Sioro* ['Sir/Madam'], and if the corre-
> spondent has any title, office, or profession which is considered in our correspondence, let
> us write: *Sioro Prezidero* ['President'], *Sioro Profesoro* ['Professor'], etc. For those who
> have an office or situation for which the simple *Sioro* does not seem sufficient, we have the
> word *Sinioro* ['Lord, Highness']: *Sinioro Episkopo* ['Bishop'], *Sinioro Ministro*
> ['Minister'].
>
> At the end of a letter, let us generally use *Kun sincera saluto* ['With sincere greeting']. For
> nobles and people whom we term "noble" let us say: *Kun respektoza saluto* ['With
> respectful greeting]'.)

Finally he states that these instructions do not apply to informal situations: "Komprenende ta reguli ne koncernas la korespondado kun amiki, kamaradi, parenti, qua admisas tre granda diverseso en ta formuli." ('Of course these rules do not concern correspondence with friends, comrades, relatives, which allows very great diversity in these formulae.') (On pp. 16–18 he treats forms of address more generally.)

In his book on Anderson (n.d.: 120) also discusses how to open and close letters:

> The practice of using "Dear (Sir/Madam)" as an introductory term in general corre-
> spondence is as unfitting in its expression of quasi-affection as "Sir/Madam" is in its
> curtness. Some form of the word "Respect" is seemingly appropriate here. The intro-
> ductory "(Dvm/Made) li hail": "(Sir/Madam) in respect" can be recommended; and this
> also in public address. When a Christian name is employed, then "– li zan",[4] or other
> more or less endearing expression may be deemed appropriate. As a completory, and with
> a like discretion: "Ua qu (li) yer/yrv/aye/sain bvn si": "To you (with/in) truth/earnestness/
> love/best wishes" have a simple dignity.

Another author who deals with this matter is Martines d'Antoñana, the designer of Neoispano; he writes (in Neoispano) (1973: 91):

> O terminologia do korespondensia, elemento tan importante en o komunikasión sosial i
> komersial, se modernisa en neoispano adoptando expresiones piu práktiko ao prinsipo i fin
> d letras. O expresiones, aktualmente en uso d "Muy señoro mío" i "Su seguro servidor",
> es no solomente antikuado, sino beromente ridíkulo, porke no tiene ningún signifikado. En
> korespondensia komersial, Señor: ao prinsipio, i, Salute. ao fin, es piu apropriado.
>
> Letra ao familia i relatibos: Kerido. Pa amikos: Estimado. Extraños: Señor. Ao fin d
> letra: Afektuosamente, Kordialmente, etc. Extraños: Salute.
>
> ('The terminology of correspondence, an element so important in social and commercial
> communication, is modernized in neoispanso, adopting more practical expressions at the
> beginning and end of letters. The expressions currently in use, "Muy señoro mío" [lit. 'my
> very (much) sir/lord'] and "Su seguro servidor" ['Your certain servant'] are not only

---

[4] The word *zan* means 'dear' and so I take *li zan* to mean something like 'with endearment'.

antiquated but truly ridiculous, because they do not have any meaning. In commercial correspondence *Señor* ['Sir'], at the beginning and *Salute* ['Greeting' (?)] at the end is more appropriate.

Letter to family and relatives: *Kerido* ['Dear']. For friends: *Estimado* ['Esteemed']. Strangers: *Señor*. At the end of letter: *Afektuosamente* ['Affectionately'], *Kordialmente* ['Cordially'], etc. Strangers: *Salute*.')

It is perhaps worthy of note that these authors give such importance to a relatively minor area of pragmatics (while saying nothing about some other issues).

ApGawain et al. (2008: 76) say the following about Ido forms of address:

1. Sioro (Sro): Mr/Mrs/Miss/Master/Sir/Madam. This can be used to address either a man or a woman, married or single, irrespective of age. For example in a business letter:

   Estimata Sioro—Dear Sir/Madam

2. Siorulo (S-ulo): Mr/Master/Sir. In practice this is not often used, Sioro being sufficient.
3. Siorino (S-ino): Mrs/Miss/Madam

   In practice Sioro is not often used for women, Siorino being prefered. This is partly as a compliment to the female gender and partly to help distinguish between different members of the same family: Sro e S-ino Smith—Mr. and Mrs. Smith
   Any woman has the right to be addressed as Sioro, should she so desire it. Note that Siorino can refer to both married and single women.

4. Damzelo (Dzlo): Miss/young lady:

   Should circumstances require it, an unmarried lady may be addressed as damzelo: Damzelo Jones, Yen S-ino e Dzlo Smith—Here are Mrs. and Miss Smith. La damzelo (qua esas) ibe—The young lady over there.

It is interesting that they speak of the "right" to be addressed in a certain way.

2.3 Honorifics

Few ALs have honorifics. This is not surprising since a large proportion of ALs are based on one or more Western European languages, which lack honorifics. However, there is an honorific in Sona. Searight (1935: 33–4) says the following about it, and about the need for it:

Sona has a special form of Article called the Honorific, borrowed from J. [= Japanese]. This is the vowel *o*; … It is used before names, forms of address, and verbs as an expression of politeness. Thus we have:- *o ra* '(honorable) man', 'gentleman', *o hara* 'sir', *o tu jiko* 'your (honour's) children', *o toru* 'please pass', *o min* '(please) come in!' We meet with so many ways of address in national languages, ranging from the flowery honorifics of the East to the laconic 'Say bo' of the new West, that we must have some

such mechanism to satisfy all tastes.

The Italian has no less than three forms of address—tu, voi, lei [sic]; the haughty Pathan but one—'ty'—whether to prince or pauper. In English we write Mr. to our tailor, but Esq. to our friends—yet neither word can we use in address. We have no happy way of calling the waitress. We hover between a furtive 'Miss' or a self conscious cough. In Sona the simple little vowel *o* solves all our difficulties. The word hara, both in writing and address, means 'Mister', whether tailor or friend, while *o hara* covers all the complications of Sir, Dear Sir, Respected Sir, Your Honour, and so on. Likewise *tu* 'you' is exalted to *o tu* in polite address, and solves the problem of 'Yours Truly', 'Yours Faithfully', 'Your Obedient Servant', and all the rest.

Ardano has the same honorific marker, but apparently it is only used with questions; Elhassi (2008: Lesson 10) says:

- Respectful sentences:
  We add (O) as a prefix before [t]he interrogative
  The idea is taken from Japanese.
  Hal → Ohal
  Hal ti posna aiutije min? = Can you help me ?
  Ohal ti posna aiutije min ? = Could you help me?
  Canjan → *Ocanjan*
  Canjan ti farna? = How are you?
  Ocanjan ti farna? = How do you do?

2.4 Imperatives

Some ALs have more than one type of imperative, the difference between/among them involving politeness or something like it. This is true of Volapük; Sprague (1888: 26) says:

The ending of the simple imperativ [sic] is **öd**, following the person-ending. […] There are two modifications of the imperativ, the courteous or softened form in **-ös** and the harsh form in **-öz**; called by some grammarians the optativ [sic] and the jussiv. The former expresses a request and the latter a positiv [sic] command.

In the exercise which follows this (p. 27) and the key to it (p. 118), Sprague usually uses *please* as the English equivalent of *-ös* and an exclamation point as the equivalent of *-öz* (sentences containing forms in *-öd* and *-ös* end with a period rather than an exclamation point), e.g.:

(1a) Kömolsös al visitön obis in dom obas nulik.
'Please come to visit us in our new house.' (p. 27, tr. p. 118)

(1b) Gololöz se dom!
'Go out of the house!' (p. 27, tr. p. 118)

Couturat and Leau (1903: 377) say about Kosmos:

> Pour l'*impératif*, on emploiera le subjonctif présent (forme polie): **amösi**, *aime*; **amösis**, aimez; pour un impératif plus bref et plus pressant, on emploiera le radical verbal en **-o** (avec **-s** au pluriel): **curro**, *cours*; **venios**, *venez*.
> ('For the *imperative*, one will use the present subjunctive ([as] the polite form) *amösi* 'love'; *amösis* 'love' [pl]; for a brief and more insistent imperative on will use the verbal root in -o (with -s in the plural): *curro* 'run'; venios 'come' [pl.].')

Eichhorn's Weltsprache also has two imperative forms; Couturat and Leau (ibid.: 300) state, "L'*impératif* se forme en ajoutant le pronom personnel à l'infinitif, et en intercalant un **ü** ou un **i**, suivant que le sens est plus ou moins impérieux" ('The *imperative* is formed by adding the personal pronoun to the infinitive, and inserting an *ü* or an *i*, according to whether the meaning is more or less imperious').

2.5 Word Order

In Ande word order can have a polite function; Anderson (n.d.: 97) says:

> Standard order is: Subject, Predicate, Object, Indirect Object … However, as sentential precedence may convey a suggestion of respect, or indulgence, it may be desirable on occasion to promote the Object. For example: "Me zan qu nu aya sio" : "Darling I love you so"; (Lit.—you I love so (much)).[5]

Although of course word order has pragmatic functions in language, I know of no natural (or other artificial) language in which it has this type of function.

3 Formal Language (and Other Levels of Language)

One might be surprised to learn that in some ALs there is, or is supposed to be, more than one level of language: although it is not very common, AL designers sometimes make remarks about words or constructions being appropriate for one kind of language.

Quiles (2009) makes several remarks about what should be, or is, used in a formal style of Modern Indo-European:

> In Modern Indo-European, compounds may be written with and without hyphen, as in the different modern Indo-European languages. Nevertheless, the older, not hyphenated version is preferred for formal writings; as, *sindhueurōpājóm*, and not *sindhu-eurōpājóm ['Indo-European'] (p. 126)
> The plural **wejes** ['we'] is often used for the singular **egó** ['I']; the plural **juwes** ['you'] can also be so used for the singular **tū**. Both situations happen usually in formal contexts. (p. 183)

---

[5] The meanings of the words in this sentence are *zan* 'dear', *qu* 'you', *nu* 'I', *aya* 'love', and *sio* 'very'; *me* marks nouns referring to females but it generally occurs after its noun rather than before, and so it is not clear to me whether there is an error in this sentence.

Formal writings in Modern Indo-European should follow the patterns attested in the oldest inscriptions, i.e. (S)OV, as in Vedic Sanskrit, Ancient Greek, Old Latin and Avestan. (p. 266)

In Modern Indo-European, thus, negation should usually be preverbal, as in modern Romance languages (cf. Fr. *n'est*, Spa. *no es*, etc.), but it can be postponed in emphatic contexts, as it is usual in modern Germanic languages (cf. Eng. *is not*, Ger. *ist nicht*, etc.), as well as in very formal texts, thus imitating some of the most archaic findings of early PIE dialects. (p. 300)

Stadelmann (1945: 36) indicates that there are several registers in Voldu:

In literary, scientific or commercial language it might be desirable to omit the personal pronoun. In this case the verb is conjugated like that:
    Present: Eleh, elez, eles, elek, elec, elet.
    Past: Elah, elaz, elas, elak, elac, elat.
    Imperat[ive]: Eluh, eluz, elus, eluk, eluc, elut.
    Ex: Roma(n) regoy governat. *Kings governed Rome.*
    (Rome was governed by kings) (Latin style).
    Da noktes. *Is* [sic, presumably should be *It*] *grows night. (Poetic style).*
    Vayuk! *Let us go!*
    Eluz man! *Be a man!*
    Elus kyet! (Hi sol el kyet!) *He is supposed to be quiet!*
    […]
    Sun zaynes. *The sun shines. (Poetic style).*

Not all of this is completely clear, but the basic idea seems to be that verbs can bear agreement marking in certain kinds of language; generally they do not, thus for example the present tense form of the verb *el* 'to be' is *el*, no matter what person or number the subject is, and the past tense form is *ela*. The agreeing forms are used when the subject pronoun is dropped. Given the first and last examples, it appears that they can be used even when there is an overt subject, as *regoy* 'kings' and *Sun* 'The sun' are the subjects of these sentences.[6]

The second example seems to be of the same type; one might think that *Da* is an error for *Dag* 'day' and is thus the subject, while *noktes* is a verb form derived from *nokt* 'night', specifically the 3rd person singular present tense form. This example might therefore be more literally translated as 'The day nights (i.e. becomes night)'.

The remaining examples involve imperatives. 2nd person imperatives do not have to have (and perhaps cannot have) an overt subject, e.g. *Kam tu mi!* 'Come to me!' (ibid.:43), but they also usually do not have any suffixes, i.e. the imperative form is identical to the infinitive. However, from the passage above we see that they can have an ending. One can express 1st person plural imperatives with the auxiliary verb *lar* 'let' and the infinitive of the main verb, e.g. *Lar nun vay!* 'Let us

---

[6] The *n* in *Roma(n)* is the accusative marker, which does not always occur and may sometimes be optional, as it appears to be in this example.

go!' (ibid.:64), and this is perhaps what is done in normal style, but there is the option of having marking 1st person imperatives synthetically (as well as 1st person singular imperatives). The same general situation seems to hold for 3rd person imperatives, as shown by the alternatives given in the passage, *Elus kyet!* and *Hi sol el kyet!*

Stadelmann (1945) also makes some remarks about "familiar" language: on p. 24 he says, "In familiar speech the *ke* of a conjunction can be dropped" and on p. 36 he states, "Contractions can be imitated from English in the following way (Familiar style): Ex: Yu'l n't. (Yu el not). ['You are not'] Hi'av n't. (Hi hav not). ['He does not have']".

Simon (2010: sec. 3.4) says, "Due to its Indo-European heritage, Sambahsa has some verbal forms only encountered in literary usage". One might think this odd, since various other ALs based (largely) on Indo-European languages do not have such forms, and in any case Simon (the designer of Sambahsa) was under no compulsion to retain such forms. These forms include some alternative person agreement endings:

> If these are compatible with the accentuation, verbs can bear these endings in the present indicative:
>
> 1° person singular: *-mi*
> 2° person singular: *-si*
> 3° person singular: *-ti*
> 4° person plural: *-nti.*
>
> The corresponding forms for «ses» ['to be'] are *esmi, essi, esti, sonti.*
> The other conditions for the use of these forms are that the verb stands [sic] in absolute initial position in the clause, and that this clause contains [sic] no adverb. Those conditions are seldom fulfilled. (ibid.)

That is, Simon has created forms which will rarely occur, which is reminscent of the situation with Esperanto *ci*. One might wonder why he went to the trouble of doing this. In the same section Simon also brings up two sets of "old forms", imperatives and infinitives. They cannot literally be old, since Sambahsa itself does not seem to be very old (its presence on the internet dates from 2007), but perhaps they are meant to have the feel of archaic language.

There are forms from other word classes which are also used in particular circumstances; Simon (ibid.:sec. 2.2) states:

> In Sambahsa-mundialect, endings with declensions can be added to substantives and adjectives for purposes of *euphony* or literary purposes (ex: *poetry*). This system, whose native name is *euphonic vocalisation*, can only be used if it is compatible with the accentuation patterns. For example: *uno smiegdo geront* "a frail old man" instead of *un smiegd geront*. In everyday use, those endings appear only in the words *vasyo* (all of the, every) and *alyo* (another).

In Esperanto there are some "unofficial" elements (e.g. the suffix *-ator-*, which forms words for machines). Although one would not want to interpret their existence to mean that there are two levels of the language, an official and an unofficial one, there are different elements with a different status and the language

is not completely under prescriptive control. This of course is the situation with natural languages (in spite of organizations such as the French Academy), and is perhaps to be expected of an artificial language which sees a large amount of use. Such facts could be taken as support for the idea that there can be a pragmatics of an AL, as an AL can be used extensively and is to some extent under the control of its users, i.e. it is not just an abstract and sterile creation.

An interesting question about formal language in ALs is which features are considered to be formal; given that an AL designer has complete control of his language (at least at the beginning), he can label any words, contructions, etc. that he wants as formal (literary), etc. One might wonder what effect his native language (or any other language) could have on this.

4 Conversational Implicature

I have found no explicit discussion of conversational implicatures in instructional materials for ALs, and I have not found many examples of them in AL texts. Here is one example in Usik, from Palanca Gómez (2008: 126), in a pseudo-dialogue in which someone is reporting an accident[7]:

Bulki bam?
'Is there any wounded?'

Bel, ank tendi solno bo nuski leabi
'Well, a woman lying down the floor that she breathes troublesomely' [sic]

Here the maxim of relvance is involved; the answer does not directly mention anyone who is wounded (and one could have difficulty breathing without being injured). Notice that it is introduced by *bel* 'well', which apparently can be a marker of an utterance involving a conversational implicature, as *well* can be in English.

5 Non-descriptive Meaning

5.1 Conjunctions

Sources on ALs usually are not explicit about the non-truth conditional meanings of some conjunctions. For example, a word may simply be glossed as 'but'. However, Wennergren (2005: 304–5) goes into detail about the uses of the Esperanto word *sed* 'but':

---

[7] By "pseudo-dialogue" I mean a dialogue which (presumably) has not actually occurred, but which has been created by a language designer/presenter to illustrate some point of grammar or for reading practice.

Sed ligas frazojn aŭ frazpartojn, kaj montras, ke ili iel kontrastas inter si: … *Sed* povas ankaŭ enkonduki ion, kio malebligas aŭ malhelpas la antaŭan aferon: … *Sed* povas enkonduki ion, kio estas surpriza, se oni konsideras la antaŭan aferon: … Se la antaŭaj vortoj esprimas neadon de io, *sed* povas enkonduki tion, kio validas anstataŭe: … Iafoje *sed* montras kontraston al io subkomprenata

('*Sed* connects sentences and parts of sentences, and shows that they contrast among themselves: … *Sed* can also introduce something which makes impossible or hinders the previous matter: … *Sed* can introduce something which is surprising if one considers the previous matter: … If the previous words express a denial of something, *sed* can introduce something which is valid instead: … Sometimes *sed* shows a contrast with something assumed')

Wennergren then (p. 305) discusses what he calls "nuanca *sed*" ('nuanced *sed*'):

Nuanca *sed* ne ligas du aferojn, sed enkondukas frazon, kiu esprimas ŝanĝon de paroltemo aŭ interrompon:

- **Sed ni ne parolos plu pri tiu ĉi punkto**. Oni povus imagi subkomprenitan antaŭfrazon: *Eblus daŭrigi, sed…*
  […]

Nuanca *sed* povas ankaŭ enkonduki elkrion de surprizo aŭ malkonsento. Tiam oni iafoje povus diri, ke *sed* kontrastas al la eldiro de alia persono:

- *Mi donos al vi kvin eŭrojn.—**Sed** tio ne estas justa!*
- *Li venkis en la konkurso.—**Sed** tio estas ja bonega!*

('Nuanced *sed* does not connect two matters, but rather introduces a sentence which expresses a change of subject or an introduction:

- **But we shall not speak more concerning this point**. One could imagine an understood preceding sentence: *One could continue, **but** …*
  […]

Nuanced *sed* can also introduce an exclamation of surprise or disagreement. Then one could sometimes say that *sed* contrasts with the statement of another person:

- *I shall give you five euros.—**But** that is not right!*
- *He won in the contest.—But that is indeed excellent!'*)

Wennergren (2005, 300) seems to be aware of the conventional implicature conveyed by *kaj* 'and' that there is some relation between the two clauses that it connects, stating, "*Kaj* povas ligi tutajn frazojn, kiuj iamaniere kunapartenas" ('Kaj can connect whole sentences which in some way belong together'). He goes further and gives (p. 301) examples of sentences in which *kaj* indicates different ways in which clauses "belong together".

Kiam frazoj estas ligitaj per *kaj*, tiu ligo povas reprezenti multajn diversajn signiforilatojn. Kia estas la rilato, oni devas kompreni el la kunteksto. Eblas klarigi la rilaton per aldonaj esprimoj:

- *Mi lavis la vestaĵojn, **kaj** (poste) mi sekigis ilin. Kaj* montras tempan sinsekvon. La vorto *poste* je bezono povas helpi al kompreno.
- *Ŝi aŭdis teruran bruon, **kaj** (tial) ŝi telefonis al la polico. Kaj* montras sekvon aŭ rezulton. *Tial* povas helpi al kompreno.
- *Anno estas gaja persono, **kaj** (kontraste) Elizabeto estas silentema. Kaj* montras kontraston.

- *Ni klopodis per ĉiuj fortoj, **kaj** (tamen) ni malsukcesis. Kaj* montras neatenditan sekvon. *Tamen* povas pliklarigi tion. Ankoraŭ pli klara estus *sed* anstataŭ *kaj*.
- *Promesu neniam plu fari tian stultaĵon, **kaj** (tiam) mi helpos al vi reordigi la aferon. Kaj* montras kondiĉan rilaton inter la frazoj. *Tiam* pova helpi al la kompreno. Eĉ pli klare estus transformi la unuan frazon en subfrazon kun *se*: *Se vi promesas neniam plu fari tian stultaĵon, (tiam) mi helpos al vi reordigi la aferon.* […]
- *Ŝi havas brunajn okulojn, **kaj** (aldone) ŝiaj haroj estas longaj. Kaj* montras aldonan informon.
- *Li preferis foriri tre frue de la festo, **kaj** tio ne surprizas min, ĉar vere estis tre enue tie. Kaj* montras komenton.

   ('When sentences are joined with *kaj*, that connection can represent many different meaning relations. What kind of relationship there is can be understood from the context. It is possible to make the relationship clear with additional expressions:

- *I washed the clothes **and** (afterwards) I dried them. Kaj* shows a temporal sequence. The word *poste* if necessary can help with understanding.
- *She heard a terrible noise **and** (for that reason) she phoned the police. Kaj* shows a consequence or result. *Tial* can help with understanding.
- *Anna is a cheerful person, **and** (in constrast) Elizabeth tends to be quiet. Kaj* shows a contrast.
- *We endeavored with all our strength, **and** (nevertheless) we failed. Kaj* shows an unintended consequence. *Tamen* can make it clearer. Still clearer would be *sed* ['but'] instead of *kaj*.
- *Promise never to do that stupid thing again, **and** (then) I will help you put the affair in order again. Kaj* shows a conditional relation between the sentences. *Tiam* can help with understanding. It would be even clearer if one changed the first sentence into a subordinate clause with *se* ['if']: *If you promise never to do that stupid thing again, (then) I will help you put the affair in order again.* […]
- *She has brown eyes, **and** (in addition) her hair is long. Kaj* shows additional information.
- *He preferred to leave the party very early, **and** that does not surprise me, for it was really dull there. Kaj* shows a comment.')

Such detailed accounts of the meanings of a word for 'and' (or any conjunction) are very rare in grammars of either natural or artificial languages.

5.2 Interjections

The (type of) meaning contained in interjections has been difficult to describe. Some analysts have claimed that they involve conventional implicatures. Some AL designers describe the meanings and/or functions of interjections, e.g. Ruggles (1829: 58) states, "Interjections serve to express some passion or emotion. […] They are either positive or contrastive".

   Vidal, the designer of the Langue universelle et analytique has an interesting view on interjections (1844: 39):

> Le vocabulaire de la langue universelle doit commencer par les interjections, parce qu'elles représentent chacune une pensée tout entière, quelle que soit la nature des signes dont on se sert pour cet effet. L'artifice de la parole consiste à développer ces pensées en séparant l'attribut du sujet par un assemblage de mots que noun nommons proposition.
>
> ('The vocabulary of the universal language must begin with the interjections, because each of them represents a whole thought, whatever be the nature of the signs which one makes use of for this effect. The contrivance of speech consists in developing these thoughts by separating the attribute from the subject by means of a combination of words that we call a *clause*.')

His language has a relatively large number of (words which he calls) interjections; he says (ibid.:40):

> Il est des circonstances où il est tres-essentiel de pouvoir exprimer une idée au même instant qu'on la conçoit: si l'on veut, par exemple, appeler au secours dans un danger pressant. J'ai cru devoir agrandir un peu le cadre de celles que nous avons, en observant toutefois que l'on ferait peut-être bien de l'agrandir davantage
>
> ('There are circumstances when it is very essential to be able to express an idea at the same instant when one conceives it: if one wants, for example, to call for help in pressing danger. I thought that it was necessary to expand the range of those [interjections] that we have, while however observing that one would perhaps do well to expand it more')

Some of the interjections of his language are equivalent to words or sentences of French, e.g. *hol* 'merci' ['thank you'], *hap* 'que voulez-vous?' ['what do you want?'], while others express emotions.

This might lead one to reflect on how to define the class of interjections. If we require interjections to have only non-truth conditional meaning, then many of the words which Vidal calls interjections will not be such. Of course there will be other words which have some non-truth conditional meaning, e.g. *but*, but perhaps interjections have only this kind of meaning. The question then is whether Vidal and some other language designers are in conflict with linguistic thought about the nature of interjections.

Wennergren (2005: 314) classifies the Esperanto word *nu* 'well; now' among the interjections and says that it has a variety of meanings:

> *Nu* estas ĝenerala atentiga vorto. Ĝi iel montras al la alparolato, ke io speciala sekvos. La precizaj nuancoj estas tre diversaj. Alvoko: *"Nu, mia filino?"*—*"Jes, patrino."* *"Nu, Alfred,"* ekkriis la doktoro, *"kion vi diros al tio ĉi?"* Malpacienco: *Nu, iru pli rapide! Nu*, *nu malsaĝulo, ĉesu!* Konsento, koncedo: *Nu, bone, bone! Nu do, venu, se vi vere insistas.* Fino de interparolo: *Nu, bone do, tial ni iru! Nu, bonan nokton!* Surpriziĝo: *Nu! Kiu supozus ion tian? Nu, mi neniam atendis tion!* Dubo, hezito: *Nu, kiel nun, Anton Antonoviĉ?* Klarigo: *Nu! Tion kaŭzis difekto en la hejtilo.* Rezignacio: *Nu, tia estas la vivo. Nu, kion fari?* Konsolo: *Nu*, *nu! Ne ploru!*
>
> ('*Nu* is a general attention drawing word. It shows the addressee in some way that something special will follow. Its precise nuances are very diverse. Calling: *"Well, my daughter?"*—*Yes, mother."* *"Well, Alfred",* exclaimed the doctor, *"what will you say to this?"* Impatience: *Now, go more quickly! Now*, *now you fool, stop!* Agreement, concession: *Well, good, good! Well then, come, if you really insist.* End of conversation: *Well, good then, so let's go! Well, good night!* Surprise: *Well! Who would think such a thing? Well, I never expected that!* Doubt, hesitation: *Well, what now, Anton Antonovich?* Clarification: *Well! A defect in the heater caused that.* Resignation: *Well, that's life. Well, what can one do?* Consolation: *Now*, *now! Don't cry!*')

Thus, as with his discussion of *kaj* 'and', he gives a more detailed account of the meanings of this word than one will find for the equivalent words in most grammars of natural or artificial languages.

5.3 Illocutionary Force

Utterances with an indirect illocutionary force occur commonly in natural languages. One might therefore expect them to occur in ALs as well. On the other hand, one could see them as a source of possible confusion, like non-literal language (see the next section) and try to forbid their use. To my knowledge, no AL designer has taken such a step. However, what amounts to the same thing has been done in Seaspeak, a controlled language based on English for nautical purposes: Seaspeak has mandatory markers of speech acts. Weeks et al. (1988: 96) (as quoted in Kimbrough and Yang (2005: 303) state:

> Maritime messages transmitted over VHF should be short, accurate, and relevant. Furthermore, messages should be transmitted in language simple enough for a non-native speaker of English to comprehend without difficulty.
>   One useful means of making the language simpler is to indicate, at the beginning of a message, what sort of message it is going to be. Thus, if a question is going to be asked, the speaker simply says the word 'QUESTION' before the question itself. Similarly, if a piece of advice is going to be given, the speaker says the word 'ADVICE' in advance of his message. There are just seven of these Message Markers and after a little practice, learners should experience no difficulty in using them.

Presumably sentences marked as questions could not be used e.g. as requests in Seaspeak, nor could there be rhetorical questions. However, given the limited domain in which Seaspeak was intended to be used, such indirect speech acts might not occur anyway.

Consider now the following text in Konya (Sulky 2005):

Dialogue 1:
Illustrating the perils of literally translating idiomatic English expressions into Konya:

mon-misi Xenya kesati	xenye tenwi pofu toku moti yu tufu we xuxin-xeni
Ms. Shen (a visitor)	Can you tell me how to find the train station?
Lukya pan-kesati	sunye
Luka (a local)	Yep
mon-misi Xenya	lawa… toku xexenye
Ms. Shen	Well… tell me, please
Lukya	sinu mepi yo moti taunwe
Luka	Look at a map. That's how

Sulky then comments, "One hopes that in future, Ms. Shen will simply say: **xuxin-xeni wa lito kin-lunwi xexenye** *Where is the train station, please?*" He seems to be thinking incorrectly; the question "Can you tell me how to find the train station" is perhaps not what one would call "idiomatic English" but rather involves an indirect speech act, a request in the form of a question. If we take Sulky's remark in a general sense, then he is saying that there should not be indirect speech acts in Konya. However, this could make Konya unsual when compared with natural languages since one might think that indirect speech acts occur in all of the latter; according to Brown and Levinson (1987: 142), "indirect speech acts are universal and for the most part are probably constructed in essentially similar ways in all languages".

Indirect speech acts can be found at least some ALs; below are some (possible) examples:

Ido:

Kad vu voluntus pasigar la pano?

'May I trouble you for the bread?' (ApGawain et al. 2008: 34)

Here (from a group of example sentences, i.e. not one that actually occurred in speech) we have a question functioning as a request. For some reason, ApGawain's English version is not literal; a more literal version is 'Would you be willing to pass the bread?' (the English version is given before the Ido version).

Eurolengo [from a simulated "typical business letter" (Jones 1972: 63)]:

nos gustaral resevar sampels and pryses for noster consideration

'we would be pleased to receive samples and prices for our consideration' (ibid.:62, tr. 63)

This statement could be seen as an indirect request.

Interlingua (IALA)[8] The context is from a pseudo-dialogue involving a man, his son, and another man. The boy asks his father what the other man is doing; apparently he is writing a letter:

Johnny (al senior): Senior, scribe vos un littera a vostre matre?

'Johnny (to the gentleman): Sir, are you writing a letter to your mother?'

Le senior (in un tono un pacuo irritate)	No!
'The gentleman (in a tone a little irritated)	No!'
Johnny	Que face vos alora?
'Johnny	What are you doing then?'

---

[8] There are two ALs known as Interlingua, the one connected with the International Auxiliary Language Association (IALA) and the one also called Latino sine Flexione.

Le senior	Io attende
'The gentleman	I am waiting.'
Johnny	Que attende vos?
'Johnny	What are you waiting for?'
Le senior	Io attende un momento de silentio pro finir iste littera
'The gentleman	I am waiting for a moment of silence to finish this letter.'
Johnny	Papa, ille attende un momento de silentio
'Johnny	Papa, he is waiting for a moment of silence.'
Senior Smith (con un profundo suspiro)	Ah, si. Multe gente attende un momento de silentio
'Mr. Smith (with a deep sigh)	Ah, yes. Many people wait for a moment of silence.'
Johnny	E nos, que attende nos? Nos non attende un momento de silentio, nonne?
'Johnny	And us, papa, what are we waiting for? We aren't waiting for a moment of silence, are we?'
Senior Smith	No, Johnny, nos attende mama!
'Mr. Smith	No, Johnny, we are waiting for mama!' (Gode 1954: 41)

One could interpret the gentleman's statement that he is waiting for silence as an indirect request for silence. Johnny fails to understand this, perhaps due to the fact that he is a child (one should bear in mind that this passage is meant to be humorous, and without Johnny's misunderstanding it would not be so), and his father does not enlighten him. (The book in which this appears, Gode (1954), is based at least in part on a book for learning Spanish, and so this psedo-dialogue may originally have appeared in Spanish or English; one might think that the indirect speech act occurs here for that reason. However, if such speech acts were not permitted or were not possible in Interlingua, presumably this pseudo-dialogue would not have occurred in Gode (1954).)

6 Metaphor and Non-literal Language

Although metaphors, and non-literal language more generally, are a significant part of every natural language, some ALs designers frown on their use and try to eliminate them from their languages. Consider the remarks of Morneau (2006: sec. 27.0) (Morneau is the creator of Latejami):

> when speakers of natural languages use non-literal language it is almost always because they cannot avoid it either because their vocabulary does not have an appropriate literal construction available, or because it is something that the speaker is not comfortable using.
>
> This is unfortunate because the way that a non-literal construction will be interpreted will depend very much on the native language and culture of the listener. For example, metaphoric use of the word "pig" can have meanings such as "slob", "sex maniac", or "over-eater" in English, but will have different meanings to speakers of other languages. Also, as we've seen many times throughout this monograph, many metaphors, including the above examples, can be avoided by using appropriate derivations instead. For example, pejorative morphemes or more precisely derived compounds can be used to implement the above examples. In fact, I have become completely convinced that a properly derived word can replace **any** required or unavoidable metaphor, and it can never be misinterpreted by native speakers of other languages.
>
> The goal of a designer of an MT interlingua should be to provide the means to say **anything** without the need for non-literal language. In other words, metaphor, polysemy, and idiom should be optional—they should **never** be obligatory. It is also my opinion that non-literal language should be generally avoided (except where its use is obvious to all listeners or readers), since the possibility for misunderstanding is so great.

Morneau's statement that non-literal language is not employed by choice seems to be incorrect. Other AL designers share his negative view of non-literal language: in his work on Hom-idyomo Cárdenas (1923: 153) says:

> Words should be used and interpreted in their natural meanings and not with the figurative meanings they may have in other languages. Thus, *pesto* ['pest'] should not be used in the sense of "invective", nor *maɹo* ['sea'] in that of "abundance", nor *nigɹa* ['black'] in that of "sad" or "gloomy". The expression, *Gladyo sitya di sango* ['sword thisty for blood'] may be very poetical, but it is not true. A language which has not come into general use should not be employed figurative, at least for the translation of idioms and saying. That will come later.

What is probably behind such views is a desire for clarity and ease of understanding in a language. A way of preserving this while permitting non-literal language is to have an indicator of non-literal language, and several ALs have such a device. One of these is Sotos Ochando's Lengua Universal (an a priori language). Gisbert and Lorrio in their (1862) pedagogical grammar of this AL express a more tolerant view of non-literal language: (pp. 76–77):

> Se ha sentado por base que las palabras de la Lengua Universal no han de tener doble sentido, en lo cual consiste una de sus principales ventajas; pero debe comprenderse que al decir que ninguna palabra tendrá doble sentido, no excluimos ni podemos excluir las figuras que como naturales al hombre, y procedentes de su mismo espíritu, no hay nadie que pueda proscribirlas de la locucion. Su usa nunca induce á error, como acontece con las

palabras que sin usarse figuradamente tienen sentidos diversos, representando objetos distintos. Cuando de un Papa se dice que *lleva el timon de la nave de la Iglesia*, nadie se confunde, nadie toma á la Iglesia por un buque: quando se dice, *hay nubes en el horizonte politico*, nadie mira al cielo á ver si está raso. […]—En este Lengua subsistirán por consiguiente las metáforas y las figures todas del pensamiento, y subsistirian aunque la voluntad del autor fuera excluirlas; pues son efecto natural de la imaginacion que en ellas busca la explicacion mas viva y pintoresca de sus ideas, y que sin sablerlo las usa (de continuo) aun en el lenguaje vulgar, principalmente al hallarse afectada de una pasion cualquiera.

('It has been stipulated that the words of the Universal Language should not have double meanings; in this consists one of its main advantages; however, this should not be understood to mean that no word will have a double meaning: we do not exclude, nor can we exclude figures [of speech]; as they are natural to man, and proceed from his very spirit, there is no one who could proscribe them from speech. Their use never leads to error, as it does not with words which, without being used figuratively, have different meanings, representing different objects. When someone says of a Pope that "he takes the helm of the ship of the Church", no one takes the Church to be a sailing vessel; when someone says, "there are clouds on the political horizon", no one looks at the sky to see whether it is clear. In this Language consequently there will still be metaphors and all the figures of throught, and they would exist even if it were the will of the author to exclude them, since they are a natural effect of the imagination, which seeks in them the most vivid and picturesque expression of its ideas, and which, without knowing it, uses them (continuously) even in everyday language, mainly when affected by some passion.')

Oddly enough, in a work by Sotos Ochando himself (1863) there is a more restricted view of non-literal language (p. 33):

En la Lengua Universal no se admiten figuras de letras ni de diccion. Solo se exceptúan las licencias esplicadas sobre esto, porque están tan fijas sus reglas, que no dejan lugar á dudas ni equivocaiones. Sin embargo, es conveniente y aun necesario admitir tres clases de figuras ó metáforas de sentido.
1. ª La sustitucion de ciertas voces, cuando la reclaman la moral, la decencia, el decoro ó la delicadeza. Tales son las que significan las partes pudendas de ambos sexos, varios de sus actos, los objetos que excitan asco, etc., etc.
2. ª La metáforas en que se toma un objeto, una cualidad, una accion por otra, atendida la relacion que tienen entre sí los objetos por su semejanza, participacion ú otra causa, v.g., cuando decimos *la aurora de las ciencias*, *el azote de Dios*.
3. ª Las figures de retórica, á lo menos muchas de ellas, como la hyperbole, la ironia, la personificacion.

  En efecto, estas tres clases de figures están fundadas en la naturaleza del hombre, y con mas ó menos extension son communes á todos los tiempos y paises.

('In the Universal Language figurative language is allowed neither in written nor in spoken language. The only exceptions are the liberties explained concerning it, because their rules are fixed to such an extent that they do not allow room for doubts or errors. However, it is desirable and even necessary to allow three kinds of figures or metaphors of meaning.
1st: Substitution for certain words, when morality, decency, decorum, or delicacy calls for it. Such [words] are those which signify the private parts of both sexes, various of their acts, things which arouse disgust, etc., etc.
2nd: Metaphors in which a thing, a quality, an action is taken for another one, on the basis of a relationship which holds between the things because of their similarity, participation [in the same action?], or another reason, e.g. when we say *the dawn of the sciences, the scourge of God*.
3rd: Rhetorical figures, at least many of them, such as hyperbole, irony, personification.

> Indeed, these three kinds of figures are grounded in the nature of man and to a greater or less extent are common to all times and countries.')

The indicators mentioned above are not always required in Sotos Ochando's language, nor are those of another type, those which mark the technical use of a word; among the "licencias" ('liberties') mentioned by Gisbert and Lorrio (1862: 94–5) is the following (p. 95):

> Pueden tambien suprimirse muchas veces los signos del sentido metafórico; pues que este se conoce con frecuencia por el contexto, como sucede en las demas lenguas, que no tienen otro medio de indicarlo. Esta supresion será muy conveniente en algunos casos, para que aparezca la fuerza y belleza de la metáfora.
> Las mismas observaciones pueden aplicarse al uso de los signos del sentido técnico.
> ('Many times the markers of metaphorical meaning can also be omitted, since this [meaning] is often known by the context, as happens in other languages, which do not have any other means of indicating it. This omitting will be very desirable in some cases, in order that the strength and beauty of the metaphor appear.
> The same observations can apply to the use of the markers of technical meaning.')

In addition to these words Sotos Ochando's Lengua Universal has a set of indicators which seem to be heterogeneous; Sotos Ochando (1863: 37) labels them "[v]oces significativas de un sentido especial de la frase" ('words indicating a special meaning of the sentence'); the list of those that he "proposes" is (ibid.):

Ar para sentido interrogativo ('for interrogative meaning')
Er para el dubitativo ('for dubitative meaning')
Ir para el irónico ('for ironic meaning')
Or para el admirativo ('for admirative meaning')
Ur para el de sorpresa ('for surprise')
As para el de sentido optativo ('for optative meaning')
Es para el de indignacion ('for indignation')
Is para el depreciativo ('for depreciative')

I say that they may be heterogeneous because, at least in natural languages, their equivalents would not all be of the same type. Some of them might be equivalent to mood/modality markers, e.g. the Greek optative mood endings. However, this would not be true of e.g. *ir* or *es*; I do not know of any language that has an "ironic mood" or an "indignant mood". With the possible exception of *ar* (the only one of these markers that I have found in any texts in the language), I do not believe that any of these words are markers of illocutionary force along the lines of those in Seaspeak. (I am not certain whether even *ar* is such a marker, since it is not clear whether a sentence marked with it could have an indirect illocutionary force of something other than a question.[9]) I would hesitate to classify words such as *ir* and *es* as interjections, although they express emotions,

---

[9] Unlike Esperanto's *ĉu*, it can introduce *wh*-questions; it apparently is not required with either *yes–no* or *wh*-questions.

since, if they act syntactically like *ar*, they are part of a sentence, while one feature attributed to interjections is their lack of connection with other syntactic units. (It is difficult to determine whether all the words of this group act in the same way, since, as I have noted, *ar* is the only one that I have seen in context.)

Sotos Ochando (ibid.) justifies the creation of these words as follows: "Aunque el tono debe acompañar y caracterizar estos sentidos, es convenientísimo que haya palabras bien fijas y determinadas que excluyan toda duda y equivocacion." ('Even if the tone [i.e. intonation] must accompany and characterize these meanings, it is most advantageous that there be very fixed and specified words that would eliminate all doubt and error.') He also says, "Esta clase de modificativos son frecuentes en todas las lenguas" ('This type of modifiers is frequent in all languages').

Glosa is another AL which is not friendly towards figurative language, and, like Sotos Ochando's Lengua Universal, there is a marker for it, or rather, such a marker is recommended; Sect. 8.vii of Gaskell (1999) is called "Idiom and metaphor" and it reads as follows:

> For clarity of expression across cultural boundaries these should be avoided in Auxiliary Language usage; however, where it is necessary—for literary purposes—to quote a national-language idiom within Glosa, then such non-literal language ought to be marked with some 'neutral' symbol, EG pluvi ^plu feli e kani^ ['rain ^cats and dogs^'].

Grzega (2005: 67–8) says about Basic Global English, a simplified version of English which could be used both for pedigogical purposes and as an international auxiliary language, that "Native and advanced non-native speakers of English are asked to… abstain from metaphorical expressions that cannot be interpreted word-for-word (as these have shown to be problematic in lingua-franca communication)—in this respect a certain awareness competence might have to be trained". In a later paper (Grzega 2008) he states (p. 140), "Metaphors should only be used if objectively obvious and should be marked (*this is like*…)". That is, Basic Global English, like Sotos Ochando's Lengua Universal and Glosa, has an overt means of marking metaphors.

Those language designers who try to exclude the possibility of using non-literal language in their ALs may be attempting something which would make their languages unnatural, at least if they are languages intended for the same general purposes as natural languages, although this does not apply to all artificial languages, if we take "artificial languages" in a broad sense; consider the following remarks by Cohen (1993: 59):

> it is clearly characteristic of certain categories of artificial languages that they must lack any possibility of metaphor. Programming-languages for computers, like Fortran, or interpreted formal systems, like Carnap's, would be very seriously flawed in the performance of the tasks for which they are severally designed if they allowed their component words or symbols to be attributed new and unstipulated meanings in certain contexts. It is arguable, therefore, that we radically blur the difference between these kinds of artificial languages on the one hand, and ordinary natural languages, like English (or artificial languages for everyday use, like Esperanto), on the other, if we do not allow essentially for the possibility of metaphor in our analysis of the latter.

Metaphors and non-literal language do occur (without being marked as such) in at least some ALs. Consider what Gledhill (2000: 122) says about Esperanto:

> Many expressions are used non-literally in Esperanto. For example, *forpasi* 'to pass away' = to die, *zumi* 'to buzz' = to potter about, *celi* 'to aim' = to get at/mean something. Some expressions, especially compounds, involve a non-literal sense derived from the donor languages (for example, *librotenado* from English 'bookkeeping') or have emerged because of generally expressed euphemisms (*necesejo* 'the necessary place': the toilet). […]
>
> Proverbs and clichés are complex and very fixed forms of metaphor. They generally involve truth-statements and have an element of word play about them. In the *Fundamento* Zamenhof equipped the language with a large number of proverbs… […]
>
> Although most proverbs are rarely invoked in the general language, they form part of the basic repertoire of metaphors which the speaker may allude to (*fera mano* 'an iron hand', *amata ĉevaleto* 'hobby horse'). By writing down a set of proverbs, Zamenhof effectively created an oral history of the language, a corpus of expres[s]ions to dip into and cite or reformulate.

Note also the existence of the book Dahlenburg (2006), whose subtitle is *Stilfiguoroj en la poezio de esperanto* ('*Figures of style in Esperanto poetry*'); among the figures included are metaphor, irony, and euphemism.

7 Texts and Comments

I now present some texts and examine them from a pragmatic point of view. I have already mentioned the fact that ALs are generally not native languages, and thus the pragmatics in them may be affected by a speaker's first language. Such a carry-over may also occur in AL texts which are translations of texts in another language, so for pragmatic analysis it is probably better to look at original texts in an AL (unless one wants to see whether there are any pragmatic differences between the original and the AL translation). Unfortunately from this point of view, many texts in ALs are translations of well-known works in natural languages. Even if a text is not obviously a translation, or stated to be one, it may well be one: given that even creators of ALs have another language as their first language, many or most texts which they write in their ALs may be translations from their first language. That is, they may usually think of or formulate a sentence in their native language and then give its equivalent in their AL.

It may therefore be almost impossible to find texts in ALs which are not translations in some sense, with the exception of those in the major ALs such as Esperanto which some speakers have been using extensively for decades (and are thus able to think in, i.e. to use without first translating in their minds from another language), and these major ALs may be of less interest here because they are a posteriori. However, we can at least avoid the texts that we know to be translations (unless we are specifically interested in the possible transfer of pragmatic features in translation).

It may also seem best to analyze conversations or dialogues rather than passages involving only a single participant, and many works on ALs contain what

appear to be conversations (given in written form). However, they are not real conversations/dialogues in the AL, since they were created by a single person (the author of the work); rather, they are presentations of what a conversation in the language (on a given topic) would or should be like. As noted above, one could assert that they still represent a type of language use, just not the use that they appear to be (i.e. conversation). There may be real dialogues in some a priori and mixed ALs, for example postings and replies to them in internet sites devoted to these languages, but these may often be by people who are not completely fluent (or even close to fluent) in the AL, and thus their pragmatic (and other) features may be taken from their first language. A further problem with some AL texts is that it is not always clear whether a sentence is meant to be part of the same conversation as the previous sentence, or indeed whether a conversation is being portrayed at all.

I have chosen for analysis a text from the a priori language aUI which contains an apparent dialogue. However, it is not even a pseudo-dialogue of the sort just described, but a dialogue occurring within a fictional piece, one of the "[r]eadings" contained in Weilgart (1979). Nevertheless, it may be of particular interest because the participants in it are a human child and an alien who is partly animal and partly plant; if there were a situation where one might expect unusual pragmatics, this might be it. First I give the context only in English translation:

> Ever since the space-man had visited Johnny, the boy had envied the little animal-plant of quiet mind; for this creature could travel through infinite space. Why could not Johnny return the visit and come to the distant planet of the star in outer space? The spaceman had told Johnny that in machines and inventions the earthmen were just as good as the spacemen (Weilgart 1979: 165)

Below is the dialogue that follows this passage:

2) –"fUd fnu tykwerv a!", DJoN nIpaV.
2) "Then let us conquer space!", cried Johnny.

3)– yUg, a-u tygrOpAv: 'bum UI sEfU: "tykwe" Ub a, vEv fE tykwe y-twam rUt bnu. rUt-A-jAg bnu cEv jOm bnu cEv, bnu yA watAv tykwev a. pfE ć, bnu yA watAv dav ad nEn Oki akiA Ud bUt ne Ub i yt a "bijE-da" tygle at yf. fnu a-u, xnu dav fUd, yc fUIv sE: "tykwev" a. –
3) But the spaceman had smiled: "Your very word 'conquest' of space makes this conquest impossible for you. As long as you are as you are, you will never be able to conquer space. That is, you will never be able to travel through thousands of years with almost the speed of light from one 'milky-way' spiral to the next. Now we space-men who do this do not call it 'conquering' space.

4)– Fnu rykOmQ UIv: agtev a, gaf, tykev tag a; yUt ag-niO-Uj bu tyv at retgUv ypums Ub a, bu pI nEk iUv, hU u yc wav tykwev sE. am KU tykwev a. yUg fnu cEv yn-ynam ayn Ub knynE Ib, rUt fnu, a cEv y-tnak-wam.

4) We say modestly: entering space, or, submerging in space; for the more you get to learn the mystery of space, the more you understand why man cannot conquer it. Only God conquers space. But we are tiny specks of dust and, for us, space is unlimited."

II) dvU, hUd au av. (The Spaceman's method).

1. – ynDJoN tykOm tygrOpApAv at au: fu yc bav Otgu rUt bum nUm atiO Ib nykam U-gUw. yUg fu tOv at gUv, hUd bnu Ev sE: hUd bnu wav tAv avAm rUt Oki akiA yb tyv iEv-do rUt ves Ib od rUt bnufU?"
1. John had scoffed at the Spaceman: "I am not interested in your world-view and deep philosophy, space-man. But I want to know how you do it. How can you go on for thousands of years without refueling and without eating?"

– au vUtsepAv: bu OtgUv rUt fnum da Ib Ed, yUg bnum da Ub o yc fnum o. jUf, Qg fu wav-yEc UIv fE at bu, fu yc tOv-yEc sE.'
The spaceman decided: "you are interested in the ways and means but your way of life is not our life. Even if I could tell you I would not want to."

– ynDJoN hIpAv vufU: fnum uamA vEv hUm yjU, hI? "bu UIrv at fu: hUd u wav kad-ov ad Oki akiA Ib krOv tykwe ek a Ib A?"
Johnny wondered: What difference does our attitude make? "Tell me how can you survive thousands of years, and triumph over space and time?"

Au: 'jUf, Qg fu UIv-yEc fE, bu yc iUv-vEc fu. Bu ova g yga-da Ub o Ub yrkO. Nykam tyk ag gaz, fnu Utev ryko: fnum ypus cEv rykO.'
The Spaceman: "Even if I told you, you would not understand me: You live the surface way-of-life of pride. Deep down in the center we find humility: Our secret is humility."

2. —ynDJoN krOIpAv: "bu c'krOIyv-wam eb bum rOkU-Uis."
2. Johnny laughed: "You are ridiculous with your morality."

—au cpA yktrUm: 'yUt fu yc wav typev at fnum ki, yUt fu sepAv fum o, Ut ytyrAv bnum eki, fu tEvAm rUt fum banu, tyg-ytwepAm-s, sE yc pwUrm, hE bu gaf ym-u Uv UI fu. Fu UItAv fE at bu: bu ymA AiOpAv tok, hI?'
The Spaceman was serious: "Since I cannot return to our stars, since I gave my life to save your planet, becoming an outcast to my tribe, it does not matter what you or anybody thinks of me. I will tell you: Have you ever watched a tree?"

3. —ynDJoN tEpAv y-trAwm: "fu gUv nEn tok."
3. Johnny grew impatient: "I know many trees."

—'yt hE tok tnev?'
"Out of what does a tree grow?"

—"yt to."
"Out of a seed."

—'to cEv Uj nam Uj tok?'
"Is the seed as big as the tree?"

—"yr! to c yn-ynam."
"No. The seed is tiny."

—'fA, xA to tnev tag tok, to yc tyv namU, hI?'
"Now, when the seed grows into the tree, does it not gain greatness?"
(ibid.:165–8)

The first comment one might make is that there is nothing very strange here
from a pragmatic point of view, in spite of the unusual context and the odd
appearance of the language (and aUI does seem to be unlike natural languages and
many ALs in one way: it is meant to be (largely) iconic, with there being a relation
between the sound of many (components of) words and what they mean).

Johnny appears to be rude (and uses rather direct language), but that could be
attributed to him being a child and/or being an arrogant earthman. However, to
really know the level of impoliteness, one would have to have a sense of the exact
meaning, including connotations, of *krOIyv-wam* 'ridiculous', which is impossible
without analysis of many other occurrences of it (if there even are such). It is
composed of *krOIv* 'to laugh' and *wam* 'able', and thus means 'able to be laughed
at', but it is not clear whether it has the same sort of negative connotations as
English 'laughable'; one does not always laugh at something/someone because of
its/perceived negative qualities.

Next we will see a pseudo-conversation in another a priori language, Sotos
Ochando's Lengua Universal (Gisbert and Lorrio 1862: 121; I have translated their
Spanish translations into English):

Ar saban labli riolarbem?
'What family do you have?'

Riolarbel le sacan lasfie, siba lalcae bal sibi leldes.
'I have my wife, one daughter and three servants.'

Be saban afaca riaburben sacen lague?
'Where is your mother?'

Sacan lague riaburben soreboc fle sacan lamee: mu sodibi agoldirbin glo sacan
imari bal riaburbin fle sacas.

'My mother is now with my sister: in summer she will come home and will be with us.'

Riabirben ibelin sacen lague?
'Is your mother very old?'

Nan riabirben ibelin: riolarben sicra sugas bal riaburben gan ipafon.
'She is not very old: she is 60 years old and very well.'

Bal sacen lalcae ar le saban sugas riolarben?
'And your daughter, how old is she?'

Riolarben le sibra sugas.
'She is six years old'

Bal sace nan obamerbem se lasur?
'And you, aren't you thinking of getting married?'[10]

Riabirbel gan ibefon mal se lasurdel.
'I am very young to get married.'

Riolarbem sicebu sugas: se lasurfom sorogoc: sace riabirbim ol ugefon lasfi.
'You are 25 years old: get married soon: you will be a good husband.'

Again there is nothing very strange here, although some parts of the conversation involve more directness than one would see in some natural languages; in English it might not be considered to ask whether someone's mother was "very old" or to tell someone to get married, unless one knew him well (which would appear not to be the case in this pseudo-conversation).

Let us now look at part of a pseudo-conversation in Hom-idyomo, an a posteriori language (Cárdenas 1923: 66–7; note that it may not always be clear which interlocutor is speaking):

Iⱬivi nos a dineⱬiⱬ, gesseñⱬos, bite.
'Let us go to dine, gentlemen, if you please.'

[…]

Señⱬo Franklin, degnivi sedentiⱬ be en cia sejo.
'Mr. Franklin, please take this place.'

Danke.

---

[10] Although it is not clear in the source, from the context one might think that this sentence is said by the same interlocutor as the previous sentence.

'Thank you.'

Kay bi, señóɹino Danton, bite sendenti be tie.
'And you, Mrs. Danton, please sit here.'

Peɹmitivi me, demandiɹ be la menuo.
'Allow me to ask you for the bill of fare.'

Veɹi volunte, madamo.
'With pleasure, madam.'

Favoɹvimi peɹ la pano.
'Oblige me with the bread.'

Mi rogay be pasiɹ la súpeɹo.
'Pray, pass me the tureen.'

Voluntay bi miksiɹ raspata-kaseo kun la supo?
'Will you have grated cheese with your soup?'

Danke, afteɹ bi.
'Thanks, after you.'

Pɹefeɹay peɹhapse las gesseñɹos ke oni apɹiɹin las fenestɹos? La ambyento
komencay a deveniɹ veɹi kaloɹa.
'Perhaps you gentlemen would prefer to have the windows open? The weather
is beginning to be very warm.'

Kiel bi voluntin, madamo. Nos no sensacyonay multa kaloɹo, sed bia gaɹdeno
estay tante bela ke miɹiɹ je estay, en eveɹya ciɹkunstanco, una plesuɹo.
'As you like, madam. We are not very warm, but your garden is so beautiful
that it is always a pleasure to look at it.'

[…]

Gustay bi la kaɹno sanga?
'Do you like your meat underdone?'

Mi pɹefeɹay je koktata.
'I prefer it well done.'

Ambi klasos havay estite koktatas, la una plus dan la otɹa.
'Both kinds have been cooked, the one more than the other.'

Peɹhapse bi volutay diɹiɹ ke ambi estey kaloɹatas peɹ fayɹo, una til kwande ji koktey kay la otɹa nuɹ til kwande ji komencey a emaniɹ una likido sángiša, remanante kɹuda inteɹne.

'Perhaps you mean that they were both heated over the fire, the one until it was cooked, and the other until a bloodlike liquid oozes out, the inside remaining raw.'

Peɹmitay bi me, señoɹo Pasteuɹ, demandiɹ bia opino pɹi cia temo?
'May I ask your opinion about the subject, Mr. Pasteur?'

Veɹi plesuɹe, madamo. Si mi no eroɹay, la kaɹno sanga kontenay ankoɹa mikɹobyos danjeɹas poɹ la saluɹo, kay ji no estay plus alimenta dan cia wel koktata. No pensay bi mesme, señoɹo Franklin?
'With great pleasure, madam. If I am not mistaken, underdone meat still contains microbes dangerous to health, and is not more nourishing than that which is thoroughly cooked. Do you not think so, Mr. Franklin?'

Mi no estay kompetenta kiel bi pɹi cia topiko, sed mi imaginay ke mandukiɹ kaɹno insufice koktata estay una kapɹiyo de la modo. Kio mi posiblay an infoɹmiɹ be estay ke mi no gustay je.
'I am not competent in this matter, as you are, but I imagine that eating insufficiently cooked meat is a whim of fashion. All I can tell you is that I do not like it.'

This also does not appear to be a particularly unusual conversation although it goes from what is, by contemporary standards, a high level of politeness to a distasteful remark, which one might not expect to hear at the dinner table in Cárdenas' time or ours.

The last text that I shall present is in Esata (Bothi 2006: Dlog fav ('Dialogue five'))[11]:

A: Va kand cu yufila cuha?
     'What kind of food do you feel like eating?'

B: Yola cinesi, nayu?
     'I like Chinese, and you?'

A: Mitu, benotu hofi, na jelasi satd hada so
     'Me too, but not too often, and just last Saturday I had some.'

B: Derz de mesiki ples raqi, wona trayahe?
     'There's the Mexican place right here, want to try it?'

---

[11] I have added "A" and "B" before the conversational turns; this pseudo-conversation seems to have two interlocutors.

A: Wuno mayna be yo vewe yita tumuc mesiki, den noslipa gu
      'Wouldn't mind but I always eat too much Mexican, then don't sleep well.'

B: Wel wikugo tode sifud resteran. Fila yita sifud?
      'Well we could go to the seafood restaurant. Feel like eating seafood?'

A: Yokugo fone bekda fix plet, dazgu nano kosa tumuc. Hobotyu?
      'I could go for a baked fish plate, that's good and doesn't cost too much. How about you?'

B: Yola de mixi sifud platr wit sofa hevte, tune gudil
      'I like the mixed seafood platter with some of everything, also a good deal.'

A: Hok, hez sifud den, legohina. lediz fersi
      'Ok, it's seafood then, let's go in. Ladies first.'

B: Hune lediy? Yobinewu, nane rilwan tu!
      'Who's a lady? I'm a woman, and a real one too!'

A: Yubigimi negran haptit. Wona sita qi?
      'You're giving me a grand appetite. Want to sit here?'

B: Das fayni, jenosita woyukesi detivu
      'That's fine, just don't sit where you can see the TV.'

Once again there is nothing exotic here. There is some indirectness. For example, in answer to B's question of whether he likes Chinese food, which really is asking whether he wants to have Chinese food on this occasion, A does not say "no"; rather he says that he does not like it (although "not too often") but had it recently. When asked about having Mexican food, A again does not say that he does not want to, in fact he says that he "wouldn't mind" but indicates that he does not favor this possibility. (In fact, it appears that he does like Mexican food, otherwise he would not "always eat too much" of it.)

8 Conclusion

We have looked at ALs with respect to several areas of pragmatics. Although some ALs, particularly the a priori ones, on the surface seem to be quite strange, in general they are not very odd in terms of pragmatics. A small number of them have features which are rare or non-existent in natural languages, e.g. the overt markers of illocutionary force of Seaspeak and the markers of metaphors of Sotos Ochando's Lengua Universal.

The relatively small number of remarks and instructions concerning pragmatics in works on ALs indicates that designers of ALs and authors of books about them have generally not thought much about this area of their languages. This should not be surprising; textbooks on natural languages often do not mention pragmatic matters, with the exception of brief remarks about some aspects of polite language, e.g. formal and informal pronouns. Thus it would be very unusual for a French textbook to discuss conversational implicatures. One could argue that this is because conversational implicatures are a universal features of languages and so do not need to be discussed or explained for those learning e.g. French.

On the other hand, it seems that conventional implicatures may differ to some extent among natural languages, at least with certain items; while in all languages words meaning 'and' probably have the conventional implicature that the clauses connected with them have some sort of relation (unlike the & of propositional logic), and it would be hard to imagine that it could be otherwise (what would be the point of connecting two clauses if there were no relation between them?), there may be differences involving words for 'but', namely in the strength of the contrast that they signal. For example, the Russian word *a* signals a weaker contrast than English *but*. In spite of such differences among natural languages, works on ALs rarely give information about conventional implicatures of conjunctions.

One could argue that from a pragmatic point of view ALs are interesting because they are not interesting; that is, as exotic as they may seem on the surface, and as exotic as they may be in certain respects (in particular, the way in which the lexicon has been constructed in some a priori ALs), their pragmatics are often not significantly different from those of natural languages. This is perhaps because ALs simply have carried over the pragmatics of one or more natural languages, even if they are supposedly a priori; perhaps it is quite difficult to create a truly *a prior* pragmatics, at least if one is trying to design a usable language.

This brings us to the reason why the lack of strangeness of AL pragmatics might be interesting. If there are pragmatic universals, that is, if the pragmatics of language is wired into the human brain (as has been claimed for various syntactic principles), then this might limit not only competence but also language creation (unless one were trying to be perverse). I have made this point before (Libert 2000: 1) with regard to universals in general, but the argument might be even stronger with regard to pragmatics

If a language were designed with a pragmatic feature that violated some universal, i.e. if it had an "unnatural" feature, one might expect that, if it were used for long enough by enough people and were not limited in the domains that it occurred in, i.e. if it fulfilled all the roles that that natural languages do (e.g. conversation, literature), this feature might disappear. For example, one might think that markers which clearly and overtly indicated illocutionary force, and which therefore prevented the existence of indirect illocutionary force, could not exist in a natural language. Of course many languages have ways of marking e.g. questions, but perhaps in all such languages questions can have an indirect illocutionary force as something other than a question (such as a command). It would not be surprising if eventually users of a language with Seaspeak-type markers

started using sentences with question markers with some indirect illocutionary force. This will probably not happen with Seaspeak due to the very limited contexts in which it is supposed to be used, but one could imagine the process taking place with more widely used languages. Thus, an unnatural pragmatic feature might only be able to survive in a language which was designed for particular and narrow functions.

In any case one might ask what the point of creating a new pragmatics for an AL would be. ALs have often been created in an effort to improve and/or simplify natural languages, involving e.g. the elimination of irregular verbs, and such features may have been criticized by AL designers. If these designers had felt that some pragmatic features of a natural language (or natural language in genera) were better removed, then they could have done it, or at least attempted to do this. Indeed this has been done, as we have seen, with respect to metaphor and the familiar-polite pronoun distinction (recall Zamenhof's attitude toward *ci*). However, to my knowledge, no AL designer has argued for changes concerning conversational implicatures, presuppositions, or conventional implicatures, with respect to particular items in a language or in general.

References

Anderson, A. W. (no date). *A planned auxiliary language*. No place of publication given, apparently self-published.

ApGawain, N., Hugon, P. D., Moore, J. L., & de Beaufront, L. (2008). *Ido for all* (edition 1.6). http://idomondo.org/skolo.1.6.pdf. Accessed 27 Apr 2010.

Bollack, L. (1900). *Abridged grammar of the Blue Language (English version by Professor Tischer)*. Paris: Editions of the Blue Language.

Bothi, P. (2006) *The new personal international language Esata* (draft edition 0.3). http://www.beginnersgame.com/page_25.pdf. Accessed 9 May 2010 (In this work the author's name is given simply as PAFU, but an internet search indicated that his name is apparently Pafu Bothi).

Brown, P., & Levinson, S. C. (1987). *Politeness*. Cambridge: Cambridge University Press.

Cárdenas, C. (1923). *Hom-idyomo* (2nd edn). Leipzig: apparently self-published.

Cohen, L. J. (1993). The semantics of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 58–70). Cambridge: Cambridge University Press.

Couturat, L., & Leau, L. (1903). *Histoire de la langue universelle*. Paris: Librarie Hachette et Cie.

d'Antoñana, G. M. (1973). *Neoispano*. Logroño: Editorial Ochoa.

Dahlenberg, T. D. (2006). *Pli lume la mallumo zumas…*. New York: Mondial.

de Beaufront, L. (1925/2005). *Kompleta gramatiko detaloza di la linguo internaciona Ido*. http://kanaria1973.ido.li/krayono/kgd.pdf. Accessed 24 Sept 2009.

Dellert, J. (2008). *Lambda calculus on dependency structures for a wide-coverage grammar of Esperanto*. B.A. Thesis, University of Tübingen.

Elhassi, Z. (2008). *Ardano course*. http://ardano.tripod.com/sitebuildercontent/sitebuilderfiles/pronunciation.pdf. Accessed 23 Oct 2009.

Galdia, M. (2009). *Legal linguistics*. Frankfurt: Peter Lang.

Gaskell, R. (1999). *Syntax-based grammar of Glosa*. http://www.glosa.org/en/gramm.htm. Accessed 25 Apr 2011.

Gisbert, L., & Lorrio, P. (1862). *Manual de Lengua Universal*. Madrid: Imprenta de J. Martin Alegría.

Gledhill, C. (2000). *The grammar of Esperanto: a corpus-based description* (2nd ed.). Munich: Lincom Europa.

Gobbo, F. (2008). Planned languages and language planning: the contribution of interlinguistics to cross-cultural communication. In F. Boers, J. Darquennes, K. Kerremans, & R. Temmerman (Eds.), *Multilingualism and applied comparative linguistics* (Vol. II, pp. 24–42). Newcastle upon Tyne: Cambridge Scholars.

Gode, A. (1954). *Interlingua a prima vista*. New York: Storm Publishers.

Gómez, P., & Juan, R. (2008). *Usik lan hil*: *Usik language course*. http://www.usik.es/engus.pdf. Accessed 29 Sept 2010.

Grzega, J. (2005). Towards global English via Basic Global English (BGE). *Journal for EuroLinguistiX, 2*, 65–164.

Grzega, J. (2008). Lingua Franca English as a way to intercultural and transcultural competence. *Journal for EuroLinguistiX, 5*, 134–161.

Huang, Y. (2002). Typology of coreferential anaphora and neo-Gricean pragmatics: implications for a newly designed artificial language. *Journal of Universal Language, 3*, 31–56.

Jones, L. (1972). *Eurolengo*. No place of publication given: Oriel Press.

Kellerman, I. (1910). *A complete grammar of Esperanto*. New York: D. C. Heath and Company.

Kimbrough, S. O., & Yang, Y. (2005). On representing special languages with FLBC: message markers and reference fixing in SeaSpeak. In S. O. Kimbrough & D. J. Wu (Eds.), *Formal modeling in electronic commerce* (pp. 297–324). Berlin: Springer.

Libert, A. (2000). *A priori artificial languages*. Munich: Lincom Europa.

Linderfelt, K. A. (1888). *Volapük*. Milwaukee: C. N. Casper and H. H. Zahn.

Lyons, J. (2006). *Natural language and universal grammar* (Vol. 1). Cambridge: Cambridge University Press.

Martellotta, V. (1919). *Latinulus*. Bari: F. Casini & Figlio.

Morneau, R. (2006). *The lexical semantics of a machine translation interlingua*. Formerly available on the World Wide Web. Accessed 30 Apr 2010.

Ochando, B. S. (1863). *Gramática de la Lengua Universal*. Madrid: Imprenta de J. Martin Alegría.

Parke, D. (2008). *De Koncis Frenkischgrammatk/The concise Frenkisch grammar*. Formerly available on the World Wide Web. Accessed 12 July 2009. (2nd edn currently available at http://f1.grp.yahoofs.com/v1/cJkPUN4IGLcizRjL2MqbOI213r9n0UZNzwMG1j9omi4b2l-vPlEAeHa2J2BTOHWlZz3obZc4ybF3raXN3enUJzRMX0s/De%20frenkichgrammatik%20fon%20David%20Parke%20v2.01.pdf).

Quiles, C. (2009). *A grammar of Modern Indo-European* (2nd edn). http://dnghu.org/indo-european-grammar.pdf. Accessed 2 Dec 2009.

Ruggles, J. (1829). *A universal language*. Cincinnati, OH: printed by M'Calla and Davis.

Russell, B. (1966). *Suma* (3rd edn). Plainview, NY: apparently self-published.

Searight, K. (1935). *Sona*. London: Kegan Paul, Trench, Trubner & Co.

Simon, O. (2010). *The grammar of Sambahsa-mundialect in English* (2nd edn). Formerly available on the World Wide Web. Accessed 22 Apr 2011 (3rd edn currently available at http://www.scribd.com/doc/34500642/Sambahsa-Grammar).

Sprague, C. E. (1888). *Hand-book of Volapük*. New York: The Office Company.

Stadelmann, J. (1945). *Voldu Textbook*. Caracas: Lit. y Tip. Vargas.

Sulky, L. (2005). *Konya language—dialogues*. http://larrysulky.webs.com/konya/konya-dialogue.html. Accessed 18 July 2010.

Talmey, M. (1925). *Arulo*. New York: Ilo Press.

Traunmüller, H. (1991/1996). *Conversational maxims and principles of language planning*. http://www2.ling.su.se/staff/hartmut/griceil.htm. Accessed 8 May 2011.

van Cranenbugh, A., Sassoon, G. W., & Fernández, R. (2010). Invented antonyms: Esperanto as a semantic lab. In Y. N. Falk (Ed.) *Proceedings of IATL 26*. Ramat Gan, Israel: Bar Ilan University.

http://linguistics.huji.ac.il/IATL/26/vanCranenburghSassoonFernandez.pdf. Accessed 31 May 2011.

Vidal, E. T. T. (1844). *Langue universelle et analytique*. Paris: A. Sirou.

Weeks, F., Glover, A., Johnson, E., & Strevens, P. (1988). *Seaspeak training manual*. Oxford: Pergammon Press.

Weilgart, J. W. (1979). *aUI* (4th edn) Decorah, IA: Cosmic Communication Co.

Wennergren, B. (2005). *Plena manlibro de Esperanta gramatiko* (version 14.0). http://bertilow.com/pmeg/elshutebla/pmeg_14.0.pdf. Accessed 28 Oct 2010.

Implicit Propositions in an Argumentative Approach

Sorin Stati

Abstract After presenting the contexts in which researchers speak about implicit elements—a list that highlights the conceptual diversity of meanings attributed to the term 'implicit'—Stati will focus on the actual topic of this article: the property 'implicit' as it functions in argumentative texts. Or, to put it another way, how do implicit propositions manifest themselves on the argumentative discourse level. Stati dwells on interesting inferential phenomena involving the argumentative roles of portions of text. He differentiates between a casual overhearer and the intended addressee, speculating on the differences in interpretative behavior. The inferential behavior triggered by argumentative relations within a text very often involves the recovery of implicit materials.

1 A Controversial Concept

When dealing with implicit linguistic elements several preliminary questions arise:

a. the definition, first of all: what does 'implicit' mean? This problem begins to have sense only after we take into account other labels from the same semantic sphere, which seem to have similar definitions: presupposition, ellipsis, Fr. sous-entendu etc. The wealth of terminology is underlined by Kerbrat Orecchioni (1986): présupposé, implication, inférence, arrière-pensées, allusion, insinuation, valeur illocutionnaire dérivée, ce qui se dit à demi-mots, entre les lignes;

b. which are the linguistic phenomena evoked by the term 'implicit'? Do they belong to the 'linguistique de la langue' or to the 'linguistique de la parole, i.e. of the discourse?' And what is their status with respect to the 'surface structure' and to the 'deep structure'?

Sarin Stati—Deceased

S. Stati (✉)
University of Bologna, Bologna, Italy

c. how do people recognize/deduce the existence of the missing elements and on which basis or with which justification do they reconstruct them?

We shall provide a tentative answer to these fundamental issues, starting from a methodological principle: all of the above-mentioned issues are the result of applying a grid of concepts elaborated by the linguists stimulated by the perception of a 'substance' or 'matter' that exists independently of the linguists and to which the linguists give some 'form'. Consequently, we assume that issues like the ones mentioned above are 'invented by the linguists' rather than 'discovered by them in the languages'.

In the bibliography of the problem other questionable issues are highlighted, for instance, what are the properties that cause the speakers to use utterances with implicit elements? (Kerbrat Orecchioni 1986). One explanation is the following: people sometimes feel the need to say something while appearing not to have said it and consequently avoiding the responsibility for having said it. A second origin could be that "every explicit assumption becomes, by the very fact that it is explicit, the subject of a discussion; everything that is said may be contradicted". (Ducrot 1972, 13–14).

After presenting the contexts in which researchers speak about implicit elements—a list that highlights the conceptual diversity of meanings attributed to the term 'implicit'—we shall focus on the actual topic of this article: the property 'implicit' as it functions in argumentative texts. Or, to put it another way, how do implicit propositions manifest themselves on the argumentative discourse level.

From the bibliography pertaining to traditional linguistics, we may exploit the comments—though not the theory—dealing with ellipsis, one of the least elaborated concepts in traditional grammar. From the recent bibliography of implicitness the most relevant is that which deals with presuppositions. From the conspicuous material accumulated in Levinson's classic work in 1983, we may select a considerable part devoted to presuppositions with argumentative value. Scholars interested in presuppositions did not consider the argumentative function of the implicit propositions as a special topic.

2 Implicitness: the Context of its use

A proposition is called 'an implicit rejection' when it is equivalent to the (explicit) rejection of a proposal/invitation:

"A: May I have the pleasure of inviting you for this dance?
 B: I'm sorry John, I'm so tired!"

We think that in such a case it is legitimate to assume that we are dealing with an 'argumentative figure' consisting of the justification (explicit: "I'm so tired") of a rejection (implicit: "I cannot accept your invitation") which remains unexpressed.

An analogy could be established between this figure and the pragmatic concept of 'indirect speech act'.

The origin of this figure is B's intention to save the 'face' of the sender of the proposal, hence verbal politeness. Finally, we note that ''I'm sorry, John'' plays a separate argumentative role (apology).

Rather similar is the case of the missing justification in the rebuttal of an assertion, as in:

"A: John has been arrested.
  B: That's impossible! John is a Diplomat."

Here the implicit proposition is "Diplomats cannot be arrested". Some doubts may arise: is the interpretation proposed due to the propositional content of the sentences involved (the assumption which comes after the strong rebuttal "That's impossible" is most probably its justification: ("That's impossible because…, etc.") or is it due to the receiver's encyclopedia? In fact, the speakers know that Diplomats cannot be arrested during their period of activity. On the contrary, there is no ambiguity, but a unique interpretation (suggested by the sequence of propositional contents) in a case as in the following conversation excerpt:

"A: I suggest you buy a Daewoo.
  B: In our country Korean cars are not much appreciated."

In addition, we have to presume that in the considered moment of the above conversation the topic was the desire expressed by speaker B to purchase a car, so that Daewoo could be nothing other than a car manufacturer's name .

A particular case is that of the adverb "otherwise" meaning "if you do not do so":

"I'll give you two days to do the translation. Otherwise you will be not getting any money from me."

The implicit proposition is "if you do not finish the translation in two days", which is typical for the argumentative scheme "threaten". In fact, we often find the sequence of three argumentative roles/thesis - negative condition—threat of sanction/. But how should we interpret "otherwise"? As an incomplete sentence? as a quasi-proposition? as a proposition equivalent?

A proposition is called 'implicit rejection' when it is equivalent to a rejection of an impositive speech act: directive or interrogative.

"A: Open the window, please.
  B: But it is open, darling!"

The implicit proposition is "You should not be making this request", hence a criticism of the utterance in the partner's sentence; this case enters the category, already illustrated above, of an argumentative figure. The explicit proposition

which follows is a justification act: "You need not be making this request because the window is already open". Cf. also Ducrot 1972, 16 ff, who speaks about implicitness of the enunciation ("implicite de l'énonciation").

Consider the following:

"A: How old is she?
 B: That's an indiscreet question!"

The implicit proposition is something like "You should not be asking such a question", hence we have a criticism of the utterance, as in the preceding example.

In many texts, especially oral ones, the receiver has to reconstruct a missing proposition between the elements of a couple of argumentative moves. This happens when the elements of the interchange contain no clarifying lexical element (a connective, etc.), nor do they have a denotative content which could guide the right decoding. Consider:

"A: Let's go round to John's and pay him a visit
 B: He will think that you have forgotten your political differences."

It is quite likely that participant A has correctly interpreted B's utterance, but how did a casual witness interpret it? Is it a consequence? Perhaps, but the doubts arise (for this witness) about the acceptance of speaker A's exhortation: is it a favorable argument (and in this case the implicit proposition is "the proposal is wise") or a unfavorable argument (and the reconstructed proposition is "The proposal is unwise"?) This operation, a sort of calculus, is suggested by the propositional content of the members of the couple.

The tentative, more or less likely interpretation of the argumentative role of an utterance in the conditions of a total absence of signals (connectives, meta-argumentative labels, etc.) may be illustrated by the following interchange:

"A: Let's go round to John's and pay him a visit
 B: You know, Mary has decided to spend all the winter months in London."

Is the second utterance an objection? In this case it informs the partner that there is an impediment to the achievement of the action "seeing John". Or, on the contrary, is it a reason which determines A and B to accept A's proposal?

The operation we are dealing with is risky, since the addressee does not always guess the right interpretation so that his partner is obliged to correct him.

Whenever the elements of the couple of utterances do not suggest an interpretation, the solution may derive from the addressee's encyclopedia.

An interesting case is that of utterances requiring the mental reconstruction of an implicit proposition whose meaning is "You are right", "I agree with you" or the opposite "You are wrong", "I do not agree with you". Without this mental reconstruction, the understanding of the dialogue excerpt is not possible. If the receiver

gives the sender his/her 'cooperation credit', then (s)he interprets the sender's utterance as an agreement or a rebuttal, that is (s)he reconstructs the missing proposition "You are right" or "You are wrong", cf.

"A: I am sure John will take a taxi.
B: He lives in Oxford Street."

Depending on the particular situation of the interlocutors (does the dialogue take place close to (or far from John's house?) speaker B's reaction may be interpreted by a casual witness to the verbal interaction either as an argument in favor of his/her partner's hypothesis (the implicit proposition would be "I think so, too") or as a negative argument, so that the implicit proposition would be "You are mistaken." (Intonation could also play a role in disambiguating).

The implicit proposition is the major premise of a syllogism restricted to two members (an enthymeme: Tardini 1997), cf. "John has come to see us, hence he has an interest". The enthymeme was proffered with the aim of communicating the opinion that John comes only when he has an interest (Ducrot 1972, 15). The enthymemes are classified into three categories ("degrees") according to the missing proposition; for instance, if the minor premise had been omitted, the enthymeme belongs to the second degree. The enthymemes are not the unique discourse fragments in which the provision of an implicit proposition is based on inference and the reduced formulation more economical. A second, similar case, which seems to be of little interest for linguists, concerns the logical form of an argument, rich in explicit constituent propositions, as compared with its expression in ordinary language. In order to highlight the logical relations, the formulae preferred by the logicians contain propositions that have no correspondent in the "normal" (stylistically neuter) form or variant.

The text needs the reconstruction of more than one implicit proposition, but they all belong to one of the above-mentioned species. Here is an example:

"A: By tomorrow you have to finish your job!
B: I will not take orders from you."

Speaker B's reaction implies (a) a proposition such as "I interpret your utterance as an order", and (b) "It is possible that I will not finish my job by tomorrow" that is followed by its justification "I will not take orders from you". B's reaction could have been formulated differently, for instance "You know that I will not take orders from you"; this utterance is built on a different argumentative schema: it contains a criticism ("You should not give me an order") justified by the fact that B already knew that he should not give offers to speaker A. In our opinion the phrase "You know" designates an aggravating circumstance.

3 Delimitation of the Subject

In this paper the label 'implicitness' will be used in a rather restricted way. The limitations are the following:

a. We shall be dealing exclusively with propositions, although very often single words and phrases are missing and therefore we have to reconstruct them in the decodification process;
b. The propositions considered are exclusively pieces of the argumentation mechanism; therefore, in the following conversation excerpt "Give me some handkerchiefs", we do not recognize an argumentative dimension and hence such utterances are of no interest to the present paper;
c. The propositions are expressed in ordinary language (English);
d. A limited number of examples will be given in a logical meta-language, using common letters and symbols:

"If p then q, hence if not q then not p" (Marciszewski 1994).

c. We shall not deal with propositions logically deduced or inferred from an explicit proposition; for instance, we shall ignore the fact that the conc1usion of an argument is inc1uded in the premises and it may be deduced from them, for instance, from "Richard is a one-eyed killer" we deduce the proposition "Richard is a killer", whereas from "Richard is the presumed killer" we cannot deduce that Richard is a killer; for a similar reason, we shall not include in our research the lexical presuppositions like "I knew Latin" which has to be inferred from "I have forgotten my knowledge of Latin". On the contrary, we shall consider the presuppositions of the text and the presuppositions of the enunciation of the text, only if they have a bearing on the argumentative level.

d. More generally speaking, we neglect propositions that could be deduced from a context limited to a single proposition, by applying the methodological principle that the presence of a minimum of two propositions is necessary in order to be able to speak of argumentation.

4 Concrete and Abstract Structures

The first proposal is to consider implicit propositional examples of 'langue implicitness' whenever their analysis does not presuppose a knowledge of the situation, of the interlocutors and their encyclopedia etc., that is of all the pragmatic details. In our opinion, the subject matter of pragmatics is the meaning of utterances (concrete pieces of discourse), i.e. more or less the meaning of the manifestation of a text as it is proffered in a definite, concrete situation and which is conceived by its author for the benefit of a definite receiver. This subject could be also termed the 'linguistique de la parole'.

Between the propositions as abstract schemes and utterances as actualized sentences there is an intermediate structure of relations which gives the receiver the chance to accomplish a leap forward from an abstract, system phenomenon (the proposition) to the concrete 'parole' structure. The decodification is realized thanks to this intermediate structure. The argumentative level including the analysis in 'roles' and the analysis of the relations between the 'roles' (thesis, objection, evidence, etc.) has two "sides": an abstract scheme, which is essentially the logical structure, and a concrete manifestation; in general, the addressee has access to the second side, whereas a casual witness has access to the first side.

Naturally, both types of receivers are exposed to errors of interpretation. The implicit propositions designated by the labels corresponding to argumentative roles belong to the abstract structure that we have chosen to denominate 'deep structure' although the correspondence with the generative-transformational theory remains rather vague. We insist on the fact that the formulation of these propositions offers several manifestations, since it may contain letters, symbols, arrows, brackets, etc.,.. The deep structure formulae are not unique, but of different degrees of abstractness. In this paper a formulation in an ordinary language (English) was chosen.

The reader has presumably noticed our terminological choices: proposition, sentence, utterance. We prefer 'proposition' because it is the same word as used by logicians and by many linguists; in addition, it is employed in other languages (Fr. proposition, it. proposizione). But the main reason is the fact that, in our model, it designates the underlying logical structure and, as such, it belongs to the deep structure. The 'proposition' is the verbal support of an argumentative role (thesis, proof, objection, prolepsis, rectification etc.); see Stati 2002, The term 'sentence' was employed as equivalent of 'langue item' and 'utterance' as 'parole item', i.e. 'discourse unit'.

The limitation to this class of implicit propositions does not exclude the ambiguous cases; the ambiguity may concern the argumentative role, hence the nature of the logical relationships between the single utterances. Obviously, not all ambiguities are relevant at the argumentative level of the text.

For example, the anaphorical decodification of the pronoun "he" or "she" when the preceding text has two masculine or feminine referents, respectively. This case does not concern the implicitness of argumentative relations between propositions. Here is an example:

"A: I met Helen at Rome.
 B: By the way, Mary asked me to call her sister immediately."

Does "her" mean "Mary's" or "Helen's"?

As far as the number of implicit propositions reconstructed/presupposed in order to give the text a c1ear, rational structure and interpretation is concerned, we

may assume that a certain risk of exaggeration does indeed exist. There is a danger of reconstructing too much, obsessed as many grammarians are by an 'aporia' that we may summarize like this: "What is the complete form of the sentence Sx?" The risk is to ca1culate as necessary parts—perhaps implicit—too much linguistic material. This discussion recalls the opinions of Sanctius (Sanchez), the so-called "father of the ellipsis", who was active in the 16th century.

What is the function of the implicit propositions?

A very general explanation already exists: they satisfy the speakers' desire to communicate certain opinions/theses without having the responsbility of having expressed them. And the second explanation is the tendency to reduce redundancy. But the phenomenon is more complex.

Consider, for example, the following utterance:

"John can vote at the elections because he is twenty years old".

It communicates the implicit proposition "All citizens aged eighteen and over have the right to vote at political elections". The function of this reconstructed proposition is that of a "reminder" (Fr. rappel), and not that of an assertion, since the sender does not behave as someone who provides new information ("new" for the receiver). But the proposition may also be considered as informative, especially if the sender is speaking with a foreigner who is unfamiliar with the laws of the country where the dialogue is taking place. A third possibility: the sender ignores the political notions of his partner/reader and builds an ambiguous utterance with the advantage that it fits into both of the above-mentioned circumstances.

5 Final Remarks

As our illustration has shown, the majority of texts have only one implicit proposition per utterance.

The formulation of the implicit propositions is not unique; on the contrary, several variants are conceivable, of different length and with different lexical components. The difference may also regard the degree of abstractness. It is obvious that this assumption is based on the hypothesis that two or more sentences can be synonymous. By adding and/or eliminating implicit propositions we obtain a new, synonymous proposition.

The reconstruction of the implicit propositions by the receiver of a message is an essential part of the decoding process. The identification of propositions with an argumentative role is an essential part of the understanding of the argumentative discourse level.

The principal conc1usion of the present paper is that discourse furnishes to the addressee and to a casual witness a quantity of information c1assified as follows:

information produced by the system utilized—and languages use grammar and lexical features independent of the sender's desire and intentions. A considerable

amount of this information concerns the argumentative organization, and it may be communicated by means of implicit propositions;

information communicated thanks to the common encyclopedia of sender and receiver as well as the knowledge (by the interlocutors) of the circumstances of the utterance. This kind of information has no expression by means of a verbal code— it is 'per definitionem' implicit—and shows the limits of the c1assic conception of communication (the information, all the information transmitted by discourse, reaches the receiver in the shape of a sequence of linguistic signs). One part of this implicit information is argumentative and it characterizes argumentative texts and chunks of argumentative monologic and dialogic texts with different typological dominance (narrative, descriptive, etc.).

References

Allwood, J., & Dahl, O. (1961). *Logica e linguistica, Italian translation*. Bologna: Il Mulino.

Copi, I. M., & Cohen, C. (1997). *Introduzione alla logica, Italian translation*. Bologna: Il Mulino.

Ducrot, O. (1972). *Dire et ne pas dire*. Paris: Hermann.

Kerbrat Orecchioni, C. (1986). *L'implicite*. Paris: Colin.

Levinson, S. (1983). *La pragmatica, Italian Translation*. Bologna: Il Mulino.

Marciszewski, W. (1994). *Logic from a rhetorical point of view*. Berlin/New York: De Gruyter.

Stati, S. (2002). *Principi di analisi argomentativa*. Bologna: Pàtron.

Tardini, S. (1997). *L'entimema nella struttura logica del linguaggio*. In: "L'analisi linguistica e letteraria" I, (pp.419–440).

Automatic and Controlled Processes in Pragmatics

Marco Mazzone

Abstract In utterance understanding, both personal and sub-personal aspects appear to be involved. Relevance theory (starting from Sperber and Wilson 1986/1995) and Recanati (2004) have respectively explored two alternative ways to conceive of those aspects and their interaction. Here a third account is proposed, in the light of the automatic-controlled distinction in psychology, and of recent views concerning the cooperation between these two modes of processing. Compared to Recanati (2004), the account proposed here assigns a larger role to automatic, associative processes; at the same time, it rejects the view that consciousness applies only to what Recanati calls secondary pragmatic processes. Consciousness is rather held to cooperate with associative processes in any aspect of pragmatic processing, irrespective of the pragmatic distinction between explicatures and implicatures. On the other hand, a close consideration of how associative and conscious processes plausibly interact makes it appear unnecessary the hypothesis of a specialized process for utterance understanding—such as the automatic, inferential mechanism put forth by Relevance theory.

M. Mazzone (✉)
University of Catania, Catania, Italy
e-mail: mazzonem@unict.it

1 Introduction

Pragmatic processing seems to be at the same time an automatic and a personal-level affair. Humans produce and understand utterances in context quite rapidly and effortlessly, just as it is expected to occur in automatic processing; nonetheless, verbal communication is thought to require an intentional involvement on the part of the speaker, and a recognition of this speaker's intention—possibly requiring rational capacities—on the part of the addressee.

However, it is far from immediately clear how pragmatic theory should account for the coexistence of the two different features considered above; this is probably one of the major challenges that current cognitive pragmatics has to cope with. There are two main explicit attempts to address that issue within pragmatic theories of language understanding: one is Relevance theory, the other is Recanati's (2004) framework.[1] The latter solve the problem by proposing a two-level model, where automatic associative processes and personal-level, inferential processes cooperate in explaining language comprehension in context. Relevance theorists assume instead that, in a sense, a single mechanism can account for both the aspects considered. Sperber and Wilson (1986/1995) in fact conceive of language comprehension as based on a process which is said to be unconscious and automatic, but nonetheless endowed with features that are normally attributed to personal-level processes: it would be meta-representational, and inferential rather than merely associative.

In this paper I intend to propose a different way to account for coexistence of automatic and personal-level features in pragmatic understanding. My proposal preserves Recanati's intuition that two different kinds of processes are involved, while accepting Relevance theorists' criticisms to Recanati with regard to his claim that these processes apply to distinct domains of pragmatic phenomena. The account I propose is based on the well-established distinction between automatic and controlled processes in psychology, and especially on the recent literature which emphasizes the constant cooperation between the two in most of our cognitive processes. Although Recanati's account is consistent with this framework to the extent that he conceives of a cooperation between automatic and conscious processes in language understanding, his claim that these processes apply to different pragmatic phenomena has no ground in linguistic and psychological evidence. In other words, the opposition between automatic and conscious processes does not seem to parallel any traditional distinction within the domain of pragmatic phenomena—such as the distinction between (the processes involved in the

---

[1] For sure, there are other pragmatic frameworks that could be worth discussing in this context. However, the two I have chosen are amongst the most complete and explicit attempts to analyze the overall cognitive architecture of pragmatic processing. Elsewhere I extend my analysis to other theories in the field by addressing the topic of default interpretations (Mazzone, 2013a). In particular, in that paper I address the positions of Bach, Levinson, Jaszczolt, and also Capone's (2011a, b) interesting proposal of a largely associative perspective on defaults and modularity in pragmatics.

determination of) explicit and implicit sense of utterances. On the other hand, a close consideration of the plausible division of labour between automatic and controlled processes suggests also reasons to reject Relevance theorists' proposal that a single process, conceived of as both automatic and inferential, may account for pragmatic understanding.

In practice I will proceed in the following way. First, in Sect. 2, I will survey the distinct ways in which Recanati and relevance theorists propose to combine automatic and personal-level components of utterance understanding. Section 3 will be devoted to the distinction between automatic and controlled processes in psychological and neuroscientific literature, and to a scrutiny of the collaboration between these processes, along the lines of the "distributed intentionality model" put forth by Mazzone and Campisi (2013). In that model purely associative, automatic processes play a large role, although in cooperation with conscious processes. In Sect. 4 I will analyze RT's and Recanati's positions towards associative processes in pragmatics, arguing that these processes have a key role to play in that domain too (as I argue at greater length in Mazzone 2011). Finally, in Sect. 5 I will analyze RT's and Recanati's positions towards conscious processing, and I will consider how controlled processes are to complement associative processes in order to deliver a complete account of language understanding.

2 Recanati and Relevance Theory

In the recent debate between Relevance Theory (from now on, RT) and Recanati (2004) with regard to the architecture of pragmatic processing, a key role is played by the notion of inferential process. As is well known, at the core of Grice's (1989) theoretical framework there is the distinction between two layers of utterance meaning: what is said and what is implicated by an utterance—respectively referred to in the recent literature as the explicit and implicit sense of the utterance. While the former was essentially thought to depend on the linguistic information conveyed by the utterance (except for minor appeals to context in order to obtain reference assignments and disambiguations), Grice conceived of the transition from explicit to implicit sense as a sort of rational inference requiring consideration of the current goals of the speaker. Although Grice himself insisted that the enterprise he was engaged in was a matter of rational reconstruction rather than a genuine psychological thesis about the processes involved in comprehension, in one form or another his appeal to inferential processes performed by a rational agent has framed recent cognitive accounts.

The most straightforward manifestation of this influence is Recanati's (2004) conception of what he calls secondary pragmatic processes. Recanati is between those who think that Grice had underestimated the role played by context in determining explicit meaning: in Recanati's opinion, explicit meaning is the result of pragmatic processes just as implicit meaning is. However, he thinks Grice was

right in pointing at a major difference between those two layers of meaning, or more precisely, between the processes leading to them: while the processes yielding the explicit meaning (in Recanati's terms, primary pragmatic processes) should be thought of as sub-personal, associative processes, implicit meaning would be instead the result of genuine inferential processes taking place at the personal level (secondary pragmatic processes). Let us address this proposal in some more detail.

In Recanati's (2004) account, primary pragmatic processes are conceived of as local associative processes, based on the spreading of activation within conceptual networks and the consequent degree of activation of concepts in the network. In other words, a concept would be contributed to the explicit content of the utterance insofar as that concept is the most accessible (i.e. the most activated) for the system given the situation. In practice, the literal meaning of an expression:

> is accessed first and triggers the activation of associatively related representations. That literal meaning is a natural candidate for the status of semantic value, but there are others: some of the representations activated by association contribute further candidates for the status of semantic value. All candidates, whether literal or derived, are processed in parallel and compete (Recanati 2004, 28).

Although literal meanings are said to have an initial advantage over other possible candidates, this cannot imply of course that literal meanings—or, more generally, concepts endowed with an initial advantage—always win the competition. Recanati (2004) emphasizes the importance of what he calls "accessibility shifts": in the course of processing, contextual information may change the accessibility of any concept activated previously, by adding a new train of activation to the process. According to Recanati, a key role in accessibility shift is played by abstract schemata coded in our long term memory. For a very simple example (see Carston 2007), let us consider the following utterance:

(1) I'm going to the bank now to get some cash.

Since there are two possible meanings for "bank" (FINANCIAL INSTITUTION, RIVER SIDE), one problem is how the subject may come to choose the right one. Let us suppose that, for whatever reason, at the moment when the lexical form "bank" is processed the most accessible meaning is the wrong one (RIVER SIDE). However, we can expect an accessibility shift as soon as the word "cash" is processed, since this word activates its meaning, which in turn triggers a number of concepts having to do with money, and this presumably provides further activation to the concept of bank as financial institution. In particular, the activation of CASH could recall an abstract schema—in Carston's (2007) terms, a stereotypical frame or script—for GETTING MONEY FROM A $BANK_1$ (where $BANK_1$ = FINANCIAL INSTITUTION), thus strengthening the activation of $BANK_1$. In this way, schemata drive the interpretation process by promoting the search for coherence, due to an entirely associative mechanism: on the one hand, "a schema is activated by, or accessed through, an expression whose semantic value corresponds to an aspect of the schema"; on the other hand, the "schema thus activated in turn raises

the accessibility of whatever possible semantic values for other constituents of the sentence happen to fit the schema" (Recanati 2004, 37).

Once explicit content has been recovered by means of associative processes, Recanati proposes that a quite different process leads to the determination of implicit content. In Carston's (2007, 2)[2] words, "secondary pragmatic processes are to be understood as part of a more general theory of human action and interpretation and so having the philosophically central property of being rational, personal-level (as opposed to subpersonal) processes". In practice, secondary pragmatic processes are said to be "transparently or consciously inferential" in that they satisfy the "availability condition" (Recanati 2004, 44): they are accessible to consciousness, that is, the subject is aware of what is said, of the implicature, and of the inferential process leading from the former to the latter.

For an example, let us consider the following question–answer pair:

(2) (A) Could you pay back the money you owe me?
   (B) I'm going to the bank now to get some cash.

In this context, (B)—which repeats (1)–can be interpreted as a positive answer to the yes–no question (A). The explicit content of (B), in which the concept $BANK_1$ is fed thanks to associative processes, licenses a further contextual inference to this layer of implicit meaning. Recanati's assumption is that both the explicit and the implicit content are available to consciousness, in that the former conforms "to the intuitions shared by those who fully understand the utterance" (idem, 14), and the latter follows inferentially from the former insofar as the expectation for a yes–no answer raised by the previous question is taken into consideration.

It is important to emphasize that the conscious availability here appealed to is qualified by Recanati as tacit rather than fully explicit—or, to put it differently, dispositional rather than occurrent. He recognizes that conscious processes are typically effortful, slow and under voluntary control, while comprehension processes normally are not. However, although the inferences involved in comprehension cannot be conceived of as conscious in this explicit, occurrent sense, according to Recanati they are not even the sort of sub-personal inferences that are merely ascribed to a cognitive system on the grounds that the system behaves in the same way as someone who performed the relevant inferences in an explicit form (Recanati 2004, 49). An inference can also be consciously available in a tacit, dispositional sense when "the cognitive agent to which it is ascribed […] is *itself* capable of making the inference explicitly and of rationally justifying whatever methods it spontaneously uses in arriving at the 'conclusion'" (idem, 50). Therefore, in Recanati's opinion, some inferences are merely sub-personal, some are conscious in the prototypical sense—they are explicit, slow and effortful— while some others are personal-level and yet only dispositionally conscious. As it

---

[2] Here and below the page numbers refer to the online version of the paper: http://www.phon.ucl.ac.uk/home/robyn/Carston-Recanati-22August05%5B2%5D.pdf

should be clear, sub-personal inferences are—so to speak—inferential only in the eyes of an observer. This is the case with Recanati's primary pragmatic processes: they are thought to be merely associative processes although they may nonetheless "mimic" inferential processes (Recanati 2007). On the other hand, secondary pragmatic processes are claimed to be genuinely inferential, conscious processes at least in a dispositional sense.

This cognitive version of Grice's inferential account of comprehension differs from RT's proposal on two major points. First, relevance theorists assume that one single mechanism is sufficient to account for utterance understanding: that is, an unconscious, automatic process based on expectations of relevance. In particular, they believe that the whole distinction between primary and secondary pragmatic processes, respectively conceived of as sub-personal and personal processes, is not grounded: consciousness is not thought to play a significant role in normal episodes of comprehension. Second, the single process by which RT explains comprehension is conceived of as both automatic and yet genuinely inferential in its own right. Although this process is said to be unconscious and outside the control of the subject, relevance theorists describe it as a non-demonstrative inference that takes a set of premises as input and yields a set of conclusions as output. This means that in their account the inferences in terms of which pragmaticists reconstruct utterance understanding are literally part of the automatic process of comprehension: inferences are neither merely attributed to the subject (in particular, comprehension is not based on mere associative processes that just mimic inferences, as in Recanati's primary pragmatic processes), nor are they something that the subject is just capable of delivering explicitly if necessary (as in Recanati's dispositional account of secondary pragmatic processes). They are instead genuine, occurrent inferences, although automatic ones.

In general terms, in RT's account of communication an utterance conveys a presumption of its own relevance, and the hearer has to construct a hypothesis about the speaker's meaning which satisfies that presumption of relevance. This requires constructing appropriate hypotheses about explicit content, intended contextual assumptions, and implicated conclusions—with explicit content and contextual assumptions counting as premises from which implicated conclusions are to follow. Although explicit content provides one of the premises for the inference, this does not mean that it is wholly determined by means of a previous non-inferential process—for instance, by means of a purely associative process as in Recanati's model. In fact, the whole process is conceived of as circular rather than uni-directional: hypotheses about the implicated conclusions might be suggested straightforwardly by some contextual cues, so that those conclusions can contribute to determine the premises which are apt to draw the inference. In this sense, Relevance theorists speak of a "mutual adjustment" between explicit content, contextual assumptions and implicated conclusions. Therefore, the very same process based on the construction of inferential derivations is believed to be responsible for the determination of both explicit and implicit content.

In short, neither simple associative processes nor conscious, controlled processes are claimed to play any significant role in utterance understanding. RT

conceives of comprehension as a quite specific process: an automatic inferential process which is specialized for the purposes of communication. According to Carston (2007), reflective reasoning may well play a role in communication and comprehension but only with regard to rational reconstruction of spontaneous pragmatic processes: "this is not an exercise that people perform much off their own bat. Its most likely role is as a backup mechanism when something goes wrong with the automatic intuitive mechanisms of utterance understanding" (idem, 31).

Before we address the respective weaknesses of RT and Recanati's account, let us turn to the distinction between automatic and controlled processes: this distinction may form the basis for a different account of how sub-personal and personal features may coexist in comprehension.

3 Automatic and Controlled Processes

The view that human cognition involves two different types of processing, automatic and controlled, is a well-established theme in psychology at least since the writings of William James (1890). The issue has received renewed attention in the last decades after the seminal studies of Shiffrin and Schneider (1977) and Schneider and Shiffrin (1977)—see Schneider and Chein (2003) for a recent overview. Schneider and Shiffrin (1977) define an automatic process as the activation of a sequence of nodes in our knowledge representation that "nearly always become active in response to a particular input configuration" and that "is activated automatically without the necessity for active control or attention by the subject" (idem, 2). In contrast, controlled processes are defined as "a temporary sequence of nodes activated under control of, and through attention by, the subject" (idem, 2–3).

The standard tests employed to assess whether a process of interest is automatic or controlled are subliminal presentation and techniques based on cognitive load (see Satpute and Lieberman 2006, 91). The fact that an input is processed subliminally, that is, outside awareness, is treated as the most distinctive feature of automaticity. On the other hand, it is also expected that whenever the process of interest is automatic, it will not be influenced by load manipulations. This is because automatic processes are thought to occur in parallel, in contrast with controlled processes which operate serially, resulting in task-switching costs. Another important manifestation of the automatic-controlled distinction is that "extended consistent training is required in order to develop automatic processing, while controlled processes can be established in a few trials and under varied mapping conditions" (Schneider and Chein 2003, 528). As a consequence of extended training and parallel processing, automatic processes are fast and accurate. On the contrary, controlled processes are typically slow and inaccurate.

3.1 Old and New Approaches to the Automatic/Controlled Distinction

In the last decades, the basic distinction just outlined has been framed in largely similar ways by different scholars (Carver and Scheier 2009): intuitive versus conscious in Smolensky (1988), associative versus rule based in Sloman (1996), reflexive versus reflective in Shastri and Ajjanagadde (1993). Moreover, two-mode, or dual-process, models of functioning have emerged in personality psychology (experiential versus rational system: Epstein 1973), and in social psychology (reflexive versus reflective: Lieberman et al. 2002; impulsive versus reflective: Strack and Deutsch 2004; Strack et al. 2009).

In some of those developments of the distinction, the emphasis is on the fact that the second kind of process operates on "symbolic, or propositional structures" (Saptute and Lieberman, 2006, 88), enables symbolic logic (Lieberman et al. 2002), can be simulated by symbolic architectures allowing the binding of variables (Schneider and Chein 2003, 532)—in contrast to automatic processes which can be simulated by simple connectionist networks. This shift towards the subsymbolic/symbolic distinction is pushed to the point that in some cases consciousness is no more considered distinctive of the controlled type of process. For instance, Lieberman et al. (2002) distinguishes between an X-system for reflexive processes and a C-system for reflective processes, where the former results itself in a state of consciousness: the X-system is said to be a parallel-processing, sub-symbolic, pattern-matching system that produces the continuous stream of consciousness that we experience as the world out there.

As useful as it can be for some theoretical purposes, this way to recast the automatic/controlled distinction is probably misleading for various reasons. First, it is hardly coherent with the most accepted way to assess automaticity, that is, by means of tests based on subliminal processing: these tests precisely aim to ascertain whether a cognitive process occurs outside consciousness. Second, at least for a crucial class of cognitive phenomena, by claiming that a process is controlled (versus automatic) scholars mostly intend to emphasize its conscious and voluntary nature, irrespective of whether it is a symbolic (propositional) process or not. This is the case with a large amount of research in neurophysiology, where the issue is at which conditions bodily movements become actions, that is, they are under conscious control (versus merely automatic: for instance, see Jeannerod 2006; Pacherie 2006). From this point of view, neurophysiological literature on intentional bodily movement is just an instance of a larger category: that of research on intentional action in general, which also includes linguistic behavior. In research on intentional action it is consciousness rather than propositionality that is held to distinguish controlled from automatic processing. Third, by downplaying the importance of the conscious/unconscious distinction one loses the crucial theoretical connection between the automatic/controlled issue on the one hand, and the notions of selective attention and executive functions on the other. Selective attention, which is tightly connected to consciousness, has

traditionally been considered a key component of executive functions. The notion of executive function is used in psychology and neuroscience to describe a loosely defined set of capacities having to do with guidance of behavior: planning, initiating appropriate actions while inhibiting inappropriate ones, cognitive flexibility etc. Selective attention is apparently a key component of this cognitive complex. On the other hand, executive functions appear as prototypical examples of controlled processes: but this is so because of the fact that executive functions involve consciousness and selective attention, not propositionality.

For all these reasons, I will rest on the most widely accepted view of the automatic/controlled distinction, that is, the view that takes consciousness as the main line of demarcation and emphasizes the connections with theories of intentional action and executive functions. Once this general framework is settled, further qualifications are suggested by evidence that has been acquired recently. One is the observation that "automatic" and "controlled" presumably are not all-or-none notions; instead, they appear to come in degrees. A second point is that, although consciousness and controlled processes have a crucial role to play in goal-directed behavior, it is possible to have goal-directed behavior outside consciousness. Third, the emphasis in recent research is less on how automatic and controlled processes may be detected and analyzed in isolation than on how they factually cooperate for most of our cognitive activities. Let us now briefly address each of these points in turn; this will prepare the ground for an updated view of the automatic/controlled issue—a view which has been explored in the "distributed intentionality model" proposed by Mazzone and Campisi (2013).

3.2 All-or-none Notions?

The traditional view according to which there is a sharp boundary between automatic and controlled processes has been challenged in the last decades, either because automaticity has been explained in terms of properties which vary gradually, or because automaticity has been analyzed in components which can, but need not, be present together (Garrod and Pickering 2007). The latter line of thought has been pursued by Bargh (1994): in his view, a process is automatic to the extent that it is unaware, mandatory, efficient, non-interruptible.[3] However, since those features do not always covary together, there may be different degrees of automaticity as a function of the number of features involved.

As for the former line of thought, it is well exemplified by Cohen et al. (1990). They propose that automaticity is a function of what they call "strength of processing", which in turn is defined in relation to processing pathways within a

---

[3] In Bargh's (1994) own terms, the four parameters are awareness, intentionality, efficiency, and controllability. However, Garrod and Pickering suggest that "non-interruptibility" is a more proper label for what Bargh calls "controllability". Similarly, Mazzone and Campisi (2013) observe that by the term "intentionality" Bargh properly means that a process is not mandatory.

connectionist network. A strong connection leads to fast and accurate transmission of information along the pathway; moreover, the strength of processing may determine the extent to which processes are open to interference from other processes. Considerations of this sort might be framed differently by taking into consideration the apparent coexistence of two complementary learning systems in our brain, one for rapid learning based on the hippocampus and related structures, the other for slow learning of regularly repeated sequences (McClelland et al. 1995). As it seems, the same information can be moved from the former system to the latter as a consequence of repetition and practice (Aarts and Custers 2009); this also leads to different patterns of activation in the neural circuits guiding action:

> lateral prefrontal and premotor areas are activated at the beginning of the learning of a motor sequence; with practice and repetition, however, that activation subsides, while that of subcortical structures, notably the basal ganglia, increases […]. Thus, as sequences become overlearned and automatic, their representation seems to "migrate" to lower executive stages (Fuster 2001, 321–322).

Under this hypothesis, processing can be fast, accurate and unaffected by interference to the extent that a sequence has been overlearned and therefore moved to subcortical structures—where overlearning is something that may come in degrees.

A different but possibly complementary proposal has been put forth by Dehaene et al. (2006) with regard to the neural basis of consciousness. In the model they propose, the distinction between conscious and unconscious processing is not all-or-none, for two reasons. First, they assume that besides conscious and purely subliminal processing there can be intermediate conditions. Second, in their model consciousness is also a function of the strength of activation in the interested brain areas. To be more precise, Dehaene et al. (2006) distinguish four conditions of un/consciousness, depending on the degree of activation which is found respectively in posterior sensory-motor representations and higher association cortices. Pure (i.e., unattended) subliminal processing occurs when there is a weak and rapidly decaying activation in posterior sensory-motor areas, without any significant interacting activation in anterior cortices; on the other hand, attention and task set might occasionally interact with such weak posterior activations, thus resulting in attended subliminal processing. Moreover, an intense activation which is yet confined to sensory-motor processes is thought to cause occipito-temporal loops and local synchrony, and therefore a condition of preconscious processing: processes are virtually accessible to consciousness, although attention is actually oriented away from the stimulus, so that activation is blocked from accessing higher parieto-frontal areas and establishing long-distance synchrony. Finally, conscious and controlled processes require the establishing of long-distance loops between strongly activated sensory-motor representations and higher association cortices.

In sum, considerations from both psychology of learning and neurobiology of consciousness seem to suggest that the automatic/controlled distinction admits of degrees, instead of being an all-or-none affair.

3.3 Goal-Directedness without Consciousness

Traditionally, goal pursuit has been conceived of as a typical case of conscious and effortful processing. This view has been recently challenged especially by John Bargh (starting from Bargh 1989, 1990). His notion of *automatic* or *non-conscious goal pursuit* has witnessed a number of empirical demonstrations in the last decade (for recent reviews see Hassin et al. 2009; Ferguson et al. 2007). The thesis of non-conscious goal pursuit is based on the notion of habit, with habits conceived of as

> associative networks that include contexts, goals that are regularly pursued in these contexts, and means that one usually uses to attain these goals […]. These networks are shaped by one's history, and they allow for goal pursuit via spreading of activation (Hassin et al. 2009, 550–551).

Given this conception of habits as associative networks, it seems an obvious consequence that the activation of a component may spread to other components of the network, and this has been largely confirmed by experiments based on priming.[4] In particular, priming of goals appears to affect subsequent representations and behaviors in many ways. For instance, when an action is regularly selected and performed in order to obtain a goal (for instance, taking the bicycle instead of the bus to go to the university), "priming these goals automatically activates behavior representation and resultant action according to an 'if-then' rule, enabling the goal-directed behavior to occur directly and independent of conscious intentions" (Dijksterhuis et al. 2007, 105). Of particular interest is the fact that similar results have also been obtained through *unobtrusive* or *unconscious* priming. For instance, Bargh et al. (2001) unobtrusively exposed subjects to words such as "strive" and "succeed" to prime the achievement goal, and then tested their performances in an anagram puzzle task: participants primed with the achievement goal outperformed those who were not primed with the goal. Similar effects may also be obtained in more indirect ways: for instance, priming the names of significant others may lead to the automatic adoption of the goals associated with them; or for another example, thinking to a good friend may enhance the disposition to participate in a subsequent task as a possible means to help (Dijksterhuis et al. 2007, 101–102).

What these observations apparently show is that automatic processing may go deep into the guidance of behaviors which are thought of as typically intentional and controlled. That this must be the case is also shown by the fact that most of our intentional actions are nonetheless rapid and effortless. This has led to models of intentional actions where the most part of cognitive processing is thought to occur automatically. However, such models often tend to assume that automaticity does

---

[4] See also Gollwitzer et al. (2009, 605), where they suggest that goals may behave in accordance with simple associative (hebbian) principles: "Under the assumption that goals, too, are represented mentally and become automatically activated by the same [hebbian] principles, goal representations should also be capable of automatic activation through contact with features of the contexts in which those goals have been pursued often and consistently in the past.".

only concern the implementation of intentions: conscious representations of the goals to be pursued are held to be required for action control, while the specific behavioral means by which the goals are pursued would be activated automatically (see, for example, Levelt's 1989 model of language production). On the contrary, the evidence concerning non-conscious goal pursuit invites us to believe that goals may drive action without becoming conscious, that is, they can operate in an entirely automatic way.

3.4 How Automatic and Controlled Processes Cooperate

On the basis of our previous considerations, one could be tempted to think that consciousness does not play a significant role in human cognition. Our actions are mostly rapid and effortless, and this suggests a major role for automatic processing. To be sure, human action is essentially goal-directed, but, as we saw, goal-directedness does not imply conscious processing. Another relevant line of evidence is provided by the experiments of Libet (e.g., Libet 1992), which have shown that "[c]onsciousness of the goal of an action is not immediate, it takes time to appear" (Jeannerod 2006). More specifically,

> the first conscious awareness associated with the initiation of the movements […] occurs well after the start of the neural activity that culminates in the movement. […] This clearly suggests that whatever events one might reasonably consider to be the neural initiators of these movements, those events occur pre-consciously (Pockett 2006, 18–19).

Based on this sort of evidence, some have drawn the conclusion that consciousness is essentially a post hoc phenomenon, which has not to do with initiation and guidance of action. It would rather be (part of) a mechanism "for the cognitive rearrangement after the action is completed" (Jeannerod 2006, 37), in the service of our sense of agency and the distinction between our own and others' actions (Pockett 2006; Jeannerod 2006; Choudhury and Blakemore 2006). However, there are reasons to believe that conscious and controlled processes should be accorded instead a significant role in active online processing and guidance of action. First, conscious control appears to be occasionally required in the course of action when smooth automatic processing fails (Gollwitzer et al. 2009, 610; Bongers and Dijksterhuis 2009; Jeannerod 2006, 30). Second, sometimes we make conscious plans of action, or we are explicitly required to accomplish a task, and so on. In such cases, but possibly also in cases where initiation of action is automatic, consciousness seem to play a key role in top-down maintenance of goals and top-down inhibition: the execution of long-term plans cannot be accounted for solely in terms of automatic spreading of activation. This suggests that not only have both automatic and conscious processes a role to play in human cognition, they are also expected to cooperate in most of our cognitive performances. There is nowadays growing acceptance that "conscious and nonconscious goal pursuit are two collaborative partners taking turns in working towards goal attainment" (Gollwitzer et al. 2009, 620–621).

This cooperative view of automatic/controlled processes is entirely coherent with the neurobiological model proposed by Dehaene et al. (2006): as we saw, that model proposes that both in attended subliminal processing and in conscious processing frontal and pre-frontal activations can affect automatic processes, by amplifying the independent activation of certain representations (and presumably by causing the active inhibition of others) in posterior areas. An interesting way to frame attended subliminal processing is Neumann's (1990) theory of "direct parameter specification". According to this theory "[a] given attentional (or intentional) state might be necessary for unconscious stimuli to trigger further processes" (as Kiefer 2007, 293, puts it). More specifically:

> [Subjects] search for information in order to specify free parameters within the currently active intention/action plan. Unconsciously registered information that resembles this searched-for information is selected and processed to specify the free processing parameters. Therefore, unconsciously perceived information will translate into behavioural effects that are absent if the same information is sufficiently dissimilar from the searched-for features (Kiefer 2007, 300).

In other words, top-down intentional processing would cause stimuli to affect behavior even when they are not consciously perceived.

3.5 An Updated View of the Automatic-Controlled Issue

In the light of the sort of evidence we have reviewed so far, the distinction between automatic and controlled processes should be considered just the first step on the way to understanding their cooperation in most of our cognitive operations. Mazzone and Campisi (2013) have proposed a general approach to intentional actions—the "distributed intentionality model"—based on such a cooperative view of automatic and controlled processes. We propose that in order for actions to be intentional it is not required that action plans are consciously represented and then put into effect in a purely top-down manner. In the general case, actions are largely the result of automatic processes of activation, integration and competition between a huge number of goal-related representations. On the other hand, human behavior is intuitively intentional in essence, in that it never seems to occur without agents consciously attending this or that component of the complex goal-directed representation involved. However, conscious attention is not necessarily directed towards one specific component of that representation, be it an overall goal or whatever: conscious intentions should rather be conceived of as beams of light temporarily directed towards this or that goal-related component of a largely automatic flow of processing. In a word, intentionality is better thought of as dynamically *distributed* along the complex goal-directed representation involved in any single action, than concentrated in (the representation of) one single purpose of the action.

In our model, the role played by automatisms is very large. This is in line with a proposal of Morsella (2009), according to which human behavior is based on a "stream of action […] driven by a continuous series of activations stemming from various sources" (idem, p. 19). In other words, our perceptions would endlessly feed automatic processes impinging on motor representations, so that plans of action are activated automatically at each moment and then compete for behavioral expression (idem, p. 16). However, this is far from implying that consciousness is either absent or purely epiphenomenal in most of our intentional actions. First, even if consciousness takes time to appear, nonetheless it may emerge in the course of action and then play a crucial role as a mechanism for goal maintenance and shielding, for reorganization of habits, or for the management of unexpected difficulties (Mazzone and Campisi, 2013). Second, it should not be forgotten that for the most part of our lives "we live in a supraliminal world" (Satpute and Lieberman 2006, 91), that is, automatic responses to perceptual inputs occur while we are engaged in conscious monitoring of the environment and our own behavior. In a sense, then, it could be true that there are conscious representations at the instigation of most of our actions: humans often respond to situations they are conscious of, and these situations set the purposes of our forthcoming actions. For instance, in dialogue we normally attend to our interlocutor's utterances. Such a conscious representation of the input we intend to respond to can be thought to drive automatic processing by constraining the kind of information which is needed to accomplish the task—as predicted by the "direct parameter specification" theory considered above.

In sum, it seems that in principle any component of the complex goal-directed representation involved in action—including goals—can be processed automatically. Nonetheless, consciousness is far from being epiphenomenal since it may focus on this or that component when needed and, as a consequence, play a role in directing automatic processing: specifically, as in Dehaene et al.'s (2006) model, by amplifying or inhibiting representations in posterior areas of our brain.

4 Pragmatics and Associative Processes

As we saw in Sect. 2, Relevance Theory and Recanati's view are not equally compatible with psychological and neurobiological accounts of the controlled/automatic distinction. Recanati's view is closer to those accounts than RT, to the extent that the former conceives of pragmatic processing in terms of a cooperation between associative and conscious processes, while the latter does not accord a role to any of these two processes within pragmatics proper. Relevance theorists propose instead a single automatic mechanism which is specialized for language comprehension. Let us now examine in more detail the positions of both RT and Recanati with regard to associative processes (this section), and conscious processes (the next section).

Within the literature we considered in Sect. 3 there is a general consensus that automatic processing occurs by way of associative mechanisms, that is, mechanisms based on associative strength in a network mainly due to "extended consistent training" (Schneider and Chein 2003, 528) and on subsequent spreading of activation in the associative network. For instance, in Sloman (1996) the automatic pole of the dichotomy is straightforwardly called "associative" (versus rule based), and in Satpute and Lieberman (2006, 88) the reflexive (versus reflective) component is claimed to be based on associations and to deliver constraint satisfaction processes. Moreover, the thesis of automatic goal pursuit depends on the notion of habits conceived of as associative networks involving representations of contexts, goals and means.

Recanati (2004) has in fact proposed that lexical items contribute their meaning to the explicit sense of utterances by way of what he calls primary pragmatic processes, conceived of as local associative processes. In Mazzone (2011) I have argued that Recanati's associative explanation may be extended beyond his intentions—in particular, beyond the domain of lexical pragmatics. But let us proceed step by step.

4.1 Associative Accounts of Primary Pragmatic Processes

A crucial notion in Recanati's account of primary pragmatic processes is that of abstract schemata driving comprehension. Not only can schemata explain shifts in accessibility of lexical meanings, they can also account for the search of coherence in associative processes: inputs activate schemata they are component of, and schemata in turn activate (or add activation to) other inputs (and their interpretations) insofar as they fit those schemata. In our previous example (Sect. 2), the schema GETTING MONEY FROM A BANK$_1$ (where BANK$_1$ = FINANCIAL INSTITUTION) may have a key role in explaining how, in the utterance "I'm going to the bank now to get some cash", the word "bank" is given an interpretation which is coherent with the context. Interestingly, the same sort of schematic information is invoked by RT in order to ensure the assumptions that behave as premises in their inferential explanations. Thus, in this respect what essentially distinguishes RT from Recanati's account is the thesis that such a schematic information is employed within genuinely inferential processes, instead of associative ones.

However, as argued in Mazzone (2011), this thesis is both highly speculative and unnecessary. As to the first point, on epistemological grounds associative activation and automatic inferential derivation are far from having the same status: the latter is not nearly as established as the former, which is in fact the only well-established explanation–both in psychology and neurobiology—of how we detect, store and exploit information by way of automatic processes. On the other hand, it is far from clear that a genuinely inferential account is needed. In particular, RT underestimates the theoretical role that schemata can play within an associative account of automatic processes.

   This is clearly shown by the most extensive argument against associative accounts proposed by relevance theorists, which has been put forth by Wilson and Carston (2007) in the context of a discussion of lexical pragmatics. The key claim in their argument is that statistical associations provide no basis for drawing warranted conclusions, since the associates are not logically related to each other in any systematic way. To be sure, one could maintain that inferential relationships are also associations of some sort; for instance, the association between "shark" and "fish" could be used to derive the warranted conclusion that a shark is a fish.[5] However, although inferential relationships are associations, there are plenty of associations that are not inferential relationships. Therefore, according to Wilson and Carston associative accounts will vastly overgenerate, and so one is left without any principled method of filtering out unwanted associations (and unwarranted conclusions). This is why inferential accounts should be preferred.

   The first thing to notice is that the premise of the argument is false. Far from lacking any systematic structure, associations are instead essentially schematic. In other words, associations are not stored in such a way that the relationships between their elements are in need of interpretation from the outside—so to speak. Quite on the contrary, our associative coding of contingencies yields schemata preserving information both on which content are connected with each other and *how* they are connected, be it by way of taxonomic, part-whole, temporal, causal, textual relationships or whatever. As a consequence, associative networks do not require that further mechanisms be provided in order to logically constrain their dynamics of activation. Instead, they can themselves provide—just as suggested by Recanati—the abstract schemata thanks to which the process is constrained, and unwanted associations are filtered out. For instance, although in our previous example the word "bank" may activate the meaning RIVER SIDE—not to mention all the other associations potentially activated by the utterance "I'm going to the bank now to get some cash"—this meaning either will not receive further activation from, or even will be inhibited by, other associative schemata triggered by linguistic and contextual inputs.

   As it seems, the notion of schema may help to provide, after all, an associative explanation of how unwanted associations are filtered out in comprehension.

4.2 Beyond Lexical Pragmatics

It could be objected that such an explanation may only work within the limits of lexical pragmatics, where the issue is how words confer their meanings to the explicit content of utterances. Also relevance theorists grant a role to associative

---

[5] Within this argument, Wilson and Carston essentially identify associations with statistical relationships between lexical items in a corpus. As we are going to argue, there is no ground for that identification: there exist in fact a variety of different associative relationships, most of which concern concepts rather than words.

processes in that domain. However, according to RT not even explicit content can entirely be fixed by associative mechanisms: explicit content—no matter how it is prompted—has to become a line in an inferential derivation, and mutual adjustment between the components of the derivation is needed in order for any of those components to be fixed.

Why should we presume that such an inferential derivation is needed, and that simple associative processes will not do? Carston (2007) has an argument for this which deserves consideration.[6] Her idea is that associations suffice insofar as what is at issue is activation and deactivation of concepts (parts of concepts, schemas), while associations are not sufficient in order to understand genuinely constructive processes. Although Carston is here concerned with how concepts can be constructed online rather than simply re-activated, her argument also sheds new light on the previous claim that associations provide no basis for drawing warranted conclusions. Intuitively, in order to be justified in reaching a conclusion a cognitive system needs something more than activation merely passing from one content to another: it has to *construct* an inference that may count as a justification for the conclusion. I think Carston has a good point here, but the precise implications of the argument have to be assessed more accurately.

Let us first notice that in current linguistics there is a family of theories assuming that associative relations can explain cognitive phenomena which had previously been thought to require rule-based, specialized processes instead. This is the case with what are known as constraint-based models, that is, models in which parallel activation of, and competition between, representations substitute for procedural rules, in syntax and elsewhere (e.g. Trueswell et al. 1994; Ferreira et al. 2002; Jackendoff 2007; Breheny et al. 2006). Constraint-based processes and associative processes can be seen as two sides of the same coin: as a consequence of activation within an associative network, each activated representation may act as a constraint on the overall process, insofar as it contributes to selecting the outcomes which are compatible with it.

One insightful example of constraint-based model in linguistics has been put forth by Jackendoff (2007). Although his theory has its roots in Generative Grammar, Jackendoff maintains that linguistic phenomena—syntax included— may be explained by a general-domain, constraint-based mechanism. Crucially, while in the mainstream view of Generative Grammar phrase structure has been represented in terms of procedural rules, Jackendoff proposes that any linguistic information[7] including phrase structure is instead captured by regular patterns of representation essentially abstracted away from experience: words, regular affixes, idioms, constructions, and ordinary phrase structure rules are conceived of as nothing but "pieces of structure stored in long-term memory" (Jackendoff 2007, 11). As a consequence, Jackendoff's explanation does not rely on specialized

---

[6] She also proposes another interesting argument we will consider in the next section, since it concerns the role of consciousness.

[7] With the possible exception of a very restricted number of innate constraints.

linguistic (namely, syntactic) processes operating in accordance with procedural rules. Rather, linguistic representations (pieces of structure) are thought to contain within them the information on how they can be assembled with each other, and all we would need is a general-domain process which mechanically assembles representations in accordance with that information: this process is called *unification*.

Since what is at stake in unification is the building of occurrent linguistic structures, Jackendoff rightly points out that this process necessarily requires something like a "workbench", or a "blackboard" where structures are constructed online. Typically, such a workbench is what working memory is thought to provide. But Jackendoff also emphasizes that in order to accomplish the task, working memory cannot be conceived of as just the part of long-term memory that is currently activated—as it is in some connectionist architecture. In his opinion, working memory should rather be thought of as physically separate from long-term memory. For our purposes, though, the point is that building conceptual structures requires more than simple spreading of activation in a network: it requires that the cognitive system is able to keep certain pieces of structure active until the whole process of activation, competition and unification is accomplished.

What these considerations suggest is that the construction of complex conceptual structures can be accounted for within an associative framework, provided that working memory is added to simple spreading of activation. Carston's objection has the merit of calling attention to this important qualification, but it does not speak in favor of inferential processes, since we have an associative account of how complex structures can be constructed in the course of online processing. One may speculate that those complex structures possibly include exemplifications of inferential schemata. This could explain how associative processes may mimic inferential processes, as suggested by Recanati (2007). There could occur indeed processes of mutual adjustment between assumptions counting as premises and utterance interpretations counting as conclusions, insofar as those assumptions and interpretations are unified in working memory by means of inferential schemata: such schemata would activate, or strengthen the activation of, the components (premises and conclusions) which fit them.

Would that count as an inferential account of the sort recommended by RT? Not at all. A couple of things should be emphasized. First, the mechanisms Jackendoff's model makes use of (spreading activation, working memory) are domain-general. On the contrary, inferential processes hypothesized by RT are specialized for utterance understanding. Second, in constraint-based accounts a crucial role is played by structures of representation: those structures (plus simple general-domain mechanisms) substitute for rules and derivations. Analogously, in pragmatic processing a variety of schemata (together with spreading activation and working memory) could explain how warranted conclusions could be granted without genuine inferential processes. Third, there is a clear sense in which working memory is just part of the general associative dynamics of our brain: from a neurobiological perspective, working memory is generally taken to consist in self-sustained loops occurring in cortical circuits. In other words, working memory is essentially a specific modality of activation within associative networks.

These considerations support the view that associative processes can explain linguistic and pragmatic phenomena well beyond the limits of lexical pragmatics.

Before we conclude this section, one qualification is in order. The fact that we have introduced working memory in the picture does not necessarily imply that consciousness is at play as well. For sure, the idea of a global workspace—ensured by self-sustained loops in the cortex, and accounting for the active maintenance and integration of information—has traditionally been tightly associated with consciousness (e.g., Baars 1997). Nonetheless, there are reasons to think that working memory is independent from consciousness. For instance, Hassin et al. (2009) have noticed that there is an apparent contradiction between the fact that we can only engage in a very limited number of high-order cognitive processes (and specifically, conscious processes) at any given point in time, and the simple intuition that there are points in time in which we seem to be advancing multiple goals, decisions and plans. Hassin and colleagues propose to solve this apparent contradiction by adopting the thesis of an implicit working memory. In other words, they argue that working memory can operate outside of conscious awareness and therefore it may ensure parallel processing. Incidentally, they also observe that none of the major views on this issue suggests that people have conscious access to everything that goes on within working memory.

The obvious implication of the "implicit working memory" hypothesis is that conceptual integration may also occur automatically—i.e., outside consciousness. Automatic integration of spatial information has been in fact argued for by Hommel (1996, 1998, 2002). In ERP studies of language comprehension, it could be argued that a similar notion has been invoked. Many have proposed to interpret the difference between the best known event-related potentials in that domain—N400 and P600—in terms of a difference between two modes of processing. The N400 (a negative deflection having its peak 400 ms after the stimulus that elicits it) is thought to reflect a process of semantic integration (van Berkum et al. 1999; Vissers et al. 2006; Chwilla et al. 2007), while the P600 (a positive deflection with its peak at 600 ms from the stimulus) would reflect instead a process of monitoring and "continued algorithmic analysis" presumably under executive control (Kuperberg 2007, 42; Vissers et al. 2007). The sort of integration that is proposed as an explanation of the N400 is thus conceived of as different from both simple spreading of activation (Chwilla et al. 1998) and processes involving conscious monitoring and executive control.

From a neurobiological perspective, the hypothesis of implicit working memory is compatible with the model proposed by Dehaene et al. (2006) we considered above (Sect. 3). Under the assumption that working memory is ensured by self-sustained loops in the cortex, the model distinguishes between local loops located in occipito-temporal areas and long-distance loops also involving anterior association cortices. While the latter are thought to grant conscious and controlled processes, the former are claimed to cause preconscious processing. Therefore, local loops could be the neurobiological basis for implicit working memory and preconscious integration of representations.

5 Pragmatics and Consciousness

As we saw, Recanati's account assigns a significant role to consciousness within pragmatics. More precisely, the "availability condition" (Recanati 2004, 44) posits that subjects have conscious awareness of the explicit content, the implicit content, and the inferential process leading from the former to the latter. In a sense, Recanati conceives of the transition from explicit to implicit sense in terms of conscious verbal reasoning, with the important qualification that consciousness may be only dispositional: subjects are capable of making the relevant inferences explicitly, but in the normal case they reach the implicit content without any actual involvement of conscious verbal reasoning.

Carston (2007) makes two objections that are easily agreed upon. First, moving from occurrent to (merely) dispositional reflective inferences leaves us with no idea of how the real process of implicature derivation works: what we do know is just that conscious verbal reasoning is *not* the occurrent process by which implicit content is normally obtained. As we saw, Carston's own view is that conscious reasoning should better be assigned a role in post hoc rational reconstruction, which is something that people are actually engaged in only as a backup mechanism when something goes wrong with automatic processing.

Second, Carston argues that there is no ground for the distinction between primary and secondary pragmatic processes in terms of conscious availability. She makes various examples of cases in which people seems to be aware of how explicit content may depart from linguistic meaning as a function of contextual factors. Let us consider the following example:

(3) Mother to young child just before bedtime: Have you brushed your teeth?
    Child (grinning): Yes I have—[pause]—last night.

The answer clearly shows that the child is well aware of the normal pragmatic enrichment by which the relevant time is assigned to the temporal parameter of the question: in fact, the child openly violates the expectations raised by that normal enrichment. But in Recanati's terms that sort of enrichment is a case of primary pragmatic process for the determination of explicit content. Therefore, conscious availability seems not to be an exclusive property of secondary pragmatic processes.

The claim that any stage of pragmatic processing may be consciously attended accords well with our previous considerations on conscious processes (Sect. 3), with particular regard to Mazzone and Campisi's (2013) "distributed intentionality model": speaking is a prototypical case of intentional action and, as I argued above, intentional action involves complex goal-directed representations across which consciousness is dynamically distributed. In other words, there are no specific components of goal-directed representations such that consciousness is necessarily directed towards them; consciousness may be directed instead towards different aspects in different occasions, and also in the course of the same action. But Carston also claims that consciousness have a role to play in utterance

understanding only in very special circumstances. However, it should be emphasized, Carston essentially refers to the role of conscious *reasoning*, which is a quite specific sort of conscious process. Although Carston is presumably right in pointing out that we rarely resort to conscious reasoning in utterance understanding, this in itself does not speak against consciousness having a role in pragmatics. To put it differently, while verbal reasoning proper is a prototypical instance of (largely) controlled processes, consciousness may also cooperate with processes which are mainly automatic: it is this latter kind of process, not the former, that is apparently involved in normal cases of utterance understanding.

Carston (2007) has made an argument against associative processes that in my perspective can be seen as an involuntary step in this direction. She points out that mere accessibility—even coherence-based accessibility—cannot account for the fact that utterances virtually inevitably trigger attentional focus and the expenditure of some processing effort. The conclusion this argument is aimed at is RT's thesis according to which utterance understanding cannot be explained by a general-domain associative process; it would require instead a specialized automatic mechanism based on relevance. However, one may speak of attention (attentional focus, and the like) in two quite different ways. First, one may refer to the mere fact that a cognitive system has to select somehow the direction of processing. Second, one may specifically refer to conscious attention. Carston cannot presumably be interested in the latter sense, since she argues in favor of unconscious processing of utterances. Nonetheless, it is difficult to deny that utterances do normally trigger *conscious* attention in humans. And this may contribute to explain how cognitive resources are allocated in utterance understanding, beyond mere spreading of activation: automatic processing, as we saw above, can be driven by consciously attended representations, which have a role in amplifying or inhibiting other representations in posterior cortical areas, in maintaining certain representations activated, and in creating expectations about the inputs to be automatically processed.[8]

In a word, Carston calls our attention to a fact that, again, can be easily described in terms of a general and well-established mechanism—conscious attention—although she argues in favor of a highly speculative explanation—the hypothesis of an automatic, inferential, relevance-based mechanism. While one may agree with her that there is no ground for Recanati's distinction between primary and secondary pragmatic processes, that conscious verbal reasoning is not involved in normal cases of comprehension, and that comprehension is instead a largely automatic process, it seems reasonable to acknowledge nonetheless that conscious attention may play a role in utterance understanding, in cooperation with automatic processes.

---

[8] The role of conscious attention in pragmatics is further explored in Mazzone, 2013b.

6 Conclusions

We live in a supraliminal (Satpute and Lieberman 2006, 91), personal-level world. Language perception, in particular, does not normally occur outside consciousness. According to Grice, comprehension involves personal-level, rational abilities on the part of the hearer. At the same time, however, pragmaticists have not neglected that utterance understanding appears to be a spontaneous, rapid and effortless process. For that reason, Recanati has hypothesized an automatic, associative stage in utterance understanding, and has assigned only a dispositional (versus occurrent) role to conscious verbal reasoning. On the other hand, Relevance theorists have proposed that comprehension is a wholly automatic, though inferential, process, with conscious verbal reasoning being assigned only a peripheral role as a backup mechanism.

I have proposed here a different account, where consciousness plays a significant role in cooperation with automatic, associative processes. This account is based on the automatic/controlled distinction in psychology, and on recent views about the cooperation between these two kinds of process. In that perspective, not only do automatic and controlled processes cooperate, they are also closer to each other than it was previously thought. Specifically, I have argued that automatic processes are based on schemata which may also be recruited in reflective reasoning, while the main difference between reflexive and reflective processing concerns just the dynamics of activation within cortical networks. This is why automatic processes are apt to mimic inferential ones.

This is not to say that spreading activation is sufficient to account for utterance understanding. I have claimed instead that working memory is also needed, and argued that it may come in two different varieties: implicit and conscious. In the framework I propose, spreading activation, implicit working memory, and conscious attention are all present in normal episodes of utterance understanding, with the first two components doing the greatest part of the work, although conscious attention has also a key role in maintenance, amplification/inhibition and anticipation of representations.

The present account has the advantage of explaining both automaticity and personal-level, rational features of comprehension by an appeal to mechanisms that are general and well-supported in psychology and neurobiology.

References

Aarts, H., Custers, R. (2009). Habit, action, and consciousness. In W. P. Banks (Ed.), *Encyclopedia of Consciousness* (Vol. 1), 315–328. Oxford: Elsevier.

Baars, B. J. (1997). In the theatre of consciousness: Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies, 4*(4), 292–309.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence on social perception and cognition. In J. Uleman, and J. Bargh (Eds.), *Unintended thought*, 3–51. New York: Guilford Press.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins, R. M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2), 93–130. New York: Guilford Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R. S. Wyer Jr., T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed.), 1–40. Hillsdale: Erlbaum.

Bargh, J. A., Gollwitzer, P. M., Lee–Chai, A., Barndollar, K., & Troetschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology, 81*, 1014–1027.

Bongers, K. C. A., & Dijksterhuis, A. (2009). Consciousness as a trouble shooting device? The role of consciousness in goal–pursuit. In E. Morsella, J. A. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 589–604). New York: Oxford University Press.

Breheny, R., Katsos, N., & Williams, J. (2006). Are scalar implicatures generated by default? *Cognition, 100*(3), 434–463.

Capone, A. (2011a). Default Semantics and the architecture of the mind. *Journal of Pragmatics, 4*(3), 1741–1754.

Capone, A. (2011b). The Attributive/Referential distinction, pragmatics, modularity of mind and modularization. *Australian Journal of Linguistics, 31*(2), 153–186.

Carston, R. (2007). How many pragmatic systems are there? In M. J. Frapolli (Ed.), *Saying, Meaning, Referring. Essays on the Philosophy of François Recanati*, 1–17. New York: Palgrave.

Carver, C. S., & Scheier, M. F. (2009). Action, affect, and two–mode models of functioning. In E. Morsella, J. A. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 298–327). New York: Oxford University Press.

Choudhury, S., & Blakemore, S. J. (2006). Intentions, actions, and the self. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 39–51). Cambridge: MIT Press.

Chwilla, D. J., Hagoort, P., & Brown, C. M. (1998). The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *The Quarterly Journal of Experimental Psychology, 51*, 531–560.

Chwilla, D. J., Kolk, H. H. J., & Vissers, C. T. V. M. (2007). Immediate integration of novel meanings: N400 support for an embodied view of language comprehension. *Brain Research, 1183*, 109–123.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204–211.

Dijksterhuis, A., Chartrand, T. L., & Aarts, H. (2007). Effects of priming and perception on social behavior and goal pursuit. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 51–132). Philadelphia: Psychology Press.

Epstein, S. (1973). The self–concept revisited or a theory of a theory. *American Psychologist, 28*, 404–416.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good–enough representations in language comprehension. *Current Directions in Psychological Science, 11*, 11–15.

Ferguson, M. J., Hassin, R. R., & Bargh, J. A. (2007). Implicit motivation. In J. Shah & W. Wood (Eds.), *Handbook of motivation* (pp. 150–168). New York: Guilford Press.

Fuster, J. (2001). The prefrontal cortex-an update: Time is of the essence. *Neuron, 2*, 319–333.

Garrod, S., & Pickering, M. J. (2007). Automaticity in language production in monologue and dialogue. In A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (pp. 1–21). Hove: Psychology Press.

Gollwitzer, P. M., Parks-Stamm, E. J., & Oettingen, G. (2009). Living on the edge: Shifting between nonconscious and conscious goal pursuit. In E. Morsella, J. A. Bargh, & P.

Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 603–624). New York: Oxford University Press.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge MA: Harvard University Press.

Hassin, R. R., Aarts, H., Eitam, B., Custers, R., & Kleiman, T. (2009). Non–conscious goal pursuit and the effortful control of behavior. In E. Morsella, J. A. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 442–457). New York: Oxford University Press.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research, 59*, 176–186.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus–response episodes. *Visual Cognition, 5*, 183–216.

Hommel, B. (2002). Responding to object files: Automatic integration of spatial information revealed by stimulus–response compatibility effects. *Quarterly Journal of Experimental Psychology, 55A*, 567–580.

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research, 1146*, 2–22.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.

Jeannerod, M. (2006). Consciousness of action as an embodied consciousness. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 25–38). Cambridge MA: MIT Press.

Kiefer, M. (2007). Top–down modulation of unconscious 'automatic' processes: A gating framework. *Advances in Cognitive Psychology, 3*(1–2), 289–306.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23–49.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology, 34*, 199–249.

Libet, B. (1992). The neural time–factor in perception, volition and free will. *Revue de métaphysique et de morale, 97*(2), 255–271.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge MA: MIT Press.

Mazzone, M. (2011). Schemata and associative processes in pragmatics. *Journal of Pragmatics, 43*, 2148–2159.

Mazzone, M. (2013a). A pragmatic Pandora's box. Regularities and defaults in pragmatics. In F. Liedtke and C. Schulze (Eds.), *Beyond Words, Content, context, and inference* (pp. 307–329). Berlin, Boston: Mouton.

Mazzone, M. (2013b). Attention to the speaker. The conscious assessment of utterance interpretations in working memory. *Language & Communication, 33*, 106–114.

Mazzone, M., Campisi, E. (2013). Distributed intentionality. A model of intentional behavior in humans. *Philosophical Psychology, 26*, 267–290.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419–457.

Morsella, E. (2009). The mechanisms of human action: Introduction and background. In E. Morsella, J. A. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 1–32). New York: Oxford University Press.

Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research, 52*, 207–215.

Pacherie, E. (2006). Towards a dynamic theory of intentions. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 145–167). Cambridge MA: MIT Press.

Pockett, S. (2006). The Neuroscience of movement. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 9–24). Cambridge MA: MIT Press.

Recanati, F. (2004). *Literal Meaning*. Cambridge: Cambridge University Press.

Recanati, F. (2007). Reply to Carston. In M. J. Frapolli (Ed.), *Saying, Meaning, Referring. Essays on the Philosophy of François Recanati*, 49–54. New York: Palgrave.

Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processing into neurocognitive models of social cognition. *Brain Research, 1079*, 86–97.

Schneider, W., & Chein, J. M. (2003). Controlled and automatic processing: Behavior, theory and biological mechanism. *Cognitive Science, 27*, 525–559.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I. Detection, search, and attention. *Psychological Review, 84*(1), 1–66.

Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables, and dynamic binding using temporal synchrony. *Behavioral and Brain Sciences, 16*, 417–451.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127–190.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.

Smolensky, P. (1988). On the proper treatment of connectionism. *The Behavioral and Brain Sciences, 11*, 1–23.

Sperber, D., Wilson, D. (1986/1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247.

Strack, F., Deutsch, R., & Krieglmeyer, R. (2009). The two horses of behavior: Reflection and impulse. In E. Morsella, J. A. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 104–117). New York: Oxford University Press.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33*, 285–318.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*(6), 657–671.

Vissers, C. T. W. M., Chwilla, D. J., & Kolk, H. H. J. (2006). Monitoring in language perception: The effect misspellings of words in highly constrained sentences. *Brain Research, 1106*, 150–163.

Vissers, C. T. W. M., Chwilla, D. J., & Kolk, H. H. J. (2007). The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times. *Biological Psychology, 75*(1), 8–18.

Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 230–260). New York: Palgrave.

The Mechanism of the Form-Content Correlation Process in the Paradigm of Socio-Natural Sciences

Dorota Zielińska

Abstract In this chapter, I search for the mechanism correlating linguistic form with content in order to explain (in the sense of the word 'explain' used in empirical, i.e., natural and modern social sciences) how sentence meaning contributes to the utterance meaning. I do that against the background of two currently dominating positions on that issue: *minimalism* and *contextualism*. Minimalists regard language as a self-standing abstract system and claim that only weak pragmatic effects are involved in interpreting sentences. Contextualists believe that language can be described adequately only within a theory of language understanding and that strong pragmatic effects are also involved in interpreting sentences. The resultant controversy, presented in Sect. 1, has been pronounced by Michel Seymour the most important one in the 20th century. I begin Sect. 2 with Mario Bunge's argument that since abstract systems cannot change by themselves and only speakers of language do, an explanatory theory of language (one looking at language from the perspective of empirical (socionatural) sciences) must concern language understanding, i.e., view language as a bio-psycho-social phenomenon. However, language understanding needs to be incorporated in the theory of language in a more fundamental way than current contextualist models do.

D. Zielińska (✉)
Kraków, Poland
e-mail: dorotazielinska@yahoo.co.uk

These models assume the existence of language as self-standing, abstract structure with a list of symbol-reference pairings (Such assumption is legitimate as long as one regards such an abstraction as only a methodological device.) and model language understanding disregarding its psycho-social development process. Such assumptions, however, lead to a number of insurmountable problems. I conclude Sect. 2 by arguing that to solve these problems, as well as to be consistent with the evidence attesting to the fact that language self-organizes and self-regulates, (also reviewed in this section,), we need a model of language understanding and production to be coined within a developmental bio-psycho-social perspective. In Sect. 3, I propose a specific model of the form-meaning correlation process, based on a novel mechanism of a linguistic categorization, which is compatible with a bio-psycho-social developmental perspective advocated in Sect. 2. On this view, the utterance meaning is dependent both on the approximate conventional meaning of the construction components conveying it, and on the specific social function of the whole construction (a relevant *pragmeme*), which identifies feasible situation specific contents. The given construct selects one out of these options. I finish the chapter, Sect. 4, by preliminarily testing the mechanism of the form-content correlation process introduced in Sect. 3 both qualitatively and quantitatively to meet the methodological standards of empirical sciences.

> *Languages thrive in the hospitable environment of human*
> *brains and communities*
>
> Hurfort

The secret of natural language seems to be hidden in the way that conventional meaning contributes to utterance meaning. No wonder why Seymour (2010) said that the debate between contextualists and minimalists (who represent the two dominating opposing views on the issue) is the most important controversy to arise in the analytic philosophy of language in the past 30 years. This chapter is a contribution to the debate, offered from the perspective of the paradigm of empirical sciences. During this discussion, the author questions some of the philosophical assumptions taken for granted both by contextualists and minimalists and presents the issue at stake in a very new light.

The chapter starts by defining the controversy in its original form. Next, the author argues that, looking from the perspective of empirical (socio-natural) sciences (in which researchers aim at creating explanatory models implied by the functioning of material systems), language cannot be viewed as abstract, isolated structure as proposed by minimalists. Instead, a model of language must essentially involve accounting for understanding language in a social context, as postulated by contextualists. Current contextualist models, however, also fail to explain satisfactorily major linguistic phenomena such as the compositionality of meaning. The author goes on to claim that the inadequacy just mentioned is due to the fact that contextualist models currently proposed incorporate language understanding process in language interpretation models too late—as some sort of

adages to the view of language as an abstract system. Just like minimalists, contextualists start with the description of *langue*. They both treat language structure and content as independent from each other, and, most importantly, independent from the actual process of language creation, development, and understanding. This, I will argue, is the main reason of the limited adequacy of current contextualist approaches to modelling linguistic phenomena.

I claim that language, and in particular its structure, should not be viewed independently from its usage and people using it. It should be better viewed as a result of the simultaneous form and content creation and form-and-content correlation process taking place in human brains during social interaction between linguistic community members. This process is driven by (bio)psycho-social mechanisms. It is further argued that language creation, in the sense of there being statistically strongly dominant syntactic patterns and encodings in a linguistic corpus of *parole* produced by a given community, takes place through self-organization due to functional factors (like in biology) and not due to causal laws.[1] Consequently, language modelling in the empirical paradigm must imply a search for statistical patterns concerning the statistical characteristics of a corpus of *parole* and for mechanisms accounting for these patterns I conclude the discussion by proposing a specific developmental mechanism of linguistic form-and-content correlation process and indicate how to test it in a way that meets the standards of empirical (socio-natural) sciences.

1 The Contribution of the 'Minimalist Versus Contextualist' Debate to Understanding the Relation Between Conventional and Utterance Meanings

The discussion concerning the relation between conventional and utterance meanings carried on between minimalists and contexualists has brought us significantly closer to understanding that issue—the Holy Grail of linguistics. Therefore, I present my own views on the issue against the background of the debate just mentioned.

1.1 Defining the Controversy

The controversy between contextualists and minimalists requires taking a stand on what words in modern languages mean and what relation there is between word meaning and its contributed sense.[2] Once it was believed that the meaning of an

---

[1] after Altmann (1978).

[2] of which we know very little, as recently acknowledged by Recanati (2011).

utterance was essentially a token of the meaning of the uttered sentence. Nowadays, practically everybody agrees that even the conventional meaning of a sentence devoid of indexical expressions, does not determine the meaning of an utterance it helps to convey. As Seymour (2010: 2673) puts it "the fundamental question that remains is whether conventional meaning offers at least a kernel onto which the far richer intended meanings of the speaker are grafted, or if [strong— DZ] pragmatic features intrude in the very determination of what is literally expressed".

Pragmatic contribution to an utterance meaning has been divided into weak and strong pragmatic effects. We talk about weak pragmatic effects when the conventional meaning of a sentence, even one devoid of indexical expressions, contains an "algorithm" to extract additional information from the context. Such effects are acknowledged both by contextualists and minimalists. For instance, the minimalist Stanley (2007) proposes that the relevant algorithm is due to some hidden variables, which he terms 'indexes'. Stanley illustrates the relevant mechanism by offering the following examples. The sentence *It is raining* is taken to mean "it is raining here and now," because it contains hidden variables designating a place and time, which do not have specific values assigned to them before the sentence is uttered on a specific occasion. These variables (indexes) are assigned values defined by the time–space location of the relevant speech act. On the same account, the adjective *big* in the sentence *A four-year-old Jessica made a big snowman*, has a hidden index in it, which, when specified, turns 'big' into 'big for a snowman made by a four-year-old girl.'

While Stanley's proposal definitely brings us closer to describing the utterance meaning, the problem with an explanatory value of such an approach is that the number of indexes which need to be postulated cannot be limited if language is to express all possible situations ever to be encountered and therefore such parameters cannot be contained in a finite abstract system. On another occasion, for instance, the same adjective *big* may need to convey 'big for a snowman made by a four-year-old autistic girl with exceptional manual talent on a day when snow was not wet enough to make good snowmen and when she did not have a good day, either.'

Even the so-called descriptive adjectives,[3] as it is reminded below, often require similar filling in. Typically, when taking into account supposedly purely conventional context of a descriptive adjective, the number of parameters which would need to be postulated to interpret it, exceeds what the individual items intuitively encodes. Compare the meanings of the adjective *red* in the phrase *red eyes* used on the following two occasions: first by a photographer concerned with a red reflex in the photography—a portrait of a child, and second, as understood by a mother looking at her sick child with reddish whites of his eyes indicating the intensity of sickness. Thus, when interpreting the item *red* not only the shade of

---

[3] Consequently Recanati (2011) rightly questions the sense of the very division into descriptive adjectives and adjectives requiring "filling in".

the respective "rednesses" needs to be established, but so does the area it assesses. These values, in turn, both depend not only on the noun to which adjective *red* ascribes the value of colour, but also on the given situation, whose function defines possible meanings. And since the number of potential situations (defined by their functions), in which a given construct can be used depends on language use and not on the semiotic system itself, all potential meanings of a given linguistic item cannot be predefined by the given semiotic system, in particular by any set of indexes.

Recanati (2011),[4] a contextualist, proposes, in turn, that the change of conventional meaning such as the one just described is a psychologically instantiated change of conventional meaning due to context. Recanati introduces the terms 'modulation' to cover the concept just illustrated and 'syntactic flexibility' to refer to a parallel discrepancy in interpreting syntax. To illustrate that latter concept, I shall refer the reader to my analysis of the sentence "The visions of apocalypse have every right to scare us" in Zielinska (2007) and its logical interpretation stating that "we have the right to be scared by the visions of apocalypse",[5] whose syntax seems to have no conventional algorithmic relation with the original sentence. Syntactic flexibility introduced by Recanati (2011) also includes the instances of non-syntactic communication studied by Stainton (2005, 2006a, b). In any of the situations just mentioned, it is hard to think of any systematic way of grafting the syntactic structure of the sentence actually used onto the structure of the sentence expressing the content literally meant. In other words, occasionally, linguistic compositionality seems to break down without causing problems in communication.

Considering such examples as the ones mentioned above (exemplifying the issue of apparently regular compositionality), or well described cases of the novelty conveyed by compound constructions (c.f., deriving the concept of 'e-money' from the lexemes 'electronic' and 'money'), contextualists concluded that the influence of contextual information is not merely a matter of weak pragmatic effects such as saturation, (assigning contextual values to parameters determined by encoding). Instead, when words are used in expressions and sentences, their standing meanings are affected (modulated) in a fundamentally strongly context dependent way. (Therefore, as pinpointed by Recanati (2011), the controversy between contextualists and minimalists can be rephrased as providing the answer to the question whether (1) "both strong and weak", or (2) "only weak" pragmatic effects determine what is said by a sentence).

The concepts of modulation and syntactic flexibility, however, merely label, but do not explain the source of divergence between actual and encoded meanings. Such divergences have not been accounted for so far as explicitly admitted e.g., by

---

[4] Other researchers expressing similar views include: Capone (2005, 2006 and later), Carston (2002), Levinson (2000), Mey (2001, Recanati (2004, 2011), Searle (1983), Sperber and Wilson (1986), Stainton (2006) and Travis (2001).

[5] The sentence was overheard on a TV show.

Carston (2002). Carston (ibidem), explicitly gives up on accounting for the relation between what people know about word meanings and what they "literally" say, (which she illustrates analyzing selected meanings of the item *open*) the with the following words:

> The question which won't receive any answer here, is how the more general schema or indicator arises and how it comes to be the meaning of the lexical expression type. There must be some process of abstraction, or extraction, from the particular concepts associated with the phonological form /open/ to the more general 'meaning', which then functions as a gateway both to the existing concepts of opening and to the materials needed to make new 'open' concepts which may arise in the understanding of subsequent utterances. Carston (2002: 364).

1.2 The Source of the Controversy Between Minimalists and Contextualists

Saying that standing meaning is affected by context in the way unpredictable by the given standing meaning itself is equivalent to stating that something beyond the system of syntax and standing meanings influences utterance meaning (i.e., the input of the context goes beyond providing some values to encoded parameters of standing meaning). The only related entities that could be the source of unpredictability in the use of language system (itself) are language users functioning in a linguistic community. Therefore, the controversy whether only weak or both weak and strong pragmatic effects are relevant for establishing utterance meaning is really implied by the answer to the question whether we can explain language in an adequate way treating it as a set of qualitatively defined abstract sentence structures with independently defined sign-representation, lists thus disregarding language users and their lives. Minimalists assume 'we can', contextualists reply 'we cannot', and postulate that accounting for semantic non-combinatorial novelty requires modelling human psychological processing of language in context, i.e., it requires including the process of gaining knowledge in the very model of language.[6] As Dummett (1993: 12) states it "a theory of meaning must also be a theory of understanding" [cf. Searl (1983)].

The same conclusion must be reached when attempting to construct an explanatory model of language from the perspective of socio-natural sciences. Since abstract systems cannot change by themselves, to construct an explanatory model of language in this paradigm, language cannot be viewed as an abstract system isolated from its users and use. As Bunge explains:

---

[6] The debate between contextualists and minimalists covers a very similar terrain that the divide between Cartesian and non-Cartesian linguistics does as proposed in a research of Kopytko (1995, 1998, 2001a, b, 2004).

> Languages do not develop or evolve by themselves and there are no mechanisms of linguistic changes, in particular evolutionary forces. Only concrete things, such as people can develop and evolve. And, of course, as they develop or evolve, they modify, introduce, jettison linguistic expressions. The history of mathematics is parallel: mathematicians do come up with new mathematical ideas, which are adopted or rejected by the mathematical community, but mathematics does not evolve by itself Bunge (2003: 62).

From the perspective of empirical sciences, only the behaviour of material things can be explained (in the sense of providing its cause). Abstract systems alone are not capable of explaining anything in the sense of providing its cause, but can only offer summation rules—rules summarising typical experience (subsuming under generalization). Thus, from the perspective of empirical (socio-natural) sciences, the original controversy has an unambiguous answer. There cannot be an explanatory theory of a semiotic system.

1.3 Why Contextualists have Failed to Account Adequately for Strong Pragmatic Effects?

If we are convinced that contextualists are right in believing that modelling language requires modelling language understanding processes, why have they not found the mechanism of strong pragmatic effects yet? As already mentioned the reason is that they do not model language understanding deeply enough, i.e., they do not consider the understanding process involved in and leading to language creation and development. Instead, the dominant approach to language description, both among minimalists and contextualists, starts with describing *langue* in terms of patterns of symbols and lists of symbol-representation pairings, with disregard for epistemic concerns, i.e., answering the question how come we can get to understand the meaning of lexicon and of syntactically combined lexical items. Contextualists consider language understanding only at the stage of interpreting *langue as an abstract semiotic system* in a specific context, which, I shall argue below, is too late.

2 In Search of an Alternative Paradigm

In the last half of the century, independent research in linguistics, philosophy, neurology and systems theory provided us with a new way of perceiving language in comparison to the reductionist approaches explored in the 19th and 20th centuries. Below, I shall outline a new paradigm consistent with the state of the art in philosophy and sciences, and argue why it is more appropriate for modelling language.

2.1 Insufficiency of Defining Language as Sets of Patters of Symbols

A great initial success of approaches to describing language using qualitative mathematics came from the ingenuity of introducing recursive[7] rules of combining symbols. Presenting a model of language based on recursive rules operating on uninterpreted (mental) symbols allowed linguists to account for linguistic creativeness understood as combinatorial novelty, (i.e., new combinations of non-terminal, and eventually, of terminal symbols), which, at zero approximation, can be considered to be the essence of language. Consequently, first Artificial Intelligence (AI) specialists believed, as earlier Plato did, that formal logic and ideal form is the essence of thinking, and in particular—of language. This impressively fruitful step indeed allowed one to create a number of interesting linguistic models. Among others, it also guided early attempts of AI specialists to build a machine that would imitate brain functioning, including language. Yet, the results of such projects fell short of their expectations. In particular qualitative, classificatory formalisms that are used to flesh out such linguistic models cannot account for the change of a semantic category, or define the limits on category membership acceptance and thus account for novel categorization of a given exemplar; or account for non-combinatorial novelty, either syntactic or semantic. (What qualitative formalisms are good at, is writing down in an elegant fashion the regularities observed, thus describing neatly typical, common, core aspects of language).

The belief of formal linguists (shared by the main stream of AI specialists) that language can be adequately modelled as a set of patterns combining symbols and independent representation-symbol pairings is based on Putman's (1960) *Minds and Machine* article. In this article Putman concludes that since a Turing Machine can model any machine, it follows that reasoning is independent from the physical make-up of the machine that carried it. Thus, the same cognitive processes, such as thinking and reasoning, can be obtained by various physical set-ups. Therefore, mental states are functions of physical set-ups and the identification of the mental state is independent from the actual physical characteristics of the set-up used to carry them out.

Yet, after considering Goedel's theorem about the incompleteness of axiomatic systems, Putnam (1988) reconsidered his earlier views. Goedel's theorem shows that it is not possible to define every state there is with a limited number of operations specified by a given axiomatic system in a unique way. Therefore, it is impossible to prove that the same functional state of the brain was arrived at by the same procedure that the state of the machine was. Consequently Putman (1988) rejects the position that functional states (e.g., thinking) correspond to the states of

---

[7] Recursive rules were probably first proposed to capture linguistic compositionality in Ajdukiewicz's (1935) categorical grammar. The idea of recursiveness has been popularized in linguistics, however, through Chomsky's generative rules, starting with Chomsky (1957).

any physical machine that arrived at them, i.e., ontological functionalism, for the sake of epistemic interpretation of functionalism. Putnam (1988) states that computers working with different inputs and considering different analogies may have different descriptions of that data than those in a Turing machine formulae and may operate according to different rules therefore mental states cannot be equated with the states of the machine. Putnam concludes that mental states are plastic not only as far as physical machine processing them is concerned but also in terms of calculation. Consequently, the functional cognitive state of the human mind cannot be captured with a set of symbol manipulation rules.

Putnam's theses that mental states cannot be equated with the final states of the machine lets one conclude that hardware differences (differences in physical make-up) imply software differences of the physical set-up (differences in propositions arrived at by the physical set-up) and therefore mental states are constituted by the net of elements co-implying each other, not merely by isolated structures. Putnam (1988) notices, however, that equating a mental (brain) state with the state of a machine would be legitimate if that machine had the make-up of the brain.

2.2 What Type of Mechanism Can Support Language?

If we assume that mental (brain) state cannot be equated with the final state of the machine, unless the machine at stake has the make-up of the brain, then the knowledge of the development and functioning of the brain must inform successful attempts at language modelling.

As pointed out by Brook (1986) human brain with linguistic capabilities is the result of long evolution. Single cell forms of life arose about 3.5 billion years ago, photosynthetic plants a billion years later, fish about 550 million years ago. Next, at intervals of tens of millions years ago came about insects, reptiles and dinosaurs and mammals. Man appeared on Earth 2.5 million years ago and he invented agriculture less than 20,000 years ago. The oldest evidence that man could write comes from about 5,000 years ago and the expert knowledge, when symbolic thinking starts to be needed, has been accumulated for only a couple of hundreds years. If we assume that brain functioning depends on knowledge, how it evolved and how it is organized, it needs to be based on modelling skills once needed for survival such as hearing, moving around in space, interpreting signs indicating location in time. Brook (1986) concludes that "Such skills, in today's brain often unconscious, laid foundations for logical reasoning and understanding." In other words, the functioning of the brain depends on the history of its development, in particular, on specific type of survival challenges faced. As a result, as we well know, the brain is not merely reductionist in its structure and functioning. It consists of units of organization with considerable independence each, (i.e., each of them interacting with different environment) on the one hand, but on the other hand, with non-negligible interdependence between units of adjacent levels, which goes both ways up and down. Note, for instance, that if we cut out part of a tissue,

it will stop functioning—it will loose its vital characteristics. In other words, the brain is a system in the sense of empirical sciences (c.f., Bunge 2003)—its every level is built out of units of the lower level, each higher level is characterized by fulfilling a new function, which both depends on and conditions the characteristics of the units of the lower level.

Importantly, one part of the brain system is hardwired,[8] (as a result of the evolution that has been taking place over thousands of years), another, plastic part of the brain—that hosting language and many other cognitive skills—is a currently developing system. Such hypothesis has been corroborated by the results of recent neurological studies (cf. Karmiloff-Smith 1995; Karmiloff and Karmiloff-Smith 2010; Elman et al. 1997; Gopnik 2007, 2009) who have concluded that children are not merely unfinished adults, but are designed by evolution to change, create, learn and explore. We (and other altricial species) have much longer childhoods than precocial species. As a result of that long childhood involving a considerable amount of learning, we may differ significantly from members of the previous generation: improve our survival, reproduction and care-giving skills over our predecessors.

Karmiloff and Karmiloff-Smith (2010) observes that at the beginning, the brains of human children develop far more connections between neurons than needed, none of them particularly efficient, which specialize in time. She goes on to say that

> Infants are not born with pre-specified modules. Indeed, the infant cognitive system is less differentiated and thus less modular than the adult system, suggesting that modularity is an emergent property of the developmental process. So, domain specificity is not a built-in property of the brain but emerges over developmental time. And even if a modular organisation of the adult brain is the emergent outcome of development, even adult modules should not be viewed in terms of the rigid, static notion of a Fodorian module as outlined above. Thus, instead of the notion that a given brain module can only process proprietary inputs from a specialised domain, neuro-constructivism argues that the brain *becomes* very gradually more specialised over developmental time whereby it narrows its response to the types of inputs a given brain circuit may process, after initially processing many different types of inputs. This is also a relative rather than rigid concept. Indeed, brain circuits that have become relatively domain-specific may still attempt to process new inputs from other domains.

Karmiloff and Karmiloff-Smith (2010) concludes that "therefore a predominant amount of learning and language acquisition goes on in the plastic area of the brain. And the brain keeps developing—organizing in response to the challenges met."

Another strong argument against the thesis that language skills nested in the brain are fully deterministic, i.e., hardwired due to genes—is the evolution of Nicaraguan sign language, which took place within less than a decade. The location of linguistic skills in plastic areas of the brain is also supported by the fact that children with brain damage to the Wernica area of the brain learn to speak

---

[8] approximately speaking.

using the opposite hemisphere. The research into the results of brain damage also shows that among bilingual patients with brain damage who acquired a second language after adolescence, there are such ones who suffer from brain damage causing loss of only one of the languages. No such patients have been found among bilinguals who acquired a second language in early childhood. Importantly, recent developments in the theory of evolution place main emphasis on its developmental aspect. Griffiths and Gray (1994) stress that even genes to operate consistently need resources of the same type at the right moment of the organism development.

The evidence such as mentioned above indicates that the part of the brain supporting cognitive skills and, in particular, language is much less genetically pre-programmed than assumed by formal linguists, advocates of Universal Grammar.[9] It seems that linguistic skills are supported in an important way by the mechanisms located in the plastic area of the brain, as opposed to the hard-wired parts of the brain where the language mechanism instantiating a Turing machine manipulating meaningless symbols representing recursive rules of language would need to be exclusively located. The resultant structures of the brain constitute a system, and are part of a larger system, which, if not genetically determined, must emerge in a natural way, i.e., through self-organization.

2.3 Emergence in Material Systems

Can we explain the functioning of emergent systems? As already mentioned, changes, (e.g., emergence) in concrete things (systems), unlike in abstract systems, can be potentially modelled (thus explained in the sense it is done in natural and modern social sciences). At this time, I would also like to clear a common misconception that emergence is tantamount with no explanation for a given phenomena. This is not true. The phenomenon of emergence and our ability to account for it are two separate issues. It is true that sometimes we do not know how to explain an emergent phenomena, but sometimes we do. For instance the change of characteristics of liquids after their transition from liquids to solids when being frozen can be accounted for in terms of quantum chemistry [for more, see Bunge (2003)].

As pointed out by Bunge (2003), the source of non-combinatorial, emergent novelty in material systems—the change of characteristics of a group of elements which combine together to form a higher level unit requires input from the environment—outside the system itself, (c.f., the input of energy to defrost ice, sucking out heat to freeze water.) A new level will interact with a new subset of the total environment, and the emergent properties of that new level will let it perform a new function. By definition, a new level in a system is characterized by

---

[9] The theory of Universal Grammar requires the existence of an organ in the brain hosting UG that had evolved before language did. This, however, is evolutionary implausible.

properties absent from its parts fulfilling a new function. These new properties account for the interaction of the new level with different environments. Therefore, a compound unit of the system interacts with the environment with which its parts do not).

As explained by Bunge, the combination of parts leading to the creation of a new level can emerge in a system either through artificial assembly, or self-assembly. Car making illustrates artificial assembly. (A car has a novel function of transportation, which is missing in any of its parts). Examples of a self-assembly process resulting in novelty are vapour freezing or the coalescence of a street gang. A typical consequence of emergent novelty, especially in self-assembling systems, is mutual inter-level interaction. (More precisely, this is additional downward interaction of a higher level unit with new properties (and thus interacting with the new environment), onto lower level units and their old, "lower-level" properties. For instance, gases are characterized by temperature and pressure, none of which characterizes the constituting atoms. Yet, under sufficient pressure applied from outside, or if temperature is lowered sufficiently due to external energy sucking out heat, vapour will combine into ice, in the process changing the characteristics of the very atoms initially forming it—the lower level elements.[10])

Thus, understanding a system (in the sense of the term in natural sciences) means knowing its structure, composition (elements), the environment with which its elements interact, and the mechanism supporting its function in that environment. The interest in mechanisms explaining self-assembling (or self–organizing systems) has recently picked up. It has been studied, among others, by a newly developed field of synergetics, a branch of system theory, which models self-organization through HOT (highly optimized tolerance) mechanism (see Haken 2010). Another mechanism of self-organization, called self-organizing criticality (SOC), has been proposed by Bak et al. (1988). Both types of models allow one to form testable hypothesis and corroborate the self-organizing origin of the objects under study.

2.4 Has Language Got Characteristics of a (self-organizing) System in the Sense Just Defined?

Based on the assumption that functioning of software depends on hardware, as proved in Puntam (1988), and in view of undeniably systemic (and self-organized) nature of the brain, language can be expected to a have a system structure with considerably independent levels interacting with different environments, and with higher and lower levels influencing each other both ways, Has it really?

---

[10] The artificial and natural assemblies can also combine. The process of book production starting form farming trees can serve as an example.

Language is based on meaningless phonemes, some of which combine into morphemes and words with the emergent property of meaning. Meaningfulness and the frequency of use of specific words, in turn, along with biological capabilities of human beings, influence (interacting down) the shape of phonemes which need to be contrastive enough and easy to pronounce when one follows another. Next, some lexemes may combine together into phrases and instead of being related to a separate referent each, both may refer to the same referent and enter a new relation with each other, e.g., being a modifier and modified. A group of lexemes may also convey an emergent sense which is not a result of a simple addition of the composite lexemes' senses, i.e., the lexemes as a unit may interact within (refer to) a different environment (referent) than any of the lexemes composing a given unit. Consider, for instance, the relation between the meaning of the items *post*, the meaning of the item *card* and that of the compound *post card*. The meaning of the compound *post card* has a representation which cannot be calculated from the representations of the meanings of its components, but involves non-combinatorial semantic novelty based on information from beyond the system—the world of post offices. We may also say that the function of the phrase *post card* influences (modulates, interacts down onto) the meaning of the lexeme *post* as well as that of the lexeme *card* as used in the phrase *post card*.

Lexemes and phrases may combine into units capable of communicating something about something else, i.e., conveying propositions with the emergent property of truth value, i.e., being or not being true. Finally, a sentence used in (interacting with) a given situational or textual context acquires an emergent property of having an illocutionary force. For instance, the sentence *could you open the window*?, used in the situation when the speaker clearly would like to have the window opened, but for some reason does not want to do it himself, constitutes a request for opening the window. Because of the frequency of using similar questions in contexts with the differentiation frame indicating the desire that the addressee fulfils the action expressed by the respective verb, "could + (someone + do something)", the construction becomes correlated with a request to do the action expressed by the predicate directed to the listener, in addition to what this construction was correlated with so far, i.e., the question concerning the capabilities. Additionally, a specific sentence e.g., *Could you visit Jack at the hospital tomorrow?* uttered in a specific situation, adjust the referential content of the respective components. For instance this sentence assigns to the lexeme *visit* a specific meaning of a prototypical visit in the hospital, which differs from visiting healthy people at their own homes.

Next, sentences can be further organized into dialogs, or paragraphs, sections, chapters and texts, which again exhibit a collective purpose absent from individual sentences, which again may adjust their respective messages. Note, that this hierarchical organization is certainly an emergent phenomenon. Proto-languages must have had the form of single 'word-functional sentence'[11] correlations. With

---

[11] I use the term 'functional sentence' to refer to a sentence with a specific illocutionary force.

the growing complexity of technological and social life, the number of single 'word–functional sentences' grew to the point, that such a language became difficult to acquire and use. As pointed out by Kwapień (2010), simulations of the use of such proto-languages show that as the number of words reaches a certain threshold value, the number of mistakes in language communication grows to the point of considerably lowering its efficiency. The emergent grammatical organization allows a linguistic community to lower the number of words needed, while increasing the communicative efficiency of the language they speak.

In the above view, language is a system in the sense in which the term *system* is used by Bunge (2003). The examples just presented illustrate both relative independence of linguistic elements on each level, as well as the fact that they are influenced by higher levels. Importantly, the emergence of levels with emergent properties in language due to the interaction of higher levels with new environment and in a novel way (functional self-regulation) provides a space, in which to look for semantic novelty. There is no room for such novelty on a Turing machine view of language (i.e., language being a closed, and purely reductionist system). No wonder that although modulation and syntactic flexibility are widely observed, they remain unaccounted for in the latter framework. On the contrary, the emergentist view of language is, by definition, predisposed to model linguistic change over time (language development), as well as (in particular) the adjustment of linguistic meaning during its instantiation ('language efficiency', to use Berwise and Perry's (1983) term).

A strong argument for language being a self-organizing system comes from considering its quantitative characteristics. Recently the hypothesis that the source of self-organization in language comes from the mechanism resembling natural selection, during which process certain types of constructions and lexical meanings are selected for on economic bases has been considered again. (The hypothesis that language has been brought about by some sort of economy was earlier advocated by Zipf (1935), but later criticized on the grounds that the amount of calculation needed would exceed the capabilities of the brain[12]). Next, Zipf (1949) showed that there are a number of relations between certain quantitative characteristics of linguistic corpuses which follow power laws. Today we know [cf. Haken (2010), Bak (1988, 1996)] that Zipfian laws express[13] the characteristics of self-organizing complex structures, which cannot be derived from "first principles". The degree of conformity of Zipfian laws in a number of languages leaves no doubt that the assumption of language being a self- organizing structure makes a lot of sense. (By the way, since Zipfian laws concern self-organizing structures as such, such laws have also been attested in a number of other non-linguistic phenomena, such as the

---

[12] This argument is no longer valid if we transfer the burden of making economic choices from an individual to a natural selection process taking place in a community.

[13] Strictly speaking, Zipfian distribution could also be the result of a relatively simple, statistical processes. Yet, along with the information about the hierarchical structure of the object of study and the processes involved, establishing power laws relating some characteristics of that object is enough to indicate its self-organizing origin.

ranking of cities by size, income ranking, corporation sizes, the revenue of a company as a function of its rank, the distribution of the earthquakes as a function of the rank of their magnitude, and many more). Also, the association studies modelled by Bruza (2009, 2010) with Quantum Mechanical formalisms, which formalisms reflect some self-organizing principles, corroborate the hypothesis about the self-organizing nature of language. Interestingly, according to Bruza (ibid.), the shapes of many of distributions tested by synergetic linguists turned out to be describable even more adequately with the help of certain quantum mechanical formalisms.

2.5 Problems with Grounding Basic Encodings in the Reductionist (Saussurean) Systems

An important argument for the need to view language as a system (as it is understood in empirical sciences), comes also from considering basic encodings. Basic encodings are the atomic elements of meaning in the model of language based on the idea that language is a system in Saussure's (reductionist) under-standing of the term. Saussure refined Franz Bopp's imprecise idea of a system, and to him a system is a context-free and downward-interaction-free (purely reductionist) non-developing abstract structure. To define language adequately in such terms (as it is done in formal linguistics), such systems require that two assumptions be met. The first assumption is that the human brain can be well modelled by a Turing machine. The second assumption is that we all have access to some common meanings of basic encodings. Yet, Putnam challenged both of these assumptions. In addition to challenging the former thesis, which he did in Putnam (1988) as already mentioned, Putnam (1975) proved there is no way of explaining the representation of basic encoding meaning from person to person, no way of conveying the meaning of basic encodings.

A positive solution to the latter issue, in line with the concept of a system in empirical sciences, comes from Bickhard and Campbell (1992a, b). Bickhard and Campbell (ibid.) propose that linguistic representation be expressed in terms of a control system, i.e., what can be shared among individuals, are functionally established categories. Therefore, at least some of the original linguistic categories must have been formed prior to, or simultaneously to the formation of individual linguistic representations.

There is no doubt that such functional non-representational categorization indeed takes place during the evolution of organisms. Functional categorization, which would be quite sophisticated cognitively, is even possible in creatures without respective cognitive structures. Ants, for instance, have been long known to be capable of finding the shortest distance between their dwelling and a source of food, and obviously no one would credit them with having a representation of

the concept of the shortest route.[14] Next, such externally (functionally) defined categories may allow human community members to order their individual representations relative to each other to parallel the relations between the functionally established categories and thus ground the basic encodings. As a result, separate individuals will not necessarily share absolutely identical representations, but these somewhat idiosyncratic representations will let them refer to the same functional categories.

The hypothesis of functional substantiation of representation has been corroborated by neural evidence coming from Lin and Tsien's research. Lin et al. (2005, 2006) identified neural functional cliques underlying representational neuronal structures. Lin (2007) in turn, reports the discovery of neuronal cliques in mice responsible for mice's identifying a potential nesting object. These cliques are not activated when a real nest is covered with a transparent piece of plastic, but are activated when a mouse comes across a red, plastic cube with a sufficient indentation in its top to serve as a nest. Thus, depriving an object of its capacity to fulfil its typical function results in mice's inability to recognize its physical representation, while on the other hand, presenting a mouse with a highly a-typical object, yet one with the capacity to fulfil a given function, makes the mouse classify the given object as representing the given functional category despite its lack of appropriate physical characteristics (physical representation).

Not-surprisingly, functional interactions precede formation of basic representation both onto- and philo-genetically. In the last decade there has been a growth of interest in the way children acquire functional (pragmatic) linguistic skills along with, or even to some extent prior to acquiring semantic and syntactic ones. The point may be illustrated with the following example. A toddler has been reported to use the expression. 'This is …' as a request to have the object pointed to opened. He did it evidently without being aware of the semantic content of the words. His mother acknowledged that before opening something for the child, she used to explain what the object was using the structure *this is* "*x*". The child used the phrase *this is…* to functionally categorize the states of containers into "closed" and "open" ones, and next he correlated the phrase mentioned with the change between these states. The functional understanding of the phrase clearly preceded in this toddler the understanding of the sentential meaning and the meanings of components. (The same toddler used the phrase *Once upon a time* as a request to be read to, again apparently without being aware of the semantic content of the words).

Basic differentiation frames that initially subcategorize experience, using Campbell and Bickhard's terminology, allow one to ground basic encodings—attach labels. This position corresponds to Mey's (2001) role of *pragmemes*. Mey postulates the existence of *pragmemes*—situations motivated by social functions—

---

[14] Ants when walking, leave scent on the trail. The group of ants in search of food that has found the shortest route will cover the distance between their anthill and that source of food the largest number of times, making that trail most smelly. A new group of ants which has just left the anthill to search for food, will chose the most smelly path, which is the shortest one.

to argue that language interpretation takes place as if "outside- in" or downward, we might also say. For instance, one needs to know the relevant *pragmeme* (differentiation frame) to interpret e.g., *red eyes*, as already illustrated. In other words, knowing the functional subcategories of the relevant *pragmeme*, is prerequisite for assigning the semantic representational meaning to its parts. Capone (2005, 2006) shows us that the recognition of a *pragmeme* is necessary for the very concept of *explicature* to make sense, thus the concept of *pragmeme* cannot be excluded from forming a coherent model of linguistic encodings. [Capone (2010) further proves the utility of *pragmemes* when analyzing reported speech].

Initially, classes of pragmemic options which differed functionally (e.g., "chase it" vs. "stop chasing it") must have been differentiated by single signs. These categories, in turn, could have been used to subcategorize options defining other *pragmemes*, initially also marked with single signs, resulting in simultaneous development of a system and the temporary specification of individual meanings of elements in those *pragmemes*. As a result, the utterance subparts could assume approximate individual, temporarily encoded meanings, which are used next by these individuals to refer to functionally identifiably, thus commonly shared, *pragmemes*. These *pragmemes*, in turn, are used to specify the utterance meaning of the individually encoded subparts of the symbolic structure used on a specific occasion. At the current stage of language development, individually represented "conventional"[15] word meaning may serve to help identify *pragmemes*, which reflect back on the meaning of the constituents of the linguistic construct, which had identified the *pragmeme*. This is done by substituting the respective individualized "conventional" meanings of the components of the construct used to identify the *pragmeme* with the (utterance) reference meaning, which is determined with help from options identified by the function of the *pragmeme*. Thus, structure, representation, and function are inseparably interrelated to form language.

Note that the basic encodings identified through basic interactions will correspond not to semantic primitives assumed in formal semantics, i.e., the set of the most atomic meaning primitives needed to code all the lexicon, e.g., Katz componential analysis, but rather to the ground breaking idea of the universal lexicon posed by Wierzbicka (1972, 1985). Wierzbicka postulates the existence of "inborn *lingua mentalis*", an innate mini language of cognitive concepts, both lexicon and syntax, that allows a child to make a functional sense of a situation and speaker's intentions. The approach advocated, although denying the innateness part of Wierzbicka's hypothesis, shares with it its essence—the emphasis on functional origin of basic linguistic categories. The functional categorisation, along with general perceptual skills, and basic situations of usage results in the type of universal *ligua mentalis* identified by Wierzbicka. Importantly, however,

---

[15] I use the term "conventional" here in the sense of some sort of representation which is an average of the individual representations of the same functional category in a given linguistic society. .

Wierzbicka's empirical cross-linguistic investigations does not *per se* imply that the universal *lingua mentalis* is indeed inborn. It could just as well be derivative as postulated by the "outside-in" approaches, which is supported also by recent developments in the theory of evolution. The authors of the Developmental Systems and Evolutionary Explanation, Griffiths and Gray (1994) explain that there is no formal difference between inborn and developed.

> Developmental systems theory rejects the dichotomous approach to development: The genes are just one resource that is available to the developmental process. There is a fundamental symmetry between the role of the genes and that of the maternal cytoplasm, or of childhood exposure to language. The full range of developmental resources represents a complex system that is replicated in development. There is much to be said about the different roles of particular resources. But there is nothing that divides the resources into two fundamental kinds. The role of the genes is no more unique than the role of many other factors.

To sum up, fully subscribing to the resultant universal lexical units found by Wierzbicka, I see the proposition postulated in this chapter as expressing the gist of Wierzbicka's insight and intuition (initially formulated 40 years ago) in the language of the contemporary paradigm of empirical science. What also supports the functional focus of the proposition presented in this chapter is Wierzbicka's (2010) postulate of linguistic molecules, i.e., the lexemes containing more than one semantic prime combined together due to a specific *pragmeme* in which these primes originated. Wierzbicka explains the concept of a linguistic molecule with the following words:

> In addition to semantic primes ('atoms of meaning'), many NSM explications also rely (in a limited way) on 'semantic molecules', built from primes, especially in the area of concrete vocabulary. In particular, body part concepts often function as 'semantic molecules' in the meaning of verbs of physical activity, such as *walk* ('legs', 'feet'), *lick* ('tongue'), *bite* ('teeth'), and *eat* and *drink* ('mouth') ....[while -DZ] color words rely, to a considerable extent, on environmental and bodily molecules such as 'sky', 'sun', 'day', and 'blood', as well as on the molecule 'color'.

Grounding language in a material system categorized by its function (supported by some mechanism) and environment solves yet another cornerstone problem in linguistics—that of linguistic categorization. Aristotelian definition of the linguistic category, in addition to requiring everyone to share a common representation of a category, imposes limits on future applications of the symbol of a given category and precludes metaphoric usage. Introducing a prototype as the pattern defining a given category does not solve that problem. Therefore, Roch [in Lakoff (1987)] renounced her earlier claim that prototypes define category membership and stated that a model of a linguistic category must reflect the phenomenon of prototypes. The proposition that individual category members fit the category well, or not so well, as proposed by Lakoff (1987) and Langacker (1987), results in every item being at least a bad exemplar of any category, which is not good either. A proposition to model linguistic category with fuzzy sets must be rejected on the same grounds as the previous one. Therefore, these models of a linguistic category cannot model linguistic compositionality adequately, either. In Sect. 3, I propose a

category model within the empirical sciences paradigm advocated here, one coined as a developmental (emergentist, self-organizing) system, which does not suffer from the above difficulties.

2.6 How to Model Language in the Empirical (socio-natural) Paradigm?

So far we have concluded that natural language as a semiotic system is closely (integrally) related to evolving material systems (part of the plastic areas of the brain), and therefore language needs to be treated as a result of a process in which the evolution of form and representation are inseparable and determined by a mechanism dependent on the social function of language. In other words, we argued that language as a semiotic system is inseparable from its function mainly as a tool for communication. Therefore, we may repeat after Bunge (2003) that the form and content of language are the integral[16] parts of the history of specific purposeful interactions between linguistic community members involved in speech acts. During that process certain aspects of participant's brain state which was evoked as a response to bio-socio-environmental conditions, become correlated with the symbolic elements of language through changes to some plastic areas of the speaker's brain. So natural language as a semiotic system is a reflection of a dynamic system of individual speech acts produced by human agents located in and interacting with social situations. Grzybek (2006: 12) expressed that idea by saying: "Genesis and evolution of these systems must be attributed to repercussions of communication upon structure." In other words, an explanation of the existence, properties, and changes of linguistic, (more generally speaking, of a semiotic) system is not possible without treating it as an aspect of the (dynamic) interdependence between structure and function, or in Bunge's (2003) language, without understanding the mechanism supporting that function of the given semiotic system in a specific environment.

This is so because, as Bunge (ibid.) stresses, the changes in material system components involved in their combining to become a higher level unit **are always the result of some input from their environment—outside of the system.** Consequently, in an empirical paradigm, the material system (neural connections) supporting linguistic behaviour of an agent is determined not only by bio-psychological principles (as implied by Chomsky (1986) style view of language), but also by external, socially established principles. In other words, the perspective on language just advocated makes linguistics necessarily an inter-science that

---

[16] Bunge (2003) opposes an integral structure, such as the one present in a cardiovascular system, to a combinatory structure, such as that present in a car. While the latter one was put together from parts, the former one evolved by subsequent evolutionary steps and cannot be substituted fully by plastic elements. It is integrally related to all other elements in a human body and specific history of evolution. .

straddles biology, psychology, and equally importantly, social sciences (Bunge 2003: 63).

In view of the above, linguistic explanation is not likely to be possible by means of casual relations. Instead, Altmann (1978) proposes that a likely strategy is looking for a functional explanation similar to ones offered in biology. Therefore, as Altmann continues, it is plausible to assume that language is an aspect of a self-organizing and self-regulating system of members of a linguistic community engaged in communication promoting their co-existence—a special kind of a dynamic system with particular properties brought in line as a result of some sort of economy related to the communicative behaviour of its participants. Or as Grzybek (2006) puts it—"the economic result of communicative processes".[17]

In the empirical paradigm, the assumption of the self-organizing nature of language implies the fact that the laws modelling language must not have reference to specific objects but be statistical. Another reason for statistical nature of linguistic laws in the empirical paradigm is that since the formation of linguistic objects depends strongly on the history of contingencies and exact data is not available, only statistical hypothesis can be formed. Similarly, Bak (1996: 10) talks about life, (which is also characterized by the variability of its exemplars resulting from its self-organizational character): "A theory of life is likely to be a theory of a process, not a detailed account of utterly accidental details of that process such as the emergence of humans."

2.7 Epistemic Concerns

Note also that the perspective outlined so far is the result of considering the manner of gaining knowledge about the phenomenon studied (language), i.e., in establishing (linguistic) facts—(recall grounding basic encodings and inadequacy of interpreting the brain states as the final states of a Turing machine). We initially learn about the environment, which consists of physical and social components, via physical interaction. Such knowledge can be shared whenever functional interactions with the environment can be shared.

The previous approaches with explanatory ambitions, (such as Chomsky's initial proposal that the biological make-up of man generates sets of syntactically restricted uninterpreted strings of symbols along with a separate list of symbol-representation pairings) required taking God's eye view into linguistic meaning on the one hand, and disregarding the make-up of the machine using language (when equating its final states with mental states), on the other hand. Both of these assumptions have been proven to be false, as reviewed earlier.

---

[17] This view may remind of 19th century concepts in linguistics, but here, language is not viewed as an independent organism. The "organism" considered here is not the semiotic system *per se*, but a linguistic community with the semiotic system (language) being an aspect of its behavior.

The epistemic problem with models based on God's view of basic encodings in linguistics resembles the situation faced by physicists who proposed a cosmological model of the atom. It turned out that the parameters needed to test the model could not be in fact measured. (The measurement of the position of an electron with an already measured momentum on the orbit around an atom, which will typically be established by shooting a photon with a known velocity into the electron, will affect the momentum of that electron to a degree which cannot be considered negligible. Thus, the measurement of the position of an electron will deprive us of the knowledge of its momentum). The epistemic impasse in physics just mentioned resulted in a search for a totally new paradigm of the description of the micro-world—quantum mechanics—an approach guided by epistemic concerns of gaining knowledge, a theory built on new measurable concepts. The same seems to be taking place when proposing a developmental psycho-social model of language. Thus, it would be hard not to agree with the following words:

> Most reasonably, language lends itself to being viewed as a specific cultural sign system. Culture, in turn, offers itself to be interpreted in the framework of the evolutionary theory of cognition or of evolutionary cultural semiotics, respectively. Culture, thus, is defined as a cognitive and semiotic device for the adaptation of human being to nature. In this sense, culture is a continuation of nature on the one hand and simultaneously a reflection of nature, on the other—consequently, culture stands in isologic relation to nature and can be studied as such.
>
> Therefore, *langue* viewed as a cultural sign system cannot be seen as being ontologically different from nature because the nature we know can be observed only through culturally biased theories and perspectives. ….Thus, both culture and nature are cultural constructs co-determining each other Grzybek (2006:8).

3 Modelling Language in the Empirical Paradigm: A Dual Model of Linguistic Form and Content Correlation

Within the empirical (socio-natural) paradigm outlined in Bunge (2003), a model of language, in addition to specifying linguistic composition and structure, must describe the mechanism that (creates and) supports it. Relevant mechanisms must account for both the process of correlating symbols with representations (lexicon) (and constructs with representations) and for ordering symbols (account for syntax). In this section, I propose, first, a general mechanism of correlating form with representation, which in fact must reflect modelling the process of linguistic categorization. Second, I shall illustrate the essence of the mechanism proposed in this chapter by illustrating briefly "the soft way" of approaching a wide range of selected linguistic problems from the perspective advocated.

3.1 A Qualitative, Developmental Theory of the Form-Meaning Correlation Process

To account for both the stability and flexibility of meaning-form correlation in natural language, I postulate that the form and content of linguistic items are correlated via two largely independent, but co-dependent mechanisms situated in the central nervous system. The first mechanism, which I call the 'encoding mode of language use', ensures the stability of the linguistic system. *Via* this mechanism, with every use of a given linguistic item, the cases of attested similarity between the brain state considered and the brain states correlated with the given sign so far are added up (wired up) in the plastic brain area. The brain states reflect the socio-environmental stimuli filtered through the human perception and cognitive systems—represent both the object referred to and the relevant aspects of the respective situation. This encoding mechanism of categorization can be modelled in a sort of Aristotelian way.

With time features encoded more often, statistically speaking, correlate with the given linguistic sign/pattern strongly enough to be recalled by the next prompt of the given sign/pattern. The encoded representations are idiosyncratic, resembling Aristotelian representations of conventional meaning (sense). They differ somewhat between individuals, depending on their individual history of language acquisition defined by functionally established categories. Although no identical meaning corresponding to the same sign can be represented in the brain of another individual, yet the second mechanism, (which I call a selective mode of language use, and which I introduce below), ensures that the individually encoded meanings correlated to the same sign can be functionally equivalent for different individuals, i.e., different individuals using their idiosyncratic representations correlated with the same sign will make predominantly the same choices between functionally provided categories. The set of properties occurring in the sum of the representations correlated with a given sign for all members of a given community frequently will correspond best to social, conventional meanings assumed as the encoded value of that sign both by minimalists and contextualists.

The proposition outlined so far (the encoding mode of language use) presents a generally accepted rough picture of conventional linguistic meaning. The representation of the sense of a linguistic item reflects, roughly, an Aristotelian model of a category and the mechanism of its installation in the brain is compatible with a received view that learning a fact is equivalent to the emergence of a specialised system of neurons held together by excitatory plastic synaptic junctions, which arises after a sufficient number of respective excitations. That rather uncontroversial mechanism is responsible for the creation of a core, prototypical part of language. It is hypothesised to be grounded in the plasticity of the brain, which is the key to behavioural and social plasticity.

What makes my proposal different from such classical propositions is that the encoding mechanism of form-content correlation described above is integrally combined with, (i.e., results from) and influences, the second mechanism of form-

and-content creation, which I call a selective mode of language use. This mechanism is presented below. The two mechanisms jointly allow one to model the form-meaning correlation process in a more satisfactory way than previous models did. They let us avoid the major pitfalls of conventional approaches based exclusively on the former, encoding mechanism of sense installation.

The major problems faced by mechanisms of categorisation having solely Aristotelian roots are the following ones. First, the Aristotelian type of definition requires precisely delimited and ideally shared by linguistic community members core meanings, as well as requires establishing the limits on the allowed departure of the item being classified as a given category member from the set of properties defining that category. None of these requirements can be met in the case of typical linguistic categories. Second, Aristotelian definitions of a category do not model the development of meaning and cannot account for modulation and flexibility involved in linguistic compositionality[18] or instantiation. (However, liberal we decide to be in posing the limits on the allowable departure from the prototype, we may eliminate a future use of the given sign that will require an even more considerable departure from the standard. On the other hand, if we allow any degree of departure from the set of properties defining the given category as proposed by a fuzzy set approach, or Langacker's' 'schematicity' or Lakoff's 'motivation', then virtually anything can be considered to be a bad member of any other category).

The essence of the second mechanism—the selective mode of language use, (the second component of the mechanism supporting the meaning-form correlation process)—builds on a commonly accepted observation that the human brain encodes **relational** (co-occurring, associative) meaning in addition to encoding core meaning, (i.e., meaning that enumerates the properties and functions of the named object). The co-occurring, (relational) meaning that I talk about, (similar to that studied by Leibniz), is close to what psychologists refer to by the term "association", or more recently "cueing", or "priming", (which can be said to reflect the co-occurrence of elements in *pragmemes*).

Unlike the case of encoding mechanism, which operates as if "in a vacuum", (i.e., is context free), and as a result provides a representation of the concept that is correlated statistically most frequently with a given linguistic item, the selective mode of language use starts from what we already know about a given situation ["a situated speech act" to use Mey's (2001) terminology]. The relevant situational information, associations formed from cues in the verbal text, let one make predictions about the content, (interpretation) including the function, of the item to come next in the linguistic construction being formed, (or interpreted) before even considering the encoded value of the form actually used). For instance, consider interpreting the item *shrimps* in the utterance *I love shrimps* uttered when sitting with friends around a table during a party. The sentence fragment as | I love… |

---

[18] For instance, the Polish for *guinea pig* is "*swinka morska*" (literally: a sea piglet), which animal, of course, is neither a pig, nor has anything to do with marine life.

pronounced in the given situation lets us guess that the next item in the linguistic construct being formed will be a name of a food item on the table. It could also be, however, a food name qualified with *fresh*, *cooked by the host*, etc. Consequently, we end up with a set of orthogonal possibilities defined by the function of a given element within a given situated speech act, (which could potentially fill up the slot of the item under interpretation), each with a specific probability of its occurrence. A set of such options, along with the probabilities of their occurrences, will be further referred to as a communicative field. Now, in its selective mode, the function of a linguistic construct whose contribution we are assessing will not be to add its encoded information to the interpretation of the sentence being revealed (act encodingly), but **to use its encoded information to select among the options in the communicative field.**

It is postulated that one selects with a given form among the available options (i.e., out of the communicative field), the option whose encoded content is more like the encoded content of the form used for selection than the encoded content of any other option. Since the selective mechanism selects options out of any set of data, including sets of novel data, it overcomes the major problems of the Aristotelian model of linguistic categorisation mentioned earlier. The selective mode of language use can select an item, which is not a given category member according to the Aristotelian definition. This way it does not impose any limits on category memberships, and yet gives an unambiguous result in any specific situation. Consequently, it can model non-combinatorial novelty, instantiation, modulation and syntactic flexibility, as well as allow change of an encoded meaning of a given linguistic category Besides, while Langacker's (1987) or Lakoff's (1987) propositions, in which category members may resemble a category pattern only partially, leads to the problem that any item may be a bad member of any category and categorization breaks down, the selective mode of language use poses no such problem. This is so because the communicative field enumerates all possible meanings, and the selection can be carried out among these meanings in a rigorous manner, i.e., with help of a Supervised Learning (SL) technique. (These latter algorithms are used for instance, for classifying medical images). Finally, a selective mode of language use concerns the organization of individual representational data, thus avoiding the problem of deriving shared basic encodings—it does not require different individuals to share identical category patterns for them to make the same selection out of a set of predefined options. For instance, a person for whom a prototypical dog is a German Shepard Dog and a prototypical cat is a Siamese will classify 'a Poodle', or 'a mongrel' as a [dog], and not as a [cat], just as a person for whom a prototypical dog is 'a dachshund' and a prototypical cat is 'a tabby cat'.

As already mentioned, the two mechanisms just proposed are integrally related because the content that was identified selectively with a given symbolic item, (which can be quite novel and not-fitting (in the Aristotelian sense) the current encoding representation of the symbolic item considered), is added in the encoded (statistical) meaning of that item, thus affecting it. Therefore, after every occurrence the selected (utterance) content affects encoded content of the linguistic

items used. The current encoded content, in turn, influences the result of the selection among the option of the forthcoming communicative field.

Importantly, there need not be a direct, surface, correlation between syntactic and semantic structure of a given sentence uttered and the options of the communicative field on which the selection process operates, because the options are generated for a specific *pragmeme*. As elucidated in *the Extended Functional Analysis*, (Zielinska 1997), the sentence *She is a ski instructor* written in a letter of application for a job requiring a reasonable, but not outstanding, physical strength (such as a summer camp councillor), i.e., when a physical fitness of the candidate needs to be assessed, will mean roughly { [the degree of physical fitness] = that of a ski instructor}. In other words, the communicative field stipulates that the predicate *is a ski instructor* selects among [possible degrees of fitness relevant for the job at stake]. Note also that selecting and assessing may take place simultaneously. (see Zielinska 2007b for examples).

To recap, one might say that the model of categorization proposed combines what Stainton's (2010) terms System Perspective with Use Perspective, (thus in a way combines early with late Wittgenstein (1961, 1963) views). While, as Stainton (2010) points out, there is an ontological gap between the two Perspectives derived within language as an abstract system view, the model proposed derived within an empirical paradigm seals that gap.

3.2 A Preliminary Qualitative Illustration of the Mechanism Postulated

To illustrate qualitatively (in "a soft way") the mechanism just postulated, let's consider the contribution of the item *red* as used in the phrase *red car,* assuming that we have never seen a red car before. In this case, we recall what we identified as 'blue cars', 'black cars', and 'green cars' in the past, to form expectations, as to which parts of the car can be of different colours. Only now can we interpret the contribution of the adjective *red* to the meaning of the phrase *red car*. Note, that a Martian who speaks English, but who has never been to Earth, (he has observed cars from such a distance from his space ship that everything looked grey to him), will not be able to understand, what 'a red car' is even if we show him a red paint in a jar and point to a car on Earth. Neither will he draw a 'red rose' properly, not having known what flowers on the Earth look like colour-wise.

Next, let us also have a look at the contribution of the item *red* to the phrase *red barszcz,* which designates a type of Polish soup made of beetroots (a soup of certain taste, which has a crimson, or sometimes brownish colour). On the selective mode of language use 'red barszcz' is pointed out due to the fact that it is more red than the other Polish 'barszcz' is. The other type of Polish 'barszcz', 'white barszcz', is yellowish/grey in colour. Thus, the encoded value of the item *red* allows one to easily differentiate between these two options defined

functionally in the Polish culinary world. Redness, on this occasion, has a referential content of crimson, but even more importantly, indicates a certain taste provided by the respective culinary *pragmeme*, in which the parameter of taste is crucial.

Note, that this time the taste has not been encoded for the future uses of the item *red* with other nouns for statistical reasons. Since a similar taste is not shared by other red food items, it is statistically insignificant in other contexts. If some selected characteristics were more common, it could have gotten encoded, as it happened with the item *green* to encode 'not ripe' in relation to fruit and vegetables via the selective mechanism postulated above.

Obviously, on some occasions both selective and encoding modes of language use may offer a separate relevant interpretation each. For instance, when my son was recovering after his appendix had been removed, a new doctor approached me asking "Are you the mother of the boy with appendix?". "No"—I answered. "I am the mother of the boy without an appendix"—I joked. Selectively, in the situation described, the phrase *a boy with appendix* obviously chooses the patient being treated for appendicitis, which singles him out from other patients in the ward. Encodingly, however, a person after surgery no longer has an appendix, and he is not a person with an appendix.

To finish the illustration of the functioning of the mechanism postulated, let me apply it to motivating the thesis that epistemic modality originated from deontic modality and compare its effectiveness with some alternative account. The alternative account that supports the above claim, and which I am going to refer to, comes from Sweetser (1990). Sweetser argues for that claim by stating that the concept of "forcing someone to do something", present in deontic modals, is metaphorically extended onto the content "some evidence forces the subject to reach a conclusion, to have a thought". A problem with such an explanation is that it does not say a word about the mechanism that causes the metaphorical extension postulated. In other words, Sweetser describes what happened without explaining why, i.e., by what means it happened. Thus, Sweetser's account is not an explanation in the sense of the word "explanation" used in sciences, but simply a description, or to use Bunge's (2003) term: "subsumption of particulars under generalization" at best. I propose that that the claim needs to be argued in a different manner.

On the model proposed, the mechanism explaining how deontic modality gave rise to epistemic modality can be illustrated in the following way. Telling someone to do something happens most often in situations in which one has power to enforce the order. Therefore, most of the time, the action at stake will be actually performed and thus what will also became pragmatically correlated with the modal construction "A must do y", is "a big likelihood that A will do y".

Now, let us look at the following illustration of such a *pragmeme*. Parents are worried about their daughter's being late when coming back home from a university for a week-end. Mom says: *Jane hasn't arrived yet*. Now, the *pragmeme* under consideration is defined by a discussion concerning parental worries. The major issue, the function of the *pragmeme*, is to decide whether Jane is all right, or

whether something happened to her. Her parents want to consider all likely situations, which could have retained her and which do not involve any mishap before calling the police and hospitals.

In such a context as described, when we mention 'trains' and 'lateness' in a modal construction "The train must be late", since the option of forcing the train to be late is unavailable as a plausible interpretation, given the correlated issue of the likelihood that the subject will perform the action expressed by the predicate in the construction considered, there are two feasible options having to do with "trains, lateness and likelihood" and thus building up the communicative field: (1) it is likely that the train was late. (2) It is not likely that the train was late. Since, as just recalled the construction "A must do y" expresses the likelihood of action y taking place, that likelihood-feature of the construction considered will select the former one between the options 1 and 2 specified externally, i.e., it will select the option "It is likely that the train is late." This unconventional usage must have happened for the first time purely selectively.

The following situation may serve as another example of a purely selective usage of language. A four year old boy was reported to use the sentence "Open the light" in a garage without any windows. He used that sentence, which encodingly lacks any logical meaning, to select the message "open the garage door to let the light in." Unlike that latter construction that does not stand a chance of being commonly needed, the novelty involved in the pattern 'The train must be late' to indicate "the likelihood of the action being performed" was frequently useful and spread by being repeated (imitated), thus eventually inscribed in the neural system of the brain.

The above reasoning shows qualitatively, what could have happened. It shows that according to the mechanisms postulated, it is possible for epistemic modality to have arisen from deontic modality, but not precisely how it happened. In other words, the above example shows that it is logically possible for the transition to have happened via a selective mode of language use—but of course not that it did take place precisely in the circumstances presented. This is, however considerably more than labelling that process as a metaphorical extension.

3.3 Analogy

The essence of the mechanisms postulated can be also presented in the following way. While in the encoding mode of language use speakers use pre-established meanings, as if Lego Blocks, and place them together one next to another on an empty table, in the selective mode of language use speakers use encoded content to select out of options. The second mechanism assumes that the table is not empty, but that the non-verbal situation accompanying the utterance, as well as the building blocks placed so far, fill the table top up with shapes expressing viable situations, which form a specific discrete space of options. Now, instead of gluing the new building blocks to be interpreted to the ones already placed on the table,

on the selective mode of language use, the new blocks serve to identify elements of the space just described, very much the way two points will identify a specific line out of all possible lines. Consequently, the content of the lexeme used selects a part of that table top space and as a result of selecting it, the value of that linguistic construct on that occasion gets adjusted to the relevant part of the table top space.

The "building block" employed in the selective mode of language use, in addition to carrying "an encoded" "core"[19] content, carries contextual information from past uses—sort of the memory of *pragmemes*, in which it was used previously. This contextual information fills up (modifies, enriches) the space, in which the succeeding "block" will be placed. Thus, "building blocks" used selectively behave more like electrons in an electric field rather than pieces of plastic. Every electron generates an electric field around itself and therefore, if we place an additional electron in its vicinity, the behaviour of that new electron will be affected by the field (as well as the second electron will influence the former one via the field it generates).

The difference between the two modes of language use introduced in this chapter also resembles the difference between Newtonian dynamics and that proposed by the general theory of relativity. Newtonian dynamics assumes the existence of an abstract endless space, in which material bodies are placed and interact with each other. The existential hypothesis of the existence of an empty space, however, is not testable, thus philosophically cumbersome. In the general theory of relativity material objects generate time–space with testable parameters. Therefore, an object placed next to another one will interact with the field of that other object. (The reverse effect will take place, too).

Mind you that the last two analogies concern the respective sets of relations only, and not the mechanisms.

3.4 Supporting the Hypothesis of the Existence of the Communicative Field and a Selective Mode of Language Use

What is crucial when arguing for the feasibility of the mechanism proposed is motivating the existence of biological mechanisms generating a communicative field and accounting for the existence of a selective mode of language use. Below I shall introduce briefly some of arguments corroborating the hypothesis of the existence of the communicative field and a selective mode of language use.

First, the evidence from eye tracing experiments shows that the scope of our attention shifts when we proceed with decoding utterances—i.e, we might say—along with the change of the respective communicative field (c-field). Second, as

---

[19] The core meaning can be defined as the part which is statistically "significant". This is done in elementary particle physics, when identifying short living particles (it is possible to give statistically an approximate cut off point and state what is 'core').

reported in Bunge (2003), an important feature of the nervous system, which if assumed to support a communicative field will account for its potential usefulness for communication is lateral inhibition. Lateral inhibition means that neuronal excitations remain confined and do not propagate far as is the case with, e.g., electromagnetic field propagating extremely far at the speed of light. And if a communicative field is to be of any use, it needs to generate only a restricted number of options.

Third, Bruza (2010), models association patterns evoked in response to groups of linguistic items with the help of Quantum Mechanical formalism, which additionally testifies to the hypothesis that self-organization processes structure our associative knowledge. The research done by Horst et al. 2006), in turn, shows that fast mapping (identification from context), which seems to be an example of the mechanism of selection proposed here, does not result in memorization. This fact corroborates the hypothesis that the selective mode of language use performing on single occasions only does not influence the encoded meaning.

Fourth, the final group of evidence concerns the assumption that the communicative field is arranged around functionally differentiable *pragmemes*, which implies that in addition to representational meaning, the brain records functional meaning. It turns out that indeed a relevant property of the nervous tissues [in Bunge (2003)] has been found. Mountcastle (1998) discovered that neurons group into systems acting as wholes with emergent properties (functions) and with relative independence. The same hypothesis is supported by Lin et al. (2005, 2006, 2007) research concerning functional cliques mentioned earlier. Priming research, in turn, shows beyond doubt that much of the information related to a given linguistic item is hidden in the context, thus, the other way round, can be retrieved from that context, is "encoded" by it. A copious amount of psychological data, started with Heart's frames, supports the claim that people have knowledge of whole functional structures generated by prior knowledge and expectations and that people utilize these functional structures in creating representations of an incoming individual perception. It has been demonstrated, for instance, that people "remember" false facts which fit common frames.

Fifth, the everyday observations presented below also support the hypothesis of the existence of the selective mode of language use in communicative space (thus indirectly the existence of a communicative field itself). Note, for instance, how much more difficult it is to read nonce words, or simply words new to us, (cf. deciphering doctor's prescriptions) than words composing meaningful texts, e.g., letters even when written in the same handwriting. I propose that on the latter occasion, we are helped by a task of distinguishing from few feasible options with considerable differing forms, which is faster than recognizing all individual letters and decoding their respective pronunciations. In turn, subjects reading texts including words whose spelling has been slightly altered (some letters transposed) often do not even notice the mistakes. Another common phenomenon that can serve well to illustrate the operation of the mechanism of selection out of limited options is the metaphoric use of language. A given construct can refer to an item it does not encode, i.e., metaphorically, because the relevant item has already been

partially identified by the remaining contextual information and the item used metaphorically merely points out one of these options

Sixth, considerably different functioning of the left and right hemispheres gives support to the very possibility of there being two modes of language use as postulated here. While the left hemisphere tends to focus on details, i.e., decoding and logic; the right one looks for the global picture, is action oriented and, we might say, it looks for the best available fit in potential *pragmemes*. (Note, that having two ways of organizing information, e.g., as is the case with separate types of operation of the left and right hemispheres, opens the possibility of modelling the change of the correlation between the form and representation, while at the same time maintaining its short-time stability.) Additionally, the right hemisphere selects from options, which are often functionally motivated. It relies on emotions, feelings, and intuition and it is non-verbal. Note also that, historically speaking, the hardwired processes concerning, e.g., emotions, took place before those leading to the formation of plastic areas in the brain, which found more advanced cognitive processes. This all gives additional support to the claim that pre-representational early categorization was functionally defined and thus allowed one next to represent and label the categories which had already been singled out.

Finally, the model proposed is compatible with a widely accepted model concerning our memory operation (Grzybek 2006:160), which says:

> After having extracted the meaning of an actual clause, its verbatim form (words and syntax) is rapidly lost from memory, while the meaning is preserved and affects the interpretation of the following clause.[20]

3.5 Relating the Proposal to the Contemporary Mainstream Linguistic Scene

How does the proposal advocated in this chapter differ from and/or resemble major received views? Let us look at several best known proposals.

Let us begin with Chomsky's grammars. Chomsky's models, starting with Chomsky (1965), were all purely reductionist and nested exclusively in biology. The truly explanatory aspects of these propositions were to be found in psychology, which lied beyond the interests of the author. Chomsky and his followers also looked for some descriptive generalizations among data. Yet, such practices is not what is meant by explanation in empirical sciences.

Capone and Mey, in turn, have recently reintroduced the foundational role of social dimension into the main stream language modelling, when they introduced and developed, respectively, the concept of a *pragmeme*. Mey (2010: 2884) says:

---

[20] This hypothesis is motivated by Sachs (1967) in Grzybek (2006), while Luther and Fenk (1984) in Grzybek (2006) further showed that this strategy operates under "normal condition" i.e. when there is no motivation to concentrate on the form.

> Nomenclature aside, it is clear that the final question is to determine what things mean in a situation. But this meaning can be conceptualized in different ways. For a pragmaticist, this is not first and foremost a matter of determining the generalized (or even particularized meaning of an utterance), or of its individual segments; what these segments mean is always a function of their use in the situation, and in how they merge towards constituting the situation's pragmatic relevance.

In other words Mey (2010) (cf. also Mey 2001) proposes, as he terms it, an "outside-in" approach, i.e., that semantic (representational) meaning depends on functionally (originally interactively) defined *pragmemes* in which it is used. The functional (interactive) aspect of pragmatic meaning is crucial for epistemic reasons, which were already explained in the section on basic encodings. Interestingly, skin (the most primitive sensory organ) and nervous system hosting cognitive skills, originate from the same part of an embryo, which at least suggests a close correlation between the two. The crucial difference between Mey's proposal and mine is that the qualitative theory outlined here is evolutionary/developmental and both the pragmatic and semantic content continually co-develop and co-define each other—make each other more and more precise. Thus, I propose that the first imprecise representation was possible due to purely interactive categorization, and current representations and *pragmemes* modify each other with every use. Mey's proposal will coincide with mine under the assumption that the time span considered is short enough to allow one to regard *pragmemes* as being stable. This assumption holds true e.g., during the interpretation of a given text at a given moment. And this is precisely the assumption I make in this chapter when talking about qualitative (plausible) explanation and hypothetical interpretations of sentences meant to illustrate the mechanism of interpretation. Yet, neither Capone, not May is interested in quantitative tests of their proposals, i.e., tests in the empirical paradigm, which I illustrate in Sect. 4.

Mey3 (2010) points out Jaszczolt[21] (2005) as the approach to pragmatics reflecting his "outside in" perspective on language best. Jaszczolt (ibid.) proposes that encoded, default and pragmatic/cultural information all contribute simultaneously and merge into a final representation. She coins her hypothesis in the paradigm of formal linguistics based on qualitative tools, which she handles with great dexterity and sophistication. As a result, her proposition reflects the spirit that both *nature* and *nurture* matter equally, as stressed by Bunge's *systemism cum emergetism* introduced earlier. Yet, despite the same conclusion, the underling philosophical assumptions made by Jaszczolt are completely different from those of Bunge and mine. So are Jaszczolt's goals and testing methodology. For instance, as the author admits herself, the concept of default meaning lacks any proposal of a mechanisms supporting it. (On the contrary, a functionally similar concept of a communicative field has a clear psychological and philosophical justification, and a clear mechanism relating it both to pragmatic and to semantic

---

[21] Another interesting proposal crossing the boundary of semantics in the process of interpreting language in a novel way is Distributive Grammar proposed by Andre and Helene Włodarczyk, cf. www.celta.paris-sorbonne.fr/anasem/indexASMIC.html.

meanings). Therefore, unlike in Jaszczolt's, in my proposal cultural knowledge and pragmatic knowledge are not simply postulated as independent sources of information known to a language user, (which makes sense on a descriptive level and is appropriate given the goals set by Jaszczolt), but constitute an integral part of language creation process. For epistemic reasons, pragmatic and semantic knowledge in my proposal co-determine and co-create each other. Besides, while Jaszczolt's tools are suitable for descriptive analysis only, the model proposed is capable of posing some explanatory hypothesis concerning the relationship between encoded and functional (situated) meanings (see Sect. 4). These differences, however, are the differences not between the proposal advocated here and specifically that of Jaszczolt's (2005), but between the proposals coined within the developmental, empirical paradigm and all the approaches based on tackling language as a self-standing abstract structure describable with qualitative formalisms.

Another proposal with an established reputation, which covers similar ground to that covered by my proposal, is presented in Recanati (2011) and is based on the concepts of flexibility and modulation, The mechanism of selection[22] propounded in this chapter seems to meet the goals put forward by these concepts. It also seems to meet the goals set out by Carston (2002) quoted earlier. Similarly, I believe that when Stainton (2005, 2006) argues for cases of non-syntactic, or elliptical usage of words and phrases, these, too, can be seen as an instance of selective use of language, where a word or phrase serves to identify an option form parameters of potentially relevant instantiated speech acts. As already mentioned, non syntactic communication resembles, in turn, very much an early stage of language acquisition in children—when they develop a proto language. It is not unlikely then that such non elliptical communication described by Stainton (2005, 2006b) as mentioned above is the essence of children's speech, preceding their acquisition of syntactic language and eventually gives rise to linguistic structure through self-organization.

The concept of selection seems also to flesh out the concept of replacement postulated by Ariel (2008: 308). Ariel (ibid.) proposes that one of the mechanisms of explication of linguistic constructions on many occasions must be that of replacement. This is needed to cover situations of partial replacement of meaning, such as that which took place in the case of meaning change in the lexeme *gourment* when used in *gourment garage*. While Ariel mentions the concept of replacement only in passing (she is not interested in its mechanism), selection is clearly able to model it. Since we select the most similar items out of the given set, the item selected does not need to share all of the content of the item used for selection. In this way the item selected replaces the content of the item used for selection, rather than adds to it.

Ariel (ibid) concludes her book by posing the hypothesis that the same processes which are responsible for synchronic meaning creation have been *en force*

---

[22] An early version of the present model appeared in Zielinska (1997, 1999).

during diachronic change., i.e., that a diachronic change is the summation of pragmatic history of the same processes and does not involve any additional laws.

> … the very same grammar pragmatics interface representations functional in the ephemeral discourse time are also the input for the diachronic transfer of the pragmatic into grammatical. Ariel (2008: 308)

And this is precisely the position assumed here. Only one type of law—statistical laws driving the behaviour of linguistic community members during their interaction with each other are postulated. All diachronic linguistic changes are the result of self-organization processes taking place in the material system made up of members of a given linguistic community, as explained in the section discussing the concept of emergence.

As to the research concerning specific applications of the encoding and selective modes of the use of language, the position advocated can explicate the difference between descriptive and classificatory uses of adjectives, respectively. This difference has been known at least since Bolinger (1967) and recently was nicely summarized in a formal language by Kennedy (2007). The selective mode of language use can account for the semantic novelty appearing in categorizing (classificatory) usage of adjectives, while the encoding mode of language use can account for modelling the descriptive use of adjectives. The difference between the previous propositions and the account advocated here is that the selective and descriptive uses of adjectives do not exclude each other. Both can and often do function simultaneously. The evidence from Polish, where categorizing use of adjectives is marked by merely statistical preference for their postpositional usage in noun phrases, supports my view and undermines Kennedy's "either, or" stand.

Finally, the proposal presented in this chapter fits Mey's observation concerning the transition in perspective on the concept of class in pragmatic research over the first 25 years of Mey and Haberland (2002) observe that while earlier "class was essentially thought of as a product of historical developments (which it certainly is, too) the shift of perspective has to do with a more developmental view of class, class as 'work in progress'." While encoding reflects the product of historical development, selective mode of language use reflects that 'work in progress' as Mey put it—creating classes *ad hoc* for a given pragmeme. And selective mode of language use postulated in this chapter models the creation of such an *ad hoc* class.

In its general spirit, I find the proposal introduced to be somewhat similar to RT. The major similarity is that both rely heavily on context believing, as Wilson and Sperber state, that "semantics of natural language might be too weak to encode human thought" and both propose the mechanism of reaching that thought despite the shortcomings of semantics. RT does it with the help of the theory of relevance, the proposition advocated in this article does it by postulating the selection mode of language use to be modelled formally by an Supervised Learning technique. RT covers both the area of traditional Gricean pragmatics and of recovering the semantics of sentences. The proposal presented here, on the other hand, focuses primarily on accounting for the enrichment stage of RT, but the

same mechanism may operate on already explicated propositions, too. While RT worked out a very detailed mechanism of inferencing and finding relevant implicatures considering interpreted sentences, it is much less specific when applied to modelling enrichment (interpreting sentences). As already stated, the proposition presented in this chapter focuses on finding the enriched content of the utterance, to use RT terminology.

To illustrate the difference in the methodology between the two approaches, let me recall the analysis of the sentence: *It will take some time to repair your watch* uttered by a watch-maker responding to a customer who has brought a watch out of order for repair, offered by Sperber and Wilson (Sperber and Wilson 1986: 178). Sperber and Wilson (1986) say "It goes without saying that watch-repairing is a process with a temporal duration and a speaker aiming at optimal relevance must have intended to express something more then what goes without saying. In general the utterance of the form in 23 [the sentence just mentioned–DZ] should be interpreted as conveying not the truism that the job in question will take some time but that it will take an amount of time it would be relevant to remark on. i.e., longer than it would otherwise be expected."

But the authors offer not a word on why the amount of time it would be relevant to remark on is "longer than it would otherwise be expected"? RT leaves this unaccounted for—proposing no mechanism for drawing that conclusion. On the approach advocated, a qualitative explanation indicating a relevant mechanism goes as follows. In the situated speech act of a customer talking to a watchmaker when depositing his watch, the relevant options concerning the amount of time taken by the repair could have the following rough, functionally distinct (established by the *pragmeme*) values: {Done on the spot, done within a shorter time than usually, done within a usual amount of time, done within a longer than usual time, needs a lot of time to be completed}. Now, the adjective 'some' will select the "longer then usual option" by virtue of its meaning taken in comparison to the meanings of other common terms denoting values, let's assume: {no, little, some, a lot} resembling the one to the last of the five options assumed above most. (We proceed by taking the scale of conventional terms encoding values (the latter scale), next by creating the *pragmeme* relevant scale of the relevant values (the five possibilities described earlier) and next we correlate items from both scales. This way, we will temporarily assign the terms correlated with typically encoded values to actual values).

Importantly for the proposition presented, however, such an explanation such as just suggested is given only as a pre-theoretical one, offered at a descriptive stage of the research, which indicates that the proposal seems to be sensible enough to merit scientific investigations. The final goal for the proposal is to describe linguistic phenomena in an objective, quantitative fashion. Therefore, we shall be looking for theories allowing us to construct models having quantitative implications, which could be squarely born out, or rejected, by measuring relevant statistical characteristics of available corpuses. This step will be preliminarily illustrated in Sect. 4.

4 Qualitative and Quantitative Substantiation of the Hypothesis of the Mechanism of Form and Content Correlation in Natural Language

Any proposition to be evaluated within an empirical framework needs to be tested (verified.). In the preceding section, I have described the mechanism of a form-content correlation process in most general terms. Yet, (it will be explained more thoroughly in 4.1 below), following Bunge (1972), even a detailed theory of a form-content correlation mechanism will be too general to be tested *per se*. A general theory like the one postulated can only be confirmed by testing its application to models of specific phenomena. These models are described with the help of that general theory, as well as of some additional assumptions and laws concerning the specific phenomenon at stake. Therefore, to finish the presentation of the mechanism of the form—content correlation process, I shall choose a specific linguistic phenomenon to investigate, which has characteristics of its own; propose a law which could account for such specific characteristics, formulate a hypothesis this law along with the theory of form-content correlation mechanism implies, and finally test that hypothesis. The linguistic phenomenon I have chosen for this purpose is the ordering of selected classes of adjectives in noun phrases.

In other words, I shall propose a specific hypothesis that will allow me to account for some aspects of the ordering of adjectives found in corpuses. This time, however, to meet the "hard" standards of the methodology of empirical sciences, I shall account for the observations not only in a qualitative fashion to make the hypothesis proposed plausible, but I shall also propose and check the validity of some of its quantitative, statistical implications expressible in terms of objectively measurable parameters. Before doing that, however, let me recap briefly the essence of empirical linguistics advocated here and the essence of the methodology of empirical sciences.

4.1 The Foundations of Empirical Linguistics

The formal branches of linguistics treat language as a set of sentences with structures assigned to them, and, accordingly, use the formalisms of qualitative mathematical (algebra, set theory) and logics to model structural and semantic properties of language. Therefore, on the one hand, such formal approaches to language cannot hope to account for quantitative observations, on the other hand, structure and meaning being totally independent, none of them can serve to investigate the other one. Consequently meaning is left out from rigorous inspection. Since, for formal linguistics language is a closed system, neither can it hope to account for non-combinatorial novelty in language.

The perspective advocated here, (shared, among others,[23] with synergetic linguists working in the paradigm developed by Bunge (2003); with Bak (1996), who models self-organizing phenomena; and 'The Fife Graces Group'), is that language is a semiotic system,[24] which is an aspect (the result) of the communicative processes in linguistic communities. Therefore, language is subject to evolutionary processes in analogy to biological organisms. As the 'Fife Graces Group' puts it:

> Language change is a cultural evolutionary process, (Christiansen and Chapter in press, Croft 2000). According to the general Analysis of Selection (Hull 1988, 2001), evolutionary processes take place on two linked levels: replication and selection. Replicators are replicated, but with culminating errors resultant from mutation and recombination, and this way variation is generated. Selection is a process by which interactors in interaction with their environment cause replication to be differential: that is some replicators are replicated more than others, which in extreme case leads to fixation of the former and extinction of the latter. In language, linguistic structures—sounds, words, and constructions are replicated in utterances every time we open our mouths. That is, replication and variation occurs when we use language in the service of joined actions between human beings in a community. Due in part to the indeterminacy of communication described above, the replication process produces variation. Speakers differentially replicate certain structures through interaction with the environment, namely the situation being communicated and their interlocutors. In the former case, changes in life styles (e.g. the rise of *cell* (phone) and the fall of *harquebus*). In the latter case, the social identity and the social contexts of the interaction lead to the rise and fall of linguistic forms that are associated with various social values by speakers.

On the above assumption, what is especially relevant for modeling language—the development and the functioning of a linguistic system—is considering the multitude of its quantitative properties very much the way that statistical characteristics of biological groups allowed one to approach biology within the paradigm of empirical sciences. As mentioned, many quantitative characteristics of language testify to a self-organizational character of language.

> Moreover, it can be shown that these properties of linguistic elements and their interrelations abide by universal *LAWS OF LANGUAGE*, which can be formulated in a strict mathematical way—in analogy to the laws of the natural sciences. Emphasis has to be put on the fact that these laws are stochastic; they do not capture single cases (this would neither be expected nor possible), they rather predict the probabilities of certain events or certain conditions in a whole. It is easy to find counter-examples with respect to any of the examples cited above. However, this does not mean that they contradict the corresponding laws. Divergences from a statistical average are not only admissible but even lawful—they are themselves determined with quantitative exactness. This situation is, in principle, not different from that in the natural sciences, where the old deterministic ideas have been replaced by modern statistical/probabilistic models. The role of QL is to unveil corresponding phenomena, to systematically describe them, and to find and formulate the laws which explain the observed and described facts.

---

[23] The members of the Fife Graces Group, the proponents of the thesis that language is a Complex Adaptive System view, are, among others, Clay Beckner of the University of New Mexico, Richard Blythe, Edinburgh University, Joan Bybee, University of New Mexico, Morten, H Christensen, University of Cornell, William Croft, Universityu of New Mexico, John Holland Santa Fe Institute, Nick N. Ellis, University of Michigan, and others.

[24] in Bunge's (2003) understanding of the concept 'system'.

[…] the development and the application of quantitative models and methods is indispensable in all cases where purely formal (algebraic, set-theoretical, and logical) methods fail, i.e. where the variability and vagueness of natural languages cannot be neglected, where gradual changes debar the application of static/structural models. Briefly, quantitative approaches must be applied whenever the dramatic simplification caused by the qualitative yes/no scale is inappropriate for a given investigation. (Koehler, Lectures in Quantitative Linguistics, Trier University)

And such is a general philosophy of modeling language in the empirical paradigm as advocated here. Yet, before proceeding to test the form-meaning correlation law postulated in this chapter, I would like to clarify the understanding of the concepts of a theory, (law), and of a model in empirical sciences. Empirical sciences are built around theories, i.e., sets of compatible laws. The laws concern general fundamental characteristics of an aspect of a type of phenomena, and as already said, are too general to be tested *per se*. A theory is tested by its application to models of specific phenomena, which have additional characteristics and properties implied solely by the respective models. (Therefore, if the test does not support the theoretical predictions, we can never be sure whether it was the theory, or a model that failed.)

Let me illustrate what is meant by applying a law to a model. For instance, one of the Newton laws states that two material points attract each other with the force that is proportional to the product of their respective masses, to some constant G, and inversely proportional to the distance between the two points. This law, however, is not testable *per se*. It can be tested only indirectly by applying it to some model. This Newton's law can be applied, for instance, to model the movement of the Earth around the Sun. To this end, however, we will need to construct a model of the Earth orbiting the Sun, which could be written down in terms of Newton's law, i.e., we need to approximate the Sun and the Earth as two material points with specific masses. This assumption, however, is part of the cosmological planetary model, not of the Newton's law. To measure the distance between these planets, in turn, we assume that the planets are placed in empty Euclidian space, which is another assumption independent from Newton's Third Law.

Thus, if we want to test some law, we need to identify and test some of its implications for modelling a specific phenomenon—i.e., a hypothesis concerning a specific model, which model is characterized by some additional constrains absent from those defining the law. To meet the criteria of empirical sciences, such hypotheses need to be expressible in terms of data gathered with objective, received measuring techniques.

To describe a phenomenon under consideration in terms of measurable characteristics, we may need to resort also to theories and hypothesis other than the ones being tested, i.e., to the hypothesis which have already been well tested and which can assess objectively the aspects of the situation described by the new theory. We shall illustrate what that last statement means by considering a test of Hook's law (the law stating that the force exerted by a squeezed spring (coil) is proportional to the deformation caused by that force.) The relevant force can be

measured only in such a situation which describable not only with the law pos-
tulated, but also independently from the law being tested. In the case of a Hook's
law, this will take place, if we model a situation in which some weight is placed on
top of a vertically positioned spring placed on a table. In this case Hook's force is
balanced by the well tested force of gravity, the measuring of which involves
conventional and objective rules i.e., is repeatable by others following the same
instructions.

To repeat, the above example illustrates that we test theories (laws) indirectly
by testing some of their predictions in relation to a specific situation, here a spring
squeezed by a weight placed on top of it (or more exactly, to the model of that
situation.) A given situation to be of use must have characteristics measurable in
objective fashion—both with the help of the hypothesis being tested and of some
other well tested theory.

Before proceeding to propose a model of a specific linguistic phenomena that
would validate the mechanism of the form-meaning correlation process hypothe-
sized in this chapter, one more comment is in place. A given model is not expected
to cover the phenomenon at stake perfectly. "The best theory of a theoretical
model is not a copy but a theoretical model or conceptual reconstruction con-
taining concepts without a concrete counterpart (such as logical concepts), as well
as hypothesis that at best, are approximately true." (Bunge 1972: 171). Therefore,
when proposing some linguistic laws in next subsection, I do not mean to claim
that no other laws are in operation. Yet, any model creation can start only from
postulating some characteristics, which will be confirmed or rejected by experi-
ment. Only later the model can be fine-tuned and tested again, and so on.[25] (This is
a stand taken by critical realism.) Bunge (1972:171) writes about such a continual
process in the following words

> If neither experience nor reason were necessary to conduct scientific research, we could
> resort to wild intuition or to mystic communion. If theory were sufficient we would waste
> no time with empirical tests, and would give the triumph to idealism. If scientific theories
> were of no need of theory, empirism would win. In other words, critical realism assumes
> that the thing in itself is not knowable as such without any distortion. It necessarily
> involves proposing some traits and next keeping correcting them to the point of giving
> some of them up completely. As it is, factual knowledge consists of a set of theories and a
> set of data, such that the former must be compatible with at least some of the latter, while
> the data must be sought and processed with the help of some theory. Moreover, data are in
> principle as corrigible as theories, in the light of both further data and other theories.

An interesting category of object models are ones based on assumptions sim-
plifying reality to the point of significantly distorting it. This is done in situations
in which the application of the received theory to a more realistic model results in
equations which cannot be solved. Such models can be exemplified by those used
in solid state physics, in which specific phenomena are modelled in 1D (one
dimension). The resultant models cannot answer all problems, but allow one to

---

[25] Those conversant in Polish are advised to see Grabińska (1993, 1994, 1998) for especially
illuminating presentation of the relation between models, theories, and reality.

solve a certain class of problems. The usefulness of such conclusions is testified by having a special journal devoted solely to such 1D models of solid states. Apparently, the failure of a well put, good idea is often more telling than the success of an imprecise one full of *ad hoc* hypothesis mending it.

Note, also that the status of a law as concerning the most basic properties of the reality studied is relative. It is not uncommon that two laws seen at firs as independent ones later on turn out to be instances of a more fundamental one. For instance, the theory of magnetism and the theory of electricity have turned out to be special cases of the theory of electromagnetism.

4.2 Coining a Model to Verify Quantitatively the Mechanism of Form and Meaning Correlation Process

The mechanism of the form-content correlation process introduced in Sect. 3 was outlined in most general terms there. In Sect. 4, we have been arguing that form and content correlation in natural language is the result of self-organization and self-regulation and therefore the likely processes of form-meaning correlation could be constrained by laws optimizing linguistic effort and effect.

Linguistic laws proposing the optimization of speaker and listener's effort have been present at least since Zipf (1935, 1949). Zipf, for instance, accounted for diachronically observed phonological reduction in the following way. He said that since frequently occurring items are generally more predictable than the ones rarely occurring, listeners can decode the message coded by a frequently occurring item, even if it is not very carefully articulated. Consequently, speakers may afford to articulate these items less carefully, and as a result these items will undergo faster phonological change (e.g., reduction) than the ones used less frequently. This will lead to a more economical semiotic system.

These days, many linguists agree, [cf. Keller (1994), Kirby (1999), Haspelmath (2006, 2008)], in addition to those already mentioned earlier, that "a diachronic change is the necessary link between patterns of language use and grammatical structures" (Haspelmath 2006). Many researchers further agree that the diachronic change leads to the optimization of language, although such an optimization need not be a conscious goal of language users. Haspelmath (ibid: 18), for instance, elaborates on that saying

> Speakers do not intend to create well-designed grammars, but they behave purposefully and rationally in selecting from available variants and in creating new variants—they mostly opt for the most useful variants for their particular purposes. Through an invisible-hand process in language change, the cumulative effect of many individuals' behavior leads to useful language structures (cf. Keller 1994) …. So how do economical patterns arise in language change? There are two rather different routes by which this can happen: differential phonological reduction (§6.1) and differential expansion of a new construction (§6.2). Moreover, a minor route, morphological analogy, must also be recognized (§6.3).

When processes such as differential phonological reduction, differential expansion of a new construction, or morphological analogy concern purely local phenomena, (i.e., do not influence other subsystems), the hypothesis that a language user chooses an optimal variants is plausible. (Never mind whether that decision how to choose the optimal variant is rational, as Haspelmath has it, or there is an unconscious[26] mechanism ensuring such a choice.) Yet, first, the development and optimization of language is not limited to the reduction of existing forms and their propagation but, importantly, involves the creation of novel forms and meanings. Second, the optimization in language in addition to being opportunistic and in some respects optimized in local niches, also has a global character. On the other hand, plausibility of a purely cognitive mechanism being capable of making global optimal decisions relevant for specific situations, which would optimize the whole system, is unlikely, because such a task would require unrealistically complex calculations involving an extensive consideration of the whole system. Such global optimizing would mean, for instance, deciding to keep the English item *her* in the situation, where it does not convey any new content because of the benefits of the resultant pattern in other constructions. And so, in the following example: *Mary arrived late. She parked her car right in front of the main entrance and entered the school building,* the item *her* is highly predictable. Note, that in a Polish translation, the lexeme *her* would be skipped. Yet, Poles are not better at economizing natural language than the English are, one would think. The reason for the difference in the decision weather to keep the pronoun *her* or not in English and Polish, respectively, is that keeping the pronoun in English in the situation illustrated allows a given linguistic community to save more effort "across the board", among others, by not needing to add verbal suffices informing of the gender of the speaker, which is the case in Polish. However, this is not the end of the story, because languages, which inflect verbs tend to add suffixes to nouns, marking the case as well as the gender, and this all allows these languages to have a relatively free word order. That, in turn, allows them to mark 'new' and 'old' by the positioning of a given phrase in the sentence, thus do without 'a/the' type of articles, and so on. But such optimizing decisions concerning complex, interrelated parameters cannot be calculated exclusively locally. Therefore, likely some global optimizing mechanisms not requiring carrying out complex cognitive calculations concerning the whole system on the part of the speaker are also at stake during a language formation process.

I shall argue, as brought up already when quoting Altmann (1978), that these global mechanisms resemble natural selection. Note, that such global laws (based on mutation and selection of the fittest) are likely not only to account for global optimizing effects, but also for novelty in content and form (both combinatorial and non-combinatorial). And for any reduction or analogical differential spreading to take place, first some variants, e.g., morpho-syntactic patterns, need to be

---

[26] Linguistic research employing analogical modeling introduced by Skousen, puts me in that second camp.

established some way or other, thus a relevant mechanism of their creation needs to be additionally explicated.

Thus, for instance in reference to the emergence of the order of some existing elements, at the stage when there was no preferred order for that set of items, the order of such items could be selected in the following "evolutionary" way. Some sub-groups of speakers, on purely statistical grounds, start selecting one of the possible orders chosen, which results in the development of at least two sub-varieties of the current linguistic system, possibly one more efficient, the other less efficient. Next, speakers of the more instead of most efficient sub-variety being better communicators, became, statistically speaking, more successful in life. As a result, the order present in that most effective sub-variety of that system, will be repeated relatively most often due to the fact that their speech will reach a wider population, more often, and possibly will also be more esteemed thus additionally easier to memorize. Next, the decision which pattern to follow, made by an individual speaker can be purely local, i.e., determined exclusively by the relative frequencies of the options heard by that speaker. Eventually, the community will adopt the ordering of the most efficient sub-variety along with the remaining elements of that most efficient sub-variety. Establishing a new ordering of some items may lead to further unconscious reorganization because of the change in the distribution of data. Such mechanism of syntax formation could have been first implemented already on a proto-language stage, when the first combinations of previously single referent items were being tried out. Later on, this mechanism could be still taking place in relation to the ordering of items still at free variation. Among others, such too, must have once been the situation when two adjectives were used to modify the same noun for the first time.

Now, let us go back to accounting for the form-content correlation mechanism outlined in Sect. 3. As already mentioned, I have chosen to apply the proposed mechanism of form-content correlation process to explaining the order of adjective categories in 'adjective adjective noun' phrases. The phenomenon of the adjectives order in noun phrases is what I have chosen to apply the proposed evolutionary mechanism of form-content correlation to—in order to allow for its quantitative tests. More precisely, I intend to account for a statistical preference in the adjectives order between adjectives expressing semantic categories of (1) "size or shape" and (2) "age or colour" and (3) "origin or material" which is observed in English noun phrases of the A_1A_2N type. A similar statistically observable preferred order of adjectives in noun phrases has been attested, among others, in such diverse languages as Chinese, Hungarian, German, Polish, and, in a mirror reflection, in Italian and French, which suggests a universal mechanism for the phenomenon considered.

The main empirical hypothesis

In line with what has been said so far, I propose that, other factors being equal, statistically speaking, a more efficient language variant is one, which allows its speakers to express messages more precisely[27] both on specific occasions, and as

---

[27] Until the precision arrived at is sufficiently good and further increase in the precision of encoded lexemes (or their intended content) does not increase the functionality of language.

far as the resultant encoded content of lexemes is concerned. In particular, the ordering of adjectives is such that, statistically speaking, it optimizes locally the precision of the messages expressed with A_1A_2N on specific occasions, and the precision of the resultant encoded[28] values of the adjectives used. For reasons explained below, this will take place if, statistically speaking, the head noun in A_1A_2N phrases is modified with the adjective, which is more categorizing[29] first (A_2 is supposed to be more categorizing) and, by the one, which is most relative— last (A_1 is supposed to be more relative).

Qualitative justification of the hypothesis

The reason why the above order increases the precision of the encoded content of adjectives is the following. The second adjective applied (A_1) is more likely to apply to an atypical situation defined by A_2N, as a result selecting an atypical value for A_1 and thus skewing the average value of an encoded meaning of A_1. For instance consider the reference of the phrase *a red big bird* used by a visitor to a Krakow zoo. The lexeme *red* when selecting only among 'big birds', will select a flamingo, which is pink. As a result, the referred colour will add the value 'pink' to the colours encoded with the item *red* so far. In the reverse order, if we looked for a *big red bird*, i.e., for a big one among red birds, we will end up selecting a 'red bird' out of all birds thus ending up with the redness of an Ara Parrot, whose red colour represents a focal red. Therefore if we let non-relational adjectives act first, this will result in their statistical average being less dispersed, without affecting negatively the relative adjectives, which have operatorial character thus, with every use, their values depend on the category to which they are applied.

Note also, that the above conclusion is also consistent with the predictions of the theory of complexity stating that lowering of the complexity (such as that caused by limiting the number of allowable syntactic patterns) is accompanied by the increase of the informational content of components.

The hypothesised order will also ensure that the situation specific values of adjectives convey a more precise message, because the adjective that is more categorizing establishes the scale for gradable adjectives to operate on. Consequently, e.g., the gradable adjective *long* used in *long wooden bridge* will have different meaning from *long* used *a long steel bridge*, because steel bridges can be much longer than stone bridges. For the same reason, *old stone bridge* can be much older than a *wooden stone bridge* thus the value provided by the adjective *old* is

---

[28] Encoded value is understood as a statistical average of past values.

[29] What needs to be made clear here is the definition of the terms "categorizing" and "relative". By a categorizing adjective I mean one which when applied to the noun results in the selection of a distinct subcategory. For instance 'a blue crayon' differences from 'a red crayons' in colour only thus the adjectives *red* and *blue* as used in the examples above are not categorizing. On the contrary, the adjective *high* when modifying the noun *chair* selects 'a high chair', which item has a number of characteristics (including its novel function) singling out the subcategory of 'high chairs' from among all chairs, therefore the adjective *high* in the phrase *high chair* can be termed 'categorizing'. By 'a relative adjective', in turn I mean one whose actual value depends on the range of the given property in the items modified with it. For instance, the value of an adjective *big* changes depending whether it modifies a star or a mouse.

different in both cases. If we reversed the order of adjectives application and said, e.g., *a wooden old bridge*,* the lexeme *old* would select from the scale of ages of all bridges thus would not be as precise as when selecting from the scale of ages of wooden bridges only. To sum up, according to the postulated mechanism of form-content correlation presented in this chapter, if the categorizing adjective is applied to the noun first, the resultant A_cN generates a communicative field of options defined with parameters of certain range each. The relative adjectives A_r used second serves to establish an appropriate point on the scale established by the options generated with A_cN. To recap, the order of adjectives A $_{relative\ (gradable)}$ A $_{caterorizing}$ optimize the system, at least locally.

Quantitative (hard) confirmation of the hypothesis

The model of the adjectives order in noun phrases proposed above implies that the preferred order of categories (1) "size and shape", (2) "colour and age", (3) "material and origin" categories in English AAN phrases reflects the fact that the relativity (gradability) of the successive categories decreases with the number of that category, while their respective categoriability increases.

Tapping into semantic intuition and having manually tagged the adjectives, Zielinska (2007) found that when considering AAN phrases consisting of two adjectives, each belonging to one of the above defined categories, the dominance of the phrases in which the adjective further from the noun belongs to the category designated by a lower number than the adjective closer to the noun was highly statistically significant both in Polish and English corpuses. Yet, this test relied partially on human semantic intuition when categorizing adjectives, which causes some discomfort. Could we avoid semantic classification by a human completely?

If we were able to formalize the concepts of relativity and categoriability, the hypothesis we want to confirm would imply that there is a positive correlation between the degree of the difference in relativity between the first and second adjectives and the initial position of the first adjective in AAN phrases, and a negative correlation between the degree of the difference in categorizability between those adjectives and the initial position of the first of them. Or there should be a correlation between the degree of relativity and being the first adjective, and a negative correlation between the degree of categoriability and being the first of the adjectives in a AAN phrase.

Wulf (2003) set out to confirm practically the same hypothesis by providing corpus-linguistic operationalizations of concepts close to that of "relativity" and to that of the degree of being "non-categorizing". More precisely, she operationalized a concept complementary to relativity, i.e., the independence from comparison index (IndComp), which she defined as the ration of the number of occurrences of a given adjective in non-comparative degree to the number of all occurrences of the given adjective in a given corpus. By analogy, we can define the degree of relativity as the ration of the number of occurrences of a given adjective in the comparative (and superlative) degrees to the total number of its occurrences in the given corpus.

Wulf (2003) finds out that the mean IndComp values for adjective$_1$ and adjective$_2$ in her study differ highly significantly ($p < 0.001$). Adjectives standing further from their head noun occur with more forms of degree than adjectives

directly preceding the head noun, which supports the results from previous works. Wulf (ibidem) also reports that knowing the adjectives' IndComp values improves the prediction accuracy of the order of a given pair of adjectives by 35.78 % (her total prediction accuracy is 67.89 %).

Now, moving on to the concept of categoriability, the idea that adjectives expressing concepts with a high degree of categoriability are placed closer to the noun is similar to that expressed by one of Behaghel's Laws which states that things belonging close together in mind are also put closely together in communication. The concept of the degree of being categorizing can also be related to Ziff's (1960) concept of the adjectives' different "privilege of occurrence:—the degree to which adjectives may occur in different contexts", as well as to Wulf's (2003) concept of semantic closeness. Wulf gives the following corpus-linguistic operationalizations to semantic closeness:

> Accordingly, the semantic closeness of the adjectives in the present data sample was measured via the number of different head nouns that the adjective in question collocates with. For all 1,154 adjectives in the present data sample, it was checked in the whole BNC how often they occurred with any noun. More precisely, three concordances and corresponding frequency lists were produced, as not only the adjective in its positive form, but also in its comparative and superlative forms had to be included to achieve a representative picture of the span of nouns the adjective collocates with. The resulting frequency lists had to be checked manually for potential double counts of nouns, i.e. cases where a noun collocates with an adjective in its positive as well as with any/both of its compared forms. The resulting number of different noun collocates was relativized against the corpus frequency of the adjective in question because adjectives which are generally more frequent than others will automatically have a greater number of different noun collocates.

Surprisingly, Wulf (2003) finds practically no influence of the above corpus linguistic operationalization (CLO) of the degree of semantic closeness onto the adjective's being positioned closer to the noun. Yet, in the same study, she finds a strong correlation between the closeness to the noun and a membership in Dixon semantic category of 'origin, and composition (material)', which are intuitively category-forming thus, semantically close. This latter correlation, tapping also into one's intuition when categorizing adjectives, in addition to many technical arguments brought up by Wulf herself[30] against the operationalization proposed, makes one suspect that this particular CLO is not adequate. It also shows that capturing the semantic concepts purely in numerical terms is not easy and thus tapping somewhat to semantic intuition operationalized through psychological tests' results might sometimes be a better solution.

Nonetheless, Wulf (2003) comes up with statistically significant results showing there being inter-dependences between frequency information and symbolic data. Such a correlation certainly cannot be captured by qualitative formalisms.

---

[30] One more, intuitively better way to try when CLO-ing 'semantic closeness' would be to consider the actual frequency of the occurrence of specific collocates, and not the number of types, as Wulf did.

Wulf has shown that AO, although superficially a phenomenon with purely syntactic, i.e., generatively definable patterns, in fact depends on quantitative characteristics of a variety of variables from different levels of linguistic analysis. (Wulf considered also the influence of other factors such as the length of words.) These quantitative dependencies clearly point to the insufficiency of purely qualitative descriptions. Importantly, the more influential, by far, of these variables is the degree of relativity. Relativity is also the variable with the least controversial operationization. The second part of the hypothesis concerning the influence of the degree of categoriability of a given adjective on its position in an AAN string has not been clearly supported by Wulf's (2003) data, but given a questionable operatorization of what might come closest to the concept of the degree of categoriability presented in the hypothesis posed, Wulf's (2003) data does not rule out a more significant influence of that latter variable on the AO order in AANs, either.

The problems with the operationalization of the categoriability concept reported above, prompted Zielinska (2007b) to choose an approach to the issue of adjectives order based partially on tapping into semantic intuition, (as Wulf (2003) did that too, when considering the Dixon's semantic categories). Zielinska (2007b) set out to confirm the hypothesis concerning the dependence of the position of adjectives on the degree of their categoriability and relativity in the following way. She demonstrated the role of the degree of relativity (= gradability) and or categoriability of an adjective in an AAN phrase by subcategorizing the Dixon's categories into more and lees gradable (relative) subcategories and by checking the influence on the relative order between such subcategories. For instance, Zielinska (ibid.) shows both for English and for Polish that the division of the colour category into a category of intuitively highly relative colour terms, such as *light, pale, vivid, dark* and intuitively less relative ones such as *red, blue, yellow*, results in the category 'relative colour' being statistically more likely to precede other semantic categories than the category 'non-relative colour', (or let's call these 'descriptive colour' terms) does. Similarly, the subdivision of a given category of adjectives containing the information about age into a subcategory of the adjectives more and less categorizing, respectively, e.g., into {*pre-war, renaissance, baroque,*etc.} and such ones as {*one-year old, 20 year old,* etc.}, results in more categorizing subcategory following other selected categories statistically more frequently than the other subset of the category 'Age' does. As a matter of fact, in the BNC, the subcategory "descriptive colour" follows statistically the category of the "less categorizing colour" in spite of the category 'Age' as a whole preceding the category 'Colour' as a whole. (By saying that the category 'Age' precedes the category 'Colour', it means there is statistically highly significant difference in the number of AAN phrases in which 'Age' precedes 'Colour' and the number of AAN phrases in which 'Colour' precedes 'Age'.)

As reported by Zielinska (ibid.), similar correspondences have been noted also for Polish—based on the IPN corpus, (although in Polish the effect was weaker for Dixon's "middle categories" than in English). The results presented by Zielinska concerning Polish are especially interesting because so far, since Polish is a

language with a considerable free word order, there had been no[31] prior sugges-
tions as to the corresponding ordering among Polish adjectives in noun phrases.

The biggest difference between the studies of Wulf (2003) and of Zielinska
(2007), is that while Wulf (2003) was merely interested in establishing quantitative
correlations in her data, Zielinska is also interested in looking for laws that could
imply these correlations. While Wulf (2003) collected previous qualitative anal-
yses of AO, the descriptions of dependencies on various linguistic properties and
"CLO-ed" them, Zielinska searched for an explanation to these quantitative
observations. She posed a hypothesis how (via what sort of mechanism) they could
have arisen. Thus, looking at it from the perspective of an empirical paradigm, if
we note that the qualitative studies of AO may resemble data collection by Tycho
Brache, then Wulf's studies correspond to finding empirical principles (finding
patterns) such as those proposed by Kepler in relation to the data gathered by
Tycho Brache, and the current study [as well as Zielinska (2007b)] corresponds to
looking for laws explaining (implying) the numerical relations observed (the way
Newton's laws are in relation to Kepler's laws).

Zielinska (2007b) proposed that the mechanism resulting in AAN constructions
used in the right order conveying more precise messages, and in increasing the
precision of encoded adjectives values, is grounded exclusively in cognitive
human capabilities, i.e., these are speakers who calculate, consciously or not,
optimal solutions. In this study, in view of the discussion in Sect. 4, I see that the
specific mechanism leading to the optimalization could also be different than
suggested in Zielinska (2007b), i.e., as presented below.

The order of lexemes A and B, which has not been fixed yet (grammaticalized),
self-organizes through the following mechanism resembling natural selection,
which increases the efficiency of language. At first, the majority of speakers order
these items at random, on a purely statistical bases.[32] Next, on purely statistical
grounds, there may form two subgroups of people, whose idiolects show a strong
preference for one of the possible two orders, AB or BA, respectively, in addition
to those who still place these items at random. If so, the speakers of a more
efficient dialect (let us say AB vs. BA vs. (AB vs. BA)), i.e., better communicators,
will be statistically selected in the sense that they will become more influential in
life and thus their speech will receive wider reception, statistically speaking. In
other words, their speech will become an input to the corpora of a larger than

---

[31] At the same time when I published my results, Tabakowska (2007) also published a study
concerning the ordering the adjectives in Polish. Tabakowska, however does not consider any
hard frequency based evidence, but carries out a purely intuition based (cognitive) analysis of
meaning of selected AAO phrases. She concludes "that there is a preferred order of adjectives in
Polish AAN phrases, which, however, can be overridden by stylistic reasons" thus, her study is
unrelated to the empirical paradigm.

[32] In many statistical models, cf. Skousen's (1989) analogical modeling, after the system has
reached a certain level of preference of a given type, self-organization takes place. Speakers are
predicted not to chose between options at random any more, but to select one of them. Such
regularizations have been long described in language –e.g. the regularization of past tense in
Finnish modeled by Skousen (ibid.).

average number of speakers—(e.g., through mass media, education, also because the forms they use are associated with various positive social values, etc.). Consequently, the more efficient dialect (let us say that with AB order) will propagate due to the mechanism of the form content correlation process introduced in this chapter.

The explanation just proposed looks appealing because it is global in character. The self-organization of language will take place this way only if it makes the whole system more efficient, yet it is not an individual speaker who needs to carry out such complex global calculations. The speaker does not need to be aware of all benefits involved in a given choice, i.e., what let him respond more efficiently, because, in addition to making purely local judgements (based on his individual expertise), he is influenced by the frequency of the same or similar forms that he hears. Note, that the more efficient sub-variety of speech brings in "the wisdom of the crowd[33]"—its collective experience derived from independent individual experience and independent[34] individual expertise, yet evaluated statistically.

Finally, the mechanism postulated allows one also to introduce novelty—convey functions, which did not exist at an earlier stage of language development. Importantly, the quantitative data presented clearly corroborated the hypothesis that the ordering of adjective in noun phrases is not random, but characteristic of a more self-organized system. The lack of a fully deterministic ordering of adjectives in AAN phrases found by Wulf likely results simply from the dispersion in the relevant values calculated for adjectives: we resort to global parameters to form hypothesis concerning individual items, (as if we resorted to relating the volume, temperature, and pressure of gasses to form hypothesis concerning an individual particle characterized by mass and velocity). Additionally, there is also a possibility that not fully deterministic ordering is beneficial for the system as a whole—comes from the interaction with other elements of the semiotic system.

5 Conclusions

The purpose of this chapter was to position linguistics in the empirical paradigm (as a socio-natural science) by drawing conclusions from the relevant research done by others, as well as to contribute to the research in that paradigm by proposing some specific solutions. These two types of contribution can be summed up in the following way.

1. **Linguistics in the paradigm of empirical sciences**

Linguistics in the empirical paradigm is characterized by the search for explanatory laws concerning linguistic data (observed pattern) which are implied

---

[33] to use the title of Surowiecki's book.

[34] This independence is crucial for the wisdom of the crowd to be efficient.

by relevant material systems. The trend to look for explanatory laws in linguistics took place for the first time on a big scale, when Chomsky (1957, 1965) proposed that generative grammar reflects the genetic makeup of a man. It has been argued in this chapter that the reason this approach failed to account for linguistic facts well enough and faces a number of insurmountable philosophical problems is that Chomsky's approach is purely reductionist, limited to considering the biological/psychological make up of man with disregard of the crucial role of social influences (emergent phenomena). It was also argued that to arrive at better, explanatory laws concerning linguistic data from the perspective of empirical paradigm a cognitive-social approach to language is needed, just as advocated early on by Bronisław Malinowski and contemporarily by some main stream linguists, notably Capone (2005, 2006, 2009), Jaszczolt (2005), Kecskes (2010), Kopytko (1995, 2001a, b, 2004), Mey (2001, 2010), Włodarczyk (2011). However, I argue that for constructing better linguistic models, pragmatics cannot be treated merely as some independent, separate source of data from that constituted by language as an abstract semiotic system, as propounded by main stream linguistics resorting exclusively to formal qualitative tools, cf. Jaszczolt's default semantics. What is indispensable is integrally "connecting individual features with societal features", to use Kecskes' words. This can be done by assuming that language is a self-organizing semiotic system, an integral result of the history of communication processes in a given linguistic community, (cf. Altmann (1978), Grzybek (2006), Koehler (2005), Zielinska (2007), the Fife Graces Group)—a result of some sort of economy between the speaker and listener. In this vein, Altmann (1978) suggested that a possible self-organizational mechanism could resemble of natural selection. Any way, the validity of the assumption of self-organizational nature of language was corroborated both by considering the epistemic concerns presented in this chapter, as well as by pointing out the existence of quantitative data attesting to the fact that power laws characterize language.[35] The wealth of additional quantitative data copiously gathered, e.g., in the *Journal of Quantitative Linguistics*, or, for instance, in Koeler (2005), Grzybek (2006), Wulf (2003), Zielinska (2007), etc., indicates the existence of undisputable connexions of many kinds between symbolic data and frequency information. In an empirical paradigm, such connections can also be used to operationalize linguistic concepts through statistical characteristics of corpora, leading to an objective measurement also of selected semantic characteristics of language.

To sum up, we may quote Heylighen (2008) and state that according to the current philosophical understanding of the world, the cutting-edge results

---

[35] Originally, Zip inferred the power laws in language from the principle of minimal effort. Later on this principle was expressed as the optimalization of the effort involved in information transfer between the speaker and the addressee. It must be acknowledged, however, that the possibility of there being also some other sources of power laws, which would not imply the self-organizational character of language, has also been considered. Yet, these other models, cf, an overview in Kwapień (2010), require making a number of assumptions contradictory to what we know about language.

concerning language modelling should be expected in the research of language as a self-organizing system.

2. Modelling a form—meaning correlation mechanism within the paradigm of empirical sciences

My contribution to modelling language within an empirical paradigm consists primarily of two hypotheses. First, I have proposed a general mechanism of the form-meaning correlation process (categorization) that would fit a self-organizational perspective onto the nature of language. The mechanism postulated relies on two modes of language use: encoding and selective ones. (Selection takes place from options generated by expectations, goals and associations related to a specific situated speech act.) This model, as demonstrated, overcomes a number of major problems faced by an Aristotelian category relating form with content, (and its extensions, such as a prototype model of a category, or a fuzzy set model, etc.). The crucial problems present in the classical models of a linguistic category, which are avoided in the approach proposed, can be exemplified by the impossibility of crossing with basic encodings between separate individuals, and by the impossibility of defining the limits on allowable departure from the definition of a given linguistic category to verify the membership of a given exemplar in that category. Importantly, the mechanism proposed has been corroborated both in a qualitative terms and in a quantitative way. The quantitative verification took place by showing that the mechanism, (along with some additional law governing the word order in AAN phrases), has some statistically significant implications for the statistical order of adjectives in the AAN phrases.

Second, I proposed an additional mechanism of the form-content correlation formation, which is effected not only by purely cognitive processes, but also by social ones. Postulating such a mechanism is especially important in order to account realistically for the possibility of arriving at globally optimal solutions when creating language, i.e., ones valid for the whole system of interrelated phenomena. This latter mechanism shifts the burden of carrying complicated estimations resulting in creating a globally optimal semiotic system, (identical for the majority of community members faced with individual input data each) from the brain alone to the brain along with "the wisdom of the crowd" (a social process similar to natural selection, where gene transmission is substituted by linguistic replication and where the idiolects of socially successful individuals affect the decisions of individual speakers due to the frequency of the respective input data they face.) This can happen in the following way.

The idiolects of successful individuals reach a proportionally wide audience,[36] thus become highly represented in the corpuses of a larger than average number of members of a given linguistic community. This translates into a relatively high probability of the choice of a given variant by these individual speakers. (Cf.

---

[36] Additionally, the community members are motivated to remember the successful idiolect better due to prestige involved.

Skousen's (1989) modelling of past tense in Finnish). As a result, complex cal-culations concerning the influence of a given variant onto a global economy of a given semiotic system to be carried out by an individual are substituted by his assessment of the relative frequency of a given option. Importantly, by observing a new statistically significant correlation between some, possibly modified, lin-guistic construction and a new type of situated speech acts, functional novelty can be identified and established.

References

Ajdukiewicz, K. (1935). Die syntaktische Konnexit. In S. McCall (Ed.) *Polish logic 1920–1939* (pp. 207–231). Oxford University Press. (Translated from *Studia Philosophica, 1*, 1–27).

Altmann, G. (1978). Towards a theory of language. *Glottometrica, 1*, 1–25.

Ariel, M. (2008). *Pragmatics and grammar. In the series Cambridge textbooks in linguistics.* Cambridge: Cambridge University Press.

Bak, P. (1996). *How nature works. The science of self-organized criticality*. New York: Springer.

Bak, P., Tang, C., & Weisenfeld, K. (1988). Self-organized criticality. *Physical Review E, 38*, 364–374.

Berwise, J., & Perry, J. (1983). *Situations and attitudes*. In J. Barwise, & J. Perry. Cambridge, Mass: Brad-ford/M.I.T. Press.

Bolinger, D. (1967). Adjectives in English. Attribution and predication. *Lingua, 18*, 1–34.

Brickhard, M. H., & Campbell, R. L. (1992a). Some foundational studies concerning language Studies. *Journal of Pragmatics, 17*, 404–458.

Brickhard, M. H., & Campbell, R. L. (1992b). Clearing the grounds. *Journal of Pragmatics, 17*, 472–486.

Brook, R. (1986). Intelligence without representation. Unpublished research report, Cambridge Mass MIT Artificial Intelligence laboratory, 3.

Bruza, P. D., Kitto, K., Nelson, D., Cathy, L., & McEvoy, C. (2009). Is there something quantum-like about the human mental lexicon? *Journal of Mathematical Psychology, 53*, 362–377.

Bruza, P. D., & Peter, D. (2010). Quantum memory. *Australasian Science, 31*(1), 34–35.

Bunge, M. (1972). *Method, model and matter*. Heidelberg: Springer.

Bunge, M. (2003). Emergence and convergence. Qualitative Novelty and the Unity of Knowledge (Toronto Studies in Philosophy) Toronto: UTP inc.

Capone, A. (2005). Pragmemes (a study with reference to English and Italian) *Journal of Pragmatics 37*, 1355–1371.

Capone, A. (2006). On Grice's circle (a theory-internal problem in linguistic theories of the Gricean type.) *Journal of Pragmatics 38*, 645–669.

Capone, A. (2010). On the social practice of indirect reports (further advances in the theory of pragmemes). *Journal of Pragmatics, 42*(2), 377–391.

Carston, R. (2002). *Thoughts and utterances. The pragmatics of explicit communication*. Oxford: Blackwell Publishers Ltd.

Chomsky, N. (1957). *Syntactic structure*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.

Chomsky, N. (1986). *Knowledge of language: its nature, origins, and use.* New York: Praeger.

Dummett, M. (1993). *The seas of language*: Oxford: Oxford University Press.

Elman, J. L., Bates, E. A., Johnson, M. H., & Karmiloff-Smith, A. (1997). *Rethinking innateness: A connectionist perspective on development (neural networks and connectionist modelling)*. Cambridge: The MIT Press.

Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love, and the meaning of life*. New York: Farrar, Straus and Giroux.

Gopnik, A. (2007). Bayesian network, Bayesian learning and cognitive development. *Development Science, 10*(3), 281–364.

Grabińska, T. (1993). Teoria, Model, Rzeczywistość. Wrocław: Wrocławska Oficyna Wydawnicza Politechniki Wrocławskiej.

Grabińska, T. (1994). Poznanie i modelowanie. Wrocław: Wrocławska Oficyna Wydawnicza Politechniki Wrocławskiej.

Grabińska, T. (1998). Od Nauki do metafizyki. Warszawa: PWN.

Griffiths, P. E., & Gray, R. D. (1994). Developmental systems and evolutionary explanation. *The Journal of Philosophy 91*(6), 277–304.

Grzybek, P. (2006). *Contributions to the science of text and language. Word length studies and related issues*. Netherlands: Springer.

Haken, H. (2010). *Synergetics: Introduction and advanced topics*. Berlin: Springer.

Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics 19*(1), 1–33.

Haspelmath, M. (2006). Creating economical morphosyntactic patterns in language change. In J. Good (Ed.) *Language universals and language change.* Oxford: Oxford University Press.

Heylighen, F. (2008). Complexity and self-organization. In M. J. Bates, & M. N. Mack (Eds.) *Encyclopedia of library and information sciences* (Taylor & Francis).

Horst, JS., McMurray, B., & Samuelson, L. K. (2006). Online processing is essential for learning: Understanding fast mapping and word learning in a dynamic connectionist architecture. *The Poceedings of the 28th Meeting of the Cognitive Science Society.*

Hull, D. L. (1988). Replication and interactors in his Science and Selection. Cambridge: Cambridge University Press.

Jaszczolt, K. (2005). *Default semantics: Foundations of a compositional theory of acts of communication*. Oxford: OUP.

Karmiloff-Smith, A. (1995). *Beyond modularity: A developmental perspective on cognitive science*. Cambrigde: Bradford Books.

Karmiloff, K., & Karmiloff-Smith, A. (2010). *Getting to know your baby*. London: Carroll & Brown.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy, 30*, 1–45.

Keller, R. (1994). *On language change: The invisible hand in language.* London and New York: Routledge.

Kecskes, I. (2010). The paradox of communication: socio-cognitive approach to Pragmatics. *Pragmatics and Society, 1*(1), 55–73.

Kirby, S. (1999) . Syntax out of Learning: The cultural evolution of structured communication in a population of induction algorithms. In D. Floreano, J. D. Nicond, & F. Mondado (Eds.), ECAL99, 694–703.

Koehler, R. Synergetic Linguistics. In R. Koehler, G. Altmann, Rajmund G. Piotrowski (Eds.), *Quantitative Linguistik- Quantitative Linquistics*. Ein internationales Handbuch, Berlin/New York: de Gruyter, 760–774.

Kopytko, R. (2001a). From Cartesian toward non-Cartesian pragmatics. *Journal of Pragmatics, 6*, 783–804.

Kopytko, R. (2001b). The great dissolution a la Michel Foucault and the non-Cartesian pragmatics. *Journal of Pragmatics, 10*, 1637–1642.

Kopytko, R. (2004). The affective context in non-Cartesian pragmatics: A theoretical grounding. *Journal of Pragmatics, 3*, 521–548.

Kopytko, R. (1998). Relational pragmatics: Towards a holistic view of pragmatic Phenomena. *Studia Anglica Posnaniensia, 33*, 195–211.

Kopytko, R. (1995). Against rationalistic pragmatics. *Journal of Pragmatics, 23*, 475–491.

Kwapień, J. (2010). Fizyczne charakterystyki złożoności. The Institute of Nuclear Physics of Polish Academy of Science in Kraków. Raport 2036/AP.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: MIT Press.

Lin, L., Osan, R., Shoham, S., Jin, W., Zuo, W., & Tsien, J. Z. (2005). Identification of network level coding units for real time representation of episodic experiences in the hippocampus. In *Proceedings of the National Academy of Science* USA, 102 no 17, pp. 6125–6130.

Lin, L., Osan, R., & Tsien, J. Z. (2006). Organizing principles of real-time memory encoding neural cliques assemblies and universal neural codes. *Trends in Neurosciences, 29*(1), 48–57.

Lin, L., Chen, G., Kuang, H., Wang, D., & Tsien, J. Z. (2007). Neural encoding and the concept of nest in the mouse brain. In *Proceedings of the National Academy of Science*, USA, 104 no 14, pp. 6066–6071.

Lakoff, G. (1987). *Woman fire and dangerous things*. Chicago: CUP.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford: SUP.

Luther, P., & Fenk, A. (1984). Wird der Wortlaut von Saetzen zwangslaeufig schneller vergessen als ihr Inhalt? Zeitschrift fuer Experimentelle und Angewandte Psychologie, *31*(1), 101–123

Mey, J. L. (2001). *Pragmatics*. Oxford: Blackwell.

Mey, J. L., & Haberland, H. (2002). Linguistics and pragmatics, 25 years after. *Journal of Pragmatics, 34*(12), 1671–1682.

Mey, J. L. (2010). Reference and pragmeme. *Journal of Pragmatics, 42*, 2882–2888.

Mey, J. L. Forthcoming. Modular, cellular, integral: A pragmatic elephant? In A. Capone, F. Lo Piparo, & M. Carapezza (Eds.) *Perspectives on pragmatics and philosophy*. Dordrecht: Springer.

Mountcastle, V. (1998). *Perceptual neuroscience; the cerebral cortex*. Cambridge: Harvard UP.

Putnam, H. (1960). Minds and machines. In S. Hook (Ed.) *Dimensions of mind* (pp. 148–180). New York: New York University Press.

Putnam, H. (1975). The meaning of meaning. In K. Gunderson (Ed.) *Language, mind, and knowledge* Minneapolis: University of Minnesota Press.

Putnam, H. (1988). *Representation and reality*. Cambridge: MIT Press.

Recanati, F. (2011). *Truth-conditional pragmatics*. Oxford: OUP.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Searl, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Percept Psychophys. 2:*437–442.

Seymour, M. (2010). Speech act pluralism, minimal content, and pragmemes. *Journal of pragmatics, 42*(11), 2870–2881.

Skausan, R. (1989). *Analogical modelling of language*. Dordrecht: Kluwer Academic Publishers.

Stainton, R. J. (2010). Philosophy of language. In L. Cummings (Ed.), *The pragmatics encyclopedia* (pp. 348–353). London: Routledge.

Stainton, R. J. (2005). In defense of non-sentential assertion. In Z. Szabó (Ed.) *Semantics vs. pragmatics* (pp. 383–457). Oxford: Oxford University Press.

Stainton, R. J. (2006a). *Words and thoughts: Subsentences, ellipsis, and the philosophy of language*. Oxford: Oxford University Press.

Stainton, R. J. (2006b). Neither fragments nor ellipsis. In L. Progovac et al. (Eds.) *The syntax of nonsententials.* (pp. 93–116). Philadelphia: John Benjamins.

Stanley, J. (2007). *Language in context*. Oxford: Clarendon Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sweetser, E. E. (1990). *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: CUP.

Travis, C. (2001). *Unshadowed thought: Representation in thought and language*. Cambridge: Harvard University Press.

Wierzbicka, A. (1972). *Semantic primitives. Linguistische Forschungen no 23*. Frankfurt: Athenaeum.

Wierzbicka, A. (1985). *Lexicography and conceptual analysis*. Ann Arbor: Karoma.

Wierzbicka, A. (2010). Bilingualism and cognition: The perspective from semantics. In V. Cook, & B. Bassetti (Eds.) *Language and bilingual cognition* New York, NY: Psychology Press.

Włodarczyk, A. (2011). Distributed grammar. www.celta.paris-sorbonne.fr/anasem/intexASMIC.html.

Wulf, S. (2003). A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics, 8*(2), 245–282.

Wittgenstein, L. (1963). Philosophical investigations. Oxford: Basil Blackwell. Originally 1953.

Wittgenstein, L. (1961). Tractatus Logico-Philosophicus, London: Routledge and Kegan Paul. Originally 1921.

Zielinska, D. (1997). A note on the extended functional analysis. *Journal of Pragmatics, 27*, 841–843.

Zielinska, D. (1999). The selective mode of language use; the way nature language adopted itself to describing the word around us. *Zeszyty Naukowe UJ, 119*, 173–176.

Zielinska, D. (2007). The selective mode of language use and the quantized communicative field. *Journal of Pragmatics 39*(5), 813–830.

Zielinska, D, (2007b), Proceduralny model języka. Językoznawstwo z pozycji teorii modeli nauk empirycznych. WUJ. Kraków.

Ziff, P. (1960). *Semantic Analysis.* Ithaca, NY: Cornell University Press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Boston: Addison-Wesley.

Zipf, G. K. (1935). *The psycho-biology of language*. London: Routledge & Sons Ltd.

Language-Game: Calculus or Pragmatic Act?

Marco Carapezza and Pierluigi Biancini

Abstract We have tried to make the potentiality inherent in the concept of the linguistic game evident by taking it back to its original context in the work of Wittgenstein. This paper aims to re-examine some features of Wittgenstein's thought, considering in particular the notion of 'language-game'. We believe that the language-game might play a role in overcoming once and for all the classic distinction between semantics and pragmatics. We deal with the exegetical discussion of the notion 'language-game' as it was interpreted in two different senses: as a synonym of calculus or as a minimal unit of linguistic activity that is directed to obtaining certain pragmatic effects in a societal context. The latter, broader interpretation, is characterized by three different features: topicality, broader normativity and multimodality. Starting from an interpretation of language game as a pragmatic act, we work out a possible parallel between language games and the notion of pragmeme as presented by Mey. Both language game and pragmeme refer to an extended notion of the linguistic symbol seen as a non-linear, multi-modal concept that overlaps the mere verbal unit of expression and is now considered as a set of diverse expressive resources (such as gesture, tone of voice and so on). This comparison will also work for a problem common to both language-game and pragmeme, that is the need to set a boundary to these units of analysis thanks to which they could be identified. We advance a possible solution to this problem, which is rooted in a rethinking of Wittgenstein's notions. The proposal consists in focusing on the topic for which the language game is played. The topic is taken to be the organizing aspect of understanding of the game. The societal

M. Carapezza (✉) · P. Biancini
University of Palermo, Palermo, Italy
e-mail: marco.carapezza@unipa.it

rules, the worldly knowledge, often taken to be the ground of understanding in our discourse are considered as merged together in a holistic unit called language game.

The relationship between Wittgenstein's work and pragmatic literature can be summed up in two commonly held truisms: the first is that Ludwig Wittgenstein is one of the key figures who, with his *Philosophical Investigations*, created the conditions for the pragmatic turn in language studies giving birth to 'use theories' of meaning. The second of these truisms is that due, to their philosophical fragmentation and obscurity, Wittgenstein' remarks, especially those that are devoted to the relationship between meaning and use, cannot be transformed into a proper semantic theory.

Starting from this point of view, Wittgenstein's remarks about language and meaning are often considered to be old-fashioned reflections that are not useful in the current debate between scholars belonging to different poles of the field, as in delineating a theoretical background that makes identification of the problems of the discipline possible.

As scholars arriving from the exegetical analysis of Wittgenstein work, our starting problem could be summed up with the following questions: What could Wittgenstein still say to language analysis? Could his thought be of some interest in building a new perspective on language? In both cases our answer would be: yes!

This paper aims to rethink some features of Wittgenstein's thought, examining in particular the notion of 'language-game'. We believe that the language-game might play a role in a definitive transition from the classic distinction between semantics and pragmatics to a unified theory as it is proposed by various contemporary authors (e.g. Mey (2001); Recanati (2006) and others). A discussion of the concept of 'language-game' is a way of examining the relationship between semantics and pragmatics, in particular this famous Wittgensteinian notion could be the key to shed light on the role played by context in linguistic processes.

The paper is divided into three sections: in the first section we deal with the exegetical discussion of the notion 'language-game' as it has been interpreted in two different senses as a synonym of calculus or as a minimal unit of linguistic activity that is directed to obtaining certain pragmatic effects in a societal context (Duranti 1997). The latter, broader interpretation, is largely characterized by three different features: topicality, broader normativity and multimodality.

In the second section of the paper we work out a possible parallel between language-games and the notion of pragmeme as presented by Mey (2001) and Capone (2010b). The pragmeme is presented as a type of a certain pragmatic act and it is meant to bring into prominence the idea that meaning is always to be situated in a widened context, constituted by social norms and different systems of signs. The same section faces a problem common to both language-game and pragmeme, that is the need to draw a boundary to these units of analysis thanks to which they could

be identified. Our proposal consists in looking at *Philosophical Investigation* (§69) as a crucial point in which Wittgenstein gives a solution to the problem.

The third section applies what we discuss in the previous sections and describes two polysemous language-games that could be understood as completely different pragmatic acts (such as greeting versus making an assertion). The polysemy of these games may be resolved thanks to the contextual knowledge shaped by the goal to which the act is directed.

1 The Language Game

In this section, we discuss two different conceptions of what should be considered a *language-game*: the first being a narrow conception, which considers a language game to be a set of words held together by strict grammatical rules, and the second being a broader conception that considers a language game as the entire field of possible linguistic interactions that are held together by the existence of a general point or purpose in which rules are progressively (re)constructed by the subjects involved.

1.1 The Language Game as Calculus

Wittgenstein began to use the term *Sprachspiel* in the second part of his philosophical career, after his return to Cambridge.

> Here, for example, is a simple language-game [*Sprachspiel*]: Turning on the electric light in a room, you say «light» to a child […] then turning it off you say «dark»; and you might do that several times, emphasizing your words and doing it for varying lenghts of time. Then you might go into the adjoining room, from there turn on the light in the first room and get the child to tell you «light» or «dark» (MS 113: 45r).

The first occurrence of the term is dated Feb-March 1932 and is contained in manuscript 113 at page 45 (MS 113, 45r), then transposed into one of the typescripts now published as the *Big Typescript* (§ 46). Just like many other typically Wittgenstenian terms, the language-game does not receive a strict definition and its use is very heterogeneous as its functions in Wittgenstein's philosophy.

The game metaphor was used by Wittgenstein to criticise the Logical Atomism contained in his former work, the *Tractatus Logico-Philosophicus*, in which he maintained that any proposition was logically independent because it is a picture of a fact. The *Tractatus* was written with the idea that the act of speaking a language is a calculus operation in which the rules are hidden and deeply connected with the inner logical form of the world.

Rethinking this idea of language as calculus after his return to Cambridge, Wittgenstein should have realised that this autonomy of the logical form with

respect to the actual realisation of a speech act was a mistake. However, he did realise that elementary propositions could not be logically independent. In these first years, the language-game was seen by Wittgenstein as an attempt to rethink the old idea of the calculus without abandoning it altogether. Some of the remarks contained in the *Philosophical Grammar* are clear examples of the persistence of these references to calculi:

> I can only describe language-games or calculi (PG: 62).
> For us language is a calculus; it is characterised by linguistic activities (PG:193).

According to these remarks, Wittgenstein did not give up his former attempt to give an account of language in terms of operations that obey strict rules, but only the atomistic tenet about the relationship between propositions and facts. In this new framework, propositions are compared to reality in groups instead of individually, like the graduating marks of a ruler. The colour octahedron is one of the paradigmatic pictures of this idea of what it means to have a synoptic representation of the entire grammatical system involved in representing colours. Another example is the cube seen as the representation of a geometrical rule, thanks to which he concluded that propositions form 'systems of propositions' (*System von Sätzen*) i.e. sets of propositions in which the members exclude each other thanks to grammatical relationships that occur between the words in them:

> But how can the cube (or the drawing) serve as a notation for a geometrical rule? Only if it belongs, as a proposition or part of a proposition, to a system of propositions (PG: 55).

In the *Blue Book* Wittgenstein will refer to the understanding of a proposition, a sign, as a matter of understanding a system of signs (BlB: 5), again allowing the interpretation of language and the language-game in formalist terms.

Since in the *Blue Book*, and more clearly in the *Brown Book*, the use of the language-game became more heterogeneous, only in the *Blue Book* do we find a different characterisation of the term which points to a different use:

> Language-games are the forms of language with which a child begins to make use of words (BlB:17).

Then, in manuscript MS 115, containing his attempt to translate the *Brown Book* from English to German, he used the term «… *Ganzen Praxis der Sprache* (The whole practice of language)» as a synonym for the original «whole language game» (BrB: 108). These examples show a progressive widening of the functions to which the language-game was dedicated. At first, it was intended to undermine logical atomism, then it even became evidently useful for an entire criticism of the referentialist image of language (PI: §1). The exegetical point is not addressed here. Suffice it to say that Lo Piparo (2009) has recently proposed an interesting interpretation of the change that occurred in Wittgenstein's vision of language-games: this change could be due to the indirect influence of Antonio Gramsci's doctrine of Grammar and Praxis on Wittgenstein, via Sraffa, a mutual friend.

The following remark clearly explains how deeply the concept of the language-game was embedded in Wittgenstein's philosophy:

> We can also think of the whole process of using words in (2)[1] as one of those games by means of which children learn their native language. I will call these games "language-games" and will sometimes speak of a primitive language as a language-game.
> And the processes of naming the stones and repeating words after someone might also be called language-games. Think of much of the use in games like a ring-a-ring-a-roses.
> I shall also call the whole, consisting of language and the actions into which it is woven, a language-game (PI: §7).

In the first two sections, the language-game is devoted to undermining the role of the ostensive definition in language, arguing in support of the thesis that a referent is only identifiable in a linguistic normative sequence and not in accord with an inner mental representation in the mind of the subject. We could say that the function played is primarily psychological, or at least anti-psychologistic, and for our purposes is less important.

We are interested in the following issues: Is the calculus model still present in the PI? And what does it mean from a semantic point of view? Why might it be of interest today?

1.1.1 The Chess Metaphor

Max Black discussed these questions some years ago in a well-known article (1979) distinguishing between a narrow and a broader interpretation of the language-game. Even in the last period, the narrow interpretation maintains the persistence of the calculus model, focusing on the existence of a nucleus of constitutive rules characteristic of each game. Any kind of language-game is constituted by a set of intrinsic semantic properties, given by those rules that precede the activity and make the game a certain kind of game (e.g. the game of promising is constituted by a proper set of rules and actions), and by another set of semantic properties depending on the situation of proposition, and are contextually driven, depending on a set of pragmatic rules.

In the remarks devoted to family resemblances in PI (§§66–72) Wittgenstein deals with the concept of 'game', showing how different games have different characteristics that make each game belong to a larger family of GAMES.

The persistent use of the chess metaphor by Wittgenstein is another possible mark of continuity that could be taken as supporting the narrow interpretation of the language-game: during the first period, chess is probably the favourite example in his remarks. He compared the understanding of a language to the ability to play a game of chess stating: "the meaning of a word is its role in the calculus of

---

[1] The game (2) to which Wittgenstein is referring is the famous example of builders contained in PI §2. In the game a builder A calls B, his assistant, using only 4 words: block, pillar, slab, beam.

language. (I compared it to a piece of chess)" (PG: 63). What interest might chess have in our discussion?

First of all, any chess piece is characterised by a set of possible moves given to it apart from the context of use. In the same way Wittgenstein could have thought that any sign (be that the word or the proposition) has a literal meaning, apart from any kind of context (Cappelen and Lepore 2005: 2). A similar idea to what was argued by Kaplan (1978) with the notion of *character* and by Perry (1978) with the notion of *role* for indexicals follows. The idea is more or less this: every statement in which an indexical appears should be interpreted relatively to a context through a restructuring of the Fregean meaning in two stratified components: the character (or the role) and the content, i.e. the truth conditions of a statement. The character is the linguistic meaning of an indexical expression established by the rules, while its content should be its referent. The term, "I", has an I character: it refers to the person who speaks independently of who might be speaking and this is the rule that guarantees the application and a content, the person who is effectively saying I. In this case M.C. is one of the authors.

The first consequence of the chess metaphor is that it compels us to take a minimalist stance towards language-games.

As a second consequence, chess players have to follow pragmatic rules based on their ability to make inferences about the state of the game while moving the pieces on the chessboard appropriately. This kind of pragmatic appropriateness depends on the context, i.e. on the disposition of the pieces on the chessboard, but does not alter the constitutive rules of the game. Thus if I move the queen in a way that gives my opponent the opportunity to make a winning move, my move should be considered inappropriate but not devoid of meaning, not nonsensical. In language, this would mean that in the language-game we should consider the situation of use of words and sentences as an external apparatus thanks to which we can take decisions about the appropriate use of the signs but not about their sense, which is given semantically and not pragmatically.

This is roughly what happens with certain interpretations of Austinian theory of speech-acts (Bach 2005). According to these interpretations, we should take the partition into locutionary, illocutionary and perlocutionary acts literally and the decision about the contextual conditions of appropriateness of a speech act are part of the pragmatic representation of the communicated meaning of the locutionary content, i.e. what is said in communication.

The third point of the chess metaphor is constituted by the narrow representation of what it means to be a subject involved in the calculus, or communication. The chess player has to be able to make inferences that are only based on the pieces arranged on the chess-board. He has to compositionally calculate each piece, its possible uses, and to infer what the best move is. This is an operation that any kind of machine could perform, if supplied with the appropriate input. The game is narrowly inter-subjective, that is to say it is a game played by two distinct mechanisms each of which has to do primarily with the chess-board (linguistic input/output).

Communication is just a matter of creating the output to a given input. A corollary given by this image of the chess-board is the modality of communication: the chess-board is monodimensional, i.e. it only has one dimension with a binary code (black/white). Chess moves are ordered in a linear succession; each move is followed by another move in a predetermined space and time. In fact, a chess match is described using simple well-formed formulas such as those used in chess magazines. In a chess match, the story of the game is hardly or not at all important. Instead, what really counts is the configuration of the pieces at a given moment. The same kind of linearity was illustrated by Saussure when he talked about the linearity of the sign as one of the characteristic principles of language.

The linearity or mono-dimensionality assumption (the terms are used in this discussion as synonyms) seems to be modelled on writing. In fact, writing is a model of speech based on a linear succession (albeit with small interferences from some punctuation signs such as the question mark) of alphabetic characters (Albano Leoni 2009). In a written text, even though there is the minimum irregularity that we mentioned, each letter has its predetermined role in the alphabetic system and the letters are arranged linearly along a dimension.

This is a not very Wittgensteinian model that is, in fact, critically described in an (undervalued) passage of the *Brown Book*:

> Though from a certain point of view we should call the linear character of the sentence merely external and inessential, this character and similar ones play a great role in what as logicians we are inclined to say about sentences and propositions (BrB: 98).[2]

Accordingly, language is to be taken only in its linguistic/verbal dimension, to which everything else is external and added, and cannot change its inherent, literal meaning.

Before going on with the discussion, a summary of the characteristics of the first narrow interpretation of the language game at which we have arrived is useful:

Narrow normativity: the language-game is played obeying a number of constitutive rules, thanks to which any sign/sentence has a proper literal core meaning.

Narrow inter-subjectivity: the game is played by subjects (speakers/hearers) that have to make inferences on the literal meanings of words taken in a particular situation. The operation of inference is made from the type to the token in its appropriate situation.

Monomodality: the language-game is intended as a set of linguistic types to which contextual material only adds external properties. Context does not affect the type of game that is being played.

---

[2] The binomium «sentence and proposition» is worth noting; it will be conveyed in the EPB simply with *Satz* meant in the most open sense possible.

1.1.2 Some Problems for the Calculus Model

This interpretation of the language-game is very restrictive and surely it could be considered as an attempt to take the concept to its limits, but it is also consistent with a familiar way of reading the following famous passage in *Philosophical Investigations*:

> But how many kinds of sentences are there? Say assertion, question, and command?— There are *countless* kinds: countless different kinds of use of what we call «symbols», «words», «sentences». And this multiplicity is not something fixed, given once for all; but new types of language, new language-games, as we may say, come into existence, and others become obsolete and get forgotten […].
> 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 form of life (*Lebensform*) (PI: §23).

There is a traditional, standard reading of this passage according to which Wittgenstein undermines the concept of proposition, as it was presented in the *Tractatus* in addition to the works of Frege and Russell (Baker and Hacker 1983). In the traditional view, the proposition is something that could be treated in terms of truth conditions; in particular the *Tractatus* advanced the idea that the essence of a proposition was its bipolarity: the possibility of its being either true or false. The standard reading maintains that in PI §23 Wittgenstein completely gives up the old idea of bipolarity as the essence of the proposition embracing a form of contextualism in which the sentence was still the minimal unit of semantic analysis (Glock 1996: 89). In this way the second question is intended as a list of possible different speech acts: assertion, question and command. In this list, only the first one is compatible with an evaluation in terms of truth conditions, while the others have to be valued in terms of other properties given by their use «in a form of life». Here the form of life is intended as the external context that is added to the language-game. In this picture the language-game plays the role of the set of linguistic activities (all that is literally asserted, questioned, commanded) and the form of life is the entire environment in which this activity takes place (Cavell 1979: 168–180).

1.2 The Language Game as a Total Pragmatic Act

The second, broader, interpretation considers language-games as being the total act of communication in which language, gesture, tone of voice and all features belonging to the context of a sentence are embedded together, since it is impossible to divide the meaning of a single verbal sign from the other components connected to it. This account takes the "and" in the final sentence of PI §7 ("the whole consisting of language *and* its actions") not as a logical conjunction between different types but as an enumeration of the elements belonging to the same set, which only has to do with the different ways in which they are carried out.

From an exegetical point of view, this broader interpretation, contained in the works of Wittgenstein's second period, has the merit of giving a more exhaustive account of PI: §23. The standard reading by Baker-Hacker finds some obscure points in this remark and accuses Wittgenstein of being incoherent.

The following is the first problematic point: «There are *countless* kinds: countless different kinds of use of what we call "symbols", "words" "sentences"». According to the standard readers, this sentence contains a logical gap between the first part, in which the sentence is the subject, and the second in which «different ways of use» becomes the matter: talking about sentences is not the same as talking about different usages of signs (Baker and Hacker 1983: 87). But this is correct only if we presume that Wittgenstein is trying to open up the concepts of sentence and proposition to show their inherent vagueness. But why don't we take Wittgenstein literally? Why can't we read this passage as a way of substituting the units of semantic analysis? We could take this passage as the point of departure of a new vision of semantics. The minimal unit of semantics becomes the usage of the sign.

Thanks to this first step, the broader interpretation is able to illuminate another passage that seems to be incoherent in the standard reading. In PI: §23, Wittgenstein writes a list of language-games: some of these are played with words (giving orders, describing, reporting an event, cursing, thanking) while others are played with other systems of notation (the mathematical code for solving a problem in arithmetic) and some are played with more than a system of signs (obeying orders, forming and testing a hypothesis, making a joke and telling it) or could simply be played with gestures and intonations without the use of words (play-acting, requesting). Baker and Hacker (1983: 87) do not find a single criterion for classification of the games, highlighting the opacity of the passage: there is no clear and unique criterion of classification because not all of the games could be reduced to sentences; not all of the games are speech acts. But all of the language-games are indeed pragmatic acts, Wittgenstein's intention was not to reduce the language-game to a sentence in verbal language but instead to take a larger point of view on language in its being primarily a tool for communication and not a code for calculation. The most important feature of this broader notion of game consists in taking the whole constituted both by words and context as the minimal unit of analysis.

1.3 Some Consequences of the Broader Interpretation

The most important features of the widened version of language game are the following: topicality, the property of game of being directed to a certain goal; broader normativity, meaning the game property to create its own rules of understanding; multimodality, taken as the property of symbols to mingle different systems of notation in a holistic union.

1.3.1 Topicality

Let us consider the following remark from PI: §69 in which Wittgenstein is advancing his family resemblance doctrine:

> How should we explain to someone what a game is? Imagine that we should describe *games* to him, and we might add: "This and *similar things* are called 'games'." And do we know any more about it ourselves? Is it only other people whom we cannot tell exactly what a game is?—But this is not ignorance. We do not know boundaries because none have been drawn. To repeat, we can draw a boundary— for a special purpose (*Zweck*) (PI §69).

The concept could be defined only contingently for a special purpose, only when it is taken in a particular context. According to this idea, the semantic properties of a certain element can only be given starting from its token occurrence in a certain context characterised by a specific point. Michael Pelczar has dubbed this feature of the language-game as *topicality* and defined a topical expression as «one that expresses different contents in different contexts of use, in such a way that *which* content is expressed in any given use depends on what is under discussion, in the context of that use» (Pelczar 2000: 487). One example of a topical expression is the use of a polyseme such as 'to get'. We could use an expression like «Smith has got the virus» that could be taken as expressing different kinds of contents: in a conversational setting it could be understood as a synonym of «Smith has contracted the virus», but if the same conversation is going on between the members of a biological research team it could be understood as «Smith had successfully collected a sample of the virus» (Pelczar 2000: 492). The topic plays the role of cutting the edges of the game and in a way it affords the same definition of the game as a certain kind of game. But at the same time it is just that *kind* because of the particular game we are playing. To put it briefly: we have to think of the topic as a particular type that precedes the game and gives it its form; for it is the particular game in its context of use that creates its peculiar topic, namely the conditions for being understood.

1.3.2 The Normativity of Bunga Bunga

Consider for a moment a certain expression that has been quite trendy recently in everyday conversations and newspapers: *bunga bunga*. Silvio Berlusconi, Italy's prime minister, was recorded in a tapped phone conversation with some young girls talking about a particular practice called *bunga bunga*. Previously, this peculiar expression never entered into Italian use and no-one was acquainted with it before its first occurrence in the media, and almost no one had any understanding of what *bunga bunga* could be—also because up to that point there was no clear definition of what it was and in what it consisted. Here, we have an expression that creates its rules of understanding on the grounds of some internal characteristics. First, the iteration of *bunga* gives a certain *rhythm* to the expression, in which the

repetition of the word *bunga* is like a dance. Second, the word *bunga* is related *phonetically* with other terms like bongo that, in the usual trade of language, are associated with something primitive and akin to animal instinct. A third kind of knowledge helping to understand the expression concerns Berlusoni's passion for women and his unscrupulousness.

All of these marks are internal to the game played with the expression *bunga bunga* and constitute the field of conditions for its understanding. The game constitutes its peculiar topic in the sense that it has the proper points necessary to appreciating it, provided internally. It is not so much the question of the reference to the expression *bunga bunga*, but the fact that a language expression, without any precise reference, can become perfectly meaningful. In this there consists this particular language game, i.e. in the significant use of something that previously had no meaning. The *bunga bunga* case is a borderline one able to highlight the working of other language-games in which the norms are established in a linguistic community. As in the case of assertions and performatives.

This first characteristic of the language-game is also illuminating for a second point: the revision of the concept of normativity. The fact is that it is not a rule that is applied for understanding a speech act, but on the contrary the rule is determined during the interaction of the speakers. What has been effectively understood can only be described successively.

For a moment, let us take the following sequences of numbers: "1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 19, etc.", and a second series: "1, 3, 5, 7, 9, 11, 12, 24, 36, 48".

What is the rule followed in the first and second series? Suppose that we are still learning to count in a foreign language—for example we are Italians learning English and our teacher has asked us to count. Did we understand the order correctly? Surely the most intuitive and plausible answer to our questions would be: NO! But imagine instead that we are on an Alitalia flight from Rome to Palermo: if we are uttering the first series we are following the series of numbers of the air-plane's rows of seats. The numbers 13 and 17 are not counted in the series due to a superstitious practice.

Regarding the second series, we could have a different language-game, played with roses: in Italy there are rules according to which roses have to be given as a present only in odd numbers up to a dozen.

We are dealing with the understanding of different rules of two different language-games in which competence is given not by simple obedience to a material input but by the ability *to guess* the topic of the game. This kind of normativity gives form to the entire game and resembles what Sellars dubbed pattern-governed behaviour. Sellars maintained that to learn this kind of behaviour "is to become conditioned to arrange perceptible elements into patterns and to form these, in turn, into more complex patterns and sequences of actions" (Sellars 1954: 209). This kind of rule-following behaviour is affected by any kind of perceptible element in the environment, be it a verbal sign or just a tone of voice or gesture. As ethnographic approaches have suggested, linguistic competence should be extended to a larger set of abilities then the calculus model of language suggested.

The rules of the language-games are not given in advance but are always work in progress: it is our understanding the rose series, and our *agreeing* to play that particular language-game, that is constitutive of the rule, and not vice versa.

From a semantic point of view, this means that there is no intrinsic set of semantic properties stated a priori, and any kind of property depends on the pragmatic decision to play a certain kind of game. Wittgenstein enumerates the characteristic accompaniments of a word (gesture, faces, tones of voice) and then concludes that a decisive judgement on a word could be given only considering its «field of force (*Feld* eines Wortes)» (PI: 186). Semantic judgement can be passed only once the totality of the language-game has been considered. The rules constitutive of this game are dynamic and are not always made explicit: how could a verbal tip be expressed? Or how could a certain tone of voice be translated logically?

The second feature of broadly conceived language-games is broader normativity, as is clearly expressed in the following passage:

> Can one learn this human knowledge (*Menschenkenntnis*[3])? Yes, some can. Not, however, by taking a course in it, but through '*experience*'.—Can someone else be a man's teacher in this? Certainly. From time to time he gives him the right *tip*.—This is what 'learning' and 'teaching' are like here.—What one acquires here is not a technique; one learns correct judgements. There are also rules, but they do not form a system and only experienced people can apply them rightly. Unlike calculating rules (PI: 193).

The subject involved in the game is more than a speaker/hearer: he/she is the partner of a communicative interaction based on his ability to act in a pattern in which the rules are to be guessed thanks to a process of abduction taken from the material present in the game. Take the second series: there is no verbal rule, not even any kind of situational feature that could solve the problem of how to interpret the sentence. There is no inference in understanding that series but a process of abducting the point from the co-text and the common and un-explicit ground occurring between the partners.

1.3.3 Multimodality

The third characteristic of the language-game is multi-modality. Although we tend to think of the speech act as a sentence which can be expressed through an utterance or writing, speech acts use different expressive resources embedded within each other: spoken language, tone, gesture and so on. There is no utterance that does not use the resources afforded by prosody or by the gestures that always accompany vocal emissions or the use of sign languages.

---

[3] The German expression is curiously translated with a certain range of terms: Anscombe chooses «knowledge of mankind», Hacker and Sculte «knowledge of people», while the Italian translator «conoscitori degli uomini». We believe that each of these translations in a way deprives the original German term of its capacities of characterizing the knowledge about which Wittgenstein is talking as specific to a certain animal species.

The role of the gesture and of supra- and sub-segmental traits are often neglected (Gumperz 1992; Albano Leoni 2009), not only by semantic studies but also by the pragmatic approach to linguistic analysis; the idea of context seen as a background in which the linguistic act is inserted has, in fact, neglected characteristic traits of the linguistic act (prosody, gestures and so on). If we really want to consider an anthropologically oriented paradigm, we must keep in mind the original multidimensionality of the linguistic act, synaesthetic and multi-modal, and that it is in this way that the linguistic act presents itself in human interactions, strictly interrelated with gesture and intrinsically prosodic:

> What is the primitive reaction with which the language game begins[…] The primitive reaction may have been a glance or a gesture, but it may also have been a word (PI, II:185).

Gestures carry out functions similar and complementary to those carried out by words, if we consider the role played by words with the necessary openness. Wittgenstein goes beyond the identification of elements belonging to a new paradigm, and focuses his attention on a very widespread (but volatile by nature and perhaps ungraspable) phenomenon, interrogating himself about the situation of the «right word». A volatile concept, but not for this reason unimportant; the word has in fact an atmosphere. In *Last Writings on the Philosophy of Psychology,* Wittgenstein asks (LW: 726) how we can say that two terms, even though they refer to the very same kind of object, have the same meaning. Johnston (1993: 101) writes: one might talk of each word having its own 'face' or atmosphere, so that the non standard word 'knoif' has a different atmosphere from the word 'knife'»

Take the following passage (PI, II: 186)

> How do I find the 'right' word? How do I choose among words? Without doubt as if I were comparing them by fine differences of smell: *That* is too…, *that* is too…,—*this* is the right one—. But I do not always have to make a judgement, give explanation; often I might only say: "it simply isn't right yet".

In these cases Wittgenstein continues to ask what can I do to find that word? «I act it.—But *what* can I learn in this way? What do I reproduce?—Characteristic accompaniments. Primarily: gesture, faces, tone of voice». Again, the gesture and all the stuff are taken to be part of the symbol. But what is a characteristic accompaniment? Once, it was said that an accompaniment is part of non-verbal communication, a pretty vague term with which the nonessential background colouring of verbal communication is indicated, or in a larger sense with non-verbal communication, the involuntary communication of bodies is intended, for example the physical attraction between two individuals, or proxemic phenomena. Instead, for Wittgenstein, characteristic accompaniments are traits internal to the linguistic game.

The sign language equips us with an extraordinary element for reflection on how language is always situated in a context and a body. The sign languages used by deaf people are also always accompanied by gestures that have an important

function, which is not limited to colouring the linguistic sign, but is an integral part of any communicative interaction. McNeil (2005) has made the role of *catchment* evident: during a conversation gestures are made with recurring characteristics, which have the function of showing and at the same time reaffirming the cohesion of the discourse, underlining a particular theme. In this sense, McNeil's proposal of a *growth point,* a minimum speech-gesture unit which we need to account for an effective linguistic interaction (even if, as McNeil maintains, gestures and speech refer to different semiotic systems with different expressive possibilities) appears to be of great interest for every linguistic analysis. Vocal languages, like sign language, should be understood as multi-modal systems, that is to say as systems of systems, including the gesture in the concept of language. Taking the step from this presupposition, we account for how languages in general are systems of heterogeneous values, of which, because of the strong influence of writing, we can only see a linear level (Fontana 2010; Cuccio and Fontana 2011).

Yet again, we realise how the model underlying our linguistic analysis is the textual corpus of written language in which linguistic expressions follow one another in an ordered fashion, which begins and ends with the «.», each of them articulated in words and then in letters to supply a great model of reflection. But this is a use situated in a particular context and certainly not the way in which we use language during the majority of our verbal interactions.

Let us consider some examples. Two lovers say sweet nothings to each other with that strange tone of voice used in these cases, a sort of baby-talk in which the configuration that the mouth assumes to express those sounds is accentuated. They are words that are certainly not interchangeable with other expressive forms. And Wittgenstein asks himself (LW: 712) «Isn't it perhaps because they are gestures?» This is an interesting case in which the context does not come before the action, but it is the action that generates the right context of understanding. It is the language full of feeling that activates the context of reference through which it is then possible to understand the linguistic game.

Or again, let us take the example of the verb «to get the virus» already used. Two speakers, say X and Y, are working together in a biological team and X says to Y: «Z has got the virus» with a vexed tone of voice, whispering it in Y's ear. All of these gestures are not a simple accompaniment for the game, but are integral parts of the pattern which they are (creating) following. All of the gestures, words and tones of expression used by the two speakers codetermine the context of the game that we could take as defined by its particular topic: talking about how Z contracted the virus. The sentence could be determined only insofar as it was just determined by the context of the game and its peculiar purpose.

2 Beyond the Sentence: Pragmeme and Language Game

The broader interpretation of the language game leads linguistic analysis to embrace an ethno-anthropological stance toward language that, in current semantic studies, could be found in the notion of pragmeme as it is presented by Mey. In both cases the usage of the sign becomes the minimal unit of semantics. Now this is something familiar for the pragmeme theorists,[4] who, following Mey, maintain the need to abandon the speech act for the totality of the pragmatic act (Capone 2009, 2010a, b), that is a holistic unity of a textual part, the so-called co-text, and contextual features given by the activities of the participants in the interaction. As Mey writes, what counts as a pragmeme "depends on the understanding that the participants have of the situation and on the outcome of the act in a given context" (Mey 2009: 752).

This approach takes the speech act and the assumption that «no speech act is viable by itself, speech acts as such do not exist, unless they are situated» (Mey 2010: 2882) as its point of departure. Starting from this point, Mey changes the relationship between pragmatics and semantics, showing how any speech act should be understood in the context of the concrete situation of use in which it is rooted. Mey's criticism of the traditional speech act analysis is based on the assumption that semantic analysis is generally based on speech, or at least on verbal output. Starting from this assumption of language linearity or mono-modality, the context could only be something added to the words. The pragmatic turn proposed by Mey and other contextualists consists in taking language as a multi-modal activity in which speech acts contribute to the creation of the situation in which they are carried out (Mey 2009: 750): the speech act is the totality constituted by words, gestures, body movements, tone of voice and the entire social setting or social environment of the conversation. This deep switch authorizes Mey to change the unit of semantic analysis, which is no longer the sentence, but the pragmeme, seen as the prototype of a certain speech act; to give a semantic description means to characterize the «typical, pragmatic act as it is carried out in a given situation» (ibid: 751).

The pragmemic interpretation of utterances applies an ethnographical methodology to the study of meaning according to which the role played by an utterance could only be understood as embedded in a socio-cultural context in which it co-occurs together with societal rules, practical norms of behaviour and any kind of knowledge that the participants in the conversation share with each other (Capone 2005). This integrated approach to the study of meaning in context defines the pragmeme as «a situated speech act in which the rules of language and society synergize to determine meaning» (Capone 2005: 1357). To give an example, according to Capone a single utterance like «I saw you!» could be

---

[4] But not only for them, an attempt to think about semantics with a different unit of analysis is also in Jaszczolt's *Default Semantics*, a research program based on discourse analysis that hold, acts of communication to be the proper subject of analysis (Jaszczolt 2005: xv).

understood in a different way according to the context of utterance: as an 'assertion' or as a 'threat' to someone. In any case we have two pragmemes, that is two different types of pragmatic acts. The pragmeme is considered as a sort of frame or schema that organizes the understanding of a pragmatic act, or as Capone (2005: 1360) puts it, pragmemes are «sequentially organized algorithms that reshape the original illocutionary force of a speech act by adding contextual layers of meaning, or even may change the illocutionary value of the speech act».

The interpretation of the pragmeme seems to fluctuate between two different poles: from a certain point of view the pragmeme is situated and creates its own situation, from the other pragmeme is the class in which individual speech acts are to be placed. An example of the first interpretation is given by Mey's analysis of a Marlboro advertisement. The message works in a multimodal way; gestures, landscape and objects are all connected to give a unique message that is situationally marked: «(the pragmatic act) is all contained in the way the situation is set up» (2010: 2883). An example of the other 'typological' interpretation is given by Capone's definition mentioned above according to which the pragmatics of the acts can in any case be given in a standardized way.

The pragmeme theory and the Wittgensteinian notion of the language-game both share an intuition of the multimodal aspect of language, which has to be taken as a non-linear totality in which speech, gestures and situation are to be taken together; but at the same time Mey's approach has to be reconsidered in the light of its assumption of pragmemes as proto-types to which a certain speech act, taken as a token, could be linked, or, as Capone puts it, as a formal schema of rules that is useful for mingling together contextual features and language. This distinction between a certain type and its tokens is very familiar in traditional approaches to semantic analysis but is also very problematic for its essentialist stance, which clashes with the so-called family-resemblance doctrine contained in Wittgenstein's *Philosophical Investigations*. We do not have enough space here to go into detail about the rejection of the essentialist approach by Wittgenstein, but we take this rejection as a fundament for the appreciation of the concept of the language-game and its semantic interpretation.

2.1 The Limit of the Pragmatic Act

Our aim is to work out the broader conception in a dialectical relationship with Mey's view of the pragmeme. In particular, we are trying to solve the problem of the relationship between type and token by presenting a reading of PI § 69, contained in the family resemblance discussion, thanks to which it will be possible to have a new idea of the normativity of language. According to our reading, the language-game overlaps the distinction between type and token because it is either taking place in the present moment or in the potential one: at the same time as it is being carried out in the present with its concrete use, it has a normative status derived from its being related to other possible ways of language use.

This reading of the language-game in broader terms could enable us to solve a problem through the use of pragmemes in semantic analysis: how can we draw a limit around the pragmatic act? Where does it begin and where does it end? This is a common problem both for pragmeme and language-game.

Our answer to these questions is that the semantics of the game, i.e. the properties expressed by the game, the kind of knowledge that is earned by playing that particular game is gained thanks to a process of abduction made using symbolic material in which cotextual and contextual features are taken together. This abduction is afforded by the existence of a particular purpose that is an 'edge-drawer.' The game is played as long as that kind of purpose has to be pursued.

To be able to define the semantics of a sentence, we should be able to appreciate the pragmatic dimension of the whole language-game to which a particular verbal sentence belongs: either by understanding its appropriate truth conditions or by giving a sketch of its contextual properties.

3 How to Draw a Boundary: A Proposal

This part is devoted to showing how the language game could be applied in describing the pragmatic effort required by each of the partners in the communication process in arriving at the meaning of certain act. The first is a polysemous Italian expression that is intended to show the impossibility of stating the meaning by a certain expression autonomously of the context of utterance. The second game develops the insight contained in Pelczar's theory of topicality and is meant to shed light on the problem of the boundary of a certain language game. Both the former and the latter require a certain background of contextual knowledge and institutionalized systems of notations that are merged with the situational set of properties in which the partners of communication take part.

Let us consider the following Italian sentence:

(1) *Ci vediamo*!

A sentence of (1) in Italian could be used as a typical form of greeting equivalent to '*see you soon*' in English and could be taken as belonging to the pragmeme GREETING, constituted by other allopracts like: *ciao* (informal), *buongiorno*, *buonasera* (more or less formal). The language-game of greeting can't be understood in terms of an evaluation of truth conditions, but instead should be evaluated on the basis of the appropriateness of the words spoken in the language game. For example, in daily Italian life, these words are used informally between friends that are used to each other's company. These properties belong to the game and are grammatically triggered and inferred by any competent speaker, and in a sense these properties are literally associated with the words in (1). We would

define these properties as word properties triggered by the kind of words used in a certain language game.

The example (1) could also be taken as a language-game of asserting something. For example, consider a situation in which the whole city has suffered a power cut while I'm in a village just outside the city and my sister calls me asking if we have 'light' (in Italian, *luce*, 'light', is commonly used to refer to electricity). Now I could assert something like *Ci vediamo!*, literally 'We can see', which is the equivalent of saying *We have light!*. Literally the language-game I'm playing is that of asserting something that has a definite content that is evaluable as true or false. So sentence (1) has a first set of semantic properties, which are the set of truth conditions of the sentence that constitutes the minimal unit of analysis:

a) *We have light*!

This first semantic nucleus of properties, roughly corresponding to what is said by the sentence, has to be enriched with another set of properties because the language-game has to be situated in a certain scene or form of life. So (a) contains a sort of hidden indexical reference:

(a1) *We have light* [here, in the village near the city]!

The hidden indexical contained in square brackets is triggered by the same structure and composition as sentence (a) so that it belongs to the language game: it is a property of the game triggered by the words spoken in it, but at the same time it is a different kind of property with respect to the truth T. We could call this kind of property, scene properties: these are the properties of a certain linguistic act associated *ceteris paribus* with a certain prototypical scene.

What is important about this language-game, is that it appears polysemous only when we consider the sentence *Ci vediamo* as distinct from its context, as if words taken by themselves had a special meaning on their own, but taken in a larger context of use the language game is uncontrovertibly clear to the partners that take part in it and understand the point of the game in which they are involved.

Take the example of the verb «to get the virus» already used. Two speakers, say John and Mary, are working together in a biological team and we imagine that John says to Mary «Peter has got the virus» with a vexed tone of voice, whispering it in Mary's ear. All of these gestures are not a simple accompaniment for the game, but they are an integral parts of the pattern which they are (creating) following. All of the gestures, words and tones of expression used by the two speakers codetermine the context of the game that we could take as defined by its particular topic: talking about how Peter contracted the virus. The sentence could be determined only insofar as it was determined by the context of the game and its peculiar purpose.

We could represent our language-game with the verb «to get a virus» with a schema that tries to rethink the model of representation of a pragmeme proposed by Mey (2009: 752). In that model, Mey draws a sharp distinction between two different parts of a pragmatic act: on the left side he puts the activity part (which includes any

kind of element belonging to the conversational settings) and a textual part on which inferences could be drawn. The pragmeme is presented as being simply a blend of these two parts, but two issues are still open: How is it possible to move from these two parts to the unity of the act, and what kind of process is going on? And a second point: How do we draw a boundary between a pragmeme and another?

Partners:{John, Mary}
Cotext: {*Peter*, *to get the virus*; whispering, tone of voice}
Situation point: talking about contracting a virus
[Abduction]
Semantics of the game: T + C:
T (truth value): {Mary; had contracted the virus}
C (contextual properties): {general knowledge about the world; possess of the common sensory capacities; particular knowledge of a laboratory situation}.

In this proposal, which is just a way of trying to advance a possible solution to the open problem of the limit of a pragmatic act—and consequently of the possibility of identifying a certain act as *that* pragmatic act—the key feature is represented by the focus on the point or goal of the linguistic activity. This is something that should be worked out better and here we only give a first step towards a future development of the pragmatics of language-games.

4 Conclusions

Our work has tried to make the potentiality inherent in the concept of the linguistic game evident by taking it back to its original context in the work of Wittgenstein. The widened concept of the language-game as we have presented it could help to arrive at a better understanding of the problems left open by the pragmeme, as it is defined by Mey; but most of all it demonstrates the opportunity for a total revision of the instruments with which the problem of meaning is currently addressed. This is the most appealing part of Wittgenstein's heritage and at the same time the most complex: the necessity of finding appropriate instruments for describing the meaning of multimodal, non-linear symbols that break with the traditional analytic vision of the study of semantics inherent in the identification of the minimum unit of discourse. Language-games as a pragmatic acts are not ideal type of linguistic action but every time bring into play the partners' ability to recognise or create their own rules. The holism of the linguistic symbol, its being a totality constituted by various parts, represents Wittgenstein's principle lesson to us with his treatment of linguistic games.

References

Albano Leoni, F. (2009). *Dei sensi e dei suoni. Il volto fonico delle parole*. Bologna: Il Mulino.

Bach, K. (2005). Context ex machine. In Zoltan Szabó (Ed.), *Semantics versus Pragmatics* (pp. 15–44). Oxford: Oxford University Press.

Baker, G., & Hacker, P. M. S. (1983). *Analytical Commentary of the Philosophical Ingestions. Understanding and Meaning vol.1. Exegesis*. Oxford: Blackwell.

Black, M. (1979). Wittgenstein's language-games. *Dialectica, 33*, 337–357.

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics. A defence of semantic minimalism and speech act pluralism*. Oxford: Blackwell.

Capone, A. (2005). Pragmemes (a study with reference to English and Italian). *Journal of Pragmatics, 37*, 1355–1371.

Capone, A. (Ed.). (2009). *Perspective on language use and pragmatics*. Munchen: Lincom.

Capone, A. (2010a). On the social practice of indirect reports (further advances in the theory of pragmemes). *Journal of Pragmatics, 42*, 377–391.

Capone, A. (2010b). Pragmeme special issue. *Journal of Pragmatics, 42*, 2861–3162.

Cavell, S. (1979). *The claim of reason*. Oxford: Oxford University Press.

Cuccio, V., & Fontana, S. (2011). Spazio cognitive e spazio pragmatico: riflessioni su lingue vocali e lingue dei segni. *Esercizi Filosofici, 6*, 133–148.

Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.

Fontana, S. (2010). *Linguaggio e multimodalità*. Pisa: ETS.

Glock, H.-J. (1996). *A Wittgenstein dictionary*. Oxford: Blackwell.

Gumperz, J. (1992). Contextualization and understanding. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: language as an interactive phenomenon* (pp. 229–259). Cambridge: Cambridge University Press.

Jaszczolt, K. (2005). *Default semantics. Foundations of a compositional theory of act of communication*. Oxford: Oxford University Press.

Johnston, P. (1993). *Wittgenstein. Rethinking the inner*. London: Routledge.

Kaplan, D. (1978). On the logic of demonstratives. *Journal of Philosophical Logic, 8*, 81–98.

Lo Piparo, F. (2009). Gramsci and Wittgenstein: An intriguing connection. In A. Capone (Ed.), *Perspective on language use and pragmatics* (pp. 285–320). Munchen: Lincom.

McNeil, D. (2005). *Gesture and thought*. Chicago: The University of Chicago Press.

Mey, J. (2001). *Pragmatics*. Oxford: Blackwell.

Mey, J. (Ed.). (2009). *Concise encyclopaedia of pragmatics*. Oxford: Elsevier.

Mey, J. (2010). Reference and the pragmeme. *Journal of Pragmatics, 42*, 2882–2888.

Pelczar, M. (2000). Wittgensteinian semantics. *NOUS, 34*, 483–516.

Perry, J. (1978). The problem of essential *Indexical*. *NOUS, 13*, 3–21.

Recanati, F. (2006). Pragmatics and semantics. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 442–462). Malden, Oxford: Blackwell.

Sellars, W. (1954). Some reflections on language-game. *Philosophy of Science, 21*, 204–228.

Works of Ludwig Wittgenstein

BlB. The Blue Book in *The Blue and Brown Book*.

BrB. The Brown Books in *The Blue and Brown Book*.

The Blue and Brown Books, ed. R. Rhees, Oxford: Blackwell, 1958.

BT. The Big typescript TS 213: German-English Scholars' Edition, edited and translated by C. G. Luckhardt and M. E. Aue. Oxford: Blackwell.

EPB. Eine Philosophische Betrachtung in *Das Blaue Buch. Werkausgabe Band 5*, Frankfurt am Main: Suhrkamp, 1969.

LWPP. Last Writings on the Philosophy of Psychology, eds. G.H. von Wright and H Nyman, Oxford: Blackwell, 1982.

PG. Philosophical Grammar, ed. Rush Rhees, Oxford: Blackwell, 1969.

PI. Philosophical Investigations, eds. G.E.M. Anscombe and R. Rhees, Oxford: Blackwell, 1953, 2001[3] (ed. Hacker-Schulte, 2009).

TLP. Tractatus Logico-Philosophicus (1921), translated by D. F. Pears and B. Mc Guinness, London: Routledge, 1975. Critical edition: *Logish-philosophische abhandlung/Tractatus Logico-Philosophicus*, ed. By B. Mc Guinness and J. Schulte, Frankfurt: Surkamp, 2001.

Z. Zettel, ed. G.E.M. Anscombe and G.H. Von Wright, Oxford: Blackwell, 1967.