Studies in Theoretical Psycholinguistics 45

Florian Schwarz *Editor*

Experimental Perspectives on Presuppositions

Springer

STUDIES IN THEORETICAL PSYCHOLINGUISTICS

Volume 45

The goal of this series is to bring evidence from many psychological domains to the classic questions of linguistic theory. The fundamental question from which the others flow is: What is the mental representation of grammar? Evidence from all aspects of language are relevant. How is the grammar aquired? How is language produced and comprehended? How is the grammer instantiated in the brain and how does language breakdown occur in cases of brain damage? How does second language acquisition and processing differ from first language acquisition and processing? A satisfactory theory of language calls for articulated connections or interfaces between grammar and other psychological domains. The series presents volumes that both develop theoretical proposals in each of these areas and present the empirical evidence needed to evaluate them.

More information about this series at http://www.springer.com/series/6555

Florian Schwarz
Editor

Experimental Perspectives on Presuppositions

Springer

Editor
Florian Schwarz
Dept. of Linguistics
University of Pennsylvania
Philadelphia
Pennsylvania
USA

ISSN 1873-0043
ISBN 978-3-319-07979-0 ISBN 978-3-319-07980-6 (eBook)
DOI 10.1007/978-3-319-07980-6
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014946623

Printed on acid-free paper

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

Contents

Contributors

Patrícia Amaral Department of Romance Languages, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

David Beaver Department of Linguistics, University of Texas at Austin, Austin, USA

Sigrid Beck Descriptive and Theoretical Linguistics, University of Tuebingen, Tuebingen, Germany

Dylan Bumford Department of Linguistics, NewYork University, NewYork, NY, USA

Chris Cummins Department of Linguistics and English Language, The University of Edinburgh, Edinburgh, UK

Emilie Destruel Department of French and Italian & FLARE, University of Iowa, Iowa City, USA

Rachel Dudley University of Maryland, College Park, MD, USA

Valentine Hacquard University of Maryland, College Park, MD, USA

Jean-Baptiste van der Henst L2C2, CNRS, Lyon, France

Ingo Hertrich Ingo Hertrich, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tuebingen, Germany

Jacques Jayez ENS de Lyon and L2C2, CNRS, Lyon, France

Manizeh Khan Department of Psychology, Harvard University, Cambridge, USA

Christina S. Kim University of Kent, Canterbury, UK

Mareike Kirsten Collaborative Research Centre 833, University of Tuebingen, Tuebingen, Germany

Jeffrey Lidz University of Maryland, College Park, MD, USA

Valeria Mongelli Institut du Cerveau et de la Moelle Épinière Research Center, Institut National de la Santé et de la Recherche Médicale, Paris, France

Edgar Onea Courant Research Centre "Text structures", University of Góttingen, Nikolausberger Weg 23, Góttingen, Germany

Naho Orita University of Maryland, College Park, MD, USA

Anne Reboul L2C2, CNRS, Lyon, France

Bettina Rolke Evolutionary Cognition, University of Tuebingen, Tuebingen, Germany

Jacopo Romoli School of Communication, University of Ulster, Jordanstown Campus, Newtownabbey, UK

Department of Cognitive Science, Macquarie University, Sydney, Australia

Florian Schwarz Department of Linguistics, University of Pennsylvania, Philadelphia, PA, USA

Jesse Snedeker Department of Psychology, Harvard University, Cambridge, USA

Yasutada Sudo Linguistics, UCL, London, United Kingdom

Linguistics and Philosophy department, MIT, Cambridge, USA

Sonja Tiemann Collaborative Research Centre 833, University of Tuebingen, Tuebingen, Germany

Daniel Velleman Department of Linguistics, University of Texas at Austin, Austin, USA

Jingyang Xue Courant Research Centre "Text structures", University of Góttingen, Nikolausberger Weg 23, Góttingen, Germany

Introduction: Presuppositions in Context—Theoretical Issues and Experimental Perspectives

Florian Schwarz

Abstract A central issue in semantics and pragmatics is to understand how various different aspects of meaning contribute to the overall conveyed meaning of an utterance. Asserted content, implicatures, and presuppositions are commonly assumed to differ in terms of their source, their status, and their interaction with the context in the theoretical literature. This chapter starts with a brief introduction of this theoretical background, and then reviews the experimental literature on related phenomena in some detail, with a focus on presuppositions. In the course of this, the contributions to the present volume are situated in the context of previous work. The conclusion provides an outlook on future directions.

Keywords Presuppositions · Processing · Implicatures · Experimental pragmatics · Triggering · Presupposition projection · Acquis · Accommodation · Meaning in context

1 Introduction

The study of natural language meaning within formal linguistics has its roots in philosophy of language and logic, and this tradition sees truth-conditions at the center of the study of semantics. As Heim and Kratzer (1998) put it in the first sentence of their influential textbook: 'To know the meaning of a sentence is to know its truth conditions.' One of the central concerns of current semantic research then is to characterize natural language phenomena in truth-conditional terms. While most theorists might agree that truth-conditions are the core of linguistic meaning, it is also clear—and widely accepted—that there are further aspects of the overall conveyed meaning of a sentence uttered in a context that go beyond what I will refer to as the conventionally encoded (or literal) *at-issue* content. The latter is directly contributed by the lexical entries of the expressions involved and constitutes the central message the speaker wishes to convey (see also Roberts's (1996) notion of

F. Schwarz (✉)
Department of Linguistics, University of Pennsylvania, 619 Williams Hall,
255 S. 36th Street, Philadelphia, PA 19104, USA
e-mail: florians@ling.upenn.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_1,

proffered content). The additional layers of meaning come in various sorts. They can be a part of the conventionally encoded meaning for a given lexical item but not contribute to the *at-issue* meaning. Alternatively, they can result from general reasoning about communicative situations (and be *at-issue* or not). Finally, they can arise through an interaction of conventional content and general reasoning. In characterizing these phenomena in these terms, it is common to divide up the work in one way or another between related but distinct components of the language comprehension system, namely semantics and pragmatics, where the former crucially involves conventionally encoded content, whereas the latter depends (at least in large part) on information from the context of utterance. The lines between the two are by no means agreed upon, and for many phenomena, key arguments in the literature are precisely about the way in which a given expression gives rise to contributions to the overall meaning as well as the nature of that contribution.

The 'overall conveyed meaning' of a given sentence uttered in context can then be seen as a conglomerate of inferences. For naive speakers, there is no simple way of divvying up those inference in terms of what their source or status is. It is the job of the theorist to come up with criteria for differentiating what inferences are introduced in what way, and to identify the corresponding properties of the relevant aspects of meaning. Providing theoretical arguments for differentiating between distinct aspects of meaning and identifying their key properties—both on their own and in contrast with other aspects—has always been a central concern in the field. However, until very recently, the empirical scope of these investigations has been limited in a number of ways, and few efforts were undertaken to study the relevant questions systematically with tools from experimental psycholinguistics. Similarly, much of the research has focused on English (and a couple of other languages), without much of a perspective on cross-linguistic comparisons. But over the course of the last decade or so, a shift has started to take place, with more and more researchers bringing together experimental approaches and theoretical questions about linguistic meaning, as well as exploring details of semantic phenomena across a wider set of languages.

A large share of experimental work on meaning from a linguistic perspective indeed has focused on comparing different aspects of meaning.[1] The primary focus has been on implicatures (and amongst those, primarily scalar implicatures) in the sense of Grice (1975) (see discussion below). But other aspects of meaning lend themselves to similar investigations, and ultimately we will want to inform our theoretical considerations using empirical and experimental investigations that explore the relevant phenomena to the fullest extent possible. The present volume presents recent work that extends experimental approaches to another central aspect of meaning, namely that of presuppositions, and it is—to the best of my knowledge—the first to focus on that topic specifically. While by now there is a small but growing body of experimental work on presuppositions, the hope is that a focused presentation of

[1] For surveys of psycholinguistic work on semantics more generally, see Frazier (2012) and Pylkkänen and McElree (2006).

current results within one volume will help to galvanize efforts in this area further, both by serving as a point of reference on the current state of research and as a starting point for future investigations.

This introductory chapter is intended to situate the overall endeavor that the later chapters contribute to in a broader context by providing a brief overview of the basic theoretical background and reviewing existing experimental work from the literature and related theoretical issues. I conclude by taking stock of the current state of research and reflecting on future directions.

2 Background: Ingredients of Meaning

Early work in the tradition of philosophy of language from a logical perspective regarded natural language as deficient and messy, as many aspects of the way it conveys meaning did not squarely fit into simple logical analyses. However, at least since the middle of the twentieth century, the prima facie non-logical aspects of language have been taken to be the subject of rigorous formal analysis themselves. One of the earliest relevant discussions involved the analysis of definite descriptions. One line of thought, originating with Frege (1892) and later taken up by Strawson (1950), argued that definite descriptions come with an existence and uniqueness requirement. If either requirement was not met (as in *The king of France is bald.*), the sentence would not receive a truth-value. In other words, these requirements were taken to be a type of pre-condition, or presupposition, for truth-evaluable sentences that could felicitously be uttered in a given context. In contrast, Russell (1905) argued them to simply be part of the literal at-issue content, so that failure to meet them would result in simple falsity.

While subsequent debates about definite descriptions continue to this day, it soon became clear that there is a wide range of expressions that systematically exhibit properties similar to those found with definites. In particular, presuppositions are commonly assumed to have two key characteristics (Karttunen 1973; Stalnaker 1973, 1974). First, they (at least typically) do not convey any new information, but rather consist of backgrounded information that the interlocutors take for granted (at least for purposes of the conversation). Secondly, they remain present at the level of the overall sentence even when introduced under embedding operators that (loosely speaking) cancel the literal *at-issue* content. Consider the example in (1), where the verb *stop* introduces the *presupposition* that whatever activity is expressed by the verb phrase is something that has gone on prior to some contextually salient time in the past (introduced by the past tense), and the *assertion* that this activity was not going on after the relevant point in time.

(1) John stopped smoking.

Presupposed and asserted content behave very differently in so-called family-of-sentences environments (Chierchia and McConnell-Ginet 1990), which include variations such as the following:

(2) a. John didn't stop smoking.
b. If John stopped smoking, then he should be healthier now.
c. John might have stopped smoking.
d. Did John stop smoking?

None of the sentences in (2) convey what's asserted in (1)—that John's smoking did not go on after a salient point in time. They either convey the opposite (in the case of negation), or remain neutral in that regard. However, the presupposition—that prior to the relevant point in time John was smoking—remains constant across all variations. This global presence of the relevant inference, also referred to as 'presupposition projection', is a hallmark of presupposed content (But note that projection does not always result in the global presence of inferences; see Sect. 4.3).

Another important aspect of meaning is that of conversational implicatures (Grice 1975). A prototypical example is that of the scalar implicature associated with the quantifier *some*, illustrated in (3):

(3) John ate some of the cookies.
a. John ate some, and possibly all, of the cookies.
b. John did not eat all of the cookies.

Although naive speakers might commonly take (3) to entail (3b), it turns out that the literal meaning of *some* is best characterized as (3a), e.g., because it can be cancelled without any sense of contradiction (3i), in contrast to literal meaning (3ii):[2]

(3′) i. . . . In fact, he ate all of them.
ii. . . . # In fact, he ate none of them.

So how does the standard interpretation of (3) as implying (3b) come about? Grice (1975) proposed that it is based on the hearer's reasoning about the speaker's role as a rational, cooperative interlocutor that adheres to certain conversational maxims. In the present case, the maxim of Quantity requires speakers to be as informative as necessary, and the maxim of Quality requires them to speak truthfully. Based on this, a hearer can then reason as follows: the speaker did not make the logically stronger claim that John ate all of the cookies, despite being required to be as informative as possible. Assuming she is cooperative, there must have been something else that kept her from asserting the universal claim. The maxim of Quality is a prime candidate—she must have not been in a position to make the stronger claim while also observing this maxim. Thus, the hearer concludes that the stronger claim is not true.[3]

Yet another class of inferences distinct from the literal, truth-conditional content is that of conventional implicatures, a term also introduced by Grice (1975) (See Potts (2005), for an influential recent perspective). Conventional implicatures share some properties with presuppositions and some with implicatures, as can be illustrated with the examples in (4).

[2] Also see Destruel et al. (2014) for further discussion of cancellation tests.
[3] The actual reasoning is more involved than that, and as I hinted at, other maxims (such as Relevance) could come into play as well.

(4) a. Sue, a pianist, teaches music lessons on the weekends.
b. The damn cat knocked over a glass of water.

The notion that Sue is a pianist introduced by the appositive in (4a) is by no means something that had to be established prior to the utterance, but at the same time, it does not contribute to the main point of the assertion (in denying (4a), for example, one would not deny that Sue is a pianist). Rather, it seems to introduce a side-comment of sorts. At the same time, this information is clearly conventionally encoded, and not inferred via general reasoning. In a way similar to presuppositions, the relevant inference remains constant under various types of embeddings. In fact, conventional implicatures have been argued to (almost) always be present at a global level, unlike presuppositions. Similar considerations apply to the expressive *damn* in (4b).

While the presentation of the three aspects of overall meaning above more or less represents a traditional perspective, note that the boundaries between them are by no means clear. As will be seen both below and throughout several of the contributions to this volume, there are many cases of expressions where theorists have argued about the precise nature of certain related inferences. Furthermore, it is not even clear that the distinctions made here are exactly at the right level. A more fine-grained classification may ultimately be required. Tonhauser et al. (2013), for example, propose a view based on an overall class of projective content (i.e., content that remains unaffected in certain embedding environments) that would encompass both presuppositions and conventional implicatures, with more fine-grained distinctions amongst different sub-classes. The overview articles by Beaver and Geurts (2012) and Potts (to appear) also suggest a likely need for more fine-grained classifications. That being said, the distinctions laid out above clearly serve as the general starting point in the literature and provide a useful basis for further investigations.

There are, of course, numerous further ways in which linguistic utterances can give rise to inferences beyond what is literally expressed, e.g., through metaphor and irony. While these also merit more in-depth investigation, we will restrict our focus here on phenomena that are more closely related to the literal meanings of linguistic expressions and which have been explored extensively in formal terms.

3 Issues in Presupposition Theory

To set the stage in theoretical terms, I now provide a brief sketch of the central approaches to presuppositions, both traditional and current (for a comprehensive recent survey, see Beaver and Geurts (2012)). The next section discusses more specific issues in connection with the relevant experimental work, both in the existing literature and in this volume, with the aim of situating the results from the chapters to come in their broader context. Theoretical approaches to presuppositions can be divided along various fault-lines. For purposes of presentation, I will draw the main line between dynamic views of semantics, which see sentential meanings as intimately interwoven with their impact on the context, and static ones, which take a purely truth-conditional approach that does not incorporate effects on the context in the

semantics proper. I begin with the dynamic tradition, as it has long occupied the perhaps most prominent place in linguistics. The main focus in the latter category will be on recent developments of so-called modular accounts, which maintain a non-dynamic semantics but incorporate the incremental unfolding of discourse at another level.

3.1 Dynamic Approaches

Starting with the seminal work by Robert Stalnaker (Stalnaker 1970, 1973, 1974, 1978), presuppositions have commonly been seen as imposing requirements on possible contexts of utterance. For Stalnaker, this is a fundamentally pragmatic notion. His approach is based on a notion of contexts of utterance in terms of the Common Ground—the set of worlds compatible with what is mutually supposed for purposes of communication. Assertions serve to add information to the Common Ground, thereby reducing the corresponding set of worlds. Presuppositions, on the other hand, correspond to what is already entailed by the Common Ground. This view encodes the notion that presuppositions are an aspect of meaning that is taken for granted by the discourse participants. An utterance that comes with a presupposition requires that the Common Ground entail that presupposition in order to be felicitous. On Stalnaker's pragmatic perspective, it is speakers that presuppose, and the connection between certain expressions and speakers' general tendency to presuppose a corresponding presupposition has to be spelled out further. It's possible to combine a conventional encoding with this pragmatic view (see, for example, von Fintel's (2004) notion of 'Stalnaker's bridge'), though Stalnaker himself remains relatively non-committal in this regard.

Crucially, the dynamics of interpreting a given phrase in relation to the context extends to the intra-sentential context. This is at the heart of the Stalnakerian approach to presupposition projection. To illustrate with the simplest example, presupposition triggers introduced in the second clause of a conjunction are evaluated relative to a context that already includes the first conjunct. This provides an explanation for the specific projection patterns we find with conjunction: Based on the choice of the first conjunct, the entire sentence either does (5a-i) or does not (5a-ii) presuppose that John has a wife:

(5) a. i. John has the week off . . .
ii. John is married. . .
b. . . . and he's on a trip with his wife.

The fact that the first conjunct in (5a-ii) (plus the assumption of traditional marriage laws!) establishes that he has a wife suffices to satisfy the presupposition introduced by the possessive description *his wife*. In contrast, variations of the sentence where the first conjunct does not entail that John has a wife, such as (5a-i) will presupposing that he has one. This general dynamic approach to projection provided the starting point for much following work.

The dynamic semantic proposal of Heim (1982, 1983) adopts Stalnaker's idea of capturing projection in terms of incremental context update, but it builds this dynamic aspect directly into the semantics. Instead of the traditional truth-conditional notion of sentence meanings, it takes sentence meanings to encode the sentence's potential to change any given context when uttered—its context change potential. Since contexts still are construed as sets of worlds, at least in their basic form,[4] the truth-conditional contribution of a sentence can easily be reconstructed from the context change potential. But this approach also makes it possible to directly encode presuppositions as introducing requirements on contexts by utilizing partial context update functions, i.e., functions that would only be able to update contexts with certain properties (namely ones that entail the relevant presupposed proposition). What was truly novel in Heim's approach was that presupposition projection phenomena could now be characterized in terms of the context change potentials of the relevant operators, including conjunction and conditionals, leading to a unified semantic representation of entailed and presupposed content. To illustrate, the context update procedure for the conjunction of p and q would be as follows:

(6) c + p *and* q = (c+p) + q

The update of context c with sentence meaning p (i.e., $c + p$) amounts to set intersection. The potential partiality of p and q encodes the requirements introduced by the corresponding presuppositions: a context c can only be updated with p if it entails the presupposition(s) of p. Crucially, the context for q is the result of updating the original c with p, so that any presuppositions of q need not necessarily be entailed by c, but only by the combination of c and p. Parallel definitions can be given for other operators in such a way that they generally capture projection data successfully.[5]

Another prominent proposal in the dynamic rubric is that of discourse representation theory (DRT; Kamp 1981). It was developed independently of Heim's proposal, and largely aimed to capture the same anaphoric phenomena that Heim (1982) addressed. Later work by van der Sandt and Geurts (van der Sandt and Geurts 1991; van der Sandt 1992; Geurts 1999) developed an explicit theory of presuppositions in this framework, which fundamentally sees them as a form of anaphora. The basic idea is that there is a representational level that keeps track of the discourse as it develops, in the form of variables that represent discourse referents and conditions that the referents have to meet. Operators such as negation and conditionals introduce hierarchical embeddings within this discourse structure, and both anaphora and presupposition resolution involve searching for an antecedent along a pre-defined search path. For a presupposition introduced in the second conjunct of a conjunction, this search would first check in the first conjunct, and then in the discourse context to find a suitable antecedent. By and large, the predictions are quite similar to those of Heim's (1982)

[4] The full system of Heim (1982), which provides an analysis of anaphoric interpretations of definites, extends this basic view of contexts to include assignment functions.

[5] Note that both the extent of this success and the explanatory adequacy of the proposal have been questioned in the subsequent literature (see below).

approach, though there are some crucial differences, some of which we will turn to below (see Sect. 4.3).

3.2 Static Approaches

While dynamic accounts enjoyed a relatively dominant position in the literature throughout the 1980s and 1990s, there also are alternative approaches that capture the special status of presupposed information in different ways, namely via adding a third truth-value to the logic they work with (using systems based on Kleene 1952), or by assuming a supervaluation-based framework (van Fraasen 1968, 1971). The initial proposals pre-dated the dynamic ones, of course, and were attempts to formalize Strawson's observation that sentences whose presuppositions are not met are neither true nor false. Recent years have seen various revivals of variations of these approaches, and the differences in predictions between them and the dynamic ones turn out to be more subtle than it might initially seem (see Beaver and Krahmer 2001; Fox 2008; George 2008a, b among others, and Schwarz (2014b) for some further discussion in this volume).

One of the most influential recent developments in this area has emerged from work by Philippe Schlenker (Schlenker 2008a, b, c, 2009, 2010a, b). In its latest form, this line of work has become a compromise of sorts of the types of accounts considered so far, by proposing a non-dynamic recasting of a theory that is very much in the spirit of Stalnaker and Heim. This theory makes do with a semantics that is classical (i.e., without a third truth-value) and static. It represents both presupposed and asserted content in these classical terms, but assumes presupposed content to introduce additional pragmatic requirements. In particular, such content has to be entailed by its local context. In its simplest form, a local context roughly consists of the pre-utterance context and all parts of the sentence that precede the presuppositional expression. Formally, the evaluation of a presupposition involves consideration of all possible continuations of the sentence. This captures the incremental effects of presupposition interpretation, in the same spirit as the original Stalnakerian approach.

However, an important aspect of this type of theory is that it opens up the possibility of considering the basic nature of presuppositional requirements (of being established in the discourse context), on the one hand, and the role of the incremental unfolding of spoken language, on the other hand, as separate components. While the basic version of local contexts in Schlenker (2009) requires any possible continuation of the sentence to be felicitous, the incremental aspect can be weakened by demoting it to a processing preference. This makes it possible, in principle, for presuppositions to be supported by expressions that follow it, though that is dispreferred, because of the preference for incremental interpretation. The resulting type of modular account along these lines is investigated experimentally by Chemla and Schlenker (2012). They argue their data to support the modular approach, in that introducing support

for a presupposition after the trigger, thus violating incrementality, seems to be preferred over having no support for the presupposition in the context at all (see also Sect. 4.3 and Schwarz 2014b).

On the theoretical level, Schlenker's theory of 'Local Contexts' arguably improves on previous dynamic approaches in several ways. First, it avoids a problem of explanatory adequacy that had been acknowledged in the literature for some time. In short, the formal setup of context change semantics does not require the context change potential for conjunction discussed above (parallel concerns arise with other connectives). Based on the formalism, an alternative entry for *and* would be just as possible that reverses the update procedure. Schlenker's approach avoids this issue by basing the incremental effects (yielding asymmetries between the first and second conjunct) on the left-to-right linearization of linguistic structure. This derives (almost entirely) the same results as Heim's theory, but without having to stipulate anything further about the lexical entries for connectives (which are understood as in classical logic). Secondly, unlike Stalnaker's original approach, Schlenker's theory is versatile enough to apply the notion of local context to non-propositional nodes, which allows him to capture presupposition projection with quantifiers.

3.3 Key Questions about Presuppositions

To conclude our initial overview of theoretical approaches to presuppositions, it is generally accepted that presuppositions crucially relate to the context, but theories differ in terms of the level at which context comes into play. Stepping back from specific theoretical proposals and reflecting on the general questions that are crucial in investigating presuppositions, there are three key issues that are central in the theoretical literature. First, there's the question of the source of presupposed content, commonly called the triggering problem: Where do presuppositions come from? Are they specified in the lexical entry of a given expression, or are they derived in purely pragmatic terms? Furthermore, if they are lexically encoded, how do the corresponding pragmatic effects come about? Secondly, we need to account for the interaction of presuppositions with their linguistic context, specifically a host of embedding environments that give rise to projection phenomena. Finally, there is the issue of what status presuppositions have relative to other aspects of meaning, and what effects arise when presuppositions occur in contexts that do not support them. Different triggers may require different answers to these questions, which adds another potential layer of complexity. Theoretical issues and empirical questions related to these questions will be considered in more detail in the next section, in direct connection with corresponding experimental investigations.

4 Experimental Investigations of Presuppositions

In turning to experimental investigations of presuppositional phenomena and their relation to theoretical proposals, it is prudent to start out by acknowledging that presupposition theories put forward within philosophy and formal semantics do not generally commit themselves to making any specific claims about the actual cognitive processes involved in interpreting presuppositions (though there are some exceptions, perhaps most notably DRT). Nonetheless, it seems clear that everybody's shared ultimate goal is to advance our understanding of how the minds of human beings deal with language. But we have to be careful that in doing so, we do not mistake abstract, theoretical characterizations as making immediate claims about mental processes. When the evaluation of predictions about contextual acceptability or the presence or absence of certain inferences is concerned, this may not seem to be an issue at first sight. After all, even the most abstract theories in this area claim to model actual people's understanding of the linguistic expressions they analyze. However, even at that level, it is clear that judgment data need not correspond directly to the theoretical constructs that the theory in question make predictions about. Take an example from syntax: as has been evident right from the start of modern linguistics, not all structures that most accounts would want to see as grammatically legitimate are judged as acceptable (the classical example of course are multiple center-embedded clauses), and there are examples of the reverse case as well (e.g. Gibson and Thomas 1999). In the realm of presuppositions, a parallel point has been made by von Fintel (2004), who argues that speakers' judgments about truth-values (or lack there-of) for sentences containing non-referring definites should not be expected to be in a one-to-one correspondence with our theoretical notions of truth-values and infelicity. In short, even in simple judgment tasks, people may be influenced by a variety of considerations that are not directly related to the theoretical dimensions of the experimental manipulation. Obviously, parallel considerations apply *a fortiori* to investigations of aspects of cognitive processes that do not directly involve conscious decisions by subjects, such as timing measures for processes involved in linguistic comprehension.

While these considerations are important to keep in mind, they should not hold us back from setting out to pursue experimental tests of predictions from non-cognitive theories. We will generally need some linking assumptions about how the theory could possibly be amended so as to make specific predictions for experimental settings. But in many cases, these may be fairly straightforward and simple, at least at the outset of the enterprise. In the absence of other, more cognitively real, explanations of the same type of phenomena, theoretical proposals constitute the starting place for asking questions about the actual cognitive processes involved. As actual results are evaluated and interpreted, it needs to be clear that what is tested is the combination of the theoretical claims and the linking assumptions, which in turn means that any evidence against a specific proposal could be due to either one of them being wrong. But that is the nature of experimental investigations, and no particular issues arise

for the area of interest here, as far as I can tell. As long as new results lead to new testable hypotheses, there is progress.

In the following, I provide an overview of experimental work on presuppositions, starting with a brief review of earlier work that is related in general terms but doesn't directly focus on presupposition theory, and then turning to more recent work, including the contributions to this volume. The latter part in particular is organized thematically, roughly following the key questions identified in Sect. 3.3. For reasons of space, I cannot provide a more detailed introduction to the experimental methodologies used in the relevant lines of work. See the later chapters for relevant details in this regard.

4.1 Experimental Work on Related Phenomena

While experimental investigations directly tied to presupposition theory have only emerged relatively recently, there are several lines of related work that I will review here briefly to set the stage, focusing on two themes: reference resolution and implicature processing.[6]

As mentioned above, definite descriptions were the original poster-child for presuppositions. Figuring out, in a given context, whether their presuppositions are met and what individual is accordingly referred to by the definite involves a variety of processes and factors, various aspects of which have been studied experimentally in quite some detail in relation to online processing.[7] For example, the seminal study by Crain and Steedman (1985) showed that the context for definite noun phrases affects syntactic parsing decisions, e.g., by varying whether a *that*-clause following a noun phrase would need to be interpreted as a restrictive relative to ensure reference, or whether it could be taken to be a clausal complement (of a preceding verb) instead. Their contextual manipulation led to strong variation in acceptability judgments, and even garden path sentences were ameliorated by contextual biases towards otherwise dispreferred parses. Based on their findings, Crain and Steedman proposed a principle of parsimony, which guides the selection between different syntactic parses in their parallel parsing architecture, so that the reading carrying the fewest unsatisfied presuppositions will be the preferred one. Similar designs are used in more recent work by van Berkum and colleagues (van Berkum et al. 1999, 2003), which shows that there are ERP-effects related to whether the definite description can refer successfully or not.

In a similar vein, much of the work within the visual world paradigm (Tanenhaus et al. 1995), where eye movements relative to visually presented scenes are recorded while linguistic stimuli are played back auditorily, utilizes definite descriptions in critical parts of the sentences. Typical stimulus types involve instructions for clicking

[6] For a more general recent review of a similar range of topics, see Noveck and Reboul (2008).
[7] See Gibson and Pearlmutter (2011) and Gundel and Hedberg (2008) for recent collections of relevant work.

on one of several presented pictures (or objects) of the form *Click on the [NP]*. In line with the findings by Crain and Steedman (1985), work within this paradigm has revealed various context-related factors influencing parsing decisions both in adults and children (e.g. Sedivy et al. 1999; Hurewitz et al. 2000; Sedivy 2003; Chambers et al. 2002 among many others), which crucially rely on the uniqueness presupposition of definites. A further related area of research takes a more interactive perspective by looking at conversations between two people, who jointly have to manipulate some set of objects. Crucially, the experimental setup occludes some of the objects from one of the interlocutors, thus implementing a distinction between common ground (including the objects visible to both) and privileged ground (including objects only visible to one). Once again, definite descriptions are used as part of instructions, and the identity of the referents depends on which ground the listener is considering. While some studies provide clear evidence for the general and immediate ability to focus on the shared ground (e.g., Hanna et al. 2003; Nadig and Sedivy 2002), other studies report delays in homing in on the referent of the definite as interpreted relative to the common ground (Keysar et al. 2000, 2003).

The second line of research I want to highlight here directly concerns the more general issue of studying different types of inferences and investigating their properties in processing. In particular, there is a by now substantial body of experimental work on scalar implicatures. Central questions in this area have concerned the time-course of computing implicatures in online processing as well as the factors that affect whether or not implicatures are computed in the first place. Initial experimental investigations by Noveck and Posada (2003) and Bott and Noveck (2004) provided evidence based on reaction times and data from event-related potentials (ERP) that interpretations of *some* as 'some but not all', based on a scalar implicature, arise later than literal 'some and possibly all' interpretations, which has been replicated across various other methodologies, including reading times and visual world eye tracking (Breheny et al. 2006; Huang and Snedeker 2009, 2011; Bott et al. 2012). These results have generally been taken to be inconsistent with theoretical models where implicature-based inferences are present by default (Levinson 2000; Chierchia 2004) and instead have been interpreted as reflexes of some form of online Gricean reasoning (see, e.g., the 'literal first' model discussed in Huang and Snedeker (2009)). Early work in this realm from the perspective of language acquisition points in a similar direction(Teresa Guasti et al. 2005; Noveck 2000; Chierchia et al. 2004), because children are generally found to be more willing to accept 'logical' interpretations (without any implicatures) than adults, though the exact timing in acquisition and the extent to which this holds depends on a host of factors (see, e.g., Papafragou and Tantalou 2004; Gualmini et al. 2008). However, there also are studies that suggest a more rapid availability of content based on pragmatic enrichment: Grodner et al. (2010) and Breheny et al. (2013) report results based on the visual world paradigm that are argued to show immediate availability of implicatures in online processing. Similarly, Atanassov (2014) finds rapid implicature effects for the modal *might* (which implicates non-necessity). Degen and Tanenhaus (2012) investigate various contextual factors that play a role for the availability of implicatures (as well as the felicity of implicature triggers) and argue that independent factors are responsible for

delays found in other studies. All in all, these studies suggest that the previous results supporting generally delayed computation of implicatures have to be integrated into a broader perspective that allows us to understand factors affecting the availability of implicatures in online processing as well as the precise timing thereof.

Both of these lines of work are directly relevant to the present enterprise, in more than one respect. First, the methodological approaches used there can, in many cases, be straightforwardly extended to the study of presuppositions in general. Second, they are part of the overall perspective of studying how various aspects of meaning relate to one another and how they interact with the broader context. We will ultimately want to conduct comparisons of all types of meaning in context, i.e., not only compare implicatures and presuppositions to conventionally encoded *at-issue* content respectively, but also compare, say, presuppositions to implicatures directly. Such comparisons are highly relevant theoretically as various recent proposals assimilate the analysis of (certain) presupposition triggers to scalar implicature (Simons 2001; Abusch 2010; Romoli to appear.)

4.2 Triggering

Beginning with the question of how presuppositions arise—the triggering problem—let us now turn to two central issues: the time course of presuppositional content becoming available in online processing, and potential differences between (classes of) presupposition triggers that help to shed light on the nature of (possibly different forms of) triggering.

4.2.1 The Time Course of Triggering in Online Processing

In studying different aspects of meaning in terms of their processing properties, one central question concerns the timing of the availability of a given type of meaning. Much of the experimental literature on scalar implicatures focuses on this aspect, as reviewed above, and the commonly (but not universally) found delays are often taken to support a literal-first processing model, with costly pragmatic reasoning taking place online to give rise to implicatures.

With respect to presuppositions, similar questions arise, and information about the timing of the availability of presupposed content is crucial both for a theory of presuppositions as well as a processing model for them. For example, different theoretical perspectives suggest different temporal orderings of presupposed and asserted content. If we assume that presuppositions are conventionally encoded and constitute conditions on context updates, as on dynamic approaches to presupposition, we would expect that they will be checked immediately, possibly even before the asserted content is computed. From the perspective of pragmatic accounts of presuppositions, on the other hand, which see presuppositions as inferences that arise via pragmatic reasoning in a way at least broadly parallel to scalar implicatures, we might expect

delays that are comparable to delays found in the implicature literature. Note, however, that some caution is in order in interpreting any results in this regard. First of all, more needs to be said about how the various theoretical approaches can be translated into corresponding processing hypotheses. Furthermore, while the presence of delays would indeed be problematic for conventional and dynamic accounts (supplemented with appropriate linking assumptions to derive processing predictions), the absence of a delay will not necessarily be evidence against pragmatic accounts. There is ample evidence for rapid availability of various types of pragmatic information (largely from work within the visual world paradigm, such as Sedivy (2003)), so such an outcome could just as well be seen as falling into this category. Nonetheless, insights into the time-course of presupposition interpretation in processing will be crucial for a processing model of different types of meaning, and various studies have begun to shed light on this issue.

Overall, there is mounting evidence that presuppositions are available rapidly during online processing, much of it using paradigms based on reading times. In two self-paced reading studies on German *auch* and its English correlate *also*, Schwarz (2014a) finds increases in reading time for the clause containing the presupposition trigger when it was not supported by the (intra-sentential) context. This delay is attributed to the infelicity of the sentence (assuming that *also* resists accommodation), which in turn can only give rise to reading time effects if the presupposition is indeed available and evaluated relative to the context. Tiemann et al. (2011) expand this general approach to several other presupposition triggers and narrow down the time-window during which the presupposition is available by using word-by-word self-paced reading. One of their experiments indeed finds delays on the presupposition trigger itself, relative to non-presuppositional controls. Furthermore, they find immediate effects on 'critical words', which contribute the information needed to determine that a given presupposition is not supported by the context. Using the same general approach of contextual manipulations of felicity, but using eye tracking during reading, Schwarz and Tiemann (2012) look at German *wieder* ('again') and find delays in early processing measures, such as first fixation duration and regression proportions, on the verb following *wieder* (at least when *wieder* is not in the scope of negation; see below for effects of embedding under negation). Finally, in this volume, Tiemann et al. (2014) use word-by-word self-paced reading to look at *wieder* in both supporting and neutral contexts, and again find reading time increases at the critical word for the latter. They also present a basic processing model for presuppositions, which includes a proposal for the steps the processor goes through when a presupposition is not supported by the context (see the discussion on accommodation in Sect. 4.4.2).

While there have been a number of reading studies on the topic, researchers in linguistics only recently have turned to the visual world paradigm to investigate the time course of presupposition interpretation at an even more fine-grained level. Indeed, apart from a couple of earlier related studies by Craig Chambers and colleagues (see Chambers and Juan (2005) on *another* and Chambers and Juan (2008) on *return*.) and related work on reference resolution (e.g. Heller et al. 2008), the group of papers using this methodology in the present volume constitutes the first major step towards

systematically investigating standard presupposition triggers in comparison with asserted content.[8] They look at English *also* in comparison with the asserted part of *only*. Romoli et al. (2012) (see also Romoli et al. 2014) present two experiments using a paradigm building on that used by Kim et al. (2009) to investigate focus alternatives in the interpretation of *only*. The linguistic stimuli consisted of a context sentence establishing one individual as having two types of objects, followed by a target sentence that either did or did not include *also* or *only* associating with the subject and a new or old noun in predicate position respectively. They find shifts in eye movements already 400 ms after the trigger onset *also* condition, indicating that the presupposition of *also* is utilized in determining the referent before further disambiguating information is introduced. Interestingly, this effect is independent of whether the antecedent of *also* is introduced linguistically or merely visually present in the context.

In the second contribution to this volume on this topic, Schwarz (2014a), results from two visual world experiments with a similar approach are reported. In the first experiment, the time course of interpreting *also* associating with an object noun phrase was investigated in visual contexts where its presupposition either did or did not disambiguate the referent during an otherwise ambiguous time window. A shift in fixations towards the target referent was observed in the disambiguating condition as early as 200–300 ms after the onset of *also*, suggesting that the presupposition introduced by *also* is immediately available and utilized in identifying the referent. In a second experiment, the interpretation of stressed *also*, which associated with the subject of the sentence, was compared to the exclusivity asserted by *only*. While *also* again gave rise to an essentially immediate shift in fixations towards the target in the critical condition (starting at 300 ms after the onset of *also*), the exclusive inference introduced by *only* did not give rise to a parallel shift until 700 ms after its onset. But in contrast to Romoli et al.'s (2014) findings, this information did yield a significant effect during the otherwise ambiguous time window, in line with Kim et al.'s (2009) findings, who found even more rapid integration of the information introduced by *only*. One difference between the studies is that the ones yielding online effects involved association of *only* with an object, whereas Romoli and colleagues used stimuli where it associated with the subject. Further work is needed to fully understand the factors affecting the ease of interpreting *only*. As far as the presuppositional contribution of *also* is concerned, the results from these studies argue against a delay in availability for presupposed material, and thus may be most naturally compatible with accounts that assume presupposed content is encoded conventionally. But as noted above, it is also possible that we are looking at rapid pragmatic effects, so the results do not per se settle the question about the source of presupposed content. Nonetheless, they provide the most direct and time-sensitive evidence yet that presupposed information is available and utilized as soon as the presupposition trigger is introduced.

[8] For even more recent results extending this approach to the triggers *stop* and *again*, see Schwarz (2014c).

The third study in this volume that looks at *also* using the visual world paradigm, Kim (2014), takes a different angle, in that it investigates the effects of discourse structure on the selection of an antecedent for *also*. This is done by presenting multi-sentence discourses, which provide various possible antecedents for *also* in the final target sentence. In two initial comprehension studies, Kim asked subjects to choose one of several descriptions of what the sentence with *also* conveyed, which reflects how they resolve its presupposition in the discourse. While there was a general preference for linearly local antecedents in the comprehension studies (where *also* was understood relative to the immediately preceding sentence), a structurally (but not linearly) local interpretation also became available when the discourse structure was manipulated. In a visual world eye tracking experiment, Kim also found a preference for structurally local interpretations. The eye movement results for the condition that involves a structurally local antecedent furthermore add to the evidence from the two studies above, showing that the presupposition of *also* is available immediately in online processing.

4.2.2 Types of Presupposition Triggers

A crucial issue concerning the triggering of presuppositions is that there may be distinct classes of triggers with different properties, possibly both with regards to how they are triggered and how they behave in embedded environments. Various considerations have been brought fourth in the theoretical literature to argue for such distinctions. For example, Karttunen (1971) already noted that factive verbs seem to vary in how strongly they give rise to inferences based on global presupposition projection. Taking an example from Jayez et al. (2014) for illustration, the factive presupposition of *realize* in (7a) does not necessarily give rise to the inference that Paul missed the point, in contrast to that of *regret* in (7b)

(7) If Paul {(a) realizes/(b) regrets} he has missed the point, he will probably reformulate his objection.

Another proposal for a distinction, which Amaral and Cummins (2014) take as their starting point, was made by Zeevat (1992). It differentiates between resolution triggers and lexical triggers. The former are crucially anaphoric in that they directly relate back to entities (or events) in the discourse context (examples include *too* and *again*). The latter consists of cases where the presupposition is a requirement that comes with the asserted component of the trigger (examples include *stop* and *win*; see Amaral and Cummins (2014) for more detailed discussion). More recently, Abusch (2002, 2010) has proposed a distinction between 'hard' and 'soft' triggers, based on the possibility of accommodation in the antecedents of conditionals (see Jayez et al. (2014), for further illustration). Yet another difference between triggers commonly acknowledged in the literature concerns the extent to which accommodation is possible (for review, see Beaver and Zeevat 2012). For example, pronouns and *too* are generally found to be hard to accommodate, whereas factives and verbs

like *stop* accommodate easily (quantitative support for such differences comes from Spenader (2002)).[9]

At least some of these differences have been used to construct theoretical proposals that accord a different status to the assumed sub-types. For example, Romoli (to appear) argues that soft presupposition triggers really should be understood as a type of scalar implicature, building on the proposal by Abusch (2002, 2010) that soft triggers crucially involve reasoning about alternatives (see also Chemla 2009a). Whatever the verdict on the best approach for capturing the differences between types of presupposition triggers, a more thorough understanding of the actual empirical differences will be needed to flesh out the theoretical picture. This includes direct comparisons between particular pairs of triggers, detailed investigations of the properties of individual triggers, as well as cross-linguistic comparisons between (roughly) equivalent triggers.

While various initial efforts in this direction have been made and reported at conferences (Smith and Hall 2011; Amaral et al. 2011; Xue and Onea 2011; Jayez and van Tiel 2011; Velleman et al. 2011; Cummins et al. 2013), much still remains to be discovered and to be documented in the peer-reviewed literature. Three of the papers in the present volume contribute to this topic, both with new experimental results and detailed perspectives on the emerging body of empirical data more generally. Destruel et al. (2014) investigate the exhaustive inference of *it*-clefts in detail by revisiting previous results from Hungarian and German and reporting new data on English and French. They argue that while much of the previous literature has argued about the *source* of exhaustive inferences of clefts, i.e., whether they are semantic or pragmatic, the key difference between clefts and exclusives (e.g., statements with *only*) instead lies in the *status* of the inference. In particular, they argue that with clefts, it is not part of the main assertion, or *at-issue* content of the utterance, whereas it is with exclusives. Their discussion also makes a substantial methodological contribution by evaluating the test they employ, which involves choosing between various continuations of a given discourse. These differ in whether the preceding utterance is affirmed or denied (by starting with *yes* or *no*), and furthermore, whether a following statement contradicting the exclusive inference is introduced by *and* or *but* (e.g., *Yes, and . . .* vs. *Yes, but . . .*). The availability of *yes, but . . .* is argued to be indicative of the status of the inference, in contrast with standard cancellation tests for implicatures, which are indicative of the source of the inference. The data presented by Destruel et al. shows that for clefts, *yes, but . . .* continuations are preferred over *no, and . . .* ones, whereas the reverse holds for exclusives.

Amaral and Cummins (2014) use a similar technique in investigating a variety of Spanish presupposition triggers, building on previous work on English (Cummins et al. 2013). Based on Zeevat's (1992) distinction between lexical and resolution triggers, they find differences between examples from both classes of triggers with

[9] But see Singh et al. (2013) for results suggesting that accommodation of *too* may be easier than previously thought.

respect to the acceptability of *yes, but* . . . continuations and *no, because* . . . continuations. Lexical triggers, such as *lamentar* ('regret'), and *dejar de* ('stop'), among others, yielded significantly higher acceptance rates for *no, because* . . . continuations, whereas no such difference was found for resolution triggers such as *también* ('too') or *otra vez* ('again'). These results mirror those that the authors previously found for English and lend support to the distinction posited by Zeevat. Several interesting questions arise when comparing these findings with those for clefts by Destruel and colleagues. First, the existence presupposition of clefts also is one that is entailed by the asserted content of the triggering construction, which would seem to put clefts on par with the lexical triggers considered by Amaral and Cummins. Nonetheless, Destruel and colleagues find a strong preference for *yes, but* . . . continuations over *no, and* . . . continuations, in contrast with Amaral and Cummins' findings for lexical triggers. Secondly, Destruel and colleagues' claim that the availability of *yes, but* . . . continuations is indicative of non-at-issueness may have to be qualified, as the presuppositions of lexical triggers presumably still count as not being at-issue (note that Velleman et al. (2011) also find different patterns for clefts and *too* on the one hand and *find out* and *know* on the other). Further work will be needed to resolve these tensions, e.g., by refining the theoretical perspective on the continuation task in general, or by investigating possible task differences between these studies in particular.

Jayez et al. (2014) also is concerned with differences between types of triggers. It uses constructions modeled after Abusch's (2010) test for the weak vs. strong distinction, with presupposition triggers in the antecedent of conditionals and a context sentence that explicitly suspends the global interpretation of the presupposition. Looking at French *aussi* ('too'), *regretter* ('regret'), and clefts, they present evidence that the distinction is not entirely robust, as it seems to interact with certain contextual factors. Instead, they argue their results to be consistent with a three-way distinction between presupposition triggers, in line with Jayez (2013). These intriguing findings open up the way towards further investigations, ideally including a broader comparison of the various methods used in the studies discussed here, as well as methods that can shed light on the online processing of the relevant inferences.

4.3 Projection

Presupposition projection, i.e., the phenomenon that presuppositions introduced in many embedding environments are interpreted outside of that environment, constitutes one of the central theoretical challenges in presupposition theory, and much of the literature has focused on attempts to capture it. Three major approaches have mostly dominated the picture. Dynamic semantics and Discourse Representation Theory (DRT) constitute distinct approaches to projection within a general model of how information evolves in discourse dynamically. More recently, Schlenker's

(2009) theory of Local Contexts has revived a non-dynamic perspective on projection, which is nonetheless conceptually quite close to Stalnaker's original view of contexts and empirically almost equivalent to Heim's dynamic semantics.

Each of the three theories successfully accounts for a large part of the projection data, but their predictions differ in subtle ways. Assessing which prediction is correct in specific cases is a difficult empirical task. For reasons of space, I cannot review the details of the projection mechanisms of each theory in detail here, and only review some of their key properties (for an accessible overview, see Schlenker (2011a, b)). As discussed briefly above, dynamic semantics sees the meanings of sentences in terms of their context change potentials, and presuppositions are seen as restrictions on admissible contexts. Projection phenomena are dealt with in terms of the formulation of the embedding operators and connectives. As illustrated above, conjunction allows a simple illustration: the contribution of the second conjunct in a conjunction is evaluated relative to a context derived from the context for the entire sentence by updating it with the first conjunct. DRT also models the way presuppositions interact with the previous discourse context, but unlike dynamic semantics, it does so in representational terms. Discourse structure is modeled in hierarchically structured representations, and connectives introduce structures that are associated with a search path (from local to global contexts) for presuppositional antecedents.

Finally, the Local Contexts theory takes a classical, non-dynamic semantics but defines a notion of Local Context that mimics the dynamic effects in pragmatic terms. In particular, it considers all possible continuations of the linguistic material preceding a presupposition trigger and requires all of them to be presuppositionally acceptable. While this version of the theory assumes a strict role for incrementality, it allows more flexibility than traditional accounts in that incrementality can be seen as a processing bias that can in principle be violated.

Much of the experimental work on projection has focused on the question of what exactly the presupposition in a given embedded context is. For example, there is disagreement in the theoretical literature on whether a presupposition trigger in the consequent of a conditional (such as the possessive definite in (8) gives rise to a conditional presupposition (9a) or a non-conditional one (9b):

(8) If John goes swimming, he will wear his swim-suit.
(9) a. If John goes swimming, he has a swim-suit.
b. John has a swim-suit.

Both types of interpretations appear to be attested, but theories differ in terms of which one they see as the basic one. Romoli et al. (2011) provide a first experimental exploration of this topic using a covered box picture matching task (adopted from Huang et al. 2013) and argue their results to favor accounts that predict a conditional presupposition as the basic one (they also shed light on the factor of whether the presupposition intuitively can be seen as dependent on the content of the antecedent). Another line of work has been concerned with presupposition triggers in quantified sentences and the readings that they give rise to. Chemla (2009b) reports results from several judgment studies and argues that quantifiers vary in terms of the strength of the

presupposition (e.g., whether it is universal, existential, or somewhere in-between), in a way that is not predicted by any of the main theories on the market.

Another line of experimental work on projection has been concerned with the role of incrementality. Dynamic accounts, at least as they are standardly construed, assume that presuppositions have to be supported in their context by material that precedes the trigger. More recent modular accounts, on the other hand, have opened up the possibility that incrementality can be seen as a processing preference, which can be violated (see Schwarz (2014b) for some further discussion). Chemla and Schlenker (2012) home in on this prediction and test presupposition triggers in conditionals, disjunctions, and *unless*-sentences in configurations where the presupposition trigger appears either in the linearly first or second clause. In an inference judgment task, they find that subjects more strongly endorse inferences corresponding to a conditional inference, compared to a non-conditional one, regardless of where the presupposition trigger appears. They interpret this as support for a symmetric theory of presupposition satisfaction, where material introduced later on in the sentence in principle can provide support for an earlier presupposition. Such cases are still going to be considered as dispreferred on the grounds of a processing preference for incremental presuppositional support, but this preference is not hard-wired into the projection mechanism.[10]

Schwarz (2014b) further investigates this topic by looking at presuppositions introduced in the antecedent of conditionals. Unlike Chemla and Schlenker, who varied placement of the presupposition trigger by putting it either in the antecedent or the consequent, the studies reported here always place the trigger in the antecedent and reverse the order of the clauses by creating *if*-clause initial and final variants. They employ a covered box picture matching task, with a variation in whether the consequent of the conditional is true or false of the target picture. This should make a difference on a symmetric account, but not on a dynamic account. The results for the *if*-clause initial conditions suggest a strong role for incrementality, but the *if*-clause final versions are more in line with the predictions of a symmetric account. However, dynamic accounts might be able to take into account linear order in their update procedure to incorporate these results. Further work will be needed to tease apart the more subtle differences in predictions between such modified accounts.

Another important topic involving projection concerns the availability of so-called local interpretations, where presuppositions in the scope of an operator that typically would give rise to a global inference are interpreted relative to that operator. (10) provides an illustration for the case of definite descriptions:

(10) The king of France is not bald—because there is no king of France!

By and large, the consensus in the theoretical literature has been that local interpretations are dispreferred (beginning with Heim (1983)), though this was only based

[10] For another recent contribution to this topic looking at disjunction, see Hirsch and Hackl (2013).

on individual intuitions. But recently, Chemla and Bott (2013) offered the first experimental evidence to support that assessment, using a truth value judgment task with sentences such as (11) and looking at reaction time measures.

(11) Zoologists do not realize that elephants are reptiles.

The factive verb *realize* presupposes the truth of its complement clause, and on its global interpretation, this presupposition prevails even in the context of negation. However, a local interpretation would have that inference negated, as if it were introduced as part of the asserted content. In the latter case, the sentence should be judged true, whereas on the former, it should be judged false. Both types of responses are given by subjects throughout the experiment (with only a slight bias towards 'false'-responses), but the 'true' responses take significantly longer than 'false' responses. Chemla and Bott interpret this as evidence for traditional, semantic accounts that take local accommodation to be a last resort repair strategy. In contrast, the results are incompatible with pragmatic accounts a lá Schlenker (2008a), which assume that the local reading corresponds to a literal semantic reading, while the global reading requires additional pragmatic inferencing.

Romoli and Schwarz (2014) utilize a different task to investigate the speed of local interpretations of the presupposition introduced by *stop* under negation. Using a covered box picture matching task, their experiment compares cases where the overt picture supports the presupposition with ones where it does not. Their design allows them to not only look at acceptance rates for each case, but also to compare reaction times for target picture choices in both conditions. This avoids a possible confound in the study by Chemla and Bott, who compare reaction times for true vs. false responses. Acceptance rates were much lower for target pictures corresponding to the local interpretation. Furthermore, response times were slower for local target acceptances than for global ones, in line with the findings by Chemla and Bott (2013).

A final set of studies to be mentioned in connection with projection is that reported in Schwarz and Tiemann (2013a, 2013b). In two reading time studies using eye tracking, embedding of presupposition triggers is found to modulate processing effects of presuppositional acceptability. In one study, immediate eye movement effects on the critical word are found when the context was inconsistent with the presupposition, but only when the trigger (German *wieder*, 'again') was outside of the scope of negation. No effects of context emerged when it was embedded under negation, and follow-up studies suggest that this is not due to a general availability of local interpretations. In a second study, presuppositional support for *wieder* in the consequent of conditionals is introduced in varying locations, namely in the antecedent or in a context sentence. Furthermore, embedding under negation was another factor, as in the other experiment. The results from this study suggest that the hierarchical distance in terms of the projection search path assumed by DRT directly affects reading times on the critical region. This is arguably unexpected under purely semantic accounts (such as dynamic semantics).

It will be interesting to relate Kim's (2014) contribution to the present volume to these results, as her study also involves different resolution options for the presupposition of *also*, but at the level of discourse structure. A promising next step would be

to consider seeing the two results in a unified way, which would have to incorporate a broader discourse structural perspective on both intra- and inter-sentential relations between clauses. A further natural extension of such an approach would be a more general comparison between presupposition resolution and pronoun resolution. There is an extensive literature on factors influencing the search for an antecedent of pronouns, both theoretical (e.g. Grosz 1977; Grosz et al. 1995; Gundel et al. 1993) and experimental (e.g. Gordon et al. 1999). Given theoretical proposals such as Discourse Representation theory (DRT, Kamp (1981)), which consider presupposition and pronouns as essentially the same type of phenomenon, as well as the results by Schwarz & Tiemann and Kim, relating the insights on presuppositions and pronouns to one another seems timely and promising.

4.4 Presuppositions in Defective Contexts: Infelicity and Accommodation

4.4.1 Infelicity

Turning to the question of the status of presuppositions in different types of contexts, we can return to one of the starting points for the study of presupposition, namely the intuition that the use of presupposition triggers gives rise to infelicity when the context does not entail the presupposition. This goes back, of course, to Strawson's observation about the definite description *The king of France*. While that basic intuition may seem robust in certain cases, it is less clear in others. To the extent that this property is fundamental for the notion of presuppositionality, it therefore is important to assess these judgments systematically and to investigate the factors that give rise to variation in judgments. Detailed discussions of this issue have primarily focussed on definite descriptions. It was already observed by Strawson (1964) himself that not all definites seem to give rise to infelicity. More specifically, he suggested that unmet presuppositions of definites give rise to infelicity when they are topical. When they are not, as in the following example, speakers seem more inclined to simply judge the sentence as 'false':

(12) The exhibition was visited yesterday by the king of France.

Reinhart (1981) fleshed this view out further. Lasersohn (1993) and von Fintel (2004) present a different perspective and argue that definites with presupposition failure give rise to intuitions of plain falsity when they are (actually, in Lasersohn's case, and in principle, in von Fintel's case) falsifiable under the (temporary) assumption that the existence presupposition was met. Another recent proposal is that of Schoubye (2010), who argues that plain false-judgments arise when the sentence in question can be construed as being a 'consonant response' (a technical notion spelled out in detail by Schoubye) to a Question Under Discussion (Roberts 1996). This has the promise of wider empirical coverage, e.g., to explain the apparent role of focus on truth-value judgments.

While Schoubye (2010) reviews a much larger set of data, it is not based on systematic evidence from a wide range of speakers. However, two recent papers set out to address the issue experimentally: Abrusán and Szendrői (2013) and Schwarz (2014c). Abrusán and Szendrő (2013) report a truth-value judgment study investigating the effects of topicality and verifiability (actual and in principle) on speakers' willingness to judge sentences with non-referential definites as plain false. The task was a forced choice truth-value judgment with an additional option of 'can't say'. In addition to affirmative sentences, they also included negated variants of each condition. Interestingly, they found that subjects were generally quite willing to judge the affirmative versions as 'false', without any significant differences between conditions. But in the negated versions, they found significant increases of 'true' judgments for both the topicality and verifiability manipulations, which they argue supports the view that these factors affect speakers' disposition to assign a truth-value rather than being unable to do so.

However, as the authors note themselves, once negation is introduced, there is another possibility, namely that of local accommodation (see Sect. 4.3). The differences due to the experimental manipulations could then just as well be seen as modulating the availability of local accommodation. A further point to note in connection with their interpretation is that the proposals in the literature (especially by Lasersohn and von Fintel) explicitly limit their explanation to cases where sentences with non-referring definites are judged false, rather than infelicitous, by construing a more general notion of pragmatic rejection. It is therefore not clear that one would want to invoke their mechanism to account for 'true' judgments on negated sentences. With these cautionary remarks in place, it nonetheless is clear that Abrusán and Szendrői's (2013) study constitutes a welcome and much needed first step towards investigating truth-value judgments for presuppositional sentences systematically.

Schwarz (2014c) takes a different approach to the same issue, which avoids the notorious difficulties of having subjects differentiate between judgements of falsity and infelicity (witness the results for affirmative sentences by Abrusán and Szendrői (2013)). The central idea is to look at reaction time measures for false judgments, comparing cases where they are based on presuppositions on the one hand and asserted content on the other. The task uses visual displays as the basis of the truth-value judgment, and only allows 'true' and 'false' as answer choices.[11] If subjects arrive at the 'false' judgment in the presuppositional case in a way different from cases where asserted content is false, this may be reflected in differences in reaction time. The results in Schwarz (2014c) indeed suggest that there are such differences, as reflected in an interaction between definite and indefinite sentence variants and the part of the sentence that is at odds with the visual display: rejections based on presuppositional content take relatively longer than rejections based on asserted content. This approach provides an alternative methodological avenue for assessing

[11] For an earlier study along similar lines looking at asserted vs. presupposed content introduced by *only*, see Kim (2007).

judgment behavior with measures that are not directly based on subjects' conscious choices.

It is interesting to compare these response time results to the discussion of presuppositions in online processing above, as there is a certain tension between the results from reading and visual world studies on the one hand and truth-value judgment studies on the other. The response time results just reviewed, taken on their own, could be seen as suggesting that presupposed content is not readily available in processing. However, that is incompatible with the results considered in Sect. 4.2.1. It therefore seems that the delays in the judgment studies must be due to the role that presupposed content plays in the verification procedure. This is, of course, entirely in line with the standard view that presuppositions represent backgrounded information, which is taken for granted. Delays in verification might then be due to a reluctance to challenge information expressed in this way, rather than a delay in availability of the information per se, since there seem to be no delays in utilizing this type of information, e.g., in identifying a referent in a visual display.

4.4.2 Accommodation

In many cases where a presupposition is consistent with the context but not entailed by it, it is quite common for hearers to quietly accept the presupposition, even though it in fact is introducing new information. Following Lewis (1979), such cases are commonly referred to as *accommodation*. In its original conception, this is seen as a repair strategy, which hearers utilize to rescue an otherwise infelicitous discourse. They do so by simply adjusting the common ground to entail the presupposition, and then proceeding to interpret the sentence in light of this adjusted context. Numerous interesting issues arise once an attempt is made to spell out the details involved in this process. For recent discussion, see Beaver (2001), Simons (2003), von Fintel (2008), and Beaver and Zeevat (2012). An important issue to come to terms with is that accommodation often proceeds very smoothly and is quite common for certain presuppositional expressions (for corpus results, see Spenader (2002)), which seems at odds with the notion of a repair strategy. Another important question concerns the variation in accommodatabiliy between triggers, which remains poorly understood in theoretical terms.

A fair amount of experimental work has been carried out on related phenomena with definite descriptions, especially cases of *bridging*, where a definite has not been introduced explicitly but directly relates back to something else in the context. In an early study, Haviland and Clark (1974) compared contexts that required a bridging inference (e.g., mentioning picnic supplies) with ones where some entity (e.g., beer) was mentioned explicitly, and found longer reading times on a subsequent presentation of a definite (e.g., *the beer*). O'Brien et al. (1988) showed, however, that prior mention of a referent is not necessary if the context is sufficiently specific: the definite *the knife* was read more slowly when the antecedent was more general (*a weapon*) than when it directly matched the noun phrase (*a knife*), but only when the context involved a general verb such as *assault*, and not when it involved *stab*, which

is more closely associated with knives. Burkhardt (2006) (and subsequent work) used ERP-studies to identify neural correlates of bridging by looking at definites such as *the conductor* in contexts with an explicit antecedent, a bridging antecedent (*a concert*), and no antecedent. Intriguingly, the bridged cases display neural hallmarks of both new and old information in the form of an reduced N400 effect, followed by a P600 effect.

A more recent set of studies looks at the impact of plausibility on accommodation, e.g., by varying the noun phrase occurring in a definite. Frazier (2006) used passive context sentence (*My order was taken*) followed by a plausible or implausible definite or indefinite (*a/the waiter/busboy*) in a reading time study, and only found effects of plausibility, not of definiteness. Using a similar approach but a different task, namely a stops-making-sense variant of self-paced reading, Singh et al. (2013) do find an interaction of definiteness and plausibility in the frequency of stops-making sense judgments. While these response patterns do not directly shed light on the issue of what processing costs, if any, might be associated with accommodation, such systematic evidence for the role of plausibility in accommodation is an important addition to our understanding in this area.[12]

While the literature has very much focused on definite descriptions in investigating accommodation, some studies have looked at other presupposition triggers as well. It is commonly assumed that certain triggers, in particular anaphoric ones such as *also* or *too* are hard to accommodate (Kripke 1991, 2009). Accordingly, the reading time delays for the *also*-clause in Schwarz (2014a) are attributed to the unavailability of accommodation. Similarly, Chemla and Schlenker (2012) utilize *too* (or rather its French variant *aussi*) in their design to take advantage of the difficulty of accommodating an antecedent for *too*. Interestingly, however, the results from Singh et al.'s (2013) stops-making-sense task for *too* are very similar to those for definites (see above): not only do presuppositionality and plausibility interact, but for plausible contexts, subjects tend to accept the sentences at comparable rates (essentially at ceiling level). This suggests that the option of accommodating *too* is more viable than previously thought.

From a processing perspective, an important question is to what extent there is a general pressure to avoid or minimize accommodation, and this is one of the central issues that Tiemann et al. (2014) address. Crain and Steedman's (1985) study discussed above is an early example of a proposal where parsing decisions are influenced by the desire to avoid accommodation, and subsequent work, in particular within the visual world eye-tracking paradigm, has found support for this view from different angles (for an example, see Chambers et al. 2002). Additionally, in the questionnaire study in Schwarz (2014a), using ambiguous German sentences with *auch* ('also'), syntactically dispreferred interpretations become more frequent when

[12] Direct evidence for processing costs of accommodation has been hard to come by. Perhaps the most convincing result in this regard comes from the accommodation study by Tiemann et al. (2011), which finds longer readings times on critical words in neutral contexts as compared to both verifying and falsifying contexts.

this ensures that the presupposition of *auch* is met, presumably due subjects' desire to avoid accommodation. Tiemann et al. (2014) take this view even further by suggesting that for certain triggers (in their case, *again*) whose presupposition does not crucially contribute to the compositional interpretation of their sentence, accommodation may be avoided by ignoring the presupposition altogether. Their version of a *minimize accommodation* principle thus goes so far as to see accommodation as a truly last resort that is avoided whenever possible.[13]

4.5 Relation to Other Aspects of Meaning

From the overall theoretical perspective assumed throughout this chapter, it is clear that the study of presuppositions crucially takes place in relation to other aspects of meaning. Key questions are where to draw the lines between different types of meaning and how they interact with one another. Experimentally, the best angle of understanding the processing properties of one type of meaning often employs contrasts with another type of meaning. While my review in Sect. 2 assumed a traditional division between central aspects of meaning, much recent work has focused on the potential need for re-drawing the boundaries and turning towards more fine-grained distinctions.

For example, as has come up at various points throughout, certain pragmatic approaches to presuppositions raise the possibility that (at least certain) presuppositions are much closer to implicatures than previously thought. And some proposals, such as Romoli (to appear) and Chemla (2009a), go as far as seeing (certain) presuppositions as theoretically equivalent to implicatures. The contribution by Romoli and Schwarz (2014) attempts to investigate this possibility experimentally. As discussed in Sect. 4.3, their experiments look at *stop* under negation in one experimental condition. But in another condition, they also look at the strong scalar item *always* under negation. In downward entailing contexts (which include negation), strong scalar items give rise to indirect scalar implicatures (Chierchia 2004). For example, in saying that *John didn't always go to the movies this week*, the implicature arises that he sometimes went. By presenting such sentences with target pictures that either were or were not compatible with the implicature, the experiment allowed for a direct comparison of indirect scalar implicatures and the presupposition of *stop* under negation. If they behaved like other implicatures, we might expect response delays in line with prior results. Interestingly, Romoli and Schwarz's results are not straightforwardly reconcilable with any of the traditional perspectives. While presuppositions and indirect scalar implicatures are found to pattern alike, the reaction time pattern is exactly the opposite of that reported for (direct) scalar implicatures in the literature, with delays for (putatively) *literal* responses. While the result for presuppositions is in line with

[13] See Moulton (2007) for an earlier proposal of *minimize accommodation* based on experimental work on ellipsis resolution. For yet another recent study in this area that only came to my attention after writing this chapter, see Domaneschi et al. (2013).

the previous finding by Chemla and Bott (2013), the result for indirect scalars is surprising. Romoli and Schwarz explore possible explanations for this pattern, and also sketch the path forward for further research to elucidate this puzzling empirical situation.

An important recent theoretical contribution towards the issue of classifying different aspects of meaning is put forward by Tonhauser et al. (2013). They compare a variety of expressions and constructions which project out of embedded environments, including various types of presupposition triggers as well as conventional implicatures in English and Guaraní, and argue for a unified theory of projection. Crucially, the typology of projective meaning they present cuts across some of the traditional boundaries, e.g., by grouping together certain presupposition triggers with conventional implicatures. There have only been a small number of experimental investigations of conventional implicatures that relate to this discussion. In response to Amaral et al.'s (2008) criticism of Potts (2005), Harris and Potts (2009a, 2009b) investigate the conditions under which conventional implicatures can receive non-speaker oriented interpretations, using both experimental and corpus methodologies. More recently, Syrett et al. (2014) investigate conventional implicatures and presuppositions by looking at both truth value judgments and reaction times. They find an interesting effect of linguistic context, in that the contribution of sentence final appositives (which are assumed to contribute a conventional implicature) can become part of the *at-issue* content, in contrast with sentence initial or medial ones. Further experimental and cross-linguistic investigations are in order to assess the nature and variation in types of projective content in natural language more thoroughly.

Yet another important phenomenon at play in interpreting utterances in context is that of domain restriction, which directly interacts with presuppositional phenomena. This has long been known to come into play for individual phrases (e.g., noun phrases), and both choice of referents and truth of statements crucially depend on which domain is chosen (Westerstahl 1984; von Fintel 1994). For definite descriptions, the felicity and choice of referent often crucially depends on domain restriction (Neale 1990). Only a small number of studies have directly investigated questions concerning the process of domain selection and restriction experimentally. Warren (2003) finds facilitation effects in self-paced reading for both quantifiers and plural definites when a plural set antecedent that can provide the domain is present in the context. Harris et al. (2013) investigated the choice of domain for adverbial quantifiers such as *mostly*, which can quantify either over parts or times, and find a general preference for the former. And Schwarz (2012) varies the linguistic context for definites which—on their own—have multiple possible referents in a visual display, with varying effects on the choice of domain for the definite in both response behavior and eye movements. This variation is argued to be driven by listeners' reasoning about what the question under discussion is in a given linguistic context.

A better understanding of the processes involved in domain selection and restriction will be of central importance for work on presupposition processing. By way of illustration, consider the study of accommodation by Evans (2005), which manipulated the context to vary whether or not a unique individual of the right kind would likely be involved in the described setting:

(13) a. Juan drove up to the busy tollbooths. . .
b. Juan looked at the busy tollbooths. . .

. . . The toll-taker was rude.

In a fill-in-the-blank task, subjects had to provide the determiner in the target sentence, and chose the definite more frequently in the first context. Rather than merely adding some missing information, this accommodation effect can be seen in light of the assumptions about the domain of interpretation that speakers make based on contextual information on whether we are talking about a specific toll-booth (and corresponding toll-taker) or not.

In conclusion, a full picture of processing meaning in context will only be possible once we study the full variety of meaning-related phenomena, both on their own and in interaction with one another. While initial steps have been taken for various specific aspects of meaning, in particular implicatures and presuppositions, a more comprehensive perspective clearly is called for. Given the methodological advances through the work reviewed here, there is a clear path forward for a host of novel research directions.

4.6 Acquisition

While most of the work in this volume reports investigations involving comprehension by adults, the questions and perspectives presented there of course can also be considered from the angle of language acquisition. While there are several related lines of work that involve presuppositional aspects of meaning, a full-fledged investigation of presuppositions and their various properties in their own right is largely yet to be carried out. As with the earlier adult literature, much of the existing work focuses on definite descriptions, and more narrowly on those involving restrictive relative clauses (Hamburger and Crain 1982, starting with) or other postnominal modifiers, such as prepositional phrases (Trueswell et al. 1999; Hurewitz et al. 2000), specifically in connection with the contextual needs for restriction in order to satisfy the presuppositions of the definite. More recently, Syrett et al. (2010) investigated definites containing gradable adjectives and find evidence for 3-year-olds' understanding of the existence and uniqueness presuppositions of the definite article, as well as their ability to shift their contextual assumptions appropriately when accommodation is needed. In contrast with presuppositions, the acquisition of scalar implicatures has received quite a lot of attention in the literature over the last decade or so. Most of the findings here focus on non-adult-like behavior in children, who tend to accept literal interpretations (lacking the scalar implicature) much more generally than adults (Noveck 2000; Papafragou and Musolino 2002; Huang and Snedeker 2009), though it is not necessarily clear to what extent such results show that children are unable to compute scalar implicatures in the first place (Katsos and Bishop 2011).

One related area that has been investigated quite extensively is that of children's comprehension of attitude predicates, such as *think* and *know* (for references, see Dudley et al. 2014). As some of these (e.g., *know*) come with a factive presupposition, a crucial question is to what extent children are sensitive to this aspect of meaning and the corresponding differences between such verbs. Dudley et al. (2014) address this question head on by testing children's interpretation of *think* and *know*, both in unembedded contexts and in the scope of negation. They had children find hidden toys in one of two boxes, utilizing cues provided by a puppet, which were relayed to the children by means of an attitude report. Their results have important consequences for the theoretical options considered in the literature for explaining the acquisition of attitude verbs in general, and also show that at least some of the 3-year olds in their study have an exquisite understanding of the factive component of *know*, including its presuppositional property of projection out of the scope of negation.

While these initial steps towards understanding the acquisition of presuppositional content constitute important progress, it is clear that the various other theoretical issues concerning presuppositions that were discussed above merit more detailed investigation from the perspective of acquisition as well. Given the existing methodologies and findings from the literature on implicatures, as well as the emerging body of adult studies, the door should now be wide open for researchers to plow ahead and extend the empirical domain of study for presuppositions even further in this direction.

5 Conclusion and Outlook

As should be clear from the various pointers throughout the previous section, we have begun to learn a good bit about presuppositions in language comprehension, but the field is still in its initial stages and much work remains to be done. To return to the questions raised in Sect. 3.3, progress has been made in studying the time-course of presuppositions, which seem to be available rapidly and at least as early as conventionally encoded *at-issue* content. This still leaves open whether we're looking at rapid pragmatic processes, or whether presuppositions should be seen as conventionally encoded. Concerning the status of presuppositions, a fairly wide range of data is in line with the notion that they are backgrounded (e.g., based on reaction times and cancellation strategies). It likely is this backgroundedness that accounts for the interesting combination of online evidence for rapid availability of presupposed content and delays in reaction times for rejecting statements based on false presuppositions. It remains unclear precisely in what circumstances presuppositions can be accommodated, and whether accommodation comes with a processing cost. Finally, presupposition projection continues to constitute a central area of investigation, both theoretically and experimentally. One of the key questions concerns the role of incrementality and its place in the components of grammar in this regard: is it hardwired into presupposition resolution machinery, or is it merely a processing preference? Relatedly, we need to compare the factors affecting the choice of antecedents for pronouns and presupposition triggers, as well as the online processes involved

in this. Finally, as emphasized in Sect. 4.2.2, and various aspects of our discussion throughout, it is quite possible that answers to the basic questions vary for different presupposition triggers.

Consequently, there is much room for empirical extensions of research in this area, in various directions. First, all of the theoretical issues above should be studied more comprehensively by considering a broad range of triggers. Secondly, it is desirable to address one and the same question from various methodological angles to assess the generality of findings, but many of the specific issues reviewed here have only been looked at within one methodological approach, often a quite basic behavioral one. Relatedly, the range of experimental methods utilized to investigate presuppositional phenomena altogether is still relatively limited, mostly to behavioral studies and eye tracking. While some studies using neuro-imaging techniques exist (mostly on definites; see discussion and references in Sect. 4.4), a more extensive use of such methods would substantially enrich research in this area. Finally, most of the work reported here is focused on language comprehension in healthy adults. As mentioned in the previous section, extensions of these approaches to the study of children and their acquisition of knowledge about presuppositions would strongly enhance the overall enterprise. Similarly, consideration of individuals with language-related disorders would seem to open up new angles of understanding, specifically with regard to the question of which aspects of meaning are narrowly based on linguistic knowledge, and which crucially involve domain-general resources. Corresponding lines of research have been initiated for implicatures (e.g. Katsos et al. 2011), but not for presuppositions.

As our methods for investigating presuppositional phenomena (as well as other aspects of meaning in context) evolve, we will also be able to address further, and more detailed theoretical questions. On the processing side, the ever increasing amount of evidence will allow us to flesh out a more comprehensive processing model for the various ingredients of linguistic meaning. Given the growing body of work in this area and ongoing research projects both by the contributors to this volume and other researchers, the coming years promise to yield substantial progress in all of these directions.

Acknowledgement The writing of this introduction benefitted substantially from discussions in a seminar I taught at the University of Pennsylvania in the fall of 2013. Many thanks to the students in this class, as well as to Alan Munn, for extensive and insightful comments on the topics considered here.

References

Abrusán, M., and K. Szendrői. 2013. Experimenting with the king of France: Topics, verifiability, and definite descriptions. *Semantics and Pragmatics* 6 (10): 1–43.

Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presuppositions. *Proceedings of SALT* 12:1–19.

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27 (1): 37–80.

Amaral, Patrícia, and Chris Cummins. 2014. A cross-linguistic study on information backgrounding and presupposition projection. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

Amaral, Patrícia, Craige Roberts, and E. Allyn Smith. 2008. Review of the logic of conventional implicatures by Chris Potts. *Linguistics and Philosophy* 30 (6): 707–749.

Amaral, P., C. Cummins, and N. Katsos. 2011. Experimental evidence on the distinction between foregrounded and backgrounded meaning. Proceedings of the ESSLLI 2011 workshop on projective meaning, Ljubljana.

Atanassov, Dimka. 2014. Implicatures of modals and quantifiers in processing. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

Beaver, David, and Bart Geurts. 2012. Presuppositions. In *Semantics: An international handbook of natural language meaning*. vol. 3. ed. Claudia Maienborn, Klaus von Heusinger, and Paul Portner. Berlin: Mouton de Gruyter.

Beaver, David, and Emiel Krahmer. 2001. Presupposition and partiality: Back to the future. *Journal of Logic, Language and Information* 10 (2): 147–182.

Beaver, David, and Henk Zeevat. 2012. Accommodation. In *Oxford handbook of linguistic interfaces,* ed. G. Ramchand and C. Reiss. Oxford: Oxford University Press.

Bott, L., and I. A. Noveck. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* 51 (3): 437–457.

Bott, Lewis, Todd M. Bailey, and Daniel Grodner. 2012. Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language* 66 (1): 123–142.

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

Breheny, Richard, Heather J. Ferguson, and Napoleon Katsos. 2013. Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes* 28 (4): 443–467.

Burkhardt, Petra. 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language* 98 (2): 159–168.

Chambers, Craig, and Valerie San Juan. 2005. Accommodation and the interpretation of presupposition during referential processing. Poster presented at the 18th CUNY sentence processing conference.

Chambers, Craig G., and Valerie San Juan. 2008. Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition* 108 (1): 26–50.

Chambers, Craig G., M. K. Tanenhaus, K. M. Eberhard, H. Filip, and G. Carlson. 2002. Circumscribing referential domains in real-time sentence comprehension. *Journal of Memory and Language* 37:30–49.

Chemla, E. 2009a. Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. Under revision for *Semantics and Pragmatics*.

Chemla, Emmanuel. 2009b. Presuppositions of quantified sentences: Experimental data. *Natural Language Semantics* 17 (4): 299–340.

Chemla, Emmanuel, and Lewis Bott. 2013. Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and Cognitive Processes* 38 (3): 241–260.

Chemla, Emmanuel, and Philippe Schlenker. 2012. Incremental vs. symmetric accounts of presupposition projection: An experimental approach. *Natural Language Semantics* 20 (2): 177–226.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In *Structures and beyond,* ed. A. Belletti. Oxford: Oxford University Press.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and grammar. An introduction to semantics*. Cambridge: MIT Press.

Chierchia, Gennaro, Teresa Guasti, Andrea Gualmini, Luisa Meroni, Stephen Crain, and Francesca Foppolo. 2004. Semantic and pragmatic competence in children's and adults' comprehension of *Or*. In *Experimental pragmatics,* ed. Noveck Ira and Sperber Dan, 283–300. New York: Palgrave Macmillan.

Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path. The use of context by the psychological parser. In *Natural language parsing: Psychological, computational, and theoretical perspectives,* ed. D. R. Dowty, L. Karttunen, and A. Zwicky, 320–358. Cambridge: Cambridge University Press.

Cummins, Chris, Patrícia Amaral, and Napoleon Katsos. 2013. Backgrounding and accommodation of presupposition: An experimental approach. In *Proceedings of SuB 17,* ed. E. Chemla, V. Homer, and G. Winterstein, 201–218. http://semanticsarchive.net/sub2012/: Semanticsarchive.

Degen, J., and M. K. Tanenhaus. 2012. Naturalness of lexical alternative predicts time course of scalar implicature processing. Poster presented at CUNY.

Destruel, Emilie, Edgar Onea, Dan Velleman, Dylan Bumford, and David Beaver. 2014. A cross-linguistic study of the non-at-issueness of exhaustive inferences. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

Domaneschi, Filippo, Elena Carrea, Carlo Penco, and Alberto Greco. 2013. The cognitive load of presupposition triggers: Mandatory and optional repairs in presupposition failure. *Language and Cognitive Processes* 0 (0): 1–11. http://www.tandfonline.com/doi/pdf/10.1080/01690965.2013.830185.

Dudley, Rachel, Naho Orita, Valentine Hacquard, and Jeffrey Lidz. 2014. Three-year-olds' understanding of *know* and *think*. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.

Evans, William. 2005. Determining a small world: Accommodating with the definite article in plural contexts. B.A. Honors Thesis, University of Massachusetts Amherst.

Fox, Danny. 2008. Two short notes on Schlenker's theory of presupposition projection. *Theoretical Linguistics* 34 (3): 237–252.

Frazier, Lyn. 2006. The big fish in a small pond: Accommodation and the processing of novel definites. Manuscript, University of Massachusetts, Amherst.

Frazier, Lyn. 2012. Meaning in psycholinguistics. In *Semantics: An international handbook of natural language meaning*. vol. 3, chap. 102. ed. Klaus von Heusinger, Claudia Meienborn, and Paul Portner, 2703–2724. Berlin: De Gruyter.

Frege, Gottlob. 1892. On sense and reference. In *Translations from the philosophical writings of Gottlob Frege,* ed. P. Geach and M. Black, 56–78. Oxford: Blackwell.

George, Benjamin R. 2008a. Predicting presupposition projection: Some alternatives in the strong kleene tradition. Ms., UCLA. Semanticsarchive at http://semanticsarchive.net/Archive/DY0YTgxN/.

George, Benjamin R. 2008b. Presupposition repairs: A static, trivalent approach to predict projection. Master's thesis, UCLA, Los Angelos, CA.

Geurts, Bart. 1999. *Presuppositions and pronouns*. Amsterdam: Elsevier.

Gibson, Edward, and Neal J. Pearlmutter. 2011. *The processing and acquisition of reference*. Cambridge: MIT Press.

Gibson, E., and J. Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14 (3): 225–248.

Gordon, Peter C., Randall Hendrick, Kerry Ledoux, and Chin Lung Yang. 1999. Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes* 14 (4): 353–379.

Grice, H. P. 1975. Logic and conversation. In *Syntax and semantics: Speech acts,* ed. P. Cole and J. L. Morgan. New York: Academic Press.

Grodner, Daniel J., Natalie M. Klein, Kathleen M. Carbary, and Michael K. Tanenhaus. 2010. "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition* 116 (1): 42–55.

Grosz, Barbara. 1977. The representation and use of focus in dialogue and understanding. Technical Report 151. SRI International, Menlo Park: Artificial Intelligence Center.
Grosz, B. J., A. K. Joshi, and S. Weinstein. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 21 (2): 203–225.
Gualmini, Andrea, Sarah Hulsey, Valentine Hacquard, and Danny Fox. 2008. The question-answer requirement for scope assignment. *Natural Language Semantics* 16:205–237.
Gundel, Jeannette K., and Nancy Hedberg, eds. 2008. *Reference: Interdisciplinary perspectives. New directions in cognitive science*. Oxford: Oxford University Press.
Gundel, Jeannette K., Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69 (2): 274–307.
Hamburger, H., and S. Crain. 1982. Relative acquisition. *Language Development* 1:245–274.
Hanna,, J. E., M. K. Tanenhaus, and J. C. Trueswell. 2003. The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language* 49 (1): 43–61.
Harris, Jesse, and Christopher Potts. 2009a. Predicting perspectival orientation for appositives. *Proceedings from the Annual Meeting of the Chicago Linguistic Society* 45 (1): 207–221.
Harris, Jesse A., and Christopher Potts. 2009b. Perspective-shifting with appositives and expressives. *Linguistics and Philosophy* 32 (6): 523–552. ArticleType: research-article/Full publication date: 2009/Copyright 2009 Springer.
Harris, Jesse, A., Charles Clifton, and Lyn Frazier. 2013. Processing and domain selection: Quantificational variability effects. *Language and Cognitive Processes* 0 (0): 1–26.
Haviland, Susan E., and Herbert H. Clark. 1974. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior* 13 (5): 512–521.
Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Ph.D. thesis, UMass.
Heim, Irene. 1983. On the projection problem for presuppositions. In *Proceedings of the second West Coast conference on formal linguistics,* ed. Michael Barlow, Daniel P. Flickinger, and Michael T. Wescoat, 114–125. Stanford: Stanford University Press.
Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
Heller, Daphna, Daniel Grodner, and Michael K. Tanenhaus. 2008. The role of perspective in identifying domains of reference. *Cognition* 108 (3): 831–836.
Hirsch, Aaron, and Martin Hackl. 2013. Incremental presupposition evaluation in disjunction. Talk presented at NELS 44, UConn.
Huang, Y. T., and J. Snedeker. 2009. Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive psychology* 58 (3): 376–415.
Huang, Yi Ting, and Jesse Snedeker. 2011. Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes* 26 (8): 1161–1172.
Huang, Y., E. Spelke, and J. Snedeker. 2013. What exactly do number words mean? *Language Learning and Development* 9 (2): 105–129.
Hurewitz, Felicia, Sarah Brown-Schmidt, Kirsten Thorpe, Lila R. Gleitman, and John C. Trueswell. 2000. One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research* 29 (6): 597–626.
Jayez, Jacques. 2013. Presupposition triggers and orthogonality. Ms. Lyon.
Jayez, J., and B. van Tiel. 2011. Is 'only' special? Presentation at the XPRAG workshop, Pisa, Italy.
Jayez, Jacques, Valeria Mongelli, Anne Reboul, and Jean-Baptiste van der Henst. 2014. Weak and strong triggers. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.
Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language: Proceedings of the third Amsterdam colloquium.* vol. I. ed. J. Groenendijk, T. Janssen, and M. Stokhof, 227–321. Amsterdam: Mathematical Center.
Karttunen, Lauri. 1971. Some observations on factivity. *Paper in Linguistics* 4 (1): 55–69.
Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4 (2): 169–193.

Katsos, Napoleon, and Dorothy V. M. Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120 (1): 67–81.
Katsos, Napoleon, Clara Andrés Roqueta, Rosa Ana Clemente Estevan, and Chris Cummins. 2011. Are children with specific language impairment competent with the pragmatics and logic of quantification? *Cognition* 119 (1): 43–57.
Keysar, B., D. J. Barr, J. A. Balin, and J. S. Brauner. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science* 11 (1): 32.
Keysar, B., S. Lin, and D. J. Barr. 2003. Limits on theory of mind use in adults. *Cognition* 89 (1): 25–41.
Kim, Christina. 2014. Presupposition satisfaction, locality and discourse constituency. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.
Kim, Min-Joo. 2007. Formal linking in internally headed relatives. *Natural Language Semantics* 15: 279–315.
Kim, C., C. Gunlogson, M. Tanenhaus, J. Runner, and S. und Bedeutung. 2009. Focus alternatives and contextual domain restriction: A visual world eye-tracking study on the interpretation of only. In *Proceedings of Sinn und Bedeutung 13,* ed. Arndt Riester and Torgrim Solstad, 261. Stuttgart.
Kleene, Stephen. 1952. *Introduction to metamathematics*. Amsterdam: North-Holland.
Kripke, Saul. 1991. *Presupposition and anaphora: Remarks on the formulation of the projection problem*. Princeton: Ms., Princeton University.
Kripke, Saul A. 2009. Presupposition and anaphora: Remarks on the formulation of the projection problem. *Linguistic Inquiry* 40 (3): 367–386.
Lasersohn, Peter. 1993. Existence presuppositions and background knowledge. *Journal of Semantics* 10 (2): 113–122.
Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature. Language, speech, and communication*. Cambridge: MIT Press.
Lewis, David. 1979. Scorekeeping in a language game. In *Semantics from different points of view,* ed. Rainer Bäuerle, Urs Egli, and Arnim von Stechow. Berlin: Springer.
Moulton, Keir. 2007. Small antecedents: Syntax or pragmatics. In *Proceedings of NELS 37*. vol. 1. ed. E. Elfner and M. Walkow, 45–58. Amherst: GLSA.
Nadig, Aparna S., and Julie C. Sedivy. 2002. Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science* 13 (4): 329–336.
Neale, Stephen. 1990. *Descriptions*. Cambridge: MIT Press.
Noveck, Ira A. 2000. When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition* 78 (2): 165–188.
Noveck, Ira A., and Andres Posada. 2003. Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language* 85 (2): 203–210.
Noveck, I., and A. Reboul. 2008. Experimental pragmatics: A gricean turn in the study of language. *Trends in Cognitive Sciences* 12 (11): 425–431.
O'Brien, Edward J., Dolores M. Shank, Jerome L. Myers, and Keith Rayner. 1988. Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (3): 410–420.
Papafragou, Anna, and Julien Musolino. 2002. The pragmatics of number. Proceedings from the 24th annual conference of the Cognitive Science Society. Hillsdale: Erlbaum.
Papafragou, Anna, and Niki Tantalou. 2004. Children's computation of implicatures. *Language Acquisition* 12 (1): 71–82.
Potts, Christopher. 2005. *The logic of conventional implicatures. Oxford studies in theoretical linguistics*. Oxford: Oxford University Press.
Potts, Christopher. To appear. Presupposition and implicature. In *The handbook of contemporary semantic theory*. 2nd ed, ed. Shalom, Lappin and Chris Fox. Oxford: Wiley-Blackwell.
Pylkkänen, Liina, and Brian McElree. 2006. The syntax-semantics interface: On-line composition of sentence meaning. In *Handbook of psycholinguistics,* ed. M. Traxler and M. A. Gernsbacher, 537–577. New York: Elsevier.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27:53–93.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. ed. J. H. Yoon and Andreas Kathol, Papers in Semantics, number 49 in OSU Working Papers in Linguistics, 91–136. OSU.

Romoli, Jacopo. To appear. The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics*.

Romoli, Jacopo, and Florian Schwarz. 2014. An experimental comparison between presuppositions and indirect scalar implicatures. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.

Romoli, Jacopo, Manizeh Khan, Jesse Snedeker, and Yasutada Sudo. 2012. Resolving temporary referential ambiguity using presupposed content. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.

Romoli, Jacopo, Yasutada Sudo, and Jesse Snedeker. 2011. An experimental investigation of presupposition projection in conditional sentences. *Proceedings of SALT* 21:592–608.

Romoli, Jacopo, Manizeh Khan, Yasutada Sudo, and Jesse Snedeker. 2014. Resolving temporary referential ambiguity using presupposed content. Poster presented at the 26th annual CUNY sentence processing conference.

Russell, Bertrand. 1905. On denoting. *Mind* 14:479–493.

Schlenker, P. 2008a. Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics* 34 (3): 157–212.

Schlenker, P. 2008b. Presupposition projection: Explanatory strategies. *Theoretical Linguistics* 34 (3): 287–316.

Schlenker, Philipe. 2008c. Presupposition projection: The new debate. In *Proceedings of SALT 18,* ed. T. Friedman and S. Ito. Ithaca: CLC Publications.

Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 3:1–78.

Schlenker, Philippe. 2010a. Local contexts and local meanings. *Philosophical Studies* 151 (1): 115–142.

Schlenker, Philippe. 2010b. Presuppositions and local contexts. *Mind* 119 (474): 377–391.

Schlenker, Philippe. 2011a. Presupposition projection: Two theories of local contexts part I. *Language and Linguistics Compass* 5 (12): 848–857.

Schlenker, Philippe. 2011b. Presupposition projection: Two theories of local contexts part II. *Language and Linguistics Compass* 5 (12): 858–879.

Schoubye, Anders J. 2010. Descriptions, truth value intuitions, and questions. *Linguistics and Philosophy* 32 (6): 583–617.

Schwarz, F. 2014a. Processing presupposed content. *Journal of Semantics* 24 (4): 373–416.

Schwarz, F. 2012a. Presuppositions vs. asserted content in online processing. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing

Schwarz, F. 2014b. Symmetry and incrementality in conditionals. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.

Schwarz, F. 2012. Domain restriction and discourse structure-evidence from online processing. Poster presentation, 25th CUNY sentence processing conference.

Schwarz, F. 2014c. False but slow: Rejecting statements with non-referring definites. Ms., UPenn.

Schwarz, F. 2014. Presuppositions are fast, whether hard or soft-evidence from the visual world paradigm. Talk presented at SALT 24.

Schwarz, F., and Sonja Tiemann. 2012. Presupposition processing-The case of German *wieder*. In *Proceedings from the 18th Amsterdam colloquium FoLLI,* ed. M. Aloni, V. Kimmelman, F. Roelofsen, G. Weidman Sassoon, K. Schulz, and M. Westera. Berlin: Springer.

Schwarz, F., and S. Tiemann. 2013a. Presupposition projection in online processing. (Ms., submitted).

Schwarz, Florian, and Sonja Tiemann. 2013b. The path of presupposition projection in processing—the case of conditionals. In *Proceedings of SuB 17,* ed. E. Chemla, V. Homer, and G. Winterstein, 509–526. Paris: semanticsarchive.net

Sedivy, J. C. 2003. Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of psycholinguistic research* 32 (1): 3–23.

Sedivy, Julie C., Michael K. Tanenhaus, Craig G. Chambers, and Gregory N. Carlson. 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition* 71 (2): 109–147.

Simons, M. 2001. On the conversational basis of some presuppositions. *Proceedings of SALT* 11:431–448.

Simons, M. 2003. Presupposition and accommodation: Understanding the Stalnakerian picture. *Philosophical Studies* 112 (3): 251–278.

Singh, Raj, Evelina Fedorenko, and Edward Gibson. 2013. Presupposition accommodation is costly only in implausible contexts. Manuscript, Carleton University & MIT.

Smith, E. Allyn, and K. C. Hall. 2011. Projection diversity: Experimental evidence. Proceedings from the workshop on projective meaning at ESLLI 2011.

Spenader, Jennifer. 2002. Presuppositions in Spoken Discourse. Ph.D. thesis, University of Stockholm. Computational Linguistics.

Stalnaker, Robert C. 1970. Pragmatics. *Synthese* 22 (1): 272–289.

Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic* 2 (4): 447–457.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and philosophy,* ed. Milton K. Milton and Peter K. Unger. New York: New York University Press.

Stalnaker, Robert. 1978. Assertion. In *Pragmatics, volume 9 of syntax and semantics,* ed. Peter Cole, 315–322. New York: Academic Press.

Strawson, Peter. 1950. On referring. *Mind* 59:320–344.

Strawson, P. F. 1964. Intention and convention in speech acts. *The Philosophical Review* 73 (4): 439–460.

Syrett, K., C. Kennedy, and J. Lidz. 2010. Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics* 27 (1): 1.

Syrett, Kristen, Todor Koev, Nicholas Angelides, and Maxwell Kramer. 2014. *Experimental evidence for the truth conditional contribution and shifting information status of appositives*. Journal of Semantics (Advance Access). doi:10.1093/jos/ffu007. New Brunswick: Ms., Rutgers University.

Tanenhaus, M. K., M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268 (5217): 1632–1634.

Teresa Guasti, Maria, Gennaro Chierchia, Stephen Crain, Francesca Foppolo, Andrea Gualmini, and Luisa Meroni. 2005. Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes* 20 (5): 667–696.

Tiemann, Sonja, Mareike Kirsten, Sigrid Beck, Ingro Hertrich, and Bettina Rolke. 2014. Presupposition processing and accommodation: An experiment on *wieder* ('again') and consequences for other triggers. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.

Tiemann, Sonja, Mareike Schmid, Nadine Bade, Bettina Rolke, Ingo Hertrich, Hermann Ackermann, Julia Knapp, and Sigrid Beck. 2011. Psycholinguistic evidence for presuppositions: On-line and off-line data. In *Proceedings of Sinn & Bedeutung 15,* ed. I. Reich, E. Horch, and D. Pauly, 581–597. Saarbrücken: Saarland University Press.

Tonhauser, J., D. Beaver, C. Roberts, and M. Simons. 2013. Towards a taxonomy of projective content. *Language* 89 (1): 66–109.

Trueswell, John C., Irina Sekerina, Nicole M. Hill, and Marian L. Logrip. 1999. The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition* 73 (2): 89–134.

van Berkum, Jos J. A., Colin M. Brown, and Peter Hagoort. 1999. Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language* 41:147–182.

van Berkum, Jos J. A., Colin M. Brown, Peter Hagoort, and Pienie Zwitserlood. 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology* 40:235–248.

van Fraasen, Bas C. 1968. Presupposition, implication, and self-reference. *The Journal of Philosophy* 65 (5): 136–152.

van Fraasen, Bas C. 1971. *Formal semantics and logic*. New York: Macmillan.

van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.

van der Sandt, R., and B. Geurts. 1991. Presupposition, anaphora, and lexical content. In *Text understanding in LILOG,* ed. O. Herzog and C.-R. Rollinger, 259–296.

Velleman, Dan, David Beaver, Dylan Bumford, Emilie Destruel, and Edgar Onea. 2011. "Yes, but ..."—exhaustivity and at-issueness across languages. Poster presented at PEPA 2011.

von Fintel, Kai. 1994. Restrictions on quantifier domains. Ph.D. thesis, UMass.

von Fintel, Kai. 2004. Would you believe it? The king of France is back! (presuppositions and truth-value intuitions). In *Descriptions and beyond,* ed. Reimer Marga and Bezuidenhout Anne, 315–341. Oxford: Oxford University Press.

von Fintel, Kai. 2008. What is presupposition accommodation, again? *Philosophical Perspectives* 22 (1): 137–170.

Warren, Tessa. 2003. The processing complexity of quantifiers. In *UMass occasional papers in linguistics*. vol. 27. ed. Luis Alonso-Ovalle, 211–327. Amherst: GLSA.

Westerstahl, Dag. 1984. Determiners and context sets. In *Generalized quantifiers in natural language,* ed. J. van Benthem and Alice ter Meulen, 45–71. Dordrecht: Foris.

Xue, J., and E. Onea. 2011. Correlation between presupposition projection and at-issueness: An empirical study. Proceedings from the Workshop on Projective Meaning at ESLLI 2011, Ljubljana.

Zeevat, Henk. 1992. Presupposition and accommodation in update semantics. *Journal of semantics* 9 (4): 379–412.

Presupposition Processing and Accommodation: An Experiment on *wieder* ('again') and Consequences for Other Triggers

Sonja Tiemann, Mareike Kirsten, Sigrid Beck, Ingo Hertrich and Bettina Rolke

Abstract This paper investigates the processing and accommodation of the presuppositions triggered by *wieder* ('again'). We conducted a word by word self-paced reading experiment where we presented sentences containing *wieder* in a context which furnished the relevant presupposition and one which did not. We then additionally asked questions to determine whether people accommodate the presupposition of *wieder* when it is not explicitly given in the context. The results show that effects due to a missing presupposition arise very early during reading and that there is no accommodation of the presupposition introduced by *wieder*. On the basis of this, we introduce the interpretation principle *minimize accommodation* and discuss what implications this brings about for other presupposition triggers. Another interesting result is a late increase in reading times in the condition which verifies the presupposition. We argue that this is due to the referential nature of *wieder*.

Keywords Presuppositions · Psycholinguistics · Presupposition processing · Accommodation · Self-paced reading

S. Tiemann (✉) · M. Kirsten
Collaborative Research Centre 833, University of Tuebingen,
Nauklerstr. 35, Tuebingen, Germany
e-mail: sonja.tiemann@uni.tuebingen

M. Kirsten
e-mail: mareike.kirsten@uni-tuebingen.de

S. Beck
Descriptive and Theoretical Linguistics,
University of Tuebingen, Tuebingen, Germany
e-mail: sigrid.beck@uni-tuebingen.de

I. Hertrich
Ingo Hertrich, Hertie Institute for Clinical Brain Research,
University of Tuebingen, Tuebingen, Germany
e-mail: ingo.hertrich@uni-tuebingen.de,

B. Rolke
Evolutionary Cognition, University of Tuebingen,
Tuebingen, Germany
e-mail: bettina.rolke@uni-tuebingen.de

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_2,

1 Introduction

According to a theory which analyzes presuppositions as restrictions on appropriate contexts, the context set has to entail all the presuppositions of a sentence in order for the sentence to get a truth value. The context set is the set which contains all "possible worlds where all the propositions that are the background assumptions of speakers are true" (Stalnaker 1973, p. 450). What that means for a sentence like (1-a) is that the context set has to consist only of those worlds in which Susanne went skiing some time before last weekend. A semi-formal representation is given in (1-b) which states that the context set c has to be a set of worlds in which Susan went skiing at prior to last weekend.

(1) a. Last weekend, Susan went skiing again.
b. $c \Rightarrow \lambda w.\exists t[t < \textit{last weekend} \,\&\, \text{skiing}(\text{Susanne})(w)(t)]$

If the context is not such that it provides the relevant presupposition, there are only two ways out: Either, the hearer of a sentence like (1-a) accommodates the relevant presupposition, i.e., s/he adjusts the context set in such a way that it only contains the worlds in which the presupposition is true, or s/he decides that the relevant piece of background information is missing and thus the whole sentence cannot receive a well defined interpretation. However, in the case of *wieder*, the matter is even more complicated than that. It has been argued by Soames (1982), Heim (1990), and Kamp and Rossdeutscher (1994) among others that the temporal variable in the presupposition of *wieder* is not existentially bound but receives its value from the context. Consequently, (1-b) has to look like (2) where t is not existentially bound, but remains free.

(2) $c \Rightarrow \lambda w.t < \textit{last weekend} \,\&\, \text{skiing}(\text{Susanne})(w)(t)$

So from a theoretical perspective, there are two issues that have to be looked at when it comes to the processing of *wieder*: 1. How and when is the presupposition of *wieder* checked in a given context, and 2. How is the free time variable processed. The first question relates to the processing of presuppositions in general. The second question, however, is interesting from a wider perspective of sentence interpretation. For decades, psycholinguists have been concerned with how bound and free variables are processed. Recent research (e.g. Frazier and Clifton (2000), Koornneef (2008)) has shown that there is a processing advantage for bound over free variables. The issue at hand is even more interesting since previous research on the processing of variables has almost exclusively focused on individual variables. To date there is only very little work on the processing of temporal variables. In fact, Dickey (2000) seems to be the only one who makes a serious effort to combine semantic theories of tense and processing hypotheses. Moreover, to our knowledge, there is no one out there who has looked at the processing of variables which are introduced at the level of presupposition.

The experiment presented in this chapter will look at the two issues of presupposition verification and variable assignment. Concerning the former, the focus is

on what happens when the presupposition of *wieder* is given versus not given in a certain context. When is the missing presupposition recognized by the processor? What happens when it is recognized? Does accommodation of the missing presupposition apply or does the processor give up on the sentence completely? Are there alternative routes? As it will turn out, the answers to these questions are intimately connected to the second issue of assigning a value to the free time variable. Does variable assignment interfere with presupposition verification? When does the free variable receive its value during online processing?

This paper is organized as follows: First, we will lay out the possible processing predictions on the matter of the missing presupposition that we can deduct from previous experiments on presuppositions. Second, we will present a self-paced reading study which is designed to investigate the issues laid out above. Finally, the results and their implication for the semantic theory will be discussed in the last section.

2 What happens When a Presupposition is Not Entailed by the Context?

There is already an abundant amount of work on the processing of definite NPs in the psycholinguistic literature, but a lot less on other presupposition triggers. But what can these results tell us about the processing of presuppositions in general? In the following, we will try to excerpt common underlying mechanisms uncovered by these studies and what they can tell us about the processing of presuppositions.

In one of the first experiments which are relevant to the issue of presupposition processing, Altmann and Steedman (1988) investigated in a self-paced reading fashion how sentences with a definite determiner are processed when the uniqueness presupposition is not met. Even though their actual goal was to test Frazier (1978)'s *Minimal Attachment Hypothesis*, their results turn out to be very relevant for a theory of presupposition processing. In their experiment, they had subjects read sentences along the lines of either (4-a) or (4-b).[1] They presented these test sentences in two different contexts. One which introduced two safes (3-a), and another one in which only one relevant safe was mentioned (3-b). By having the PP modifying *the safe* in (4-a), the uniqueness presupposition is met even when the sentence is presented in a context like (3-a), whereas this is not the case for a sentence like (4-b).

(3) a. A burglar broke into a bank carrying some dynamite. He planned to blow open a safe. Once inside he saw that there was a safe with a new lock and a safe with an old lock.

b. A burglar broke into a bank carrying some dynamite. He planned to blow open a safe. Once inside he saw that there was a safe with a new lock and a strongbox with an old lock.

[1] Slashes indicate the strings of words that were presented together at the same time.

(4) a. The burglar/blew open/the safe/with the new lock/and made off/with the loot.
b. The burglar/blew open/the safe/with the dynamite/and made off/with the loot.

The results reveal that reading times came apart on the disambiguating region (i.e. the PP *with the new lock* or *with the dynamite*). Test sentence (4-b) was read more slowly in the context which introduced two safes than all the other conditions. Moreover there was no difference in reading times for (4-a) in either (3-a) or (3-b). There was, however, a difference between (4-a) and (4-b) when presented in a context like (3-a) with (4-b) being read more slowly. This shows that when the uniqueness presupposition of the definite determiner is not met, people experience processing difficulties quite early, i.e., before the end of the sentence. However, the results of this experiment also suggest that comprehenders do not detect an unmet presupposition immediately (on the NP). It seems that the processor rather delays the evaluation of the presupposition until later. In this experiment, such a strategy turned out to be especially useful since the constituent following the NP made it obvious whether the presupposition was met or not. If comprehenders were to check the presupposition right away, one would expect an increase in reading times on the NP already. Moreover, one would expect these inflated reading times for both test sentences when presented in a context such as (3-a). Since there were no effects of that sort, it seems that the processor delayed the decision of whether the presupposition of the definite article is met or not met to a later, disambiguating region of the sentence.

In relation to this, the studies in van Berkum et al. (1999) and van Berkum et al. (2003) are relevant. In an experiment with spoken sentences, van Berkum et al. (2003) replicated the findings obtained from a written language experiment in van Berkum et al. (1999) that definite NPs evoke early ERP effects when the uniqueness presupposition is not met. Their material comprised of discourses like (5) and (6). In (5), the uniqueness of the NP **the girl** in the last sentence is verified because there is only one girl salient in the context. In (6) on the other hand, there are two girls introduced in the context which are both equally salient and hence the uniqueness presupposition of the definite determiner is violated.

(5) David had told the boy and the girl to clean up their room before lunchtime. But the boy had stayed in bed all morning and the girl had been on the phone all the time. David told **the girl** that had been on the phone to hang up.

(6) David had told the two girls to clean up their room before lunchtime. But one of the girls had stayed in bed all morning and the other girl had been on the phone all the time. David told **the girl** that had been on the phone to hang up.

The ERP results reveal that there is an early negative deflection on the critical noun (*girl*) in (6) when compared to (5). This deflection occurs about 300–400 ms after the acoustic onset of the noun. Van Berkum et al. (2003) take this to suggest that "referential ambiguity" is detected very early during sentence comprehension. This can also be interpreted as saying that the unmet presupposition of the definite determiner leads to early effects in ERPs. What is crucial here is that even though the test items were such that there was a disambiguating relative clause after the NP, initial processing of the presupposition was not delayed until this part of the sentence had been

heard/read. In other words, interpreters checked the presupposition in the context as early as hearing/reading the noun and encountered a problem when it was not given even though they knew that the succeeding discourse could in principle still save the sentence from presupposition failure. This is an interesting result because it contrasts with the result found in Altmann and Steedman (1988) in the way that the ERP data suggest that the presupposition of *the* is checked as soon as possible regardless of what might follow afterwards. Without knowing how exactly the filler items in the three experiments looked like, it is hard to say how these different effects come about. It may well be that in the study of Altmann and Steedman (1988), subjects developed a processing strategy in which they always waited for the disambiguating PP before caring about the presupposition. The same could hold for the material in van Berkum et al. (1999) and van Berkum et al. (2003), but since these were physiological rather than behavioral studies, they can give us an even deeper understanding of what happens cognitively during sentence processing. Whatever the exact reasons for this difference in timing are, at this point it is sufficient to acknowledge that the electrophysiological experiment clearly shows that people realize the unmet presupposition of a definite NP at the earliest point possible, even though it might still be amended later on in the sentence.

Another ERP study which gives interesting insights in the processing of the presupposition of the definite determiner, is found in Burkhardt (2006). This paper reports an experiment in which a target sentence with a definite NP like (8) was presented in three different contexts. The *given* context in (7-a) explicitly introduces the individual which the definite NP refers to. The other two contexts do not explicitly verify the existence presupposition of the definite, but it can be inferred in the *bridged* condition (7-b), whereas this is not possible in the *new* condition in (7-c).

(7) a. Tobias visited a conductor in Berlin.
b. Tobias visited a concert in Berlin.
c. Tobias talked to Nina.

(8) He said that the conductor was very impressive.

The results show an early negative deflection 400 ms after the noun onset in the new condition (N400). This effect was less pronounced in the bridged condition. This shows that when the existence presupposition of the definite determiner is not given, an N400 emerges, parallel to what van Berkum et al. (2003) found for sentences where the uniqueness presupposition was violated. This effect is not as strong when the relevant individual can be easily inferred from the context. Most importantly however, Burkhardt (2006) found a late positive effect (P600) in the new and in the bridged condition when compared to the given condition. The author concludes that this "suggests that the respective discourse units are fully integrated at this point, indicating that an independent discourse referent is identified to be stored and maintained in discourse representation" (Burkhardt 2006, p. 166). In the words of a presupposition theory, the late positive deflection might be taken to signify accommodation.

Other presupposition triggers which have been studied by the means of online methods are German *auch* ('too'), English *also* and factives such as *to know*.

In a reading experiment using self-paced reading, Schwarz (2007) investigated the processing of German *auch* ('too') and English *also*. Since the results are roughly the same for both languages, we will only discuss the experiment conducted in English here for ease of presentation. In this experiment, subjects were asked to read sentence along the lines of (9).

(9) a. The congressman/who wrote to John/had also written to the mayor/to schedule a meeting/for the fundraiser.
b. The congressman/who wrote to John/had just written to the mayor/to schedule a meeting/for the fundraiser.
c. The congressman/who John wrote to/had also written to the mayor/to schedule a meeting/for the fundraiser.
d. The congressman/who John wrote to/had just written to the mayor/to schedule a meeting/for the fundraiser.

The sentences were constructed in such a way that the presupposition of *also* (that there is another x $\neq$ the mayor, such that the congressman wrote to x) is given in (9-a), but not in (9-c). The sentences in (9-b) and (9-d) do not trigger the relevant presupposition and were thus used as controls. The results show that a sentence like (9-a) was read faster than a sentence like (9-c) on the region containing *also*. This can again be seen as evidence for an early presupposition processing since the mode of presentation in this experiment was such that the string of words that *also* was presented with already made the presupposition evident. However, it fails to pinpoint the exact point in time at which presupposition processing takes place.

A similar point applies to Inhoff (1985)'s eye-tracking study. The author had subjects read a text in which some of the sentences were such that a non-factive (a)/factive (b) verb embedded a false (according to world knowledge) clause. An example is given below.

(10) . . . Today was an arithmetic test. The teacher asked little Tom. He {(a) said/(b) knew} that two and two equalled three. . . .

On the false complement, there was a significant difference in first gaze durations where the complement was read more slowly when presented after a factive verb than after a non-factive verb. Inhoff (1985) concludes that "the finding that lexical presuppositions affected the interpretation of the false complement suggests that reader's sentence interpretation was based not only on their empirical and analytical knowledge but also on their linguistic knowledge". The interesting finding is here that people did not experience processing difficulties on the false complement per se, but only when it was embedded under a verb which presupposes its complement to be true. This is just one more experiment which shows that presuppositions play an important role in sentence understanding. However, again the critical region looked at was so large (the whole complement), that we cannot determine precisely at which point the presupposition started to have an effect on people's reading times.

These experiments tell us a lot about the processing of presuppositions and the definite determiner in particular. First, that effects connected to presuppositions do arise with a wide variety of modalities (behavioral and physiological). Second, early deflections in ERPs indicate that presuppositions (at least the one of the definite determiner) are computed as soon as possible.

However, it would be premature to conclude from these experiments that all presupposition triggers are processed in the same way. We have already discussed that there are substantial semantic differences between a trigger like *again* on the one hand and the definite determiner on the other hand. Another issue which is not really answered by the experiments presented here is when the presupposition of a trigger other than the definite is processed. Those studies that investigated other triggers were set up in such a way that it is not possible to track the exact time course of presupposition processing, simply because the regions analyzed were too large. Additionally, it is not clear what happens when a presupposition is not given in the context. While Schwarz (2007) concludes that his results show that the presupposition of *also* is not accommodated, Burkhardt (2006) seems to suggest that the effects obtained in her experiment signify accommodation. The experiment in this paper is designed to provide answers to these questions.

3 Experiment: *wieder* in Supporting and Neutral Contexts

The experiment presented in this section aims to address the questions laid out above. The two central questions are: 1. How is the presupposition of *wieder* processed?, and 2. What is the exact time course of this process? An additional issue which will be looked at in this experiment is the notion of accommodation. Most theories of presuppositions assume that in case of presupposition failure, a sentence will either be undefined or the relevant presupposition will be accommodated. Since it has been argued for triggers like *again* and *too* (cf. Kripke (2009), Beaver and Zeevat (2008)) that their respective presuppositions are especially hard to accommodate, it is especially interesting to test experimentally how people deal with sentences containing *again* in a context which does not support its presupposition.

To this end, we set up an experiment using self-paced reading (SPR) with questions which targeted the presupposition of *again*. As we have seen in Schwarz (2007)'s experiments on *auch* and *also*, SPR is sensitive enough to capture presupposition related effects. The present study extends on Schwarz (2007)'s experiments in that it presents target sentences in a word by word fashion. This ensures that effects due to the presupposition will emerge as soon as they arise during sentence processing.

3.1 Method and Materials

The basic idea was to present a target sentence with one of two context sentences which differed as minimally as possible. In order to do this, we created 40 items along the lines presented below, where the presupposition of (12)-a is given in (11-b) (supporting context), but not in (11-a) (neutral context), and the presupposition of (12)-b is given in (11-a) (supporting context), but not in (11-b) (neutral context). In order to prevent anticipation, we additionally constructed target sentences where the subject of the target sentence was not the beneficiary but the agent.

(11) a. Letzte Woche hat Linda Judith eine rosa Lampe für ein Zimmer
last week has Linda Judith a pink lamp for a room
gekauft.
bought
Last week, Linda Judith a pink lamp for a room.

b. Letzte Woche hat Judith Linda eine rosa Lampe für ein Zimmer
last week has Judith Linda a pink lamp for a room
gekauft.
bought
Last week, Judith bought Linda a pink lamp for a room.

(12) Vor zwei Tagen hat {(a) Linda /(b) Judith} wieder eine rosa Lampe
ago two days has {(a) Linda /(b) Judith} again a pink lamp
erhalten, als sie mit einer Freundin unterwegs war.
received, when she with a friend out was
Two days ago, {(a) Linda /(b) Judith} received a pink lamp again, when she was out with a friend.

Apart from the 40 experimental items, we created 40 filler items in order to mask the purpose of the experiment. The filler items were constructed in a parallel fashion to the experimental items. Crucially, they did not contain the presupposition trigger *wieder*. After each (experimental and filler) sentence we asked multiple-choice comprehension questions with three possible answers to choose from. For one third of the experimental items, the question targeted the content of the presupposition directly. An example of such a question is (13-a) with the possible answers in (13-b).

(13) a. How many pink lamps did {Linda/Judith} receive?
b. Cannot be answered/one/at least two

In those cases where the question did not target the presuppositional content of the sentence, they addressed information introduced either by the context sentence or the test sentence, like the one in (14).

(14) a. Who was {Linda/Judith} out with?
b. Cannot be answered/a friend/her father

We included the option *cannot be answered* to explore the possibility that an unaccommodated presupposition leads to the uninterpretability of the sentence it occurs in. In order not to make it a viable answer only in the questions concerning the presupposition, *cannot be answered* was always one of the three possible answers

displayed. Therefore, we also included questions that really could not be answered on the basis of the information given in the context and the target sentence.

The experiment was programmed using the ERTS language. The experiment was divided in eight randomized lists. All participants came in two times. Each time they were given four of the eight randomized lists. The order of the eight lists was counterbalanced across participants. A global context was provided in the beginning of the experiment in order to introduce the relevant characters. Responses were collected via an external keyboard consisting of six separate keys.

A trial began with the presentation of a warning signal that consisted of five stars. Then the context sentence was presented. A self-paced reading paradigm was used for the critical target sentence that followed. Participants were presented the sentence word by word. Reading times were collected by having the reader press a response key after each word. Thus, reading times were measured on-line. The end of a sentence was signaled by a full stop. After that, participants had to rate the acceptability of the test sentence with respect to the context sentence on a four item rating scale (1 = very bad 2 = rather bad 3 = rather good 4 = very good). Participants were given an even number of points on the scale to choose between in order to guarantee that they had to express a tendency towards acceptability or unacceptability. They delivered their judgment by pressing a corresponding button. After the end of each trial, the comprehension question was asked.

Sixteen subjects participated in the experiment. Most of them were students of the University of Tübingen (13 women; mean age = 24,38; age range = 18–37). They were native speakers of German and had normal or corrected to normal vision.

3.2 Results

The analyses were carried out using the R programming language (R Development Core Team) as linear mixed models, using the program *lmer* (Bates 2005). The fixed factor was context (neutral/supporting). The random factors were subjects and items. Additionally, models with random slopes for both subjects and items were calculated. When an ANOVA revealed a significant difference between the models, we included the more complex model in our analysis. For each of the analyzed words, trials were deleted if they deviated by 3 SD or more from the mean reading time of the respective word.

3.2.1 Questions

All participants answered more than 75 % of the comprehension questions which did not target the presupposition correctly, suggesting that the participants paid ample attention. The Average accuracy was 91 %. The mean accuracy was not influenced by the context (supporting or neutral) the target sentence appeared in; in both contexts, mean accuracy of the answers was 91 %. Regarding the questions which targeted the presupposition, there were differences with respect to the two different contexts. In

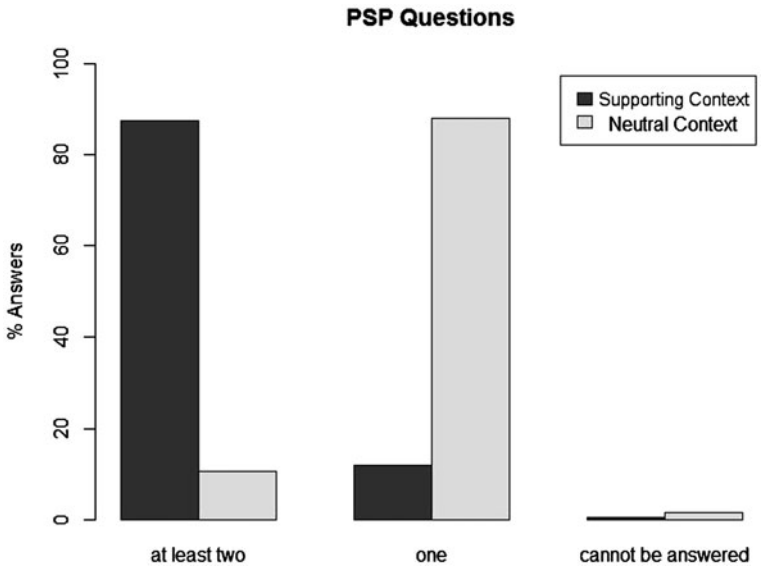


Fig. 1 Answers for questions targeting the presupposition content—mean answers

a supporting context, people answered the presupposition question with *at least two* 87.5 % of the time, with *one* 12 %, and with *cannot be answered* 0.5 % of the time. When the target sentence was presented in a neutral context, 10.5 % of the answers were *at least two*, 88 % were *one*, and 1.5 % of the time subjects chose *cannot be answered* (see Fig. 1).

3.2.2 Acceptability Judgments

Regarding the acceptability judgments elicited, sentences in a neutral context averaged around 1.9, whereas sentences in a supporting context received a mean rating of 3.6. This difference was significant ($|t| = 8.24$, $p < .001$).

3.2.3 Reading Times

The reading times for each word are listed in Table 1. On the critical word, i.e., the word at which the content of the presupposition of *wieder* was evident (*erhalten* in (12), repeated in (15)), reading times in the neutral context condition were significantly longer than in the supporting context condition ($|t| = 2.927$, $p < 0.05$). This effect was still marginally significant in the spillover region ($|t| = 2.262$, $p = 0.058$), and subsided after that on spillover +1 ($|t| = 0.609$, $p > 0.1$). On word 14 and 15 (*einer Freundin* in (15)), the sentence in the neutral context condition was read significantly faster than in the supporting context condition ($|t| = 2.51$, $p < 0.05$, and $|t| = 2.812$, $p < 0.05$ respectively) (Fig. 2).

(15) **T:** Vor zwei Tagen hat Linda wieder eine rosa Lampe erhalten, als sie
Two days ago has Linda again a pink lamp received, when she
mit einer Freundin unterwegs war.
with a friend out was

Table 1 Reading times per word in ms. Asterisks mark significant differences at the $p > 0.05$ threshold

condition	word 1	word 2	word 3	word 4	word 5	word 6 (wieder)	word 7
supporting	359.3	288.8	285.2	284.3	296.8	300	274.4
neutral	357.7	287.5	284.5	284.2	297.7	301	275.8
condition	word 8	word 9	word 10 (crit. word)		word 11	word 12	word 13
supporting	274	286.6	394*		295	269.6	262.4
neutral	277.7	286.8	423.9*		303.2	268.5	260.7
condition	word 14	word 15	word 16	word 17			
supporting	261.3*	271.1*	271.2	274			
neutral	256*	263*	268.8	272			

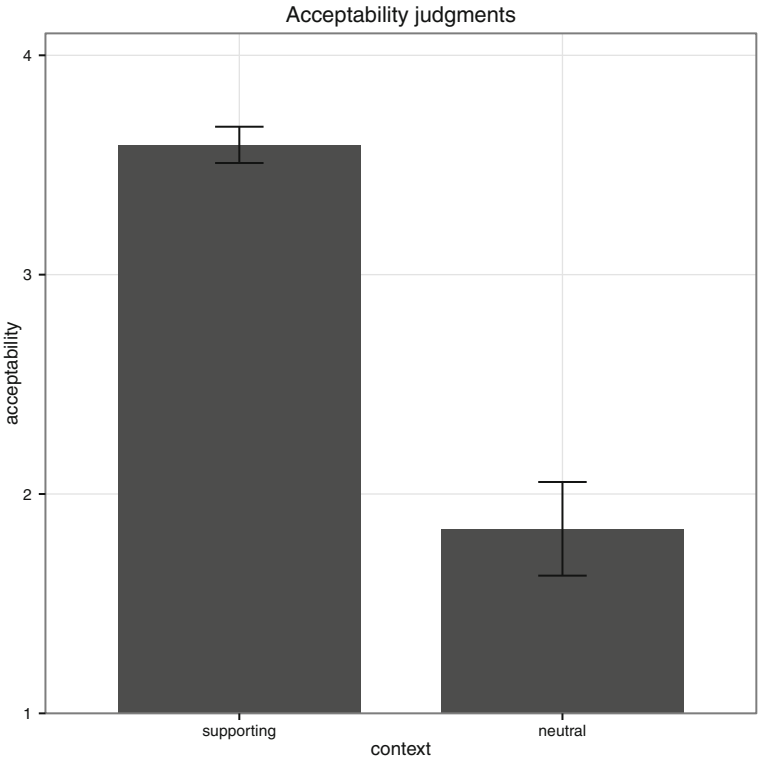


Fig. 2 Acceptability judgments for sentences in the two conditions—On a 1–4 scale where 1 = very bad and 4 = very good

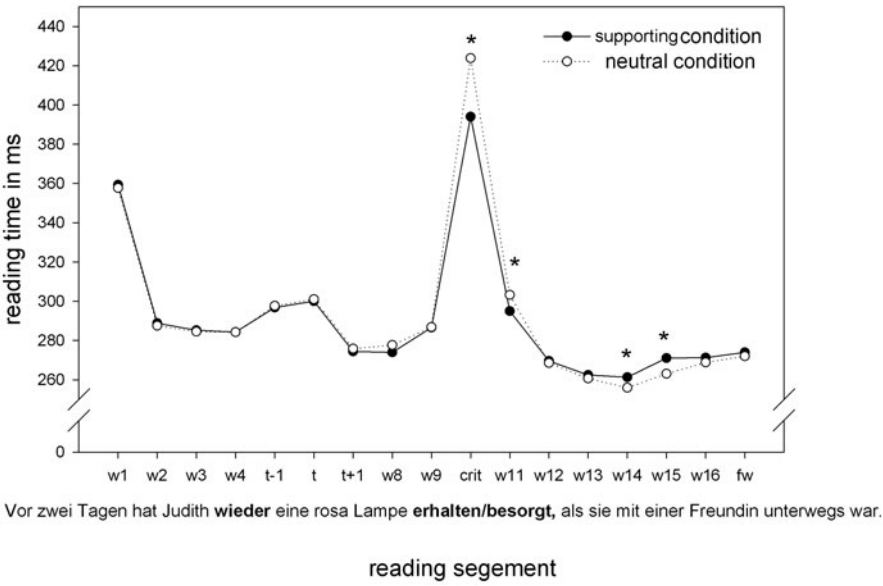


Fig. 3 Reading times for each word in the two conditions—Asterisks mark differences significant at the $p > 0.05$ threshold

Figure 3 depicts the average reading time for each word of the test sentence in either a neutral or a supporting context.

3.3 Discussion

The results of the experiment presented here show that people are sensitive to presuppositions which are not entailed by the context. This can be gathered from the lower acceptability ratings of a sentence containing a presupposition in a neutral context compared to the same sentence in a supporting context, as well as an increase in reading times in the neutral condition. The experiment was also designed to target the issue of accommodation by asking questions about the presupposed material. Reading times came apart on the word at which the content of the presupposition was fully known to the participant (the critical word), showing that presuppositions are computed immediately. Additionally, the results of the presupposition questions suggest that participants did not accommodate the presupposition of *wieder* in a neutral context. These results are interesting in several ways: First, they provide more evidence for the immediate incremental processing of presuppositions. The fact that presupposition induced effects show up as soon as the relevant content of the presupposition is known and not only at the end of the sentence for example, shows that even meaning components which are not part of the asserted meaning

are processed immediately. Secondly, these results offer new insights for linguistic theory with the apparent lack of accommodation which does not lead to complete incomprehensibility of a sentence. Most presupposition theories assume that presupposition failure will either lead to an uninterpretable sentence or to accommodation, which is not attested by the results in this experiment. The third surprising result is the late increase of reading times in the supporting condition. In the following, we will discuss the two latter findings and what they can tell us about the semantics and processing of *wieder*. From this, we will hypothesize about other presupposition triggers and discuss a novel idea about how presuppositions are accommodated.

3.3.1 Unpredicted Answer Pattern

Agent-Less Presuppositions

In order to account for the observed answer pattern, we have to take a step back and reevaluate what the presupposition of the test sentence really is. The underlying idea so far was that a sentence like (16-a) presupposes (16-b).

(16) a. Yesterday, Linda gave away a pink lamp again.
b. Linda had given away a pink lamp before yesterday.

Under this assumption, the presupposition is not met in a context that entails that Judith gave a pink lamp to Linda prior to yesterday (remember that this is how our *neutral* contexts were set up). However, the example in (17) shows that the presupposition of *again* does not necessarily have to incorporate the agent of the relevant event.[2] Even though the two rice cooking events have different agents, *again* seems to work just fine in the second sentence.

(17) a. On Monday, John cooked rice.
b. On Tuesday, Mary cooked rice again.

This could, indeed, also be a possible reading for the test sentences used in the present experiment. The resulting presupposition for (16-a) would then be the one in (18).

(18) Someone has given away a pink lamp before yesterday.

If this is how participants interpreted the presuppositions of the test sentences, we can account for the observed answer pattern. Such an interpretation of a sentence like (16-a) would guarantee that the presupposition of a sentence like (16-a) is satisfied in a context in which there was an earlier lamp giving event with an agent different from Linda. Crucially, there would be no need for the accommodation of an earlier lamp giving by Linda. Thus, the expected answer to a question like "How many lamps did Linda give away?" is *one*. And that is exactly the answers participants gave most often for sentences in the *neutral* condition. However, there is a caveat to

[2] We would like to thank an anonymous reviewer for pointing this out.

an explanation along these lines. This concerns the observed differences between the two conditions. If participants interpreted the presupposition of *again* without the agent, the presupposition is supported in the *supporting condition* as well as in the *neutral condition*. However, there are significant differences both in reading times and acceptability judgments between the two conditions. In the neutral condition, reading times on the critical word were longer and acceptability judgments were lower than in the supporting condition. If participants understood the presupposition to be verified in both contexts, the fixed factor should not have any influence. Since there are significant differences in both reading times and acceptability judgments depending on the context, we will assume that the context (neutral/supporting) has an influence on the interpretation of the presupposition. We will thus discuss another idea which can account for the observed pattern.

Missing Accommodation

One thing that is special to an experimental setting versus everyday conversations is that there is no conversation partner that people can interact with. This means that any clarification about the presupposed material is not possible. Therefore, when subjects are confronted with a presupposition in a neutral context, they face a dilemma: they can either assume something that has never been explicitly mentioned (i.e. they accommodate) or they choose a different strategy in order to make sense of the sentence. In this experiment, it seems that subjects went for the second option. In this particular case, the strategy was to disregard the presupposition of the sentence altogether and to accept the asserted part only. This seems to be possible in the case of *again* since it does not contribute anything to the asserted part of a sentence. Its only function is to introduce a presupposition, as we can gather from its lexical entry in (19) (cf. Beck (2007)). The presupposition after the colon expresses that there has to be a time t' prior to t" (which will be the reference time) at which the relevant proposition P was true[3]. The assertion after the dot is simply an identity function.

(19) $[[\text{again}]] = \lambda w.\lambda t'.\lambda P_{<s,<i,t>>}.\ \lambda t'': t' < t''\ \ \&\ \ P(w)(t').P(w)(t'')$

This analysis can be extended to other triggers like *too*, for example. When someone hears a sentence like (20) s/he is most likely inclined to understand the message that John likes French movies, even though s/he might not know which other relevant person besides John likes French movies. We assume the simplified lexical entry of *too* in (21) along the lines of Beck (2007). Notice that this entry is also such that it does not have any impact on the assertion.

(20) John likes French movies, too.

(21) $[[\text{too}]] = \lambda w.\lambda P_{<s,<e,t>>}.\lambda x: \exists y[y \neq x\ \&\ P(w)(y) = 1].P(w)(x)$

[3] The tree in (34) shows that the first temporal argument of *again* is a free pronoun. This is what we mean when we say that t' remains free.

We suggest that people who are faced with a sentence which contains a presupposition which is not in the common ground, but who are not in the position to challenge the speaker of such a sentence, go for a strategy which disregards the content of the presupposition rather than to accommodate something out of the blue. Note that there are other scenarios in which someone is not in the position to challenge a speaker than just an experimental setting. Other possible situations might be one where politeness prohibits explicitly challenging the speaker or an ongoing written correspondence via letter or email where it would presumably be too tedious to challenge an unmet presupposition.

However, this strategy is not equally available for every expression that is considered to belong to the class of presupposition triggers. In a sentence with a definite expression, for example, the presupposition of the definite cannot simply be ignored because the sentence will not make any sense. That is, for a sentence like (22) the hearer will either have to accommodate that there exists a unique artist who lives next to the speaker or else s/he cannot interpret the sentence.

(22) The artist who lives next to me holds regular yoga sessions.

We assume with Heim and Kratzer (1998) that the semantic type of *the* is <<s<s<e,t>>>e> and has the lexical entry in (23). Combining *the* with a predicate of type <s<e,t>> will give us an individual of type <e>.

(23) [[the]] = $\lambda w.\lambda f_{<s,<e,t>>}$: there is exactly one x s.t. f(w)(x) = 1. the unique y s.t. f(w)(y)=1

We see that the semantic contribution of *the* is more than just adding a presupposition. It takes a predicate as its argument and returns the unique individual of which this predicate holds. This is crucial for semantic composition. The predicate in (22) *holds regular yoga sessions* is of type <s,<e,t>> and thus wants something of type <e> as its second argument. Hence, if *the* was simply ignored, not only would this mean that the semantic contribution of *the* to the assertion would fall by the wayside, but it would also result in a semantic type mismatch. A similar point holds for the triggers *to know* and *to stop*. Both of them make a meaningful contribution to the assertion on top of introducing a presupposition:

(24) [[know]] = $\lambda w.\lambda P_{<s,t>}.\lambda x$: P(w) = 1.x believes P in w

(25) [[stop]] = $\lambda w.\ \lambda t'.\lambda P_{<s,<i,t>>}.\lambda x.\lambda t$: t'< t & P(w)(t')(x).¬ P(w)(t)(x)

The difference between presupposition triggers like factives, change of state verbs, and definite descriptions on the one hand and particles like *again*, *too*, and *even* on the other hand is that the truth of a sentence which contains an item of the latter group can be determined without the presupposition trigger whereas this is not possible for sentences with expressions that belong to the former group. This has already been mentioned by Stalnaker (1974) and is discussed at length in Zeevat (2002) and Zeevat (2004).

When faced with a sentence whose presupposition is not given in the context, but which is not needed in order to determine the truth of the assertion, people seem to

choose to ignore the presupposition trigger altogether rather than to assume something ad hoc that has never been mentioned. From this observation, we generalize the following maxim of interpretation:

> ***Minimize Accommodation***
> Do not accommodate a presupposition unless missing accommodation will lead to uninterpretability of the assertion!

A similar proposal has already been featured in Moulton (2007). He found that in sentences like (26), people preferably resolved the ellipsis to (27-b) rather than to (27-a), having to accommodate as little as possible.

(26) Jordy carefully reviewed the book that Kiley did Δ.

(27) a. Δ = carefully reviewed the book
b. Δ = reviewed the book

Based on this finding, Moulton (2007) proposes the principle *Accommodate Conservatively*:

> ***Accommodate Conservatively***
> Do not accommodate more than necessary to satisfy a presupposition.

As we can see, *Minimize Accommodation* is even more radical than *Accommodate Conservatively* insofar that it only calls for accommodation if there is no other way for the sentence to receive a truth value. We suggest that this is a principle that every interpreter adheres to when faced with a situation in which s/he cannot ask for further information regarding the presupposition.

We argue that *Minimize Accommodation* divides the class of presupposition triggers in (at least) two parts.[4] The first class comprises particles like *again*, *too*, and *even* (**class 1**). These are triggers which will be ignored rather than to accommodate their presupposition in the face of presupposition failure. Definite descriptions, factives, and change of state verbs are part of the second class of presupposition triggers (**class 2**). The presuppositions of these triggers will be accommodated because the interpretation of the assertion hinges on the semantic contribution that these trigger makes. This kind of distinction between different triggers is very similar to the one put forward in Simons (2001). Yet, it crucially differs in the way that Simons concludes that only the presuppositions of class 1 triggers are conventionally encoded in the semantics of the trigger, whereas the presuppositions of class 2 triggers are conversationally determined. For the analysis presented here, it is crucial that the presupposition is semantically rooted in the lexical entry of class 1 and class 2 triggers, because the assertive contribution that the respective trigger does or does not make, determines whether the presupposition has to be accommodated or not.

In light of the analysis presented in this section, it looks as if accommodation really is a last resort operation and even dispreferred to the reanalysis of a sentence

[4] After finishing this paper it came to our attention that a similar division has already been proposed in Glanzberg (2005). He presents his arguments in an *update semantic* framework, but the idea is very similar at heart.

where the presupposition trigger is simply ignored. A sentence with an unfulfilled presupposition cannot only end up as undefined or lead to accommodation, it can also result in a reanalysis of the sentence where people analyze the sentence as if the presupposition trigger was not there. The decision of whether to accommodate or to reanalyze the sentence is guided by *Minimize Accommodation*. If the sentence can be interpreted without the presupposition trigger, this option is always preferred to accommodation. Only when all else fails, will a cooperative interpreter apply accommodation in a last attempt to save the sentence from uninterpretability. It is important to note at this point, that even though the kind of reinterpretation envisaged here saves the sentence from being uninterpretable, it is still a dispreferred move. The low acceptability ratings obtained in the neutral condition are thus not surprising and simply reflect that people perceived a problem with the target sentence in the given context.

Interestingly, there is recent experimental work by Domaneschi et al. (2013) which suggests that the division between different triggers presented here is borne out. The authors report an experiment with auditory stimuli in which they investigated different triggers (definite article, factives, iteratives, change of state verbs, focus-sensitive particles) and how they are accommodated. Their results show that the presuppositions of the definite article and factives are more often accommodated than the ones of iteratives and focus-sensitive particles. The change of state verbs constitute a middle of the road case in that they pattern with the definite and factive verbs under 'normal' conditions. However, as soon as the cognitive load increases, accommodation in the case of change of state verbs decreases. The same is true for iteratives but not for the other triggers investigated. Domaneschi et al. (2013) argue that the presuppositions of change of state verbs and iteratives are harder to process because they presuppose "temporally displaced events". We will come back to this issue in the next section where we discuss the late increase of reading times in the supporting condition.

Based on the discussion above, we suggest that the interpretation of a sentence with a presupposition trigger is processed as depicted in Fig. 4. When the presupposition of a sentence is entailed by the context, context update with the asserted proposition can be performed without a problem. If the context does not provide the relevant presupposition, the sentence will be reanalyzed and the trigger will be ignored. This is only possible for class 1 triggers. In the case of class 2 triggers, the NP has to be accommodated to make a context update possible. If accommodation is not possible, e.g. because the presupposition is too unlikely to be true in the actual world, the sentence will be rejected as uninterpretable.

Two cautionary remarks are in order here. First, the schema above assumes that the hearer of the sentence is not in a position where s/he can openly challenge the speaker. If a presupposition is uttered in an ongoing conversation between two discourse participants of equal status, there is always the option of challenging the presupposition e.g. by a *Hey wait a minute* (von Fintel 2003) response. Right now it is not clear at which point in the interpretation process this will be the case. It could be that the presupposition is already challenged as soon as the first step in Fig. 4 does not lead to a context update. An example for such a scenario is given in (28).

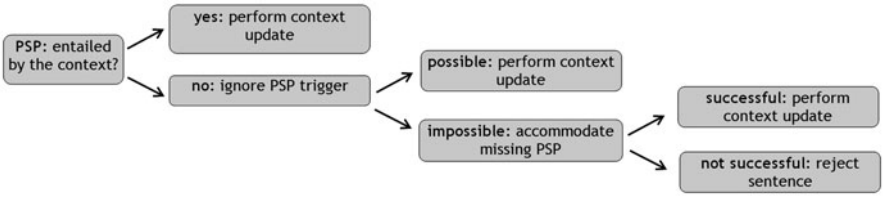


Fig. 4 Interpretation schema for a sentence containing a presupposition trigger

(28) A: Yesterday, John won the lottery again.
B: Hey wait a minute! I had no idea that John had won the lottery before.

On the other hand, the fact that many people accommodate the presupposition in (29-a) with ease seems to indicate that an unmet presupposition is only challenged if the last step depicted in Fig. 4 cannot be performed successfully. This happens e.g. in cases like (29-b) where world knowledge clashes vastly with what is presupposed.

(29) a. A: I am sorry I am late, I had to take my cat to the vet.
b. A: I am sorry I am late, I had to take my elephant to the vet.

We think that the question at which point a hearer challenges a presupposition boils down to how likely the presupposition is. There is also room for inter-hearer variation. If a hearer is extremely cooperative, s/he will probably only challenge the speaker if all else fails (that is after step three). A less cooperative speaker might already disrupt the conversation after step one fails. The second remark we would like to make here is that everything after step one should only be seen as a repair mechanism. As soon as a presupposition is not entailed by the context, the sentence is semantically undefined. This, in turn, maps onto pragmatic inappropriateness by what von Fintel (2003) calls *Stalnakers Bridge*. How strongly the inappropriateness is perceived, again, depends on the likelihood of the presupposition.

Another point we would like to discuss here is what happens when only a subpart of the presupposition of a certain trigger has to be accommodated. Many authors have pointed out that the presupposition of *wieder* essentially consists of two parts. First, that the relevant proposition was true at a time other than the reference time and second that the relevant time interval precedes the reference time (cf. Heim (1990), Kamp and Rossdeutscher (1994), van der Sandt and Huitink (2003)). An example from Heim (1990) is given in (30).

(30) a. We will have pizza on John's birthday, so we shouldn't have pizza again on Mary's birthday.
b. We will have pizza on John's birthday, so we shouldn't have pizza on Mary's birthday.

The point is that in (30-a) but not in (30-b) we derive the inference that Mary's birthday succeeds John's birthday. This is due to *again*. More precisely, it is due to the part of the presupposition of *again* which introduces the temporal relation and which is obviously accommodated without any effort. How does that fit with what

we said above about ignoring *again* rather than accommodating its presupposition? The crucial difference between the sentence in (30-a) and the sentences used in the discussed experiment is that in the former but not the latter the context already provided a part of the presupposition. We will therefore suggest that as soon as parts of the presupposition of a certain element are provided in the context, i.e. the context search for parts of the relevant information is successful, the lexical item cannot be ignored anymore. The lexical entry for *again* provides us with a presupposition along the lines of (31-b) for the experimental item in (31-a).

(31) a. Two days ago, Linda received a pink lamp again.
b. PSP: t' < *two-days-ago* & receive(*pink lamp*) (*Linda*) (t').receive(*pink lamp*) (*Linda*) (*two-days-ago*)

However, assuming that the trigger can no longer be ignored when parts of the presupposition are given in the context, one could argue that as soon as there is some kind of time interval in the context which is suitable to provide a value for the free time variable t', the whole presupposition of (31-a) has to be accommodated. This means as soon as there is another time interval which is prior to *two days ago*, a part of the presupposition is fulfilled in the context and thus the trigger can no longer be ignored. Remember that our context sentences were such that they provided a time interval of this sort, as the context in (32) illustrates.

(32) Last week, Linda bought Judith a pink lamp for a room.

In principle, since *last week* is prior to *two days ago*, one could assign the free variable the value of *last week* and consequently the first part of the presupposition would be fulfilled. But this is obviously not what happens. The interpretation of the presupposition of *wieder* really hinges on the relevant proposition being true at some other time. Only if the context furnishes this, will the temporal connection be made. What this is basically saying is that the interpretation of the presupposition of *wieder* proceeds in two parts. In a first step, it will be checked if the context provides a suitable proposition to verify the presupposition. If this is not the case, the trigger will be ignored as long as it does not make any assertoric contribution. In those cases where the context provides the relevant proposition, the temporal connection will be made and the free time variable can receive its value. When the context does not provide an explicit time which precedes the reference time, this temporal connection will be made by the means of accommodation. We will come back to the proposal that the interpretation of *again* proceeds in two steps in the next part which is devoted to the effect present in the later parts of the test sentences.

3.3.2 Late Increase in Reading Times in the Supporting Condition

The third relevant result obtained from this experiment is almost as mysterious as the missing accommodation. Downstream in the sentence, after the conflict of given versus not given presupposition has been realized by the processor, reading times increase in the supporting condition, i.e., the condition which explicitly gives the

relevant background information. This is mysterious insofar that this is the condition which should not impose any problems during processing. So what is behind this late increase in reading times? The first thing we might want to look at is the part of the sentence on which the effect became apparent.

(33) **C1:** Letzte Woche hat Linda Judith eine rosa Lampe für ein Zimmer
last week has Linda Judith a pink lamp for a room
gekauft.
bought

C2: Letzte Woche hat Judith Linda eine rosa Lampe für ein Zimmer
last week has Judith Linda a pink lamp for a room
gekauft.
bought

T: Vor zwei Tagen hat {(a) Linda /(b) Judith} wieder eine rosa Lampe
ago two days has {(a) Linda /(b) Judith} again a pink lamp
erhalten, als sie mit **einer Freundin** unterwegs war.
received, when she with **a friend** out was

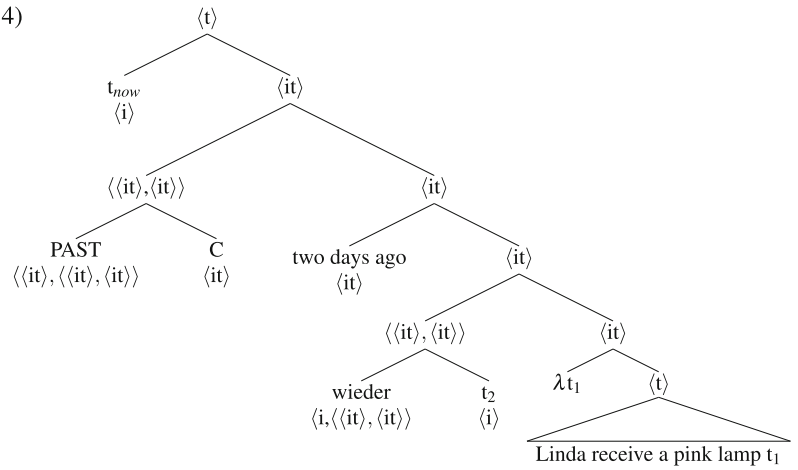
In the above example from the experimental material, the reading times in the neutral and supporting condition came apart once more on word 14 and word 15, i.e. *a friend*. Since the sentences were constructed as parallel as possible, word 14 and word 15 were always indefinite NPs. So a first guess might be that the difference in reading times has something to do with the indefinite NP. It is not clear, however, why the indefinite NP should impose more processing difficulty in the supporting than in the neutral context, since both contexts vary only with respect to who gave what to whom (or more generally: Who did what to whom). Thus, the observed effect cannot be due to the nature of the sentence material at this point.

Another very likely hypothesis is that subjects did not interpret the rest of the sentence as deeply after they realized that its presupposition was not met. The observed difference between the supporting condition and the neutral condition would thus not be an increase of reading times in the supporting condition, but rather a decrease in the neutral condition. By looking at the raw reading times alone, it is almost impossible do determine whether the reading times have risen in the supporting condition or fallen in the neutral condition. However, there are other indicators which can help us to clarify whether subjects stopped interpreting the sentence in the neutral condition after they realized the lacking presupposition. Recall the results for the comprehension questions. The comprehension questions after each sentence were such that they very often targeted the material introduced in the later part of the sentence. If it was the case that people stopped interpreting the sentence in depth after the critical word in the neutral condition, they should give more false answers compared to the supporting condition. We have seen, however, that context condition did not influence the percentage of correct answers at all. In both contexts, questions were answered accurately 91 % of the time. Assuming that decreased attention results in lower accuracy, the results concerning the comprehension questions provide a strong piece of evidence against the hypothesis that interpretation diminished in the

neutral condition. Another argument counter the explanation that subjects stopped interpreting the sentence in the neutral condition in depth comes from the fact that the difference in reading times appears to be local. If people really gave up on the sentence when the presupposition is not given in the context, it is not clear why the observed reading time difference does not persist until the end of the sentence. We will therefore conclude that the late difference in reading times is not explained by the processor "giving up" in the neutral condition.

So how can we account for this difference, then? We will argue that the answer is tightly connected to *Minimize Accommodation*. Recall that *Minimize Accommodation* basically says that you should ignore the meaning contribution of the trigger as long as it does not change the assertion of the sentence. As discussed in the last section, this is possible in the case of *wieder*. In Fig. 4, we proposed a schema according to which people interpret presuppositions. In terms of the time course of interpretation, we are assuming that people go through the individual steps outlined in the schema very quickly and as soon as they are able to calculate the presupposition of a given sentence. This is reflected in the relatively long reading times on the critical word in comparison to the other words in the sentence. Consequently, if it is already determined at the critical word that the contribution of *wieder* should be neglected, the only difference at the point the relevant effect arises is that the meaning of *wieder* still plays a role in the supporting but not in the neutral condition. After compositional interpretation of the structure in (34), the truth conditions of the sentence in (34) are the same in the supporting condition (36-a) and in the neutral condition (36-b).[5] However, while the sentence as a whole does not carry a presupposition in the neutral condition, it still carries one in the supporting condition. We will assume the quantificational entry for PAST in (35) where C is a contextual restriction which picks out the relevant temporal subset.

(34)



[5] For ease of presentation, we are leaving out the world variables in this derivation. They are still important to capture the meaning of presuppositions, of course.

(35) [[PAST]] = $\lambda C_{<it>}.\lambda P_{<it>}.\lambda t.\exists t'[C(t') \ \& \ t' < t \ \& \ P(t)]$

(36) a. $\exists\, t'[C(t') \,\&\, t' < t_{now} \,\&\, \text{receive}(t')(\textit{pink lamp})(\textit{Linda}) \,\&\, t' \subseteq \textit{two-days-ago}]$
PSP: $\exists\, t'[C(t') \,\&\, t_2 < t' \,\&\, \text{receive}(t_2)(\textit{pink lamp})(\textit{Linda})]$
b. $\exists\, t'[C(t') \,\&\, t' < t_{now} \,\&\, \text{receive}(t')(\textit{pink lamp})(\textit{Linda}) \,\&\, t' \subseteq \textit{two-days-ago}]$

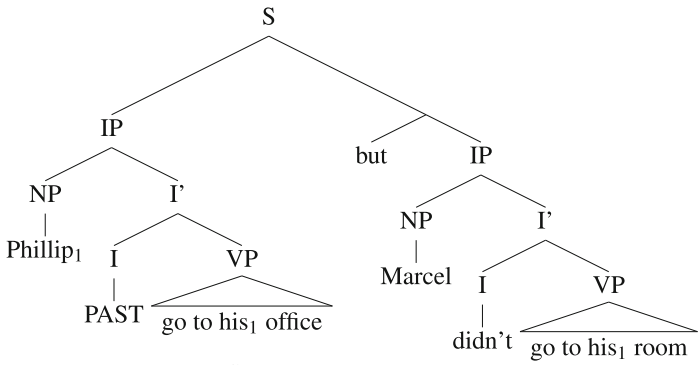
If we take a closer look at (36-a) and (36-b) we see that the only relevant difference between the two truth conditions is that (36-a) has a free variable of type <i> at the level of presupposition which (36-b) does not posses. As we have briefly mentioned in the last section, we assume that the interpretation of the presupposition of *wieder* is a two step process. The first step is initiated as soon as the presupposition can be calculated. In this step, the processor searches the context for a proposition which can potentially fulfill the presupposition (i.e. that Linda received a pink lamp). If the context does not provide such a proposition, the processor will choose to ignore the contribution of *wieder* altogether. In those contexts which provide the necessary background information, the second interpretational step is to assign a value to the free temporal variable in the presupposition of *wieder*. We suppose that this step is what is behind the late increase in reading times in the supporting condition. Assuming that the core part of the presupposition is already dealt with in an earlier step in both conditions, the free time variable is the only thing that distinguishes the sentence in the supporting condition from the sentence in the neutral condition at this point.

To our knowledge, no one has ever investigated free variables at the level of presupposition experimentally. Thus, it is very hard to find evidence backing our assumptions. There is, however, work on the processing of free versus bound individual variables at the level of LF which shows that assigning a value to a free variable is more effortful than interpreting a bound variable. In an eye-tracking experiment, Koornneef (2008) tested whether subjects exhibit a preference for sloppy over strict readings in elliptical sentences. A sentence like (37), adapted from Heim and Kratzer (1998), can have a sloppy or a strict reading.

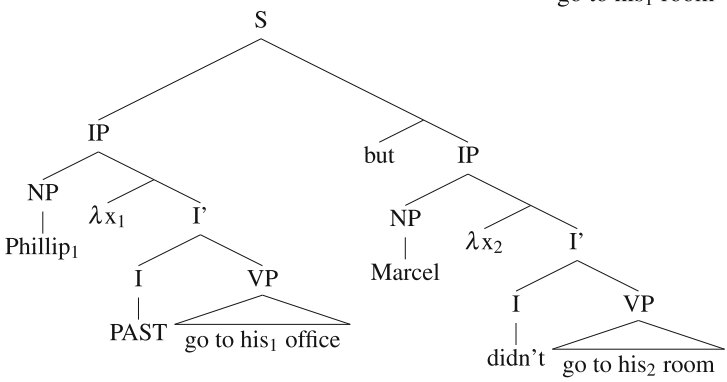
(37) Philipp went to his office but Marcel didn't.

The sloppy reading is the reading where the ellipsis in the second part of the sentence is understood as *Marcel didn't go to Marcel's office*. On the strict reading, the ellipsis is resolved to match the overt material in the sentence, i.e. to mean *Marcel didn't go to Phillip's office*. According to Heim and Kratzer (1998), the strict reading comes about by treating *him* in both the antecedent and the ellipsis as free variable which receives its value via the variable assignment function, see (38). Under the sloppy reading, both occurrences of the pronoun *his* are bound by a lambda abstractor at the level of LF (39).

(38)



(39)



In a reading experiment using eye-tracking, Koornneef (2008) presented sentences which exhibit such an ambiguity (as in (41)) in two contexts.[6] One which biased the interpretation to a strict reading (40-a) and another one which made the sloppy reading prominent (40-b).

(40) a. Lisa and Anouk love the music channel MTV. Lisa was very happy when she was selected for the show 'Pimp My Room', in which her room was redecorated.
b. Lisa and Anouk love the music channel MTV. They were very happy when they were selected for the show 'Pimp My Room', in which their rooms were redecorated.

(41) Sadly,/Lisa/thinks that/her/pimped room/has a/touch of class,/but Anouk does not.

The results reveal that there are two types of readers: the "energetic" ones and the "lazy" ones. According to Koornneef (2008), energetic readers make the decision between a free and a bound variable interpretation already upon encountering the pronoun *her*. This means that in a context which favors a strict reading, energetic readers interpret *her* as a free variable while a context with a sloppy bias leads them to bind the pronoun. Crucially, this decision results in a difference in reading

[6] Slashes indicate the regions of interest (ROIs).

times on the region after the pronoun. Here, the energetic readers exhibited longer reading times in a strict-biased context than in a sloppy-biased context, indicating that processing a free variable is more laborious than the interpretation of a bound variable. For the energetic readers, no significant difference emerges on the ellipsis region. Presumably because they have already committed to a strict versus sloppy reading before. For the lazy readers, there were no significant differences at the position of the pronoun. They did, however, exhibit a difference in reading times in the ellipsis region where the strict-biased condition took longer to read than the sloppy-biased condition. What is important for our discussion is that, neglecting different types of readers, the bound reading was always easier to process than the reading in which the variable was free. Moreover, for the energetic readers, the effect of bound versus free variable did not emerge on the pronoun itself but on the two words after the pronoun. Considering that the effect in the *again* experiment also emerges further downstream than where the free variable occurs, this finding suggest that we are dealing with a similar phenomenon. An additional factor which could be responsible for the late emergence of difficulty is the two step presupposition interpretation model which we have argued for in this paper. (42) makes this model explicit.

(42) **Two step interpretation model for *again*:**
Step one: Check if relevant proposition is entailed by the context. If not, ignore trigger
Step two: If the relevant proposition is given, assign value to the free time variable via the variable assignment function

The idea that the interpretation of *wieder* proceeds in these two steps gains additional support from other experimental work on reference processes. In the psycholinguistic literature, it has been extensively argued that reference processes happen in a two step fashion (e.g. Garrod and Sanford (1994), Garrod and Terras (2000), Sanford et al. (1983), Sturt (2003)). These authors argue that processing referential expressions like pronouns are processed first in a *bonding* phase, followed by a *resolution* phase. Garrod and Terras (2000) describe this two step process as follows:"(1) a low-level automatic process associated with establishing some kind of link between the potential role-filler and a previous verb, which we call *bonding*, and (2) a later process which tests and resolves the link with respect to the overall discourse representation, which we call *resolution*." This is very much in line with the interpretation model for *wieder* which we have proposed in this paper.

4 Conclusion

This paper set out with the question of how the presupposition of *wieder* is processed online and whether its presupposition is accommodated in a context which does not entail the relevant proposition. To this end, a self-paced reading experiment was conducted which presented a target sentence containing *wieder* in a context which

introduced the relevant information and a context which failed to do so. Additionally, questions which targeted the presupposed information were asked. The results show that in a context which does not support the presupposition of *wieder*, processing effects emerge as soon as the presupposition is known to the reader. This finding suggests that presuppositions are processed and evaluated immediately. Assuming that pragmatic reasoning applies after semantic evaluation, this result might be taken to indicate that the presupposition of *wieder* is semantically encoded in the trigger and does not arise by pragmatic reasoning as some theories suggest.

Furthermore, the experiment brought to light two additional findings which are not easily accounted for by existing theories. The first one being the apparent lack of accommodation when the presupposition of *wieder* is not given in a context. According to most theories, such a scenario should either lead to accommodation or interpretation failure. However, answers concerning the relevant presuppositions show that participants did not go either route. Instead, they chose to ignore the presupposition introduced by *wieder* (and consequently *wieder* itself) and interpret the assertion only. We proposed that this is possible because *wieder* does not contribute anything semantically to the literal content of a sentence. Based on this observation, we introduced the principle *minimize accommodation* and suggested that accommodation is only a last resort option in those cases where the presupposition trigger makes a meaningful contribution to the compositional interpretation of a sentence. This suggests that the presupposition of the definite article and change of state verbs has to be accommodated whereas this is not the case for the presuppositions of *too* and *even*. And indeed, the experiment presented in Domaneschi et al. (2013) supports this hypothesis. Moreover, their observation that change of state verbs and iteratives are harder to process because they presuppose "temporally displaced events" can be linked directly to the third important result obtained from the experiment presented in this paper.

The third interesting result of the present experiment is the late increase in reading times in the supporting condition, i.e. when the presupposition of *wieder* is given in the context. We have argued that this effect comes about by assigning the free time variable which is an argument of *wieder* a contextual value. This happens after an initial process which checks whether the context contains a proposition which could potentially fulfill the presupposition of *wieder*. If this process turns out not to be successful (i.e. in the neutral condition), the contribution of *wieder* does not play any further role for the interpretation of the sentence. If, however, the context provides the relevant information, *wieder* is interpreted as usual. In this case, the temporal variable argument of *wieder* becomes relevant. Recent experimental work by Koornneef (2008) has shown that the interpretation of a free variable is more laborious than interpreting a bound variable. Assuming that assigning a value to a free variable is effortful in general can explain the observed discrepancy between the supporting and the neutral condition in the experiment presented here.

The view presented here has interesting consequences for the theory of presuppositions. We have laid out above how *minimize accommodation* assumes two classes of presupposition triggers: Those whose presupposition has to be accommodated and those whose presupposition does not have to and eventually is not accommodated. A

similar case can be made about triggers which take a free variable as their argument and triggers which do not. This paper offers new perspectives on how presupposition triggers can be classified by the means of empirical data. Future research will have to validate the hypotheses presented here, of course. We hope to inspire further research in this direction, because individual (processing) differences between presupposition triggers have to be kept in mind when investigating more intricate data such as, for example, projection.

References

Altmann, G., and M. Steedman. 1988. Interaction with context during sentence processing. *Cognition* 30:191–238.

Bates, D. M. 2005. Fitting linear mixed models in r. *R News* 5:27–30.

Beaver, D., and H. Zeevat. 2008. Towards a general theory of presupposition and accommodation. The University of Texas at Austin and University of Amsterdam.

Beck, S. 2007. Quantifier dependent readings of anaphoric presuppositions. In *Presupposition and implicature in compositional semantics,* ed. U. Sauerland and P. Stateva. New York: Palgrave Macmillan.

Burkhardt, P. 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language* 98:159–168.

Dickey, M. W. 2000. The Processing of tense: Psycholinguistic studies on the interpretation of tense and temporal relations. Ph.D. thesis, University of Massachusetts, Amherst.

Domaneschi, F., E. Carrea, C. Penco, and A. Greco. 2013. The cognitive load of presupposition triggers: Mandatory and optional repairs in presupposition failure. *Language and Cognitive Processes* 29(1):136–146.

Frazier, L. 1978. Theories of sentence processing. In *Modularity in knowledge representation and natural language processing,* ed. J. Garfield, 493–522. Cambridge: MIT Press.

Frazier, L., and C. Clifton. 2000. On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research* 29:125–139.

Garrod, S., and A. J. Sanford. 1994. Resolving sentences in a discourse context: How discourse representation affects language understanding. In *Handbook of psycholinguistics,* ed. M. A. Gernsbacher, 675–698. San Diego: Academic Press.

Garrod, S., and M. Terras. 2000. The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language* 42:526–544.

Glanzberg, M. 2005. Presuppositions, truth values and expressing propositions. In *Contextualism in philosophy: Knowledge, meaning, and truth, chapter Glanzberg, Michael,* ed. G. Preyer and G. Peter. Oxford: Oxford University Press.

Heim, I. 1990. Presupposition projection. In *Reader for the Nijmegen workshop on presupposition, lexical meaning, and discourse processes,* ed. R. van der Sandt. University of Nijmegen.

Heim, I., and A. Kratzer. 1998. *Semantics in generative grammar, Volume 13 of Blackwell textbooks in linguistics*. Malden: Blackwell.

Inhoff, A. W. 1985. The effect of factivity on lexical retrieval and postlexical process during eye fixations in reading. *Journal of Psycholinguistic Research* 14 (14): 45–57.

Kamp, H., and A. Rossdeutscher. 1994. Drs-construction and lexically driven inference. *Theoretical Linguistics* 20:165–235.

Koornneef, A. W. 2008. Eye-catching anaphora. Ph.D. thesis, University of Utrecht.

Kripke, S. 2009. Presupposition and anaphora: Remarks on the formulation of the projection problem. *Linguistic Inquiry* 40:367–386. (Manuscript, Princeton University).

Moulton, K. 2007. Small antecedents: Syntax or pragmatics? In *Proceedings of NELS 37,* ed. E. Elfner, and M. Walkow. Amherst: GLSA Publications.

Sanford, A., S. Garrod, A. Lucas, and R. Henderson. 1983. Pronouns without explicit antecedents? *Journal of Semantics* 2:303–318.

Schwarz, F. 2007. Processing presupposed content. *Journal of Semantics* 24:373–416.

Simons, M. 2001. On the conversational basis of some presuppositions. In *Proceedings of semantics and linguistics theory 11,* ed. R. Hasting, B. Jackson, and Z. Zvolensky. Ithaca: CLC Publications.

Soames, S. 1982. How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry* 13:483–545.

Stalnaker, R. 1973. Presuppositions. *Journal of Philosophical Logic* 2: 447–457.

Stalnaker, R. 1974. Pragmatic presuppositions. In *Semantics and Philosophy,* ed. M. Munitz and P. Unger, 197–214. New York: New York University Press.

Sturt, P. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48:542–562.

van Berkum, J. J. A., C. M. Brown, and P. Hagoort. 1999. Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language* 41 (2): 147–182.

van Berkum, J. J. A., C. M. Brown, P. Hagoort, and P. Zwitserlood. 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology* 40:235–248.

van der Sandt, R., and J. Huitink. 2003. Again. In *Proceedings of the 14th Amsterdam colloquium,* ed. P. Dekker and R. van Rooij, 181–186. Amsterdam: ILLC.

von Fintel, K. 2003. *Pragmatics: Notes on presupposition*. Cambridge, MA: MIT.

Zeevat, H. 2002. Explaining presupposition triggers. In *Information sharing,* ed. K. van Deemter and R. Kibble, 61–87. Stanford: CSLI Publications.

Zeevat, H. 2004. Particles: Presupposition triggers, context markers or speech act markers. In *Optimality theory and pragmatics,* ed. R. Blutner and H. Zeevat, 91–111. Basingstoke: Palgrave Macmillan.

Resolving Temporary Referential Ambiguity Using Presupposed Content

Jacopo Romoli, Manizeh Khan, Yasutada Sudo and Jesse Snedeker

Abstract We present the results of two visual-world experiments investigating whether the presupposition of 'also' is used to predict upcoming linguistic material during sentence comprehension. We compare predictions generated by 'also' to predictions from a parallel inference generated by 'only' (i.e., that the upcoming material will be unique). The results show that adults do use the presupposition of 'also' incrementally in online sentence comprehension and they can do so within 200 to 500 ms of the onset of the presuppositional trigger. Furthermore, they use it regardless of whether contextual support is explicit or implicit. On the other hand, we did not observe effects of the inference generated by 'only' at any point during the sentence, even though this information was used in an offline task.

Keywords Presuppositions · Entailments · Processing · Visual-world · Semantics · Pragmatics · Psycholinguistics

J. Romoli (✉)
Room 17E10, School of Communication,
Jordanstown Campus University of Ulster, Shore Road,
Newtownabbey, Co. Antrim BT37 0QB, UK,
e-mail: j.romoli@ulster.ac.uk

Department of Cognitive Science, Macquarie University, Sydney, Australia
e-mail: jacopo.romoli@mq.edu.au

M. Khan
Department of Psychology, Harvard University, Cambridge, USA
e-mail: khan@wjh.harvard.edu

Y. Sudo
Linguistics, UCL, London, United Kingdom
e-mail: y.sudo@ucl.ac.uk

Linguistics and Philosophy department, MIT, Cambridge, USA
e-mail: ysudo@mit.edu

J. Snedeker
Department of Psychology, Harvard University, Cambridge, USA
e-mail: snedeker@wjh.harvard.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_3,

1 Introduction

Presuppositions are a class of inferences that we draw from utterances and are generally characterized along two dimensions: their discourse role and their behavior in embeddings.[1] To illustrate, consider a sentence like (1), from which we typically draw the two conclusions in (2a) and (2b).

(1) Fred stopped smoking.

(2) a. Fred doesn't smoke.
b. Fred used to smoke.

In the literature, inferences like (2b) are referred to as 'presuppositions,' while those like (2a) as 'entailments.' These two inferences are considered to be different for two reasons. First, they intuitively play different conversational roles: (2a) is new information added to the context, whereas (2b) is typically assumed to convey information that is given at the point of utterance of (1). We can refine this intuition by explicitly adding each of these inferences before the statement itself. When we add the presupposition to the statement, the result (shown in (3b)) is a natural discourse. In contrast, when the other inference is placed before the statement, the result (shown in (3a)) is unnatural. The standard way to account for the oddness of (3a) is to appeal to a condition that requires a speaker to not assert anything that is redundant in its context of utterance (see Stalnaker 1978 and much subsequent work). Crucially, this condition does not apply to presuppositions, resulting in the contrast between (3b) and (3a).[2]

(3) a. Fred doesn't smoke and he now stopped.
b. Fred used to smoke and he now stopped.

The second distinctive property of presuppositions is their behavior in complex sentences. Consider (4a)–(4d) below in which the statement (1) is embedded in different complex sentences.

(4) a. Fred didn't stop smoking.
b. Did Fred stop smoking?
c. If Fred stops smoking, Lisa will be happy.
d. It's possible that Fred stopped smoking.

All of these sentences still generate the inference in (2b) but they do not generate the conclusion in (2a). In other words, the presuppositions of sentences like (1), unlike entailments, appear to be 'inherited' by most of the complex sentences containing them. This pattern, traditionally called 'projection behaviour', is characteristic of presuppositions and it is generally used as the primary diagnostic for distinguishing

[1] For an introduction to presuppositions see Chierchia and McConnell-Ginet 2000 and Beaver and Geurts (2011).

[2] In fact, one approach, stemming again from the work of Stalnaker (1974, 1978), is to think that presuppositions not only *can* but *have to* be redundant in the context of utterance.

presuppositions from other inferences. Accounting for how presuppositions project is a central topic in semantics and pragmatics (Karttunen 1974; Stalnaker 1974; Gazdar 1979; Heim 1983; van der Sandt 1992 and Beaver 2001. For some recent proposals see Schlenker 2008 and Schlenker 2009).

While presuppositions have been studied extensively, we still known little about how they are used during comprehension, as the sentence unfolds. This gap is important to fill because presuppositions carry information that could be used incrementally to guide interpretation during language processing. By tracing how this information becomes available over time, we could learn more about the interplay between pragmatics and compositional semantics during language comprehension.

1.1 The Processing of Presupposed Content

Recently, people have started looking at the question of how presuppositions are processed during comprehension (Kim 2007; Schwarz 2007; Schwarz and Tiemann 2013a, 2013b). Three broad questions have guided this work. First, is the presupposed content of an utterance available as quickly as its assertive content? Or do new presuppositions systematically lag behind assertions, in a manner that parallels scalar implicatures (Bott and Noveck 2004; Huang and Snedeker 2009 among many others)? Second, is there a processing cost to presupposition violation? Observing such a cost would also inform the first question (by placing an upper bound on the time by which presupposition was calculated). Third, once presuppositions are available, how do they affect sentence processing? Are they used to resolve ambiguities at other levels or make predictions about upcoming referents? Again, data on this question would also constrain answers to the first.

Kim (2007) explored how presupposed content and asserted content are accessed during a sentence verification task. Participants were shown various visual displays and asked to judge whether sentences like (5) were true or false. Their task was to press a button corresponding to 'yes' if the sentence accurately described the visual context and 'no' if it did not.

(5) Only the girls have books.

Kim (2007) adopted the analysis of 'only' in which a sentence like (5) asserts (6a) and presupposes (6b) (see Beaver and Clark 2009 and references therein).

(6) a. No people other than the girls have books.
b. The girls have books.

There were two types of critical trials which varied in terms of the picture that was paired with the utterance. On the false assertion trials, (5) was matched with a picture in which two girls out of eight characters had books (i.e., the presupposition was satisfied) and some of the other six characters also had books (i.e., the assertion was false). On the presupposition violation trials, (5) was matched with a picture in which the two girls didn't have books and the other characters did not have them

either (i.e., presupposition false, assertion true). She found that participants were faster to reject the false assertion trials than they were to reject the presupposition violations. On the basis of these results, Kim (2007) concludes that the information conveyed by presuppositions and the information conveyed by assertions are used differently in sentence comprehension. Comprehenders do not have to verify presuppositions before evaluating the truth of a sentence. Instead, they assume that the presuppositions are true and only check them afterwards.

Schwarz (2007) used a reading time paradigm to explore the comprehension of sentences with presupposition triggers in both German and English. In one of the experiments, participants had to read sentences like (7) and (8), which differ in the content of the relative clauses. Crucially the relative clause (7) satisfies the presupposition of 'also' (i.e., that there is some other relevant person to whom the congressman wrote to), but (8) does not.

(7) The congressman, *who wrote to John*, had also written to the mayor to schedule a meeting for the fundraiser.

(8) The congressman, *who John wrote to*, had also written to the mayor to schedule a meeting for the fundraiser.

The stimuli were presented in a phrase-by-phrase self-paced reading paradigm in which multiple words were presented at once. The experimental region included the presuppositional trigger and several subsequent words (e.g., *had also written to the mayor*). The participants were slower to read the experimental region when the relative clause did not satisfy the presupposition associated with 'also' (8) than when it did (7).

This finding demonstrates that there is a processing cost in cases where there is also presupposition failure. We believe that there are at least two possible explanations for this processing cost. First, this cost could be a direct consequence of presupposition failure: participants could slow down in (8) because they fail to find any prior event of the relevant kind (one with the congressman writing someone) and thus cannot integrate the presupposition triggered by 'also'. In this case, the cost should only come after encountering the main verb in the active voice (*had also written*), since that information is needed to determine that the relative clause does not satisfy the presupposition (e.g., see (9)).[3] Alternatively, the cost could reflect the participants' predictions about the form or meaning of the verb, given the relative clause and presuppositional trigger. Specifically, in (7) the participant should expect the verb that is used (*written*), while in (8) they might expect a verb with a different argument structure (again see (9)).

[3] Schwarz seems to interpret his online data as a reflection of presupposition failure, rather than prediction (see p. 402). However, his work also provides evidence that presuppositions can be used to predict and/or revise the thematic roles in an event. Specifically, in an offline study in German, he finds that readers tend to interpret a relative clause with ambiguous case marking in a manner that would allow to satisfy the presupposition of 'auch' (*also*).

(9) The congressman who John wrote to had also received a letter from the mayor to schedule a meeting for the fundraiser.

In this paper, we report the results of two experiments that investigate the time course of presupposition calculation by looking for effects of a presupposition on the interpretation of an upcoming noun. Like Schwarz (2007), we focus on the presupposition of 'also'. Unlike Schwarz (2007), we are not concerned with the effect of presupposition failure on processing. Instead we investigate how early participants are able to use the information associated with presuppositions to determine the referents of upcoming nouns. In addition, we compare this effect to a parallel inference based on the entailments associated with 'only'. Our focus differs from that of Kim's (2007) study, as we are not looking at how presupposed versus asserted content is used in verification but whether presupposed content is used to predict an upcoming word.[4]

1.2 Experimental Paradigm

In the experiments reported in this paper, we use the visual world paradigm, because it provides a sensitive and time-locked measures of language comprehension. Prior studies have successfully used this method to study a broad range of linguistic processes including: word recognition (e.g., Allopenna et al. 1998), syntactic ambiguity resolution (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy 1995) and the calcuation of scalar implicatures (e.g., Huang and Snedeker 2009). The stimuli and task used in Experiment 1 are closely modeled on a series of experiments conduced by Kim et al. (2008).

Kim and colleagues presented participants with pairs of sentences like (10) and (11). As mentioned above, a sentence like (11) is generally assumed to presuppose (12a) and assert something along the lines of (12b).

(10) Mark has some candies and some shoes.

(11) Jane only has some candies.

(12) a. Jane has candies.
b. Jane has no relevant things other than candies.

Kim and colleagues were interested in how listeners restrict the assertion of a sentence like (11). Can context be used to determine the set of relevant things as the

[4] There are two other recent studies on presupposition processing that are not directly relevant to the question posed here (Schwarz and Tiemann 2013a, 2013b and Chemla and Bott 2013). Schwarz and Tiemann look at whether the position of the trigger in a complex sentence influences the processing of presuppositions. On the basis of their results they argue that there is a processing cost associated to the 'length of projection', that is the distance between the position of the trigger and the clause associated to the information satisfying the presupposition appears. Chemla and Bott, on the other hand, concentrates on whether the accommodation of presupposition is more costly when it occurs above or below negation (see also Romoli and Schwarz (2014)).

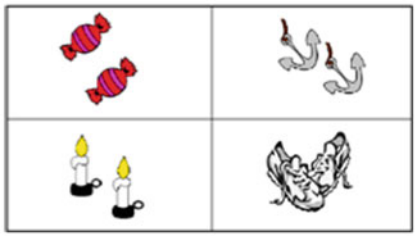


Fig. 1 An example of a display from Kim et al 2009. The relevant sentences are listed in (13). Illustration courtesy of Christina Kim

sentence unfolds and what kinds of contextual cues are relevant? They assessed this by measuring the participants' ability to predict the upcoming noun on the basis of the context.

Participants are presented with spoken instructions, asking them to choose among different objects within a visual reference world and their eye movements are measured. Each trial contained a context sentence (10) and one of four target sentences (13a-13d) below.

At the onset of the target sentence, a grid of four images appeared (Fig. 1) which consisted of the target object (for (11) 'candies'), a cohort competitor (for (11) 'candles') and two distractors (for (11) 'anchors' and 'shoes'). The competitor is crucial in this design, because it insures that the identity of the noun is unclear even after the word begins providing a longer window in which predictions generated by the context and the structure of the sentence can guide interpretation.

The target sentences were constructed by manipulating two variables: whether 'only' was present and whether the target noun was mentioned in the context sentence ('old' versus 'new'). This resulted in four conditions, as summarised in (13).

(13) a. Jane only has some candies. (only/old)
b. Jane has some candies. (control/old)
c. Jane only has some candles. (only/new)
d. Jane has some candles. (control/new)

If context is used to restrict the relevant set, then participants who heard the sentences with 'only' should assume that the upcoming noun will be one of the previously mentioned items (e.g., candies). Thus they should be faster to reach the target in sentences like (13a) than in sentences like (13b). This systematic preference for the previously mentioned object should be absent or reduced in sentences without 'only' resulting in an interaction between the two variables. This is precisely what Kim and colleagues found, leading them to conclude that context is used to rapidly restrict the comparison class. We return to these findings in the general discussion section (to explore their relevance to our 'only' control condition).

Our study employs a similar design to explore a different question. Like Kim and colleagues, we provided participants with: a context sentence that introduced two objects, a target sentence referring to one of them, and a display that contained two cohort competitors. However, our goal was to use this paradigm to explore how the presupposition of 'also' can be used to predict an upcoming referent. Our

critical items consisted of series of three sentences (14). The first (14a) established a set of two characters, making it natural to draw comparisons between them. The second sentence (14b) described what one character had. The critical third sentence (14c) contained the presupposition trigger. This sentence was intended to have the interpretation where 'also' is associated with the subject. To achieve this the utterances were produced with stress on the subject and a prosodic break after 'also' (see Appendix B).

(14) a. Mark and Jane are friends (*introduction*)
b. Mark has some candies and some shoes (*context*)
c. JANE, also, has some candies. (*target*)

We are assuming that a sentence like (14c) presupposes something along the lines of (15a) and asserts (15b) (see Kripke 2009 and Heim 1992 among others).

(15) a. A relevant individual in the context other than Jane has some candies.
b. Jane has some candies.

In both of the following experiments, we manipulated whether the intended target referent was mentioned or not in the context sentence ('old' versus 'new') whether there was a focus particle in the sentence which could allow participants to predict the discourse status of the referent before encountering it ('disambiguated' versus 'control'). For old referents the disambiguating focus particle was 'also', for new referents it was 'only.' This resulted in the following four conditions in (16).

(16) a. Jane, also, has some candies. [also/old]
b. Jane has got some candies. [control/old]
c. Only Jane has some candles. [only/new]
d. Jane has got some candles. [control/new]

The logic of the design is as follows. In the control conditions, participants have no basis on which to predict the final object, and thus they should split their gaze between the two cohort members from the onset of the final noun until phonological disambiguation, after which they should quickly converge on the target object. In contrast, in the also/old condition, the presuppositional trigger provides information that could allow participants to infer the target nouns before phonological disambiguation. Specifically, if participants are able to rapidly determine the presupposition of *also* they could potentially infer that the target is one of the previously mentioned items (candies or shoes) anytime after the trigger, allowing them to determine which noun it is (candies) after hearing the first phoneme. Thus we expect that folks will look at the correct target more often and more quickly in the also/old condition than in the control/old condition.

The only/new condition was included to ensure a balanced design. In this case the focus particle favors the novel referent: if participants incrementally recover the assertion in the 'only' sentences they should recognize that the upcoming object must be something that only Jane has, and thus cannot be one of the previously mentioned objects. This should lead them to resolve the referential ambiguity in favor of the new target before phonological disambiguation (Fig. 2).

Fig. 2 An example of a display from Experiment 1. The relevant sentences are listed in (13). Illustration–author's own

2 Experiment 1

2.1 Method

2.1.1 Participants

Thirty two participants were recruited from the Harvard community, including undergraduate students. They received either course credit or $ 5 for their participation. All participants were native speakers of English. One participant was excluded due to low accuracy with the fillers.

2.1.2 Material

Each item consisted of: a set of four pictures, an introduction sentence (17), a context sentence (18), and a critical sentence (13). The four pictures included two pictures which could readily be described by nouns sharing an onset (e.g., *candies* and *candles*). The introduction sentence presented two characters. The context sentence indicated that one character had two items: one of the cohort members (candies) and one of the non-cohort members (watches). As we noted above, there were four different variants of the target sentence.

(17) *Introduction sentence*
Mark and Jane are friends

(18) *Context sentence*
Mark has some candies and some watches.

(19) a. Jane, also, has some candies. [also/old]
b. Jane has got some candies. [control/old]

c. Only Jane has some candles. [only/new]
d. Jane has got some candles. [control/new]

Both the presence of the focus particle and the discourse status of the final noun were manipulated within subjects and within items. It is important to note that the 'only' sentences predict a target noun that was not in the context sentence, in contrast to the 'also' sentences that presuppose that the target noun was in the context sentence. The predictions for these sentences were confirmed in a norming study conducted on Amazon Mechanical Turk (Appendix A). As mentioned, we also conducted a norming task to ensure that the intonation used in the 'also' sentences evoked the expected interpretation (Appendix B). In addition, we wanted to control which of the two cohort members appeared in the context sentence. To counterbalance all three of these variables, we created eight versions of each of the 32 base items. We then created eight lists such that each base item appeared only once on a given list and in all eight cells across the lists. Thus each participant heard each of the four target sentence types eight times over the course of the experiment. Eight filler trials were included. Two fillers appeared at the beginning to make sure participants were familiar with the task before the first experimental trial. The remaining six filler trials were interspersed throughout the experiment. Trial order was pseudo-randomised, with each participant seeing base items in the same order. The positions of the different kinds of objects (target, competitor and distractors) in the visual display were counterbalanced across trials.

2.1.3 Procedure

Participants were seated at a comfortable distance from the screen of a Tobii T-60 eye-tracker. The auditory stimuli were played by the computer through external speakers. First, participants heard the introduction sentence and context sentence. Then a fixation point appeared in the middle of the screen. After the participant had fixated on this point for 500 ms, the target sentence began. This was done to ensure that participants were always looking in the same place at the beginning of the critical sentence. At the onset of the target sentence, the images appeared and participants' fixations were measured. The participants were told that their task was to pick the last object mentioned in the target sentence by clicking the relevant picture. Once the participant did this the trial ended and the next trial began.

2.1.4 Results

We analyzed the log-odds of the proportion of fixations to the old cohort item (the competitor that had been mentioned in the context sentence) versus the new cohort item (Fig. 3). Prior to the disambiguation point of the noun, the control sentences are compatible with either image. If participants incrementally incorporate the presupposed content of 'also' during sentence processing, then we should expect to

see more looks to the old cohort item in sentences with 'also' compared to control sentences.

Also Our primary analyses compared the also-old trials to the control-old trials. We conducted linear mixed-effects regressions, with the maximally appropriate random effects structure, looking at the log-odds of proportion of looks to the old cohort. We estimated p-values using the pnorm function in R. We examined four time windows: the first of these time windows coincided with the noun (offset by 200 ms), while the other three spanned the 900 ms before the noun (-700 to -400, -400 to -100, and -100 to 200). On average, the onset of the word 'also' occurred 806 ms before the onset of the target noun, so these time windows would cover saccades programmed at the very onset of 'also' (early pre-noun), as well as the period in which the presuppositional trigger could potentially generate predictions about upcoming referents (the mid and late pre-noun windows).

During the noun window, participants were significantly more likely to look at the old cohort in also-old sentences compared to control-old sentences ($t = 4.20$, $p < 0.001$). This pattern was already evident in the late pre-noun window which covered the 300 ms immediately preceding the noun ($t = 3.13$, $p < 0.002$). There were no significant differences between the 'also' and the control-old sentences in early and mid pre-noun time windows.

Only We were also interested in whether participants used the information from 'only' to anticipate unique referent. That is, did participants look at the old cohort less in the 'only' sentences compared to the control-new sentences? We conducted linear mixed-effects regressions, parallel to those above. Again we used the maximally appropriate random effects structure and looked at the log-odds of proportion of looks to the old cohort during the four time windows described earlier. There were no significant differences between looks to the old cohort in the 'only' sentences and the control-new sentences ($ts < 1$, $ps > 0.3$). This null effect is surprising given that participants correctly predicted the target noun on the basis of 'only' in an offline task (see Appendix A).

2.1.5 Discussion

Experiment 1 provides a clear answer to our original questions about the processing of presuppositions during language comprehension. First, we found that presupposed content can be used to guide lexical (or referential) predictions. In the 'also' condition, participants showed a preference to look at the previously-mentioned cohort member. This preference began well before noun onset, resulting in a robust difference between the 'also' sentences and their controls in the late pre-noun time window (-100 ms to 200 ms unshifted). Second, the timing of effect places an upper-bound on the point at which the presupposition is calculated. The effect of the presupposition became robust in a time window that began about 700 ms after the onset of 'also'. If we make the standard assumption that it takes a minimum of 200 ms for information in the speech stream to affect saccades (Matin et al. 1993; Allopenna et al. 1998),

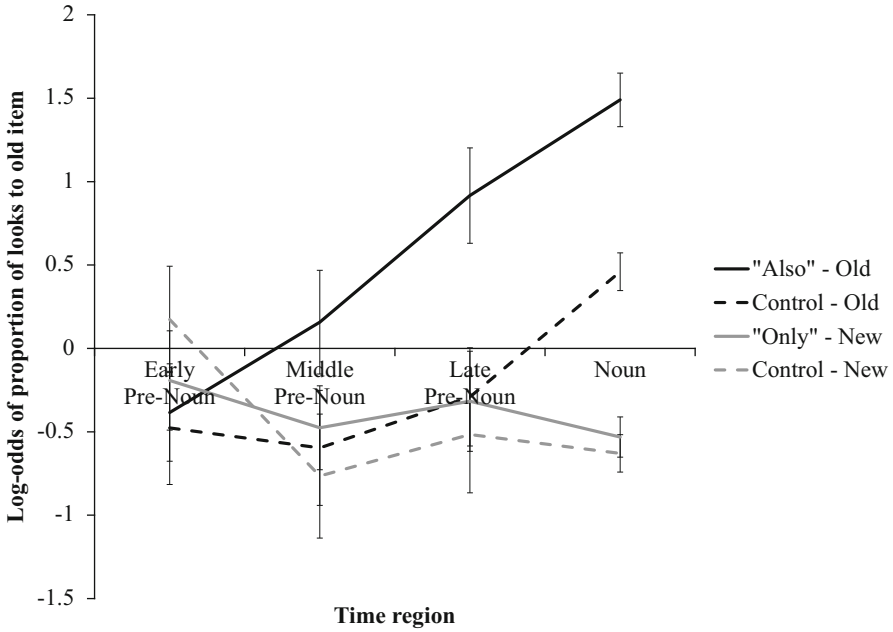


Fig. 3 Experiment 1: gaze data

then this suggests that it takes 500 ms or less, from the onset of the trigger, for the presupposition to be calculated.

However, this experiment also provided an unexpected surprise. To ensure balance in our stimuli, we had included a condition in which 'only' had subject scope (20). As we noted earlier, it is standard to assume that these utterances have the presupposition given in (21a) and express the assertion given in (21b).

(20) Only Jane has some candles.

(21) a. Jane has some candles.
b. No other relevant people have candles.

Thus we expected that the 'only' sentences would allow the participants to rule out the previously mentioned items as potential referents for the final noun, because they would make the assertion false (since the other character did have those items). But curiously, we failed to find smallest hint of this effect in the eye-movement data, even though an offline norming study confirmed that participants were sensitive to this constraint (Appendix A).

Experiment 2 had three goals. First, we wanted to replicate our critical finding that presuppositions can be used to predict upcoming words or referents. Second, we wanted to extend these results by exploring whether more indirect contextual support would be sufficient to satisfy the presupposition and guide online processing. Specifically, in Experiment 1, the presupposed content had been explicitly mentioned ('Michael has some candles'). In Experiment 2, we compare explicit mention to implicit mention ('Look at what Michael has').

Finally, we wanted to create a discourse context that would be more appropriate for the 'only' sentences, to determine whether this would allow people to use this assertion to predict the upcoming noun in real time. We reasoned that sentences like (20) are felicitous only in contexts in which there is some relevant set of people that the subject ('Jane') is being contrasted with. In Experiment 1, the discourse included just two people. Since we had already stated what the other person had, the assertion in the 'only' sentence was actually no more informative than the presupposition. Consequently, the control-new sentence was arguably a more felicitous way to express this idea. In Experiment 2, there are a total of three people in the context and the task centers around distinguishing the person in the target sentence from another person.

3 Experiment 2

3.1 Method

3.1.1 Participants

Thirty two participants were recruited from the Harvard community, including undergraduate students. Half of the participants were assigned to the explict context condition and half to the implicit context condition. They received either course credits or $ 5 for their participation. None of them had participated in Experiment 1. All participants were native speakers of English. Eight participants were excluded (2 for low response accuracy, 1 for software error, and 5 for excessive track loss).

3.1.2 Procedure and Material

Three changes were made to the materials from Experiment 1. First, two between participant conditions were created. In the explicit context condition (22), the context sentence (22a) overtly mentions the critical objects, just like the context sentence in Experiment 1.

(21) *Explicit Context*
   a. *Context Sentence*: Michael has got candies and watches.
   b. *'Also' Target*: Sarah also has some candies.
   c. *'Only' Target*: Only Sarah has some candles.

In the implicit context condition (23), the context sentence (23a) directs attention to these objects but does not mention them by name.

(22) *Implicit Context*
   a. *Context Sentence*: Look at what Michael has!
   b. *'Also' Target*: Sarah also has some candies.
   c. *'Only' Target*: Only Sarah has some candles.

Fig. 4 An example of the visual context used in Experiment 2. The relevant sentences are in (22) and (23). llustration—author's own

Second, we created new visual displays which included three new characters (Fig. 4). One character matched the gender of the person in the context sentence (the man at the top). This character had the two objects mentioned in the context sentence and was always placed on the top half of the screen in the center. The two other characters matched the gender of the person in the target sentence. One of these characters was the expected referent for the 'also' sentences because s/he had the old cohort object as well (the woman with candies on the right). The other character was the expected referent for the 'only' sentences because she had a unique object (the woman on the left with the candles). These characters appeared on the bottom half of the screen, with their relative position counterbalanced across trials.

Third, because the characters were now visible on the slide, we removed the introductory sentence which had linked the two protagonists.

The procedure was also modified. Participants were told that their task was to click on the person mentioned in each sentence. At the beginning of each trial, the visual display appeared, followed by the context sentence. After the participant clicked on the character mentioned in the context sentence, the target sentence was played. The trial ended when the participant clicked on the character mentioned in the target sentence. As in Experiment 1, each participant heard four different kinds of target sentences (also-old, control-old, only-new, and control-new) and there were eight trials per participant, in each condition.

3.1.3 Results and Discussion

As with Experiment 1, we analyzed the log-odds of the proportion of fixations to the old cohort item. We defined looks to the old cohort item as looks to the quadrant that contained the new person who had the old cohort object (lower right quadrant in Fig. 4). Similarly, we defined looks to the new cohort item as looks to the quadrant with the person who has the new cohort object (lower left quadrant in Fig. 4). We analyzed the also-old sentences in comparison to the control-old condition, and the only-new sentences in comparison with the control-new sentences. We were interested both in main effects of the markers and potential interactions between the marker and the context condition.

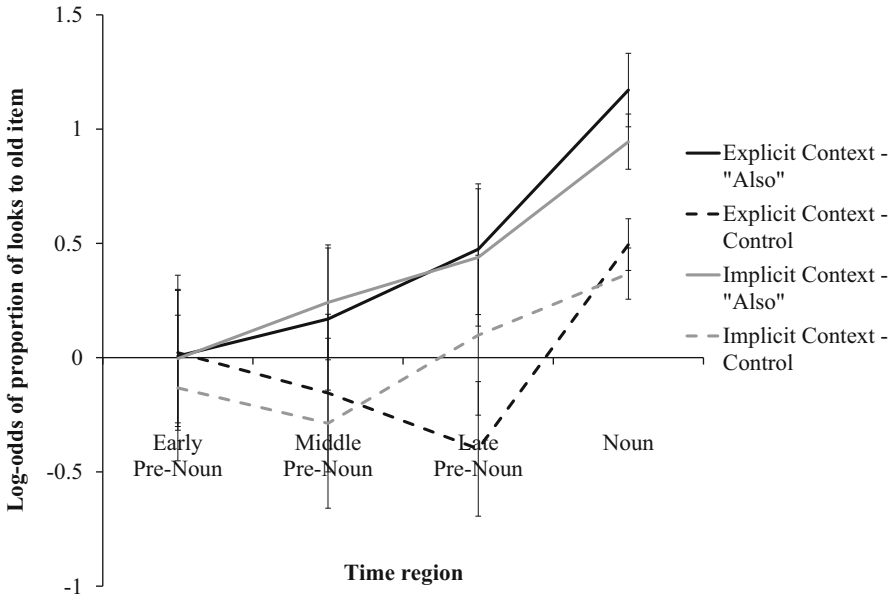


Fig. 5 Experiment 2: comparison of the also and control-old sentences

Also We conducted linear mixed-effects regressions with sentence type (also vs. control) and context condition (explicit vs. implicit) as predictors, with the maximally appropriate random effects structure. We estimated p-values using the pnorm function in R. We examined the same four time windows used in Experiment 1: the noun window and the three 300 ms timebins before the noun (early, mid, and late pre-noun). As in experiment one, on average, the onset of the word 'also' occurred around 800 ms before the onset of the target noun.

During the noun window, there was a main effect of sentence type ($t = 5.26$, $p < 0.001$) (Fig. 5). When the sentence contained 'also', participants shifted their gaze to the old cohort item. Critically, this effect was also significant in both the mid and late pre-noun windows (both t's > 3, p's < 0.002), but not in the early pre-noun time window. This pattern indicates that the presupposed information was available and able to guide reference resolution within 200 to 500 ms after encountering the trigger (allowing 200 ms to program a saccade). There was no main effects of or interactions with the context condition, suggesting that the presupposed content was integrated incrementally regardless of whether the anaphoric antecedent of 'also' was explicitly mentioned. To address this question more directly, we conducted separate analyses of each context condition. We found a reliable difference between the also and control-old sentences in both the explicit and implicit conditions ($t = 3.60$, $p < 0.001$ for the explicit context condition; $t = 3.68$, $p < 0.001$ for the implicit context condition).

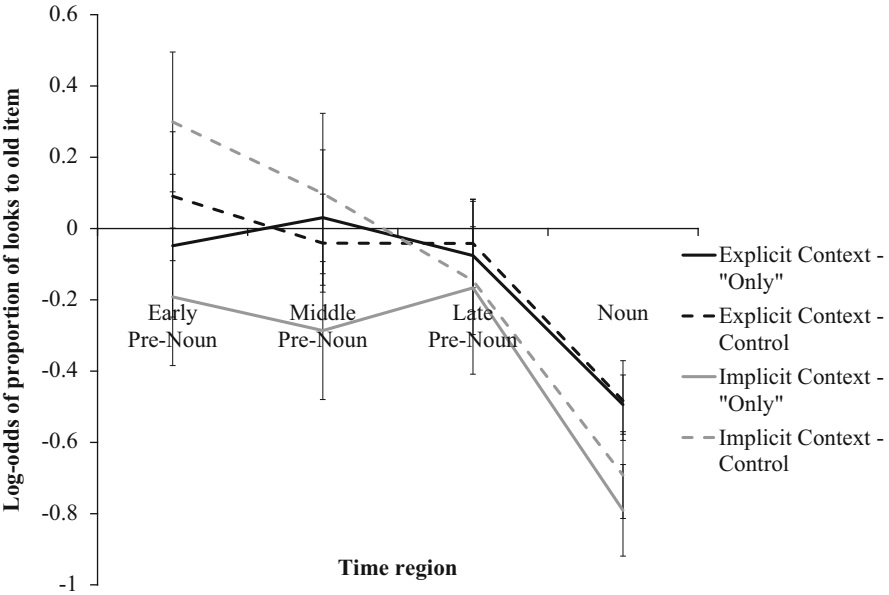


Fig. 6 Experiment 2: Comparison of only and control-new sentences

Only We conducted a parallel analysis of the 'only' and control-new sentences using the same variables and approach. If participants were using the information from 'only' to anticipate a unique referent, then we would expect them to shift away from the old cohort item well before the noun onset, resulting in a main effect of sentence type. This is not what we observed (Fig. 6). We found no significant differences between the sentence types in any of the time windows (all t's < 0.9, p's $> .3$). In addition, there was no effect of context condition and no interactions between context and sentence type (all p's $> .3$).

In sum, as in Experiments 1, hearing 'also' led to anticipatory eye movements to the target image but hearing 'only' did not, despite our attempt to create a more felicitous context for the use of 'only'. Further, participants used the information from the presupposition of 'also' regardless of whether the discourse context explicitly mentioned the repeated item or only implicitly referred to it.

4 General Discussion

The primary goal of these experiments was to explore the online processing of presupposition, and we were largely successful in doing this. Along the way, however, we encountered a strange failure in the interpretation of 'only'. We discuss these two findings in turn.

4.1 Presuppositions and Incremental Interpretation

We had three specific questions about the processing of presuppositions, which we were able to answer in these experiments.

First, can presupposed information be used to make predictions about upcoming lexical items during language comprehension? In other words, do presuppositions feed back into language comprehension creating expectations about how a sentence will end? In both Experiment 1 and Experiment 2, we found that the presupposition lead participants to identify the correct referent well before they had any information about the critical word.

Second, how soon after the trigger is the presupposition calculated? As we noted earlier, Schwarz (2007) found that readers slow down when they encounter a clause with a presuppositional violation, demonstrating that presuppositions are calculated as we read. Schwarz's data place a loose upper bound on the timing of this process: on average the presupposition must have been available at some point within 1500 ms after the clause was presented. Our findings refine that estimate. In Experiment 2, the difference between the also-old and control-old conditions was robust in the middle pre-noun time window, a period roughly 400 to 700 ms after the onset of the presuppositional trigger. If we make the standard assumption that it takes about 200 ms to launch a saccade in response to a phonological cue (Matin et al. 1993, Allopenna et al. 1998), then we can conclude that the presupposition was available within 200 to 500 ms of word onset. Put another way, it appears that the presupposition generated by 'also' can sometimes be calculated before the word itself is finished.

At first glance, our findings might seem hard to reconcile with Kim's (2007) picture verification study. Recall that she found that a sentence with 'only' was faster to evaluate when the assertion was false and the presupposition was true, than when the presupposition was false but the assertion was true. She concluded that the presupposition of a sentence is checked after the asserted content has been checked. That could be taken to suggest that presuppositions are not calculated until a late stage in verification. If that was the intended interpretation, then it would be incompatible with our findings (as well as with Schwarz's 2007). However, we see no reason to interpret the finding in that way. It is entirely possible that folks calculate presuppositions rapidly and incrementally but do not check these presuppositions immediately when verifying a statement against a single stable context. In fact, that might be a very smart thing to do: if presuppositions can generally be assumed to be true, then we ought to put highest priority on verifying the more contentious assertions.

The final question that we explored was: What forms of context can be accessed to satisfy a presupposition during online processing? The predictive inference that we saw in these experiments was based on the assumption that the presupposition of 'also' would be satisfied if the direct object of the target utterance referred to something that someone else in the discourse context also had in their possession. The results of Experiment 1 demonstrate that participants can quickly access material from the sentence immediately before the target sentence to find a suitable

prior referent. Experiment 2 refines this in several ways. First and most obviously, it shows that participants will infer that an indirect reference can satisfy a presupposition even when no direct reference is available. This is interesting in part because the presuppositional trigger that we used ('also') is typically argued to be more difficult to accommodate than soft triggers such as verbs like *win* or *stop* (Simons 2001, Abusch 2010, Romoli 2012, Romoli (to appear) among others). Our present findings do not challenge theories that propose that the presupposition of 'also' has pronominal characteristics—our indirect contexts would be sufficient to ground a pronoun (24). But they point the way toward manipulations which could explore this more thoroughly. For example, contexts like the one in (25) do not support the use of a pronoun but might allow for the predictive use of presuppositions (26).

(23) Look at what Michael has! It is good to eat.

(24) Look at Michael. ? It is good to eat.

(25) Look at Michael. Jane, also, has some candies.

The parallelism between the implicit and explicit context conditions is important because it suggest that, in this paradigm at least, there is no apparent cost to indirectness. Not only did participants use the indirect context to predictively, they did so as rapidly and efficiently as participants in the explicit context conditions. This suggests that the processes involved operate over discourse entities, rather than over lexical items. Perhaps this isn't surprising, since presuppositions are typically characterized as discourse level expectations, but it is reassuring.

4.2 A Surprising Failure

In both Experiment 1 and Experiment 2, we found absolutely no evidence that participants could use the inference associated with 'only' to predict that the upcoming object would be new to the discourse. We find this surprising for three reasons.

First, our intuition that this inference is accessible was validated in an offline study (see Appendix A). When given sufficient time, folks realize that only the new cohort item can complete the sentences with 'only'.

Second, in our task at least, the 'only' inference seems roughly comparable in complexity and constraint to inference in the 'also' condition. In both cases, participants must track the referents mentioned in the context sentence and then use this information to pick a possible referent in the target sentence. In Experiment 1, both of the critical words could potentially be used to focus in on two out of four referents prior noun onset. In Experiment 2, both the critical words could potentially allow the listener to focus on a single correct referent prior to the noun. In both cases, the critical word is associated with the subject but is being used to make an inference about the direct object. In fact, the differences that exist seem to favor 'only': the critical cue comes earlier in the sentence and the focused element is disambiguated by word order alone.

Finally, at first glance, our results seem to conflict with those of Kim and colleagues (2008) who found that participants made rapid use of 'only in sentences like 'Jane only has some candy.' Critically, these studies used essentially the same design as we used in Experiment 1. There are two critical differences between the Kim study and the present 'only' condition, which suggest two hypotheses about why adults are unable to make this inference in real time.

One difference is that in the Kim study 'only' associates with the VP, while in our study it associates with the subject. Crain and colleagues have found that children have more difficulty interpreting subject 'only' than object or verb-phrase 'only' (Crain et al. 1994), raising the possibility that this interpretation might be more difficult for adults to process. We think this is unlikely. Other developmental studies find that subject 'only' is no more difficult for children than verb-phrase 'only' and there is no evidence that we know of to suggest that adults have substantial difficulty with these forms (see Paterson et al. 2003; Paterson et al. 2006).

A second difference between the present study and that of Kim and colleagues is whether the inference in question leads one to prefer the previously mentioned item or reject it. In Kim's study, hearing 'only' leads the listener to construct a context set based on the previously mentioned items, encouraging looks to these referents. In our study, hearing 'only' should lead participants to infer that a previously mentioned items cannot be the object of the target sentence. Notice that this inference involves implicit negation (exclude the previously mentioned items) and the need to avoid looking at objects that were previously relevant. Both of these things could make processing more difficult (Wason 1965, Carpenter and Just 1975, Dale and Duran 2011 among others). One virtue of this account is that explains why we see a rapid effect in the 'also' condition: this inference, like Kim's, involves a direct preference for previously mentioned items, rather than an implicit negation.

Acknowledgement For conversations and inspiration on this project, we are very grateful to Gennaro Chierchia, Josh Hartshorne, Hazel Pearson, and the audiences of Language and Cognition at Harvard University in 2008 and 2009. A special thanks goes to Clemens Mayr and Tim O' Donnell, who contributed to the discussions that lead to these experiments. We wish to thank also Tracy Brookhyser, Carlyn Friedberg, Margarita Zeitlin, Nina Hrebenko and Miranda Fidler for assistance with data collection and preparation of the stimuli.

5 Appendix A: Off-line Norming Study on *Also* and *Only*

To ensure that inferences that we were studying were robust, we conducted an offline sentence completion study. The goal of this task was to verify that, given adequate time, participants could infer that the direct object in the 'only' sentence must be novel to the discourse, while the direct object in the 'also' sentences must have been mentioned in the context sentence. This experiment employed the materials from Experiment 1.

Sixty-four participants were recruited through the on-line crowd sourcing tool Amazon Mechanical Turk. They received $ 0.50 for their participation. Twenty eight

additional participants were excluded (26 for poor accuracy on the filler trials and 2 because English was not their native language). The design was almost identical to Experiment 1. Instead of hearing instructions, participants read the introduction, context and target sentences, except that last word of the target sentence was truncated after the first letter, as in (27).

(26) Jane also has some c__

Participants were shown the display from Experiment 1 and clicked on the picture that completed the sentence. Note that control-old and control-new trials were identical in this study, since the final word which differentiates these conditions is omitted. There were 8 'also' trials, 8 'only' trials,16 control trials, and 8 fillers which were used to filter out inattentive participants.

A logistic regression analysis was performed comparing 'old-cohort' choices for 'also' and 'only' sentences to the control sentences. Both the 'also' (M = 73.7 %, z = 18.935, $p < 0.001$) and the 'only' (M = 15.8 %, z = 3.973, $p < 0.001$) responses were significantly different from the control responses (M = 23.9 %). Therefore, we confirmed that participants were able to make the relevant inference and select the intended image for both the 'also' and the 'only' sentences.

6 Appendix B: Norming Study on the Intonation of *Also*

When 'also' appears between the subject and the verb it has two possible readings. We wanted our participants to get the reading in which 'also' associates with the subject. There is no other way to convey this reading in contemporary English. While the oldest co-author favored 'John too has some candles,' her younger colleagues insisted that no one spoke like this anymore. So we were forced to use prosody to disambiguate the intended reading. We did this by producing the sentences with the prosody given in (28). The capitalization indicates prosodic focus on the subject and the commas signal a prosodic break before the verb.

(27) JANE, also, has some candies.

It was our intuition that when the sentence was produced in this way, the only possible reading is the one where the presupposition is on the subject (Jane, in addition to someone else, has some candies). However, to make sure that folks did not get the reading where the presupposition was on the verb phrase (Jane, in addition to something else, has some candies), we conducted a norming study.

Ten undergraduate students participated in the experiment. The experiment was conducted using Psyscope X. The participants were been presented with 16 pre-recorded auditory stimuli like (28) above. These were a subset of the utterances that were used in the 'also' condition of Experiment 1. For each one they were asked to select the correct interpretation from two alternatives.

(28) a. Jane has some candies and she has something else too.
b. Jane has some candies and someone else has candies too.

Participants made their selection by pressing a keyboard button corresponding to the side of the screen where the alternatives was presented. The position of the alternatives on the screen was counterbalanced and the presentation order for the items was randomized for each subject. The stimuli were interspersed with 48 fillers. These included utterances with *too* as in (30) (followed by the same possible choices) and utterances like (31), (followed by a choice between 'Justin is Christina's neighbor' or 'Justin is Christina's neighbor and Christina is Justin's neighbor').

(29) JANE, too, has some candies.

(30) Justin and Christina are neighbors.

Participants judged the critical sentences to have the subject-association reading 84 % of the time. A Wilcoxon-Signed-Rank Test showed that this was significantly different from chance ($Z = 2.28$, $p < 0.05$). Thus the intonation pattern on the critical sentences strongly biases participants to interpret 'also' as associating with the subject.

References

Abusch, D. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27 (1): 1–44. doi:10.1093/jos/ffp009

Allopenna, P. D., J. S. Magnuson, and M. K. Tanenhaus. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38:419–439.

Beaver, D. 2001. *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

Beaver, D., and B. Z. Clark. 2009. *Sense and sensitivity. How focus determines meaning*. Oxford: Oxford University Press.

Beaver, D., and B. Geurts. To appear. Presuppositions. In *Semantics: An International Handbook of Natural Language Meaning*, vol. 3. ed. C. Maienborn, K. von Heusinger, and P. Portner. Berlin: Mouton de Gruyter.

Bott, L., and I. Noveck. 2004. Some utterances are underinformative. *Journal of Memory and Language* 51:437–457.

Carpenter, P., and M. Just. 1975. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review* 82:45–73.

Chemla, E. 2009. Presuppositions of quantified sentences: Experimental data. *Natural Language Semantics* 17 (4): 299–340. doi:10.1007/s11050-009-9043-9.

Chemla, E., and L. Bott. 2013. Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and Cognitive Processes* 38 (3): 241–260.

Chierchia, G., and S. McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics,* 2nd ed. Cambridge: MIT Press.

Crain, S., W. Ni, and L. Conway. 1994. Learning, parsing, and modularity. In *Perspective on sentence processing*, ed. J. Clifton, Charles L. Frazier, and K. Rayner, 433–467. Hillsdale: Lawrence Erlbaum Associates.

Dale, R., and N. D. Duran. 2011. The cognitive dynamics of negated sentence verification. *Cognitive Science* 35:983–996.

Gazdar, G. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.

Heim, I. 1983. On the projection problem for presuppositions. In *Proceedings of WCCFL 2*, ed. D. P. Flickinger, 114–125. Stanford: CSLI Publications.

Heim, I. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.

Huang, Y. T., and J. Snedeker. 2009. Online interpretation of scalar-quantifiers: Insight in the semantics-pragmatics interface. *Cognitive Psychology* 58 (3): 376–415.

Karttunen, L. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1:181–194.

Kim, C. 2007. Processing presupposition: Verifying sentences with 'only'. 31st Penn Linguistics Colloquium.

Kim, C., C. Gunlogson, M. Tanenhaus, and J. Runner. 2008. Focus alternatives and contextual domain restriction: A visual world eye-tracking study on the interpretation of 'only'. Proceedings of SuB 13.

Kripke, S. 2009. Presupposition and anaphora. *Linguistic Inquiry* 40 (3): 1–21.

Matin, E., K. Shao, and K. Boff. 1993. Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics* 53 (4): 372–380.

Paterson, K., S. Liversedge, C. Rowland, and R. Filik. 2003. Children's comprehension of sentences with focus particles. *Cognition* 89 (3): 263–294.

Paterson, K., S. Liversedge, F. White, Diane, and K. Jaz. 2006. Children's interpretation of ambiguous focus in sentences with 'only'. *Language Acquisition* 13 (3): 253–284.

Romoli, J. 2012. *Soft but strong: Neg-raising, soft triggers, and exhaustification*. Ph. D. thesis, Harvard University.

Romoli, J. 2014. The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics*. 0, 1–47, doi:10.1093/jos/fft017.

Romoli, J., and F. Schwarz. 2014. To appear. An experimental comparison between presuppositions and indirect scalar implicatures. In *Experimental Perspectives on Presuppositions.* ed. Florian Schwarz. Springer International Publishing.

Schlenker, P. 2008. Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics* 34 (3): 157–212.

Schlenker, P. 2009. Local contexts. *Semantics and Pragmatics* 2 (3): 1–78.

Schwarz, F. 2007. Processing presupposed content. *Journal of Semantics* 24 (4): 373–416.

Schwarz, F., and S. Tiemann. 2013a. The path of presupposition projection in processing—the case of conditionals. In *Proceedings of SuB 17*, ed. E. Chemla, V. Homer, and G. Winterstein, pp. 509–526.

Schwarz, F., and S. Tiemann. 2013b. Presupposition projection in online processing. (Ms., submitted).

Simons, M. 2001. On the conversational basis of some presuppositions. In *Semantics and Linguistic Theory (SALT) 11*, ed. R. Hastings, B. Jackson, and Z. Zvolenszky, 431–448.

Stalnaker, R. 1974. Pragmatic presuppositions. In *Semantics and Philosophy,* ed. M. Munitz and D. Unger, 197–213. New York University Press.

Stalnaker, R. 1978. Assertion. *Syntax and Semantics* 9:315–332.

Tanenhaus, M. K., M. J. Spivey-Knowlton, M. K. Eberhard, and J. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268 (5217): 1632–1634.

van der Sandt, R. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.

Wason, P. 1965. The contexts of plausible denial. *Journal of verbal learning and verbal behavior* 4:7–11.

Presuppositions vs. Asserted Content in Online Processing

Florian Schwarz

Abstract We report two experiments that investigate the time-course of the online interpretation of the presupposition of *also*, first relative to a control, and secondly relative to asserted content, namely the exclusivity of *only*, using the visual world paradigm. Both studies reveal rapid shifts in fixations to target pictures based on the presupposition expressed by *also*, after 200–300 ms after its onset. In contrast, the asserted exclusivity introduced by *only* arises roughly 400 ms later, suggesting that—if anything—presupposed content is evaluated prior to asserted content. This is as expected on semantic accounts of presuppositions, which see them as preconditions on interpreting the sentence in the first place, but somewhat surprising (though not necessarily strictly inconsistent) with pragmatic accounts that derive presuppositions via conversational reasoning, which has been found to require additional processing time in the case of scalar implicature computation.

Keywords Presuppositions · Online processing · Visual world paradigm · Eye tracking · Also · Only · Experimental pragmatics

1 Introduction

An early and crucial insight in the modern study of linguistic meaning is that what speakers and hearers generally seem to perceive as the overall conveyed meaning of a given utterance should be broken down theoretically into distinct components. The motivation for this is that upon closer inspection, they can be differentiated by their role in utterances in general as well as their behavior in different linguistic environments. To account for these differences, theorists generally appeal to distinct underlying mechanisms that give rise to these various aspects of meaning as well as to differences in how they affect the computation of the overall conveyed meaning of a complex utterance. Put very briefly, the literal, truth-conditional content of the lexical items in a sentence together with the structure they appear in is the basis for the core semantic impact of the sentence, standardly captured in terms of truth-conditions.

F. Schwarz (✉)
Department of Linguistics, University of Pennsylvania, 619 Williams Hall,
255 S. 36th Street, Philadelphia, Pennsylvania, 19104, USA
e-mail: florians@ling.upenn.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_4,

Presuppositions, our main concern here, constitute a type of backgrounded meaning that does not form part of the main point of the utterance, but which relates it to the discourse context, specifically with respect to the shared assumptions of the discourse participants. Presupposed content is typically (but not necessarily) taken for granted, i.e., it is something that all parties in the discourse are already aware of. Crucially, presuppositions display some distinct behaviors in embedded environments in that they do not seem to be affected by operators such as negation and conditionals in the same way as literal, asserted content. While these basic points are shared quite generally in the literature, there is no broad consensus about how presupposed content originates, as we will discuss in more detail below. A third major class of meaning is that of conversational implicatures, which—following Grice (1975)—are generally thought to arise via general reasoning about the role of rational agents in cooperative communication. One specific class of implicatures that has been particularly well-studied is that of scalar implicatures, which arise due to reasoning about possible alternative utterances where a given scalar expression is replaced with a logically stronger one.

While there has been decades of theoretical work on these phenomena, until recently little has been known about how the human mind constructs these various aspects of meaning in real time. However, work in experimental pragmatics over the last decade has made clear that a detailed investigation of different aspects of meaning in online processing not only enriches our psycholinguistic understanding of language comprehension in general, but also can help to inform theoretical debates about their nature and origin. In particular, a by now substantial body of work on the processing of scalar implicatures has argued that the secondary nature of implicatures, which on a Gricean view are derived by general reasoning about the literal semantic content, is reflected in delays in a number of processing measures, such as reaction times, reading times, and eye movements in the visual world paradigm, all of which are taken to indicate additional processing time and effort involved in computing implicatures (e.g. Bott and Noveck 2004; Breheny et al. 2006; Huang and Snedeker 2011; Bott et al. 2012). However, there also are studies that have found evidence for rapid computation of scalar implicatures, and the debate in the literature is far from settled (Grodner et al. 2010; Degen and Tanenhaus 2011, 2012; Breheny et al. 2013).

Turning to presuppositions, there is less evidence to date about their online processing properties, but just as with implicatures, much can be gained from such evidence. In particular, one central point of contention throughout the history of research on presuppositions is the question of whether they are primarily semantic or pragmatic in nature. In many ways, the issue comes down to the question of whether presuppositions are part of what is conventionally encoded as part of the lexical meaning of certain expressions or not. Much early work saw it as a primarily pragmatic phenomenon in terms of constraints on possible utterance contexts (e.g., Stalnaker 1970, 1973, 1974). Later linguistic analyses, such as dynamic semantics (Heim 1983; Chierchia 1995; Beaver 2001) and Discourse Representation Theory (DRT Kamp 1981; van der Sandt and Geurts 1991; van der Sandt 1992; Geurts 1999) propose semantic theories that integrate aspects of the context more closely, (e.g., by

seeing the meaning of sentences in terms of their potential for changing contexts), and thus integrate such contextual constraints into the semantics proper. But in recent years, pragmatic theories have seen a revival (Simons 2001; Abusch 2002, 2010; Schlenker 2008, 2010; Abrusán 2011). These are broadly based on the claim that presuppositions are derived via reasoning over conversational maxims (Grice 1975) and potential alternative expressions, similar to scalar implicatures (though possibly via different maxims).

Earlier work on presuppositions has begun to explore experimental methods for investigating them. Schwarz (2007) showed that presuppositions can affect the interpretation of ambiguous sentences and overrule syntactic parsing preferences, and found reading time effects (in self-paced reading) related to the presupposition of *also* when the context did not support it. Building on this paradigm, Tiemann et al. (2011) investigated a broader range of presupposition triggers and found that unsupported presuppositions gave rise to decreased acceptability and increased reading times in word-by-word self-paced reading. Schwarz and Tiemann (2012) extended this approach methodologically by using eye tracking in reading, focusing on *again* in German ('wieder'), and provided further evidence for rapid presupposition evaluation at a more fine-grained temporal resolution, namely in first fixation duration effects (at least for unembedded cases, which will be what we are concerned with in the present experiments as well).

One limitation of the reading time studies is that the effects generally arise due to lacking support or contradiction of presuppositions in the context. While the timing of these effects is indicative of the time course of presupposition interpretation, it does not allow for a direct assessment of felicitous presupposition comprehension in online processing, nor does it lend itself to a comparison with other types of content. The visual world paradigm (Tanenhaus et al. 1995), where participants' eye movements relative to a visual scene are tracked while they are listening to auditorily presented linguistic stimuli, seems better suited for addressing these issues. In particular, it allows the tracking of evolving interpretations of any type of content in real time with natural stimuli that do not involve any conflicts between the context and the presupposed information. While there are plenty of studies on reference resolution using the visual world paradigm, many of which also involve issues related to presuppositions (e.g. Keysar et al. 2000; Hanna et al. 2003; Grodner and Sedivy 2005), relatively little work has been directly focused at presupposition triggers other than referential expressions (but see Chambers and Juan 2008, for a study of the presuppositional verb *return*).[1]

This paper presents two studies using the visual world paradigm that investigate the time-course of presupposition interpretation, focussing on the presupposition of *also*. The first study, presented in Sect. 2, assesses the time course of interpreting (unstressed) *also* by comparing visual contexts where the presupposition singles out one of the displayed figures with control contexts where it does not. The second

[1] For other recent studies on presuppositions using the visual world paradigm, see Kim's and Romoli et al.'s contributions in this volume, as well as Romoli et al. 2013.

study, discussed in Sect. 3, attempts a first direct comparison of presupposed content with literal, asserted content, by including sentences with *only* and visual contexts where the exclusivity asserted by *only* singles out one of the displayed figures early on, and compares these to sentences with stressed *also* (in appropriately adjusted contexts). Control conditions in this experiment used sentence variants without *also* and *only*. The results from these studies provide clear evidence for rapid consideration of presuppositions in online processing, and the second study finds even earlier effects for the presupposition of *also* than for the asserted exclusivity of *only*. The implications of these findings are discussed in Sect. 4.

2 Experiment 1: The Time Course of Processing *Also*

The basic tenet of the visual world paradigm is that when listener's are presented with auditory linguistic stimuli while visually inspecting a display presented to them, their eye-movement behavior tends to reflect what is currently on their mind. When carefully controlling how the parts of the displayed image relate to the information conveyed by the linguistic input at a given point in time, this can be utilized to investigate what interpretation, if any, a listener is entertaining at a given point in time as the sentence in question unfolds. For present purposes, the target sentences crucially involved the additive particle *also*. When unstressed, *also* associates with the part of the sentence following it that is in focus and introduces a presupposition to the effect that there is some alternative to the interpretation of that part that yields truth in the remaining sentence frame as well (Kripke 2009; Krifka 1999; Chemla and Schlenker 2012). For example,

(1) Mary is also holding a [FORK]$_F$.

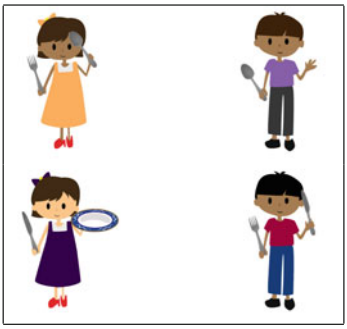
with focal accent on *fork*, presupposes that Mary is holding something else besides a fork, and furthermore requires that whatever else she is holding, say a spoon, is sufficiently salient in the discourse context (Kripke 2009). Assuming this is the case, it thus becomes in principle possible to infer that Mary is holding a spoon as well as something else at the point that *also* (or at the latest *also holding*) has been heard. It is this inference based on the presupposition of *also* that we utilize in our experimental task set within the visual world paradigm, where participants are instructed to click on one of the characters in a display. This allows us to assess the time-course of linguistic information of interest becoming available in online interpretation. In our case, this is done by manipulating whether only one or both of the characters within the display in question are in fact holding a spoon, and by framing the relevant descriptions as part of a instruction to click on one of the pictures (see (2) below). If the inference is available and reflected in eye-movement behavior, we thus expect different distributions of looks during the relevant time-window depending on whether or not one of the characters can be identified based on the information available from the linguistic input during the time-window in question. With this general characterization of the task in place, let us turn to a more detailed presentation of the actual materials and experimental design utilized.

2.1 Materials & Design

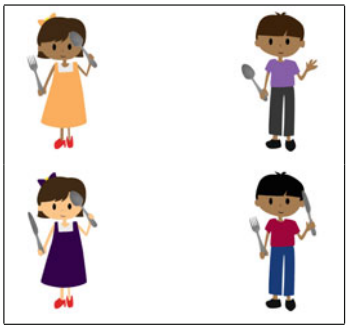
The experimental materials consisted of pairs of displays and auditory linguistic stimuli such as the following:

(2) a. *Context*:
One of the boys is holding a spoon.
b. *Target*:
Click on the girl who is also carrying a fork.

i. Critical



ii. Control



(Illustrations courtesy of Dorothy Ahn)

The target sentence, a variant of (1) above adjusted to suit the experimental task, contained a non-stressed *also* and had the main focal accent on *fork*. The initial context sentence served to ensure that the use of *also* is felicitous by providing an antecedent, in this case *a spoon*. The presupposition introduced by *also* in this context thus is that the girl to be clicked on is holding a spoon. Crucially, this presupposition can in principle already be inferred prior to encountering the final noun *fork*. The Critical display version contained only one girl that was holding a spoon (the one on the top left, henceforth referred to as the target), whereas the Control version had both of the girls in the display holding a spoon (with the other girl serving as a competitor). In the latter case, it is only at the point in time where the information provided by *fork* is available that it becomes possible to determine which girl to click on. In the former, however, this can already be done at an earlier point in time, namely as soon as *also* is encountered. This is possible if and only if the presupposition of *also*—that the girl in question be holding a spoon—is available. We thus are interested in eye movement patterns during the time window corresponding to the underlined part of the sentence in (2). If the presupposition of *also* is available in online processing, we expect a relative shift of fixations towards the target, as compared to the competitor in the Critical display condition.

A total of twelve sentence-picture pairs along the lines of the one illustrated above were created, with a variety of characters and objects. The crucial variation always involved whether the competitor figure had the object mentioned in the context sentence. In addition to the experimental items, there were a total of 32 fillers. The

first group of fillers consisted of 6 items with sentences parallel to the experimental ones, but without *also*. Furthermore, the figure to be clicked on never had what the figure in the context sentence had, to avoid general predictability of the target picture based on the context sentence. A second set of 6 fillers was again similar to the experimental items, but contained a conjunction in the target sentence (. . . *who has a fork and a spoon*). In half of these, the target had the object mentioned in the context sentence, and in the other half not. Both of these filler sets furthermore varied whether target and competitor each included two objects or whether one of them had just one object. A third set of 6 fillers included context sentences with a conjunction and target sentences with *only*, again varying whether the competitor had one or two objects between items. A final set of 12 items came from another experiment on the interpretation of *either . . . or*, with disjunctions in the target sentences and systematic variation of whether the target had one or two objects. Altogether, the fillers were designed to counter-act possible biases with respect to the likelihood of the eventual target picture containing the object mentioned in the context sentence. Furthermore, they introduced some variation in the distribution of objects, as well as in the types of target sentences subjects were to encounter.

2.2 Procedure & Participants

After signing a consent form, subjects were seated in front of a computer monitor (size: 21 inch, resolution: 1680×1050) connected to an EyeLink 1000 eye tracker by SR Research (used in desktop mode). After calibration, subjects saw instructions on the screen that they would be seeing pictures and listening to sentences that would provide instructions to click on one of the pictures. After doing a practice trial and having the chance to ask any questions they might have, the experiment began. Each trial started by the subject looking at a fixation point at the center of the screen to control for their initial eye position. Next, the visual display was previewed for 1500 ms, and subjects were free to look around in the display during this time. After the preview, an audio file with the pre-recorded linguistic stimuli was played back, after which subjects carried out the instructions by clicking on the appropriate picture using a mouse. The positions of target and competitor were counter-balanced across the experiment, though they were always vertically or horizontally aligned.

Sixteen undergraduate students from the University of Pennsylvania, all native speakers of English, participated in the experiment for course credit. Subjects were split into 2 groups, where each subject saw 6 of the experimental sentences per condition, yielding a counter-balanced data set altogether.

2.3 Results

2.3.1 Data Treatment

Fixations and Responses were coded according to which of the four figures in the display they corresponded, with interest areas of 300×400 pixels, and a 400 pixel

distance between image edges horizontally and 200 vertically. Unsurprisingly given the straightforward nature of the task, subjects's accuracy was at ceiling throughout (except for in the *or* sub-experiment, where there was some ambiguity).

To ensure proper eye tracking data without excessive amounts of track loss, we inspected the critical time window of interest, from the onset of *also* to the onset of the disambiguating noun (shifted by 200 ms to account for the time standardly assumed to be required for planning and initiating a saccade), and removed any trials where track loss exceeded 30 % of the samples taken. Only three trials had to be removed based on that standard.

2.3.2 Statistical Analysis

Given the design of the experiment, we were primarily interested in the frequency of looks to the target picture as compared to the frequency of looks to the competitor picture after hearing *also* but before encountering the noun. We therefore computed Target Advantage scores for time windows of interest by subtracting the proportion of looks to the competitor from the proportion of looks to the target. The resulting raw proportions were transformed into empirical logits (Barr 2008). Statistical analyses used mixed-effect models with subjects and items as random effects, using the *lmer* function of the *lme4* package in *R* Bates (2005). Following Barr et al. (2013), the maximal random effect structure that would converge was used in each case, with a random intercept as well as a random slope. Full random effect structures converged and were used unless otherwise noted. To assess whether inclusion of a given factor significantly improved the fit of the overall model, likelihood-ratio tests were performed that compared two minimally different models, one with the fixed effects factor in question and one without, while keeping the random effects structure identical (Barr et al. 2013). We report estimates, standard errors, and t-values for all models, as well as the χ^2 and p-value from the likelihood-ratio test for individual factors.

Figure 1 shows mean Target Advantage scores by condition as a function of time. As can be seen from the consistently higher line for the Critical condition from the onset of *also* to the mean onset of the disambiguating noun, there were relatively more looks to the target in this condition than in the control condition.

To assess this difference statistically, we first computed empirical logits for the average Target Advantage scores for the entire ambiguous region, defined as beginning 200 ms after the onset of *also* and ending 200 ms after the onset of the noun, as any significant shift in fixations to the target during this time should be attributable to the availability of the presupposition of *also*. On average, this time window lasted for 976 ms, though this varied somewhat between items. The expression *also* itself lasted 334 ms on average.

In a mixed effect model analysis, the factor manipulating the picture for the competitor contributed significantly to model fit ($\beta = 3.07$, $SE = 1.07$, $t = 2.88$; $\chi^2 = 7.09$, $p < .01$), with average empirical logit scores of 2.86 ($\approx 30.1\,\%$) for the Critical condition and -0.20 ($\approx 1.9\,\%$) for the Control condition. As the

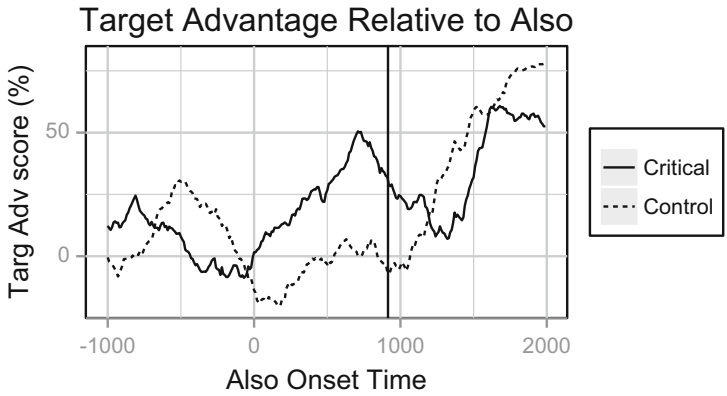


Fig. 1 Target advantage scores as a function of time aligned to the onset of *also*. The vertical black line indicates the average onset point of the disambiguating noun

Table 1 Estimates and t-values for Intercept (β_0) and effects of the picture manipulation (β_1) in 100 ms time windows (starting from 200 ms after the onset of *also*), as well as *p*-values from model comparisons for the latter

Time window:	200	300	400	500	600	700	800	900
β_0	1.68	2.45	2.73	3.24	4.36	5.09	4.09	3.03
t-value	1.51	1.92	2.26	2.91	3.669	5.11	3.89	2.45
β_1	− 3.09	− 3.01	− 2.90	− 3.20	− 3.92	− 4.78	− 3.94	− 3.48
t-value	− 2.445	− 2.31	− 2.28	− 2.45	− 3.06	− 3.65	− 2.73	− 2.28
p-value	< .05	< .05	< .05	< .05	< .05	< .01	< .05	< .05

Critical level of the factor was used as the reference level, the Intercept *t*-value of 2.81 furthermore shows that the mean for that condition was significantly different from 0, thus indicating a significant bias towards the target picture. This provides evidence that the interpretation of the presupposition of *also* is available during this time window.

To get a more fine-grained perspective on the time course of the interpretation of *also* having an effect, we furthermore divided the first 800 ms of the ambiguous period (as defined above) into 100 ms chunks and ran the same analysis for each separately. The results are summarized in Table 1.

The effect of the central competitor-picture manipulation significantly contributed to model fit right from the start in the 200–300 ms time window, and continued to do so throughout all remaining time windows we analyzed. Note, however, that the initial effect may at least in part be due to an apparent bias towards the competitor in the Control condition during the first few hundred milliseconds after the onset of *also*. It is unclear what this is due to, and analyses with Control as the reference level suggest that this is not significantly different from 0 (*t*'s < 1.5) at any point. Be this as it may, this short-lived competitor tendency ends around 400 ms, so that it cannot contribute to effects throughout. the *t*-values for the Intercept provide a

further indication of when the presupposition of *also* start affecting eye movement behavior and show that a significant preference for the target begins to emerge in the 300–400 ms time window.[2]

In order to evaluate whether the effects found for the experimental items might reflect any other general biases towards certain types of pictures in combination with the context sentences, fixation proportion plots for the three types of filler items were visually inspected as well. The only apparent effect found here was that whenever the target and competitor differed in that one involved one item and the other two, there was a slight preference for the one with two objects. Otherwise, Target Advantage scores up to the disambiguating noun fluctuated around zero, suggesting that none of the potential biases controlled for by the fillers are behind the result for our experimental manipulation.

2.4 Discussion

This first experiment investigated eye movements during a time window where relative shifts in fixation towards the target should only be possible if the presupposition of *also* is available. We find significant differences from the Control condition right from the earliest points on in the first one hundred milliseconds during which eye movements affected by *also* can arise (again assuming 200 ms for planning and implementing eye movements). This suggests that the presupposition of *also* is available right away in online processing, and that it is utilized in selecting the referent for the noun phrase in the target sentences.

However, one caveat is in order. While we controlled for several properties of the stimuli using filler sentences, there is one potential further factor that we could not assess based on the fillers, and which would be relevant for the experimental materials: it is quite plausible that subjects' looking behavior might display a preference for parts of the display that contain an object of the type mentioned in the context sentence (despite our attempts to vary the stimuli in this regard through the fillers). In the Critical picture, only the target had this property, whereas both target and competitor did in the Control condition. We can thus not rule out that such a potential preference could have at least partially contributed to the experimental effect of interest. The second experiment addresses this concern by keeping the pictures for Critical and Control conditions constant.

3 Experiment 2: Presupposition vs. Assertion—*Also* vs. *Only*

While the first experiment provides evidence for essentially immediate availability of the interpretation of the presupposition of *also* based on the fine-grained temporal resolution of the visual world paradigm, it does not provide any direct point of

[2] The Intercept can be interpreted in this way because the Critical condition was used as the reference level.

comparison to other aspects of meaning. This, however, is in many ways where much of the real theoretical interest lies, in light of the larger project of informing how different aspects of meaning relate to one another in terms of their processing properties. Most of the scalar implicature literature, for example, is concerned with trying to establish differences between asserted and implicated content. The second experiment attempts to contribute a first direct comparison between presupposed and asserted content. It does so by varying *also* with *only*, where the exclusivity asserted by the latter makes it possible to identify the target in the critical condition. There are two further changes in design from the previous experiment. First, rather than looking at unstressed *also*, we used stressed *also*, which associates with an element preceding it (Krifka 1999):

(3) John ALSO is carrying a fork.

The presupposition introduced here is that somebody other than John is carrying a fork. The main motivation for switching to stressed *also* was to allow for a maximally natural minimal variation with the *only* sentences.[3] The second change was that the control conditions used pictures identical to the critical conditions, while the sentences were manipulated by leaving out the *also*/*only*.

3.1 Materials & Design

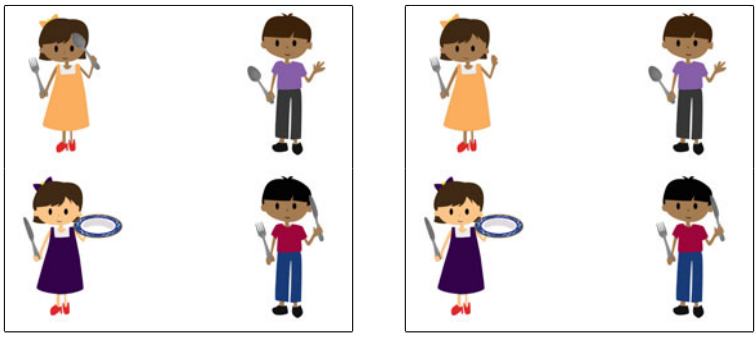
The experimental materials consisted of 24 sentence and picture pairings along the lines illustrated in (4), using a 2 × 2 interaction design. The sentences varied whether or not *also* and *only* were present in the respective conditions. The images were kept constant within the *also* and *only* Critical and Control sentence versions respectively, but were slightly different for the *also* and *only* conditions. In particular, the target in the *also* condition involved two objects, while the *only* target only involved one. The latter was necessary in order for the picture to match the exclusive condition contributed by *only*. The second object in the *also* condition was added to avoid any potential confusion arising from subjects (perhaps just temporarily) considering the interpretation associated with the unstressed *also* interpretation, which would require two objects to be involved.

(4) a. Also:
Context : One of the boys is carrying a fork.
Critical : Click on the girl who ALSO is carrying a fork.
Control : Click on the girl who is carrying a fork and a spoon.

b. Only:
Context : One of the boys is carrying a fork and a knife.
Critical : Click on the girl who only is carrying a fork.
Control : Click on the girl whois carrying a fork.

[3] Unstressed *also* in the pre-copula position was found to be non-optimal in a pilot study for Experiment 1.

c. Also Display Only Display



(Illustrations courtesy of Dorothy Ahn)

In addition to the crucial variation of the presence of *also* and *only*, several other minor variations were introduced to ensure maximal felicity of the sentences in their context and to avoid potential prosodic cues that might undermine the function of the control condition. First, the context sentence for the *only* sentences contained a conjunction to motivate the use of *only* in the target sentence while still being able to mention the object type of the noun in the the target sentence. Without the conjunction, the use of *only* here would seem unmotivated and somewhat out of place in intuitive terms. Secondly, the target sentence in the *also*-Control sentence contains a conjunction, in order to avoid a potential issue with the intonation of the verb and the noun phrase *a fork*. In particular, the verb phrase *is holding a fork* in the context *One of the boys is holding a fork* is Given (in the sense of Schwarzschild 1999) and normally would require de-accenting, which in turn would shift the main accent onto *girl*. This could provide an early prosodic cue of the verb phrase to come, thus undermining the control purpose of this condition, where disambiguation is not meant to occur until *fork* is heard. However, once we add a second conjunct, the verb phrase as a whole is no longer Given and as a result, de-accenting is not required and there is no accent shift onto *girl*.

Generally speaking, the logic of the design was parallel to the first experiment. In the Critical *also* condition, it is possible to identify the girl on the top left as the target as soon as *ALSO* is encountered, but only when its presupposition is available in online processing. Apart from this, disambiguation takes place with the (initial) noun, so its onset serves as the end-point for the time window of interest, both in the Critical and Control condition. As for the *only* conditions, it is the exclusive literal meaning expressed by *only*—that the girl is holding no more than one thing—that allows to single out the girl on the top left as soon as *only* is interpreted. Otherwise, disambiguation again comes with the introduction of the noun *fork*, which therefore also constitutes the end of the time period of interest for both *only* conditions.

Given the slight variations between sentences and contexts, we tried to ensure that the audio recordings varied as little as possible by splicing together recordings.

In particular the recording of the context up to the beginning of the relative clause (including *who*) was identical for the critical and control conditions for both *only* and *also*. This was done by recording the entire context-target sentence sequences, and subsequently replacing the context in one of the recordings with that from the other. Final items used the context from the original critical and control recordings half of the time respectively.

Yet another pre-caution we took to avoid any potential biases not related to our experimental manipulation was to counter-balance which conjunct in the *only*-context sentence occurred again in the target sentence, i.e., half of the *only* target sentences used the first conjunct from the context sentence and the other half the second.

A total of 18 fillers were included to distract from the patterns of the experimental manipulation and to control for general distributional properties of targets and competitors throughout the experiment. A first set of 6 fillers contained a conjunction in the context sentence. The subsequent target sentence picked out a target picture with just one object (while the competitor had two objects). In all these regards, the fillers were parallel to the *only*-Control sentences, the only difference being that the target sentence in the fillers introduced a new noun phrase that matched neither conjunct from the context sentence. Another set of 6 fillers was modeled after the *also*-control sentences, with a single noun phrase in the context and a conjunction in the target sentence. However, both noun phrases in the conjunction were new and did not match the one mentioned in the context sentence. Correspondingly, the target picture also did not involve the object mentioned in the context sentence. A final set of 6 fillers yet again involved a single noun phrase in the context sentence, while the target sentence involved a disjunction of the form *who is either holding a fork or a knife.* In this set, it was always the second disjunct that matched the noun phrase in the context sentence.

3.2 Procedure & Participants

The procedure was identical to that in Experiment 1 described above. Twenty-six undergraduate students from the University of Pennsylvania, all native speakers of English, participated in the experiment for course credit. For counter-balancing, subjects were split into 4 groups, where each subject saw 6 of the experimental sentences per condition.

3.3 Results

3.3.1 Data Treatment

As in Experiment 1, fixations and responses were coded according to which figure in the display they corresponded to. Response accuracy in terms of following the

instruction in the target sentence again was at ceiling, with almost no errors. Applying the same criterion for excessive track loss during the overall time window of interest of 30 %, 3 trials were removed from the full data set.

3.3.2 Statistical Analysis

The time window of interest in this experiment started with the onset of *also* and *only* in the critical conditions and ended with the onset of the noun (e.g., *fork* in (4)). In order to have a time window of equal size in the control conditions, which did not include *also* or *only*, we calculated by how much the onset of *also*/*only* preceded the onset of the copula in the Critical condition for each item and then included the same amount of time preceding the copula in the Control condition of each item in the time window. On average, the resulting time window lasted 1073 ms, with some modest variation between items and conditions, given the separate recordings that had to be used for this part of the stimuli. Target advantage scores were computed as detailed above and transformed into empirical logits.

As before, we used mixed-effects models to analyze the data. As the range of possible random effect structures (RES) becomes more complex in an interaction design, we use the following shorthand indications to facilitate presentation of results. The maximal random effects structure is represented as RES-1 and models possible variation between subjects and items for all aspects of the model, including the main effects and interaction. When the full structure did not converge, the random effect structure was simplified by removing the interaction term for items. Results are reported for the maximal random effect structure (RES-1), unless this failed to converge, in which case RES-2 was used (no further simplifications were necessary).

- RES-1: $(1 + factor1 * factor2|\, subject) + (1 + factor1 * factor2|\, item)$
- RES-2: $(1 + factor1 * factor2|\, subject) + (1 + factor1 + factor2|\, item)$

For the overall interaction analyses, predictors were centered, so as to render estimates of main effects. Planned comparisons between individual conditions were conducted using the appropriate treatment-coding.

Target Advantage scores are plotted as a function of time by condition in Fig. 2. Descriptively speaking, there are several obvious things to notice. First, the Control conditions hover in an overall more or less flat pattern up until the onset of the noun, which is as as expected since the disambiguation is only introduced by the noun. Both of the Critical conditions, on the other hand, contain a sharp rise during the time period of interest. Secondly, the rise in the Critical *also* condition appears to occur earlier than in the Critical *only* condition. Yet another observation is that the scores for the *also* conditions seem to be consistently higher than those for the *only* conditions. This is likely due to the fact that the target in the *only* conditions only contains one item, in contrast with the competitor (see discussion of fillers below).

To analyze these patterns statistically, we first looked at the means for the entire time window (from 200 ms after the onset of *also*/*only* to 200 ms after the onset of the noun) using a 2×2 interaction mixed effect model analysis. The mean Target

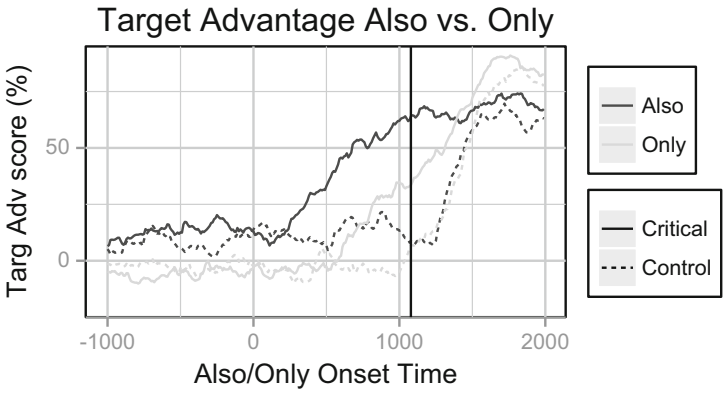


Fig. 2 Target advantage scores as a function of time aligned to the onset of *also*/*only*. The vertical black line indicates the average onset point of the disambiguating noun

Table 2 Target Advantage scores (in %) and corresponding empirical logits for the time window lasting from 200 ms after the onset of *also*/*only* to 200 ms after the onset of the noun

	Also		Only	
	Control	Critical	Control	Critical
Target Advantage	11.9	47.5	-3.7	12.8
Empirical Logit	1.39	4.22	-0.87	1.17

Advantage scores and empirical logits are provided in Table 2. There was no significant interaction, but there were significant main effects of *also* vs. *only* (RES-1: $\beta = 2.66$, $SE = 0.59$, $t = 4.50$; $\chi^2 = 14.79$, $p < .001$) and of Control vs. Critical (RES-1: $\beta = 2.43$, $SE = 0.45$, $t = 5.40$; $\chi^2 = 20.06$, $p < .001$), reflecting higher scores in the *also* conditions as well as in the Critical conditions. Planned comparisons between the Critical and Control conditions yielded significant simple effects for both *also* (RES-2: $\beta = 2.82$, $SE = 0.63$, $t = 4.46$; $\chi^2 = 16.21$, $p < .001$) and *only* (RES-2: $\beta = 2.03$, $SE = 0.63$, $t = 3.22$; $\chi^2 = 9.86$, $p < .01$), suggesting that subjects' eye movements reflected the interpretation of both expressions during the time window where no other information could differentiate between the target and the competitor.

A more fine-grained perspective of looking at eight 100 ms time-windows following the onset of *also*/*only* (shifted by 200 ms) yields a more nuanced picture of the time course of the relevant interpretations becoming available, specifically with respect to the relative timing for the two expressions investigated. Table 3 provides an overview of the results, listing estimated coefficients, t-values, and p-value ranges from model comparisons for the interaction as well as main and simple effects of Critical vs. Condition.

Table 3 Estimates, t-values, and p-values for the Interaction, the main effect ME and simple effects of Critical vs. Control (ME) for *also* (SE *also*) and *only* (SE *only*) for analyses of 100 ms time windows (starting from 200 ms after the onset of *also*)

Time:		200	300	400	500	600	700	800	900
Int.	β	− 0.32	− 1.03	− 2.64	− 3.22	− 2.10	− 1.65	− 0.55	− 0.5
	t-value	− 0.24	− 0.76	− 1.89	− 2.34	− 1.40	− 1.05	− 0.32	− 0.36
	p-value	–	–	< 0.1	< 0.05	–	–	–	–
ME	β	0.30	1.12	1.05	1.45	2.27	2.97	3.41	4.11
	t-value	0.41	1.61	1.37	1.89	3.11	3.52	4.51	5.81
	p-value	–	(0.11)	–	< 0.1	< 0.01	< 0.01	< 0.001	< 0.001
SE *also*	β	0.46	1.63	2.36	3.06	3.32	3.79	3.70	4.40
	t-value	0.49	1.68	2.32	2.97	3.37	3.64	3.67	4.01
	p-value	–	< 0.1	< 0.05	< 0.01	< 0.01	< 0.01	< 0.001	< 0.001
SE *only*	β	0.14	0.60	− 0.27	− 0.17	1.22	2.15	3.13	3.82
	t-value	0.14	0.62	− 0.26	− 0.15	1.08	1.71	2.631	3.73
	p-value	–	–	–	–	–	< 0.1	< 0.05	< 0.01

This analysis indicates that the effect of *also* emerges prior to the effect of *only*. The most solid statistical evidence for this is the significant interaction in the 400–500 and 500–600 ms time windows.[4] Furthermore, looking at the two expressions separately, we see that a simple effect of Critical vs. Control already emerges in the 300–400 ms time window for *also*, but not until the 700–800 ms time window for *only*. Nonetheless, even for the latter there is clear evidence for participants utilizing the exclusivity expressed by *only* in shifting their eye gaze to the target, prior to any further disambiguating information from the noun becoming available. Further evidence in this regard comes from the reaction times for clicking on the target, measured from the onset of the disambiguating noun. Click times were significantly faster in the Critical *only* condition (1119) than in the Control condition (1406 ms) ($\beta = 286.5$, $SE = 68.63$, $t = 4.17$; $\chi^2 = 15.14$, $p < .001$), suggesting that the earlier disambiguation based on the exclusivity expressed by *only* makes it possible for participants to carry out the clicking instruction more quickly.[5]

Target Advantage plots for the filler sentences that were similar to the experimental ones, aligned to the onset of the noun, were inspected visually to check for any effects due to properties of the stimuli separate from the crucial experimental manipulation. Data from the group of fillers modeled after the *only* control condition suggests that, as in Experiment 1, when target and competitor differed in that one had one object and the other two, there seems to be a bias towards looking at the one with two prior to the onset of the noun. This likely accounts for the main effect of *also* vs. *only* in the present experiment, since the picture types differed in this regard. However, note that this main effect is orthogonal to the critical findings of interest, namely the (temporary) interaction and the simple effects within the *also* and *only* conditions.

[4] Note also that when looking at target proportions (not relativized to looks to the competitor), the interaction already becomes significant in the 300–400 ms time window.

[5] Click times were also faster in the Critical condition for *also*, but given the different length in the target sentences between conditions here, this is not straightforwardly interpretable.

Recall that a potential concern with the interpretation of Experiment 1 was that there might be a bias towards pictures that contain an object of the type mentioned in the context sentence. The set of fillers modeled after the *also* control condition allowed us to assess this issue. In the time window of roughly 800 ms prior to the noun, such a bias was indeed found. Note that this also could have contributed to the main effect of *also* vs. *only* in the present experiment, since only the target had this property in the *also* conditions, whereas both target and competitor did in the *only* conditions. Crucially, however, within the Critical and Control conditions for each expression respectively, no variation of this sort existed. Thus, this factor cannot have contributed directly to the crucial effects of interest in the present experiment. The results from this filler type do suggest, however, that at least part of the effect in Experiment 1 may be due to it. It still seems unlikely that it is entirely due to this factor, given the size and shape of the effects found in the filler as compared to the experimental items in Experiment 1. However, since there is no straightforward way for testing this claim statistically, and since the same concern does not arise for the results from Experiment 2, we need not settle this question decisively.

3.4 Discussion

Experiment 2 sheds further light both on the time course of interpreting the presupposition of *also* in general and in comparison to the exclusivity asserted by *only* in particular. With respect to the first point, using identical pictures for the Critical and Control conditions avoided the issue that arose for Experiment 1 concerning preferences for pictures matching the context sentence in terms of including the object mentioned there. We thus have a cleaner, and more decisive result showing that the presupposition of *also* is available in online processing within about 200 ms (based on eye gaze effects within 400 ms after the onset of *also* and assuming the standard 200 ms needed for planning and implementing eye movements).

In addition to this insight into the absolute timing of interpreting *also*, an even more important aspect of this experiment was to relate the processing time line of a presupposition trigger to that of the asserted content of another expression, namely *only*. While the entire time-window between the onset of *also* and *only* did not reveal any direct differences between the two, with no interaction present and a main effect of Critical vs. Control, as well as corresponding simple effects in both the *also* and *only* conditions, a more fine grained perspective of looking at eight 100 ms time windows following the onset of *also* / *only* provided a more nuanced picture. In particular, shifts to the target did not reach significance for *only* until about 600 ms after the earliest possible point, i.e., about 400 ms later than they did for *also*. This difference in simple effects is supported most solidly in statistical terms in the 500-600 and 600-700 time windows, where there was an interaction between the type of expression and the Critical vs. Control conditions.

The difference between the two expressions suggests that not only do presuppositions not involve a delay, in contrast with what has commonly been argued for scalar implicatures, but—if anything—their availability in fact may precede that of

asserted content. The extent to which this conclusion holds of course depends on whether the exclusivity asserted by *only* indeed serves as a fair comparison. We will discuss this question in the next section.

4 General Discussion

The results from the two experiments reported here suggest that the presupposed information introduced by *also* is available early on in online processing, based on eye movement effects 200–300 ms after the onset of *also* (i.e., while still hearing *also*, which lasted 300–400 ms itself). The second experiment furthermore showed that eye movement effects based on the presupposition of *also* arise earlier than ones based on the asserted exclusivity of *only*. This suggests that presupposed content may be available even earlier than asserted content. However, the validity of such a conclusion hinges on the question of whether or not we are dealing with a fair comparison between the two types of meaning here, which we will turn to now.

Finding minimal comparisons between asserted and presupposed content is not a trivial challenge. We chose to compare (stressed) *also* to *only* based on their ability to occur in the same sentential context. Furthermore, they are similar in that they crucially relate to the focus of the sentence, and in the given contexts, they would seem to relate to the context sentence in an at least broadly parallel manner. Given that we do find online effects for *only*, albeit later than for *also*, furthermore suggests that there is nothing fundamentally wrong with our basic setup aimed at differentiating pictures in the critical conditions in terms of the exclusivity expressed by *only*. Finally, it seemed that the relation of what is linguistically expressed to the way that the contrast between the target and the competitor is implemented in the pictures is slightly more salient, if anything, in the *only* condition, since the target has just one object whereas the competitor has two (in the *also* conditions, both had two).

However, there are a differences between the two cases that we need to duly acknowledge as well, as they may in principle be relevant for the differences in the time-course of interpretation we found. First, *only* associates with a focus down-stream, later in the sentence, whereas stressed *also* associates with a preceding expression. This could mean that it is easier to flesh out the interpretation of *also* at the time the expression itself is encountered. At the same time, we found comparable effects for unstressed *also*, which also associates with a down-stream focus, so it's unclear whether this issue could really make all the difference in timing. Another potential complication present for *only* is that both the target and the competitor picture have one of the objects mentioned in the context sentence. While the competitor does contain yet another object, which excludes it from consideration if *only* is interpreted relative to a domain containing all the objects in the display, it could be considered a viable candidate if the domain for *only* is taken to just include the two objects mentioned in the context sentence. In the illustration above, a competitor holding a knife and a plate in a context where a fork and a knife have been mentioned could be thought to match the description ... *only is holding a* ... based on the

notion that out of a fork and a knife, this one is holding a knife. Consideration of this possibility could in principle contribute to a delay in deciding which picture is to be clicked on. However, even if this is the case, this option could only be considered temporarily, as we do find a significant shift in fixations to the target prior to the onset of the disambiguating noun. A final potential issue concerns the naturalness of *only* in the pre-auxiliary position in the experimental materials (which was chosen to keep the word order parallel to the *also*-sentences). At least some native English speakers have reported this word order to be somewhat degraded, and if this were true for our subjects, it could also contribute to delays in interpretation. In sum, given these issues, it seems reasonable to conclude that we are left with at least tentative evidence for presupposed content preceding asserted content in online processing, but clearly more work using other expressions, and possibly other paradigms, is needed in order to assess whether this interpretation of the results can be maintained in the long run and whether we are dealing with a phenomenon of a general nature.

Returning to the first aspect of our conclusions to be drawn from our experiments, the evidence for immediate interpretation of presupposed content is rather strong, and in line with other findings, both based on other methodologies (e.g., the reading time studies of Schwarz 2007; Tiemann et al. 2011; Schwarz and Tiemann 2012, 2013) and other recent approaches using the visual world paradigm (Romoli et al. 2013, 2014; Kim 2014). In theoretical terms, this type of result is more consistent with a semantically based notion of presuppositions, where presupposed content is part of what is conventionally encoded as the lexical meaning of the expression in question. In the final words of Beaver (2001), 'presupposition is what comes first in dynamic semantics,' for example, and other semantically based theories would lead us to expect that presuppositions are among the first things that are evaluated in the interpretation process, at least under the assumption that the theoretical procedures are mirrored more or less straightforwardly in processing. For pragmatic theories, on the other hand, the contrast of the present findings (and the other parallel results mentioned) with those for scalar implicatures are surprising, as such approaches take both types of meaning to arise based on considerations based on rational reasoning about the behavior of interlocutors in conversation. However, this point only stands to the extent that the results for scalar implicatures are (a) indeed of a general nature and (b) due to processes comparable to what is involved in interpreting presuppositions. Both points can be challenged, of course: various authors have argued recently that implicatures are not delayed after all, and despite the commonality of a basis in general conversational reasoning, there likely still are differences in the exact processes involved in computing implicatures and presuppositions, which could translate into processing differences. Finally, we need to keep in mind that it is by no way clear that presuppositions constitute a homogeneous phenomenon, as more and more attention is paid to possible differences between triggers. Assessing the more general implications of the present findings thus will require a much broader approach looking at a variety of triggers and suitable controls, and integrate the various recent developments in presupposition theory in more depth as well.

Acknowledgements I'd like to thank Robert Wilder for essential help with implementation of the design, data collection, and data treatment. Many thanks also to Dorothy Ahn for creating the visual stimuli, as well as for assistance with data collection. Part of the work reported here was supported by a grant from the University Research Foundation of the University of Pennsylvania.

References

Abruásn, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics and Philosophy* 34 (6): 491–535.

Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presuppositions. *Proceedings of SALT* 12:1–19.

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27 (1): 37–80.

Barr, Dale J. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* 59 (4): 457–474.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68 (3): 255–278.

Bates, Douglas M. 2005. Fitting linear mixed models in r. *R News* 5:27–30.

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

Bott, L., and I. A. Noveck. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language* 51 (3): 437–457.

Bott, Lewis, Todd M. Bailey, and Daniel Grodner. 2012. Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language* 66 (1): 123–142.

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

Breheny, Richard, Heather J. Ferguson, and Napoleon Katsos. 2013. Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes* 28 (4): 443–467.

Chambers, Craig G., and Valerie San Juan. 2008. Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition* 108 (1): 26–50.

Chemla, Emmanuel, and Philippe Schlenker. 2012. Incremental vs. symmetric accounts of presupposition projection: An experimental approach. *Natural Language Semantics* 20 (2): 177–226.

Chierchia, Gennaro. 1995. *Dynamics of meaning. Anaphora, presupposition and the theory of grammar*. Chicago: University of Chicago Press.

Degen, J., and M. K. Tanenhaus. 2011. Making inferences: The case of scalar implicature processing. In *Proceedings of the 33rd annual conference of the cognitive science society,* ed. L. Carlson, C. Hölscher, and T. Shipley, 3299–3304. Austin: Cognitive Science Society.

Degen, J. and M. K. Tanenhaus. 2012. Naturalness of lexical alternative predicts time course of scalar implicature processing. *Poster presented at the 25th CUNY sentence processing conference*. New York.

Geurts, Bart. 1999. *Presuppositions and pronouns*. Amsterdam: Elsevier.

Grice, H. P. 1975. Logic and conversation. In *Syntax and semantics: Speech acts,* ed. P. Cole and J. L. Morgan. New York: Academic Press.

Grodner, D., and J. Sedivy. 2005. The effect of speaker-specific information on pragmatic inferences. In *The processing and acquisition of reference,* ed. E. A. Gibson and N. J. Perlmutter, 239–272. Cambridge: MIT Press.

Grodner, Daniel J., Natalie M. Klein, Kathleen M. Carbary, and Michael K. Tanenhaus. 2010. "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition* 116 (1): 42–55.

Hanna, J. E., M. K. Tanenhaus, and J. C. Trueswell. 2003. The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language* 49 (1): 43–61.

Heim, Irene. 1983. On the projection problem for presuppositions. In *Proceedings WCCFL 2,* eds. M. Barlow, D. Flickinger, and M. Wescoat, 114–125. Stanford: CSLI.

Huang, Yi Ting, and Jesse Snedeker. 2011. Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes* 26 (8): 1161–1172.

Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal Methods in the Study of Language: Proceedings of the Third Amsterdam Colloquium,* vol. I, ed. J. Groenendijk, T. Janssen, and M. Stokhof, 227–321. Amsterdam: Mathematical Center.

Keysar, B., D. J., Barr, J. A. Balin, and J. S. Brauner. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science* 11 (1): 32.

Kim, Christina. 2014. Presupposition satisfaction, locality and discourse constituency. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

Krifka, Manfred. 1999. Additive particles under stress. In *Proceedings of SALT 8,* ed. Devon Strolovitch and Aaron Lawson, 111–128. Ithaca: CLC Publications.

Kripke, Saul. 2009. Presupposition and anaphora: Remarks on the formulation of the projection problem. Linguistic Inquiry 40 (3): 367–386.

Romoli, Jacopo, Manizeh Khan, Yasutada Sudo, and Jesse Snedeker. 2013. Resolving temporary referential ambiguity using presupposed content. Poster presented at the 26th Annual CUNY Sentence Processing Conference.

Romoli, Jacopo, Manizeh Khan, Jesse Senedeker, and Yasutada Sudo. 2014. Resolving temporary referential ambiguity using presupposed content. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

van der Sandt, R., and B. Geurts. 1991. Presupposition, anaphora, and lexical content. *Lecture Notes in Computer Science*. 546:259–296. (Text Understanding in LILOG).

van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.

Schlenker, P. 2008. Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics* 34 (3): 157–212.

Schlenker, Philippe. 2010. Local contexts and local meanings. *Philosophical Studies* 151 (1): 115–142.

Schwarz, F. 2007. Processing presupposed content. *Journal of Semantics* 24 (4): 373–416.

Schwarz, Florian, and Sonja Tiemann. 2012. Presupposition processing-the case of german wieder. In *Proceedings from the 18th Amsterdam Colloquium,* ed. M. Aloni, V. Kimmelman, F. Roelofsen, G. Weidman Sassoon, K. Schulz, and M. Westera. FoLLI: Springer.

Schwarz, Florian, and Sonja Tiemann. 2013. The path of presupposition projection in processing—the case of conditionals. In *Proceedings of SuB 17,* ed. E. Chemla, V. Homer, and G. Winterstein. 509–526. Paris: semanticsarchive.net

Schwarzschild, Roger. 1999. Givenness, avoid F and other constraints on the placement of accent. *Natural Language Semantics* 7 (2): 141–177.

Simons, M. 2001. On the conversational basis of some presuppositions. In *Proceedings of SALT,* vol 11. ed. R. Hastings, B. Jackson, and Z. Zvolensky, 431–448. Ithaca: CLC Publications.

Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic* 2 (4): 447–457.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and Philosophy,* ed. Milton K. Milton and Peter K. Unger. New York: New York University Press.

Stalnaker, Robert C. 1970. Pragmatics. *Synthese* 22 (1): 272–289.

Tanenhaus, M. K., M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268 (5217): 1632–1634.

Tiemann, Sonja, Mareike Schmid, Nadine Bade, Bettina Rolke, Ingo Hertrich, Hermann Ackermann, Julia Knapp, and Sigrid Beck. 2011. Psycholinguistic evidence for presuppositions: On-line and off-line data. In *Proceedings of Sinn & Bedeutung 15,* ed. I. Reich, E. Horch and D. Pauly, 581–597. Saarbrücken: Universaar-Saarland University Press.

Presupposition Satisfaction, Locality and Discourse Constituency

Christina S. Kim

Abstract The current study investigates presupposition-satisfying dependencies from the point of view of discourse processing. Using the presupposition trigger *also* as a case study, I ask to what extent the distance spanned by the trigger and the prior discourse content that satisfies the presupposition of *also* influences comprehenders' interpretation of the discourse—specifically, whether comprehenders have a bias toward satisfying presuppositions using material in the discourse that is closer rather than more distant.

Two offline experiments and one Visual World eye-tracking experiment provide evidence in favor of a locality bias in presupposition satisfaction. The findings support sensitivity to both linear distance and distance in terms of hierarchically structured discourse representations, consistent with an interpretation system that tracks both structure-insensitive information about discourse mention, and structured representations of larger discourse units.

Keywords Locality · Dependency length · Focus · Alternatives · Discourse processing · Eye-tracking

1 Introduction

Sentences with *also* (1-a) are standardly assumed to entail their propositional content (1-b) and presuppose a distinct proposition which differs from the sentence's propositional content in the value of the focused constituent (Horn 1969; Karttunen and Peters 1979; Rooth 1985; Atlas 1991). That is, at least one alternative in A (1-c) is true, where the alternatives are restricted by the context.[1]

(1) a. Andy also bought some NECTARINES.
b. Andy bought some nectarines.
c. A = {Andy bought bread, Andy bought some celery, Andy bought a croissant, ... }

[1]Assume the context restricts the set of possible alternatives to just the salient ones, where *salient* can be understood as *likely to be inferred by the addressee* (see e.g. Kim (2012)).

C. S. Kim (✉)
University of Kent, Canterbury, UK
e-mail: c.s.kim@kent.ac.uk

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_5,

Thus, the discourse-final sentence (2-f) entails that Andy bought nectarines and presupposes that he bought something other than nectarines. If not already contextually entailed, presuppositions must be accommodated (Lewis 1979) as background.

(2) a. The roommates often go to the farmer' s market together.
b. Beth always buys bread.
c. Andy usually buys some celery.
d. His doctor told him he needs to eat more vegetables.
e. Today Andy treated himself to a croissant.
f. He also bought some NECTARINES.

From the point of view of processing, readers have been shown to experience processing difficulty at the presupposition trigger if the presupposed information has not been processed earlier in the discourse (Moulton 2006; Schwarz 2007). Such findings can be seen as characterizing the conditions under which presupposed content be understood in discourse as intended by the speaker. However, the question of which of multiple viable interpretations—different ways for a presupposition to be satisfied—is preferred by the listener has been relatively understudied.

The current study addresses the question of preference by viewing a presupposition trigger like *also* as forming a dependency with the discourse content that satisfies the presupposition, analogous to filler-gap or anaphoric dependencies (van der Sandt 1992). By taking this view, we can ask whether such presupposition dependencies pattern like other kinds of dependencies formed during language comprehension, and are sensitive to factors known to constrain dependency formation at the sentential level. Specifically, locality (Hawkins 1994; Gibson 2000) has been shown to influence sub-sentential processing complexity. Here, I ask whether a similar bias in favor of local dependencies operates at the discourse level, using presupposition triggers and the discourse content they associate with as a case study. Asking this question opens up a way to address a related question about the mental representation of discourse—whether the relevant sense of locality is linear distance or distance relative to hierarchically-structured constituents (Webber and Joshi 1998; Kehler 2002).

The following sections outline some relevant prior findings that support structured representations of discourse (Sect. 2), including evidence from language comprehension and production that discourses hypothesized to have different structures are also processed differently (Sect. 2.1). Section 3 reviews evidence for a processing bias in favor of local dependencies. Two offline experiments assessing comprehenders' preferences for presupposition interpretation are presented in Sect. 4. Section 5 presents a Visual World eye-tracking experiment (Tanenhaus et al. 1995) that tracks the online competition of interpretations representing linear and hierarchically-structured locality. Section 6 concludes.

Fig. 1 Linear representation of discourse relations

2 Mental Representations of the Discourse

There is considerable evidence that discourse meaning cannot be adequately described using just the machinery of subsentential semantics and syntax. It is especially notable that research on discourse structure has been driven by different motivations and conducted using very different methodologies in neighboring fields of study. There is a rich history of research on narrative and text processing in psychology (Levin and Moore 1977; Kintsch and van Dijk 1978; Beaugrande and Colby 1979; Schank et al. 1982; Morrow 1985; Gee and Grosjean 1984; Graesser and Singer 1994; Simner and Pickering 2005) as well as artificial intelligence and machine learning (Hayes 1977; Cohen and Perrault 1979; Litman and Allen 1987; Hovy 1993). A temporally parallel development moved formal linguists from a static view of sentential meaning toward context-update models of discourse, which take into account the sensitivity of linguistic meaning to a representation of the context that continually updates as a discourse progresses (Heim 1982; Groenendijk and Stokhof 1984; Kamp and Reyle 1993).[2]

The major approaches to discourse structure differ along two dimensions that are relevant here. First, approaches differ in the types and number of relations that can hold between discourse units. In a number of prominent theories, the bulk of the explanatory work is done by a (presumably uni5versal) limited inventory of discourse relations (Hobbs 1979; Mann and Thompson 1988; Kehler 1995, 2002; Knott and Sanders 1998; Wolf and Gibson 2005).[3] While none of these theories in principle disallows structured representations of discourse, the emphasis remains on linear relations between adjacent segments (Fig. 1).[4]

Existing approaches also differ in the extent to which discourse representations encode hierarchical and/or non-local structural relations among discourse units (McKeown 1985; Polanyi and van den Berg 1996; Webber and Joshi 1998; Asher and Lascarides 2003). These theories typically also feature an inventory of possible relations, but the characterizations of different relations include reference to their structural properties; for example, *Narration* is a coordinating relation in Segmented Discourse Representation Theory (Asher and Lascarides 2003), while *Topic* is a

[2] See also Hintikka (1976); Lewis (1979) for early work on dialogue/discourse-based approaches to meaning.

[3] See Sanders et al. (1992, 1993) for a taxonomy of coherence relations in terms of a small set of cognitive primitives, such as causality.

[4] This is primarily in reference to the coherence relations described in Hobbs (1979) and Kehler (2002). Rhetorical Structure Theory (Mann and Thompson 1988) features both a large inventory of relations and the possibility of forming dependencies between non-atomic discourse constituents, producing hierarchical discourse structures.

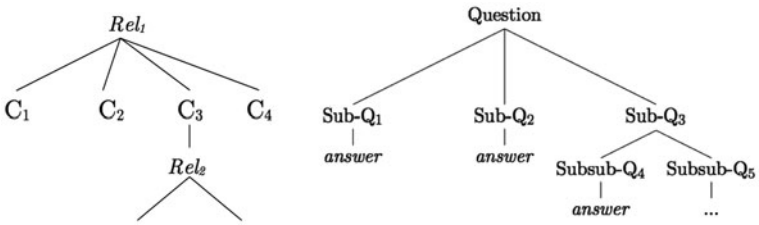


Fig. 2 Hierarchical representations of discourse

subordinating one. Of primary importance here is that such structure-based theories attempt to explain what makes larger pieces of discourse cohere, by making reference to organizing principles like topichood.

A subclass of structure-based theories that emphasizes a particular organizing principle is the question-based theories of discourse, such as Roberts (1996)'s Question Under Discussion framework (Grosz and Sidner 1986; van Kuppevelt 1995; Büring 2003; Farkas and Bruce 2009; Ginzburg 2012). These approaches have in common the claim that discourses are organized around (often implicit) questions, or to use more general language, goals. A contribution to a discourse is appropriate, or coherent, if it provides a partial answer to the question of the current discourse; because partial answers include answers to sub-questions (and sub-sub-questions), discourse representations are inherently hierarchical (as illustrated in Fig. 2). These approaches constitute a type of structure-based theory, but importantly adopt a view that questions (or goals) are what organize discourses.

2.1 Processing Consequences of Discourse Structure

Importantly for current purposes, discourse factors have been shown to clearly influence interpretation in language processing. Explanations of coreference and binding (Grosz et al. 1995; Stevenson et al. 2000; Arnold and Griffin 2007), syntactic ambiguity resolution (Frazier and Clifton Jr. 1996; Traxler et al. 1998; van Berkum et al. 1999; Hemforth et al. 2000), and ellipsis (Kehler 2000; Ginzburg and Sag 2000; Cooper and Ginzburg 2002; Frazier and Clifton Jr. 2006; Kim and Runner 2009) (among other phenomena) have appealed to properties of discourse relations or structures when interpretations vary in a way that syntactic or semantic properties cannot reliably predict.

A separate body of work in language processing collectively shows that discourse comprehension is influenced by aspects of discourse structure that go beyond sentence-level syntax and semantics. Studies in text processing (Sanders et al. 1992; Murphy and Shapiro 1994) demonstrated that memory for discourses is enhanced by integrating a text into a larger discourse structure, and by manipulating the task associated with reading a text, suggesting that the comprehension of discourse content is affected by goal structure during memory encoding. In addition, the ability

to detect contradictions, anomalies, and inconsistencies relative to discourse-level inferences is one of the signatures of depth of comprehension (Glenberg et al. 1982; Otero and Kintsch 1992; Graesser and McMahen 1993).

In online discourse processing, reading times show facilitation when a discourse resolves in a way that is expected based on prior discourse content or structure (for example, the discourse sets up an expectation for a causal continuation), and slow-downs when the actual continuation is an unexpected one. Bicknell and Rohde (2009) showed in a reading time study that comprehension in sentences containing ambiguously-attached relative clauses is influenced by the discourse relation that holds between the matrix and subordinate clauses.[5]

In language production, as well, language users appear to keep track of information related to discourse relations, and use that information to guide subsequent productions. For example, Simner and Pickering (2005) showed in a story continuation study that narratives containing information about the cause of an event generate more continuations related to consequences of the event, and narratives containing information about an event's consequences are followed by more continuations related to the event's cause. Rohde (2008) (Rohde et al. 2007; Kehler et al. 2008) show similar effects in sentence completion studies, where pronominal reference was influenced by the discourse relation that linked the clause containing the pronoun and the antecedent clause.

3 Locality Effects in Sentence Processing

The research reviewed above demonstrate systematic processing reflexes of features of the discourse, such as discourse relations; this type of evidence tells us what kinds (and what grain) of information need or need not be encoded in the mental representation of the discourse. However, it does not tell us why or how those kinds of information should matter from the perspective of facilitating processing, or making it more efficient—that is, it does not tell us anything mechanistic. A body of research in sentence processing investigates dependency formation at the sentential level with such questions in mind.

The notion of dependency length plays a central role in proposals like Gibson (1998, 2000) and Hawkins (1994, 2004). Gibson's Dependency Locality Theory (DLT), in particular, links the lengths of dependencies (e.g. between a verb and its arguments) within a sentence to complexity, as reflected by processing time. A number of previously documented empirical phenomena in sentence processing receive a unified explanation in terms of the lengths of the dependencies they require comprehenders to construct. For example, King and Just (1991) (among others) showed that object-extracted relative clauses were more complex to process than their subject-extracted counterparts. The DLT explains this asymmetry in terms of two

[5] For related studies on discourse effects in clause-level processing, see Millis and Just(1994); van Berkum et al. (1999); Roland et al. (2008); Rohde et al. (2011).

sources of complexity: a *storage cost*, which increases as the distance across which syntactic predictions of previous words must be maintained in memory increases, and an *integration cost*, incurred when a dependency is formed with a preceding word, which increases as the distance spanned by this dependency increases. In the case of subject-extracted relative clauses, the verb in the relative clause is close (often adjacent) to the relativizer (e.g. *who*); the analogous dependency in a relative clause with an object gap is longer—minimally, it must be maintained across the subject argument— and as such incurs a higher integration cost. Related work by Temperley (2007) has since shown using corpora that production data also reflect a bias to minimize dependency lengths. Hawkins (1994, 2004) ties processing complexity to universals related to constituent ordering universals.

Part of the appeal of such theories is that their characterization in terms of notions like dependency and distance makes them in principle very broad-coverage. And since a theory about dependency length minimization crucially does not make reference to dependencies among verbs, subjects, and objects, we can ask whether other kinds of dependencies involved in language interpretation are *this* kind of dependency. The experiments presented here ask this question for dependencies triggered by presuppositions.

4 Experiments 1–2: Presupposition Satisfaction and Local Dependencies

Experiments 1 and 2 asked whether comprehenders are sensitive to locality for presupposition-satisfying dependencies. Since *also* (in a discourse like (2), repeated below as (3)) presupposes only that some other alternative (of the form *Andy bought x*) is true, it might not matter to comprehenders where the content satisfying this presupposition occurs in the preceding discourse.

(3) a. The roommates often go to the farmer' s market together.
b. Beth always buys bread.
c. Andy usually buys some celery.
d. His doctor told him he needs to eat more vegetables.
e. Today Andy treated himself to a croissant.
f. He also bought some NECTARINES.

If locality does matter to comprehenders, we can ask whether linear distance matters, or distance measured with respect to hierarchically-structured constituents.[6]

Experiment 1 tested discourses like (3), in which both linear and hierarchical locality predict the final sentence is most easily interpreted as "Andy bought nectarines and a croissant," where "Today Andy treated himself to a croissant" is linearly closest

[6] For purposes of this study, I make the simplifying assumptions that sentences are atomic discourse units—i.e. they are not further decomposed, and that they are related to each other by a finite set of discourse connectives, which often but not always correspond to natural language connectives.

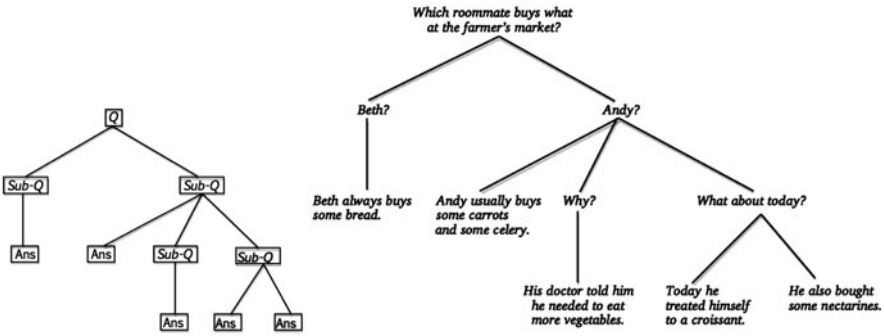


Fig. 3 Discourse tree for Experiment 1

to, and in the smallest discourse constituent containing the trigger *also* (Fig. 3 gives a hypothetical structured representation of the discourses tested in Experiment 1; tree structures are modeled after Roberts (1996) and Büring (2003)).

Experiment 2 tested discourses like (4). Here, unlike the discourses in Experiment 1, what counts as local differs depending on whether distance is measured in terms of linear distance, or in terms of structured representations. If dependencies minimize linear distance, (4-e) should show the same interpretive bias as in (3-f), yielding an interpretation where Andy bought nectarines and croissants. Alternative interpretations of (4-e) should decrease in likelihood as the linear distance spanned by the presupposed content in the preceding discourse increases: for example, the interpretation where Andy bought nectarines, croissants and carrots should be more likely than the interpretation where he bought nectarines, croissants, carrots, and bread. On the other hand, if locality is defined over hierarchical representations (as depicted in Fig. 4), the difference in discourse structure between (3) and (4) should yield different interpretations: the closest dominating discourse node is the discourse-initial topic (4-a) for (4-e), but is discourse-medial for (3-f).

(4) a. The roommates went to the farmer' s market together.
b. Beth bought some bread.
c. Frank bought some carrots.
d. When his girlfriend is there, she always gets some croissants.
e. Andy also bought some NECTARINES.

4.1 Method

Participants Twenty native English speakers recruited via Amazon Mechanical Turk participated in Experiment 1; a separate twenty participated in Experiment 2. Compensation depended on how quickly each participant completed the study, averaging a rate of $ 5.00/h.

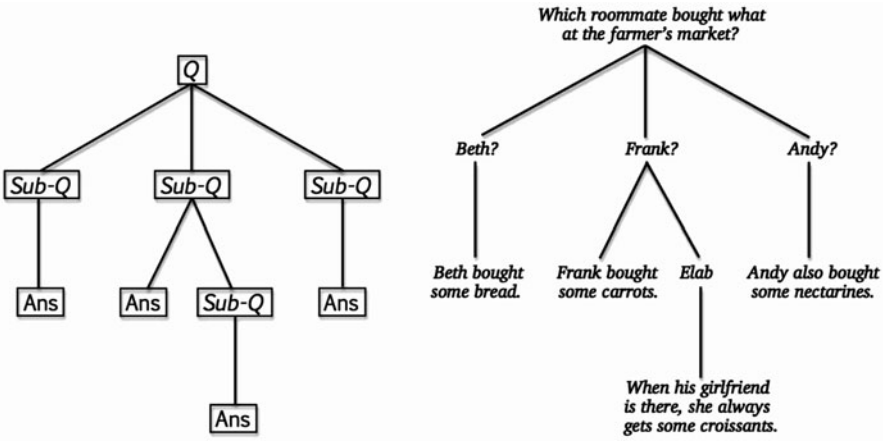


Fig. 4 Discourse tree for Experiment 2

Materials The materials for Experiment 1 consisted of ten discourses like (3). In each six-sentence discourse, the first sentence introduced the discourse topic, the second and third sentences introduced two characters, and the remaining three sentences continue being about the second character. The discourse-final sentence features *also*. Participants were instructed to indicate their interpretation of the final sentence in the context of the entire discourse by choosing one of four responses provided. An example of the response types is given in (5), for the discourse in (3).

(5) a. LOCAL INTERPRETATION:
Andy bought some nectarines and a croissant.
b. INTERMEDIATE INTERPRETATION:
Andy bought some nectarines, a croissant, and some celery.
c. GLOBAL INTERPRETATION:
Andy bought some nectarines, a croissant, some celery, and some bread.
d. FALSE INTERPRETATION:
Andy bought a croissant, some celery, and some bread.

The labels for the response types indicate either the distance spanned by the discourse content interpreted as the presupposition (LOCAL, INTERMEDIATE, GLOBAL), or that the interpretation is one that makes the discourse-final sentence false (FALSE).[7] For reference, Table 1 shows how the response types for Experiments 1–2 compare with respect to linear and structured locality.

[7] Note that because the subject of the final sentence in (3) is a pronoun, it is incompatible with the prosodic pattern associated with subject focus. This is not the case for the discourses used in Experiment 2 (4), only responses compatible with direct object focus were available in that experiment.

Table 1 Response types, experiments 1–3

Experiment	Response type	Linear locality	Structured locality
1	Local	**local**	**local**
	Intermediate	less local	(non-constituent)
	Global	least local	non-local
	False		
2	Linear local	**local**	(non-constituent)
	Intermediate wrt model	less local	(non-constituent)
	Intermediate wrt mention	less local	(non-constituent)
	Structured local	least local	**local**
3	Linear superset	**local**	(non-constituent)
	Structured superset	non-local	**local**
	Mentioned (false)		
	Mentioned-subset (false)		
	Novel (presupposition failure)		

The materials for Experiment 2 consisted of six discourses with the same form as (4). In each five-sentence discourse, the first sentence introduced the discourse topic, the second and third sentences introduced two characters, the fourth sentence elaborated on the third sentence, and the final sentence introduced a third character. As in Experiment 1, the discourse-final sentence contained *also*, and participants indicated their interpretation in the context of the discourse by choosing one of four responses. An example of the response types is given in (6), for the discourse in (4).

(6) a. LINEAR LOCAL INTERPRETATION:
Andy bought some nectarines and some croissants.
b. INTERMEDIATE INTERPRETATION (WITH RESPECT TO SITUATION MODEL):
Andy bought some nectarines and some carrots.
c. INTERMEDIATE INTERPRETATION (WITH RESPECT TO DISCOURSE MENTION):
Andy bought some nectarines, some croissants, and some carrots.
d. STRUCTURED LOCAL INTERPRETATION:
Andy bought some nectarines, some carrots, and some bread.

As explained above, the discourses in Experiment 2 were designed to pit linear locality (represented by the LINEAR LOCAL interpretation) against structured locality (represented by the STRUCTURED LOCAL interpretation).

The two INTERMEDIATE interpretations are analogous to the INTERMEDIATE interpretation in Experiment 1: the discourse content included in the presupposition is more than just the most recently mentioned item, but less than the entire set of mentioned items. The interpretations differ in that (6-b) excludes items inconsistent with the situation model of the events described (i.e. the content of the fourth sentence is excluded since it does not describe an actual event), whereas (6-c) includes this content. A bias in favor of interpretation (6-b) relative to (6-c) would suggest that comprehenders interpret presupposed content by referring to a mental model of the situation being described, rather than a discourse representation that only keeps track of whether an item has been mentioned.

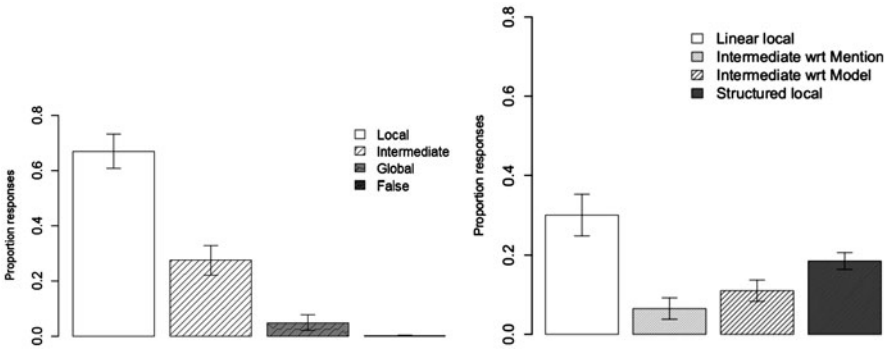


Fig. 5 Experiments 1 (*left*) and 2 (*right*) Results: proportion of responses

Table 2 Experiment 1 Pairwise comparisons of response types

Contrast	χ^2	Adjusted p
Local>Intermediate	34.04	< 0.0001
Local>Global	104.34	< 0.0001
Local>False	131.03	< 0.0001
Intermediate>Global	28.45	< 0.0001
Intermediate>False	51.07	< 0.0001
Global>False	8.33	< 0.005

4.2 Results and Discussion

The results of Experiments 1 and 2 provide support for both a structured view of discourse representations, and sensitivity to linear dependency distance. In both discourse types, comprehenders showed preferences for two interpretations: the material introduced in (i) the most recent sentence in the discourse or (ii) the smallest discourse unit dominating the sentence with the trigger were most likely to be construed as the presupposed content. Both of these interpretations were favored relative to intermediate interpretations, which neither minimized locality in terms of recency, or formed discourse constituents. For Experiment 1 (3), both (i) and (ii) correspond to the interpretation where Andy is understood to have gotten nectarines and a croissant. Figure 5 (left panel) shows the proportions of responses of each type. Pairwise comparisons showed that all pairs except for the Global and False responses differed in the frequency of responses (Table 2).

For Experiment 2 (4), minimizing the linear distance spanned by the presupposition dependency corresponds to the interpretation where Andy is understood to have gotten nectarines and croissants. This was the most frequent interpretation, as shown in Fig. 5 (right panel). In the second most frequent interpretation, the material introduced in the local discourse unit is construed as the presupposition—here, Andy is understood to have gotten every item mentioned in the prior discourse. Pairwise comparisons for the four response types showed that linear-local responses significantly exceeded both intermediate response types, and marginally exceeded structured-local responses. Structured-local responses significantly exceeded intermediate-with-

Table 3 Experiment 2 Pairwise comparisons of response types

Contrast	χ^2	Adjusted p
LinearLocal>IntermediateModel	16.49	< 0.001
LinearLocal>IntermediateMention	21.33	< 0.001
LinearLocal≈StructuredLocal	4.0	$= 0.1$
IntermediateModel≈IntermediateMention	1.64	*n.s.*
IntermediateModel≈StructuredLocal	2.63	*n.s.*
IntermediateMention<StructuredLocal	8.0	< 0.05

respect-to-mention responses, but not intermediate-with-respect-to-model responses (Table 3). Note that there was a significant number of structured-local responses ($t = 9.51$, $p < 0.0005$), despite the fact that this interpretation violates strict linear locality; by contrast, the global responses in Experiment 1 did not differ significantly from zero ($t = 1.82$, $p = 0.1$).

It is striking that, as in Experiment 1 (Fig. 5), linear-local responses predominated, considering that this interpretation is not even consistent with the situation model comprehenders must construct to interpret the discourse (in the example discourse in (4), no croissants were purchased in the sequence of events described up to the target sentence). This suggests that there is indeed a pressure to minimize linear distance for presupposition dependencies; this recency bias appears to be prioritized over interpretations where the presupposition conforms to the mental model of the situation being described by the discourse. Because the two intermediate response types do not differ reliably in the current data, no definitive conclusions can be drawn about whether presupposition dependencies care about consistency with the situation model. However, to the extent that reasoning is required to suppress readings that are inconsistent with the situation model, the results of Experiment 2 may reflect the intrusion of model-inconsistent readings due to low-level pressure to satisfy a presupposition with the closest available material.

Turning to the structured-local responses, note that the advantage over either of the intermediate responses contrasts strikingly with the pattern of responses in Experiment 1, where intermediate responses reliably exceeded global responses. Since the structured-local responses represent maximal linear distance, this contrast goes against any linear distance minimization constraint. The fact that either the linear-local or structured-local response was chosen more often than either of the intermediate responses suggests that there is both a pressure to minimize linear distance and a bias toward local interpretations that respect discourse units (though in the aggregated data we cannot tell whether both constraints were respected by all participants, or whether participants respected either linear or structured constraint).

Within the relevant local discourse constituent, exhaustive interpretations were preferred to restrictive ones: comprehenders resisted distinguishing among discourse units with the same hierarchical status. That is, they preferred to interpret sentence (4-e) as meaning that Andy bought all the items mentioned in the discourse, as opposed to e.g. carrots, croissants and nectarines (but not bread). Together, Experiments 1 and 2 provide preliminary offline data suggesting that comprehenders are sensitive to both structure-insensitive recency (minimizing linear distance) and locality in terms of structured discourse representations, even in cases where preserving discourse "constituents" may sacrifice strict linear locality.

5 Experiment 3: Different Ways of Satisfying the Same Presupposition

Experiment 3 tracks the timecourse of presupposition resolution in discourses containing *also*. In light of the results of Experiment 2, which implicates both linear and structural locality constraints, the current study asks whether competition between multiple possible interpretations of a presupposition is observed online, and additionally, whether there is evidence of a bias to preserve discourse constituency when constituency-preserving and constituency-violating interpretations are available.

5.1 Method

Participants Twenty-seven undergraduate students from the University of Rochester participated in Experiment 3. Participants were recruited from introductory Linguistics courses and flyers posted on the university campus, and were paid $ 7.50 per session. All participants were native speakers of American English, and had normal or corrected-to-normal vision.

Materials and Design Experimental materials consisted of discourses like (7).

(7) a. The roommates went to the farmer' s market together.
b. Beth bought some bread.
c. Frank bought some carrots and some apples.
d. Andy also got some NECTARINES.

Each discourse appeared with one of three display types, illustrated in Fig. 6. All displays contained at least one subset of mentioned items (8-a)–(8-b) and one set of all discourse-new items (8-c). In addition, the displays included one of the following: (i) a superset of locally-mentioned items in terms of linear distance (*Linear-local display*, (8-d)), (ii) a superset of locally-mentioned items in terms of structured discourse constituency (*Structured-local display*, (8-e)), or (iii) both linearly and structurally defined supersets (*Competition display*; (8-d)–(8-e)).[8]

(8) a. subset of mentioned (carrots, apples)
b. subset of mentioned (apples)
c. all novel (nectarines)
d. superset of mentioned (linear) (carrots, apples, nectarines)
e. superset of mentioned (structured) (carrots, apples, bread, nectarines)

[8] In order to minimize differences in complexity among display quadrants, each quadrant contained 6–8 objects, regardless of the number of object types present. For example, a *subset* quadrant would have 6–8 objects of the same type (e.g. apples), while a *superset* quadrant might have two of each of four object types (e.g. carrots, apples, bread, nectarines).

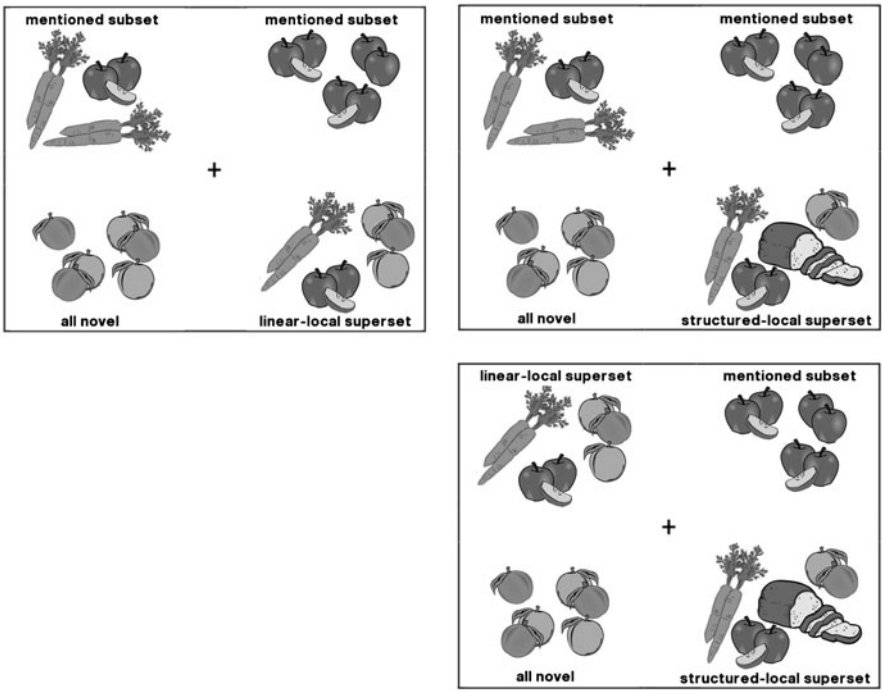


Fig. 6 Experiment 3 display types. Clockwise: *Linear-local, Structured-local* and *Competition displays* (labels for illustration only; illustration—author's own)

Each display contained four 200 × 200 pixel images located at the corners of the 1024 × 768 pixel computer screen (with images flush with screen edges).

Participants saw five tokens of each combination of display type (*Linear-local*, *Structured-local* or *Competition*) and discourse type (*Also* or No-*also*), yielding a total of 30 experimental trials. These were interspersed with 68 filler trials design to minimize statistical regularities in the materials. The trials were presented in a random order generated on each run of the experiment. The 98 trials were preceded by four practice trials, none containing *also* or featuring a target item that would appear as a target in an experimental trial.

If comprehenders construct and use hierarchical constituent structures online to constrain presupposition satisfaction, we expect a preference for the structured superset, which respects discourse constituency but violates linear locality, over the linearly defined superset, which minimizes linear distance, but breaks up a discourse constituent. The discourse trees for the interpretations corresponding to the linear and structured superset responses are shown in Fig. 7.

Procedure Each trial began with the participant fixating and clicking on a crosshair in the center of the screen. Participants listened to the discourses over headphones. The display appeared on the computer screen at the onset of the target sentence;

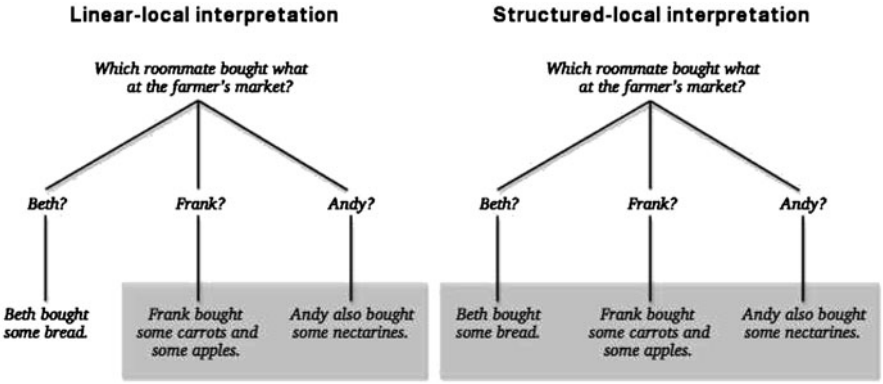


Fig. 7 Discourse tree representation of linear-local and structured-local interpretations

there was no preview. Participants were instructed to click on the items that the last-mentioned character got (e.g. "what Andy got"). The trial ended when the participant clicked on a picture in the display. Eye movements were recorded from the onset of the target sentence to the end of the trial, using a head-mounted SR EyeLink II eye-tracking system sampling at 250 Hz.

5.2 Results and Discussion

Because one of the objectives of Experiment 3 was to observe participants' behavior when more than one viable interpretation is available, I analyze response data and eye movements, and when appropriate, eye movements contingent on response.

Looking first at the Competition display condition, we can see that both of the superset interpretations (linear local or structured local) are possible interpretations of the target sentence: when both interpretations are available in the display, responses are split (albeit unevenly) between the two superset types (Fig. 8). In fact, with an online interpretation task, we see a pronounced bias in favor of the structured local interpretation; this contrasts with the offline responses in Experiment 2, where linear local interpretations predominated (Fig. 5).

The eye movement data also suggest that the two superset interpretations remain in competition after other options have been ruled out, and in fact well after the offset of the target word. Figure 9 shows the proportion of fixations for the Competition display condition; Figs. 10 and 11 break the data down by response type: Fig. 10 represents trials where the participant chose the structured local interpretation, and Fig. 11 represents trials where the linear local interpretation was chosen. When the structured interpretation was chosen (Fig. 10), fixations to the linear superset reliably exceed fixations to the subset referent in the 400–600 ms window after the target word onset ($t = 2.81$, $p < 0.05$), and this difference persists at least until the

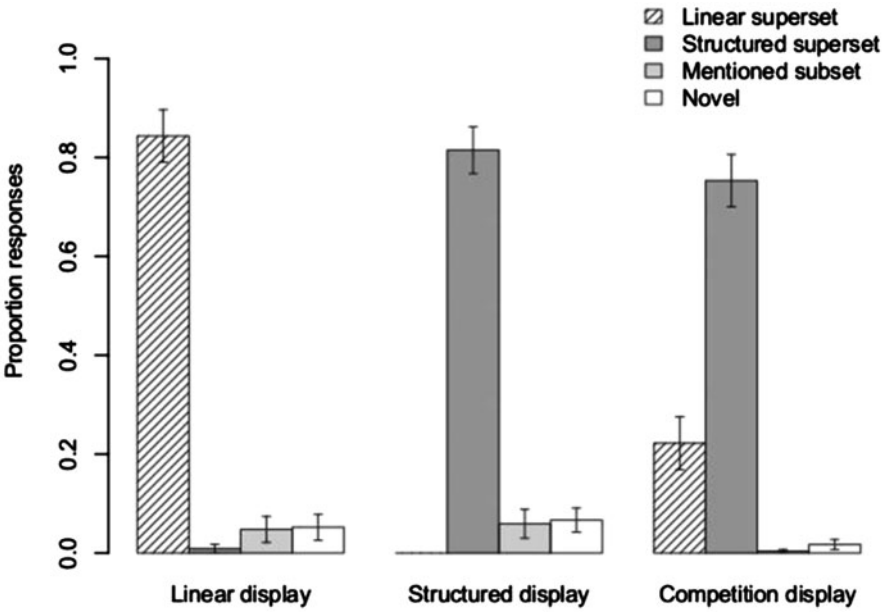


Fig. 8 Experiment 3 Results, Proportions of responses, by display type

1800–2000 ms window—well after the offset of the target word. When the linear interpretation was chosen (Fig. 11), fixations to the structured superset exceed subset fixations beginning in the 1200-1400 ms window ($t = 2.13$, $p < 0.05$); this difference persists until at least 1800–2000 ms after target onset.

Despite the fact that both interpretations appear to be considered online, one interpretation may be preferred over the other—this is suggested by the asymmetry in response types in the Competition display condition: when both interpretations were available in the visual display, participants chose the structured local interpretation more often than they chose the linear local interpretation (Fig. 8). This reflects a preference for discourse constituency-preserving interpretation at the expense of minimizing linear dependency distance. To explore timecourse differences between these two interpretations, we will look at the two conditions where only one interpretation was available (*Linear-only* and *Structured-only* display conditions).

Proportions of fixations for *Linear-only* and *Structured-only* display conditions are shown in Figs. 12 and 13. Target fixations from these two conditions (excluding the Competition conditions) were fitted using mixed-effect logistic regression models in three analysis windows, delimited by salient linguistic events in the stimuli: the pre-particle window spans the 500 ms before the onset of "also," the pre-target window starts at the onset of the particle and ends at the onset of the target word, and the post-target window starts at the target word onset and ends 500 ms later. The onsets used to delimit these windows were determined on a trial-by-trial basis. The models predicted fixations to the target referent, and included (1) Display type and

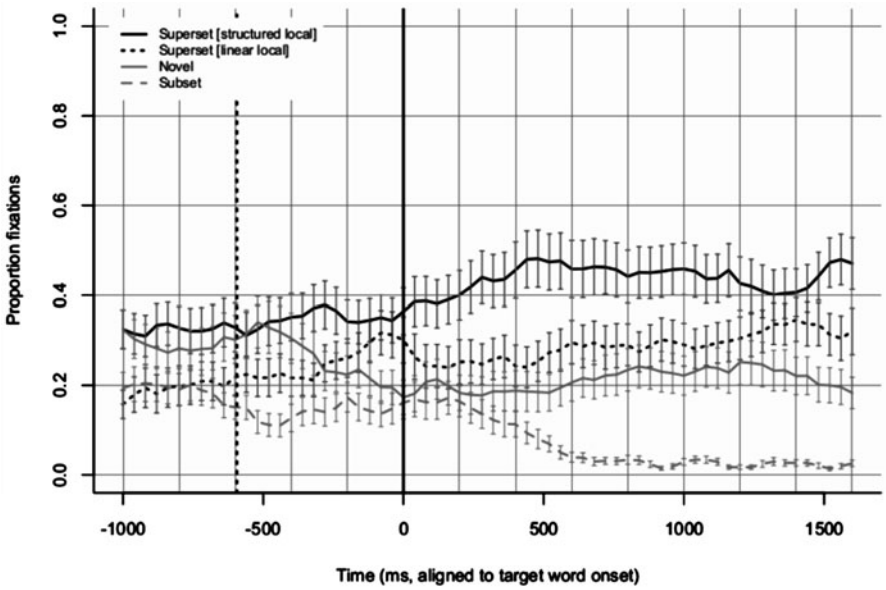


Fig. 9 Experiment 3 Results, Competition display conditions (all response types): Mean proportion of target fixations. (Dotted vertical line = average particle onset; solid vertical line = target word onset)

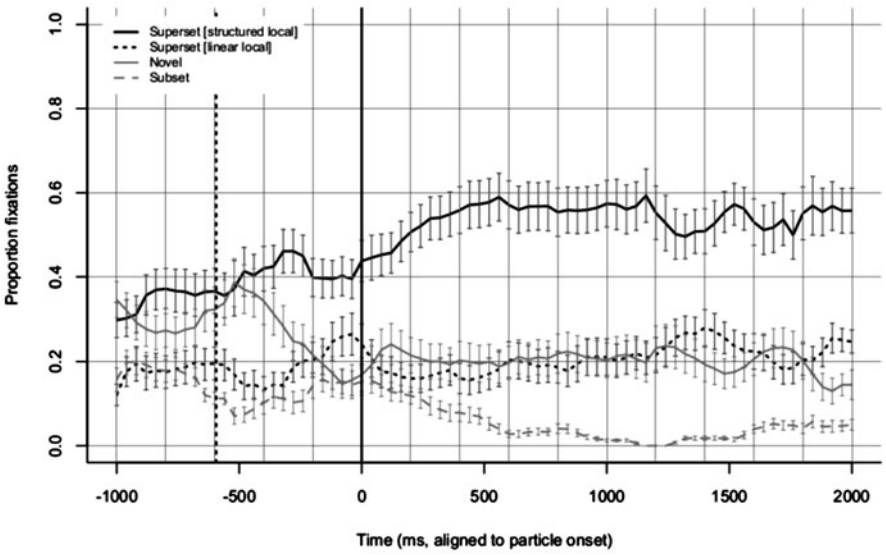


Fig. 10 Experiment 3 Results, Competition display conditions (Structured local responders): Mean proportion of target fixations. (Dotted vertical line = average particle onset; solid vertical line = target word onset)

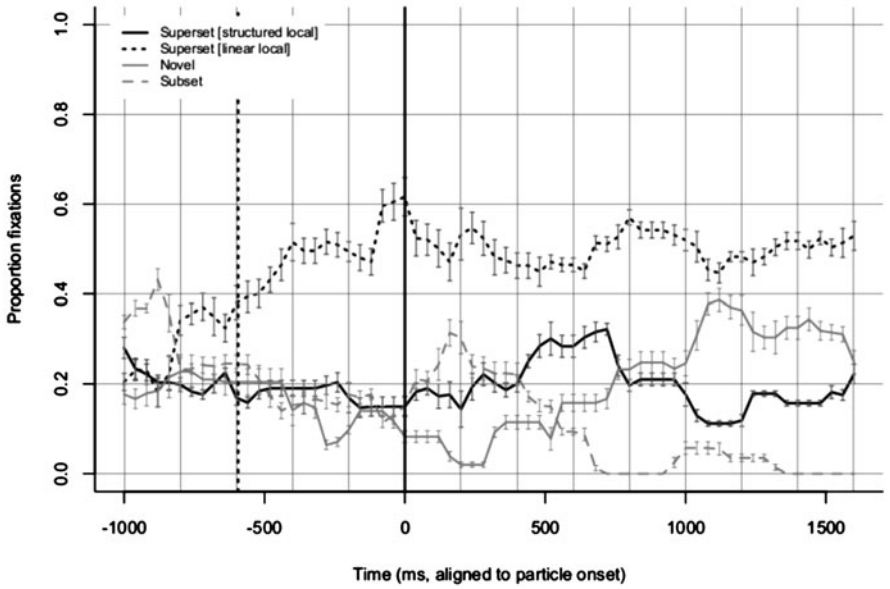


Fig. 11 Experiment 3 Results, Competition display conditions (Linear local responders): Mean proportion of target fixations. (Dotted vertical line = average particle onset; solid vertical line = target word onset)

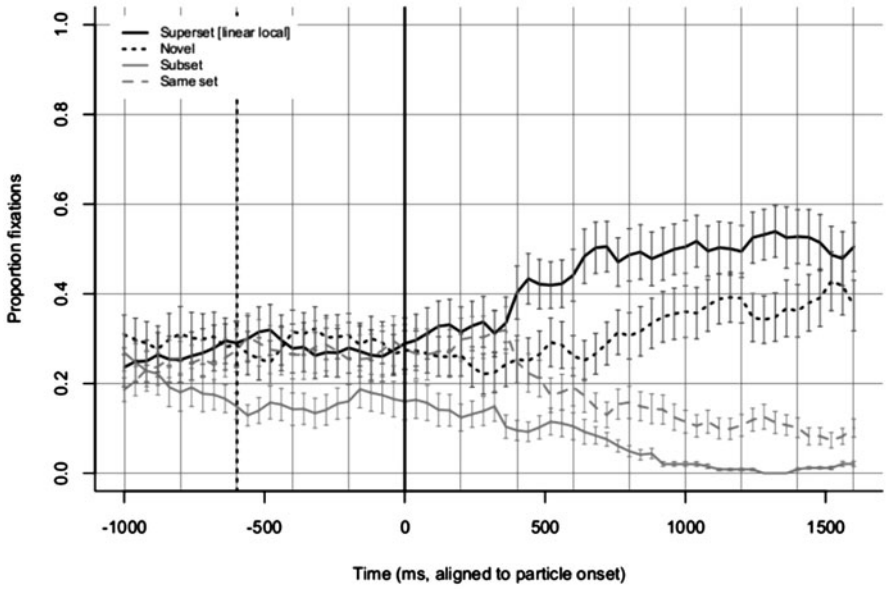


Fig. 12 Experiment 3 Results, Linear-only display conditions: Mean proportion of target fixations. (Dotted vertical line = average particle onset; solid vertical line = target word onset)

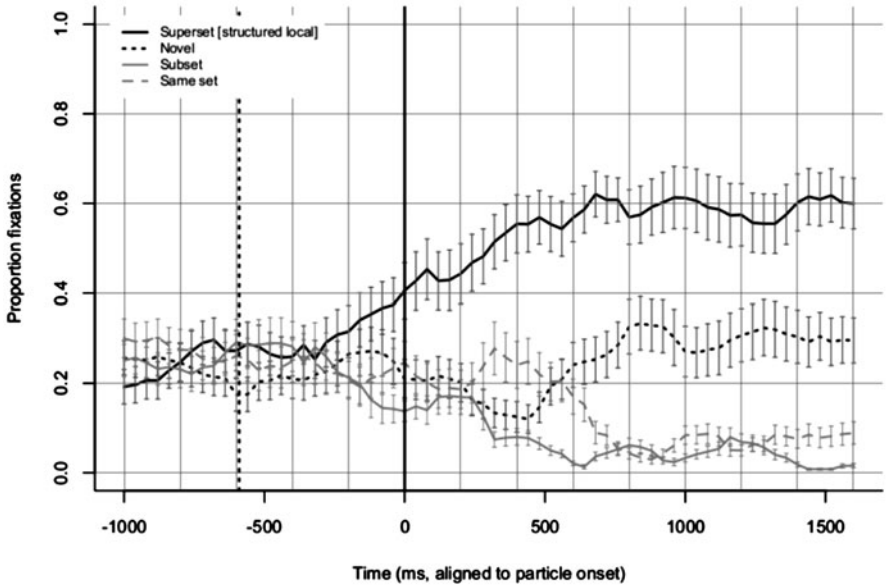


Fig. 13 Experiment 3 Results, Structured-only display conditions: Mean proportion of target fixations. (Dotted vertical line = average particle onset; solid vertical line = target word onset)

(2) Time as fixed effects, and Participant and Item as random effects (Jaeger 2008; Barr 2008). Data was unaggregated in all the regression models reported. In addition, we included the state of the previous fixation (on or off the target) as a predictor to deal with the oversampling problem that arises in analyses of Visual World fixation data (Frank et al. 2009).

Analyses began with the full model, which included the interaction between Display type and Time (interactions with State were not included, since there is no clear theoretical reason for State to interact with the variables of interest, and any such interactions would not be interpretable). The State term was left in the model, regardless of significance. All predictors were centered. Redundant terms were removed by eliminating one predictor at a time for all terms correlated with one or more other terms in the model, starting with the highest order term. Model comparison using the likelihood ratio test determined whether the model including the predictor increased the likelihood of the data relative to the model excluding that term. To determine the random effects structure, for each model with a given fixed effects structure, we performed forward step-wise model comparison, starting with the model containing just random intercepts for Participant and Item, and iteratively adding random effects to the model. Each resulting model was compared to the previous one using the likelihood ratio test. The estimated coefficients for the final models are shown in Tables 4, 5 and 6 for the initial, early and late analysis windows.

As can be seen both in the models in Tables 4 and 6, and by comparing the fixation plots in Figs. 12 and 13, there is no effect of Display type on target fixations in the

Table 4 Experiment 3 estimates of fixed effects, initial window

```
TargetFix ~ DisplayType + Time + State
  + (1+DisplayType|Participant) + (1+DisplayType|Item)
```

	Estimate	*SE*	z	*p*
Intercept	− 7.15	0.76	− 9.40	< 0.0001
StructuredLocalDisplay	0.05	0.24	0.20	*n.s.*
Time	− 1.13	0.83	− 1.37	*n.s.*
State	11.93	0.28	42.69	< 0.0001

Table 5 Experiment 3 Estimates of fixed effects, early window

```
TargetFix ~ DisplayType + Time + State + DisplayType:Time
+ (1+DisplayType|Participant) + (1+DisplayType|Item)
```

	Estimate	*SE*	z	*p*
Intercept	− 6.67	0.37	17.85	< 0.0001
StructuredLocalDisplay	0.88	0.46	1.94	0.05
Time	− 0.36	0.97	− 0.37	*n.s.*
State	12.01	0.23	51.64	< 0.0001
StructuredDisplay:Time	1.99	1.29	1.55	0.12

Table 6 Experiment 3 Estimates of fixed effects, late window

```
TargetFix ~ DisplayType + Time + State
+ (1+DisplayType|Participant) + (1+DisplayType|Item)
```

	Estimate	*SE*	z	*p*
Intercept	− 5.75	0.24	− 24.11	< 0.0001
StructuredLocalDisplay	0.14	0.20	0.69	*n.s.*
Time	0.08	0.68	− 0.11	*n.s.*
State	11.36	0.21	55.07	< 0.0001

initial (500 ms preceding particle onset) or late (500 ms following target onset) analysis windows. However, in the early window (particle onset to target onset), there is a main effect of Display type: participants were more likely to fixate the eventual superset target when the display provided only a structured-local interpretation, relative to displays providing only a linear-local interpretation (Table 5). The model also includes a positive Display type by Time which does not reach significance. This asymmetry between two interpretations—both of which are possible—is also reflected in response times. Participants were slower to respond by mouse-click when they were choosing a linear interpretation than when they were choosing a structured interpretation (Fig. 14).[9]

These results suggest that the linear interpretation is dispreferred relative to the structured interpretation, even when it is the only display item that satisfies the presupposition of *also*. Figure 15 shows a structured representation of (7-d). Since the

[9] The rightmost bars in Fig. 14 also show a numerical advantage for response times in the Competition condition when the structured interpretation was chosen, compared to when the linear interpretation was chosen.

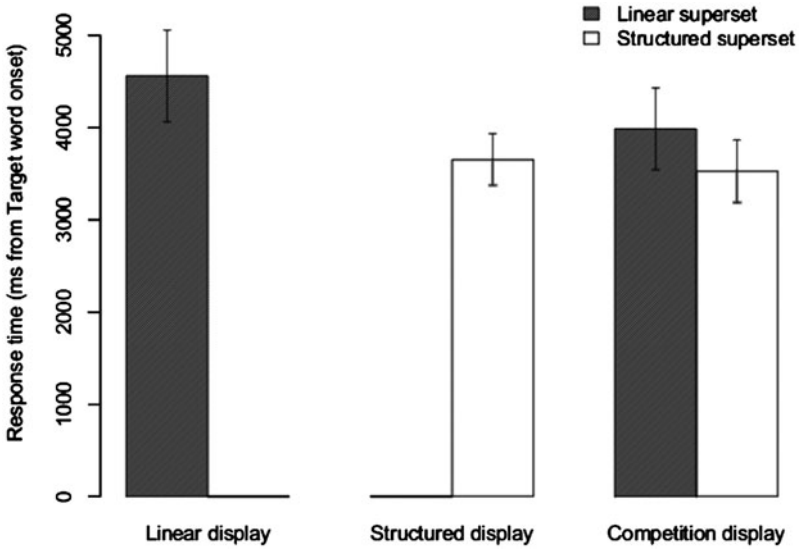


Fig. 14 Experiment 3 results, response times (mouse click), by response type and display type

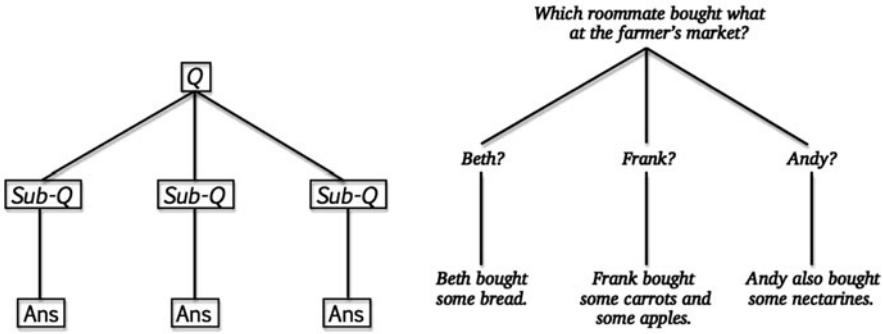


Fig. 15 Discourse tree for Experiment 3

linear and structured superset interpretations force the comprehender to interpret the immediately preceding sentence and the entire discourse, respectively, as the presupposed material, this amounts to a preference for satisfying the presupposition at the level of the smallest discourse unit containing the presupposition trigger, regardless of linear distance.

As in Experiments 1 and 2, these data show that comprehenders prefer to minimize dependency length in interpreting presupposed content. The timecourse information from Experiment 3 suggests that, in online discourse processing, interpretations that rely on structured representations may be more available than ones which rely on linear precedence alone.

6 Discussion

Together, the results of Experiments 1–3 suggest that, as in other domains of processing, comprehenders favor local dependencies. However, we see evidence both for locality defined by linear distance, and locality defined over hierarchical discourse structures. I leave for further investigation the question of how these constraints play off against each other given factors such as whether multiple viable interpretations are available, and whether the response measure is online.

Although the preceding discussion has characterized the salient interpretations as differing in terms of whether they are based on linear or structured discourse representations, the current data still allow for explanations that do not make reference to structured discourse representations. I outline one such account here, in concluding the discussion of Experiment 3.

According to a view of sentence processing as cue-based memory retrieval Lewis and Vasishth (2005); Lewis et al. (2006), each word triggers dependency formation with preceding material in the sentence. This prior material is retrieved on the basis of associative cues, where retrieval cues of the current word partially match the features of the material to be retrieved. For instance, a verb in a relative clause will trigger the retrieval of its argument, forming a syntactic dependency between the elements underlined in (9).

(9) Sameer bought *the book* that Justin *recommended*.

Such an account might be able to account for the interpretive biases observed in Experiments 1-3, if extended to discourse-level dependencies. Recall that there was evidence for both an interpretation that satisfied the presupposition of *also* using the closest available material, as well as an interpretation that satisfied the presupposition with all the material in the prior discourse, going back to a local discourse topic. While the above discussion characterized the latter interpretation as a structure-sensitive one, it might also be favored by associative cue-based retrieval. The example discourse from Experiment 2 is repeated below as (10); the material matching the focus in grammatical function, syntactic category, and conceptual features—candidates for retrieval by a cue-based retrieval mechanism—is underlined.

(10) a. The roommates went to the farmer' s market together.
b. Beth bought *some bread*.
c. Frank bought *some carrots*.
d. When his girlfriend is there, she always gets *some croissants*.
e. Andy also bought some NECTARINES.

The structured-local interpretation from Experiments 1–3 represents one where all material (in the search window) is retrieved based on featural similarity with the focused element.

The linear-local interpretation, on the other hand, is the interpretation that retrieves the minimal material from the prior discourse that is a partial featural match with the focus—if search proceeds backward from the focused element, this will be material in the immediately prior sentence (in (10)). The data from Experiments

1–3 might be interpreted as an interplay between these two pressures—to maximize feature match with the dependency trigger, and to retrieve only as little as necessary to form the dependency—rather than a difference in the discourse representations underlying the salient interpretations. In future research, studying the interpretations available in more complex discourses will help test the predictions of these contrasting explanations where they diverge.

I will close with a speculative note about cross-linguistic variation and generality. What aspects of language and how we process language should we expect to be more or less stable, irrespective of the properties of particular languages? And on the other hand, what should we expect to vary as a function of language-specific properties? The stable things will likely be things that we may think of as linguistic because they are inextricably tied to linguistic competence and performance, but which are not in and of themselves linguistic—for instance, language is affected by considerations like likelihood of communicative success, communicative efficiency (Aylett and Turk 2004; Levy and Jaeger 2007; Jaeger 2010); we plan and interpret language based on our estimations of common ground and perspective information (Clark 1992, 1996; Hanna et al. 2003), and the goals and intentions of our interlocutors (Grosz and Sidner 1986; Ballard and Hayhoe 2009); inferences we make about the knowledge states of our interlocutors then combines with features of the utterance context to determine what is salient (Brown-Schmidt et al. 2008, Heller et al. 2008; Arnold 2010). Modulo cultural differences that may influence the shape these basic processes occur, the influence they exert on language use should be stable.

On the other hand, what *should* we expect to vary as a function of language-specific properties? Take locality preferences as an example. The characterization of dependency length minimization in e.g. Hawkins (1994) suggests that it should be considered to have broad, non-language-specific application. But it has been shown that certain languages show a locality bias while others show the opposite — what looks like an anti-locality bias (Vasishth and Lewis 2006). If we assume there is a very general bias toward minimizing working memory load which applies to a wide range of cognitive processes including language processing, then language-specific properties like head directionality might systematically predict differences in local/anti-local dependency preference—a proposal along these lines is made by Vasishth and Lewis (2006). Another point to keep in mind moving forward is precisely how general processes, such as those underlying intention recognition or the integration of sensory inputs, makes contact with aspects of language use and competence that seem truly language-internal, such as mental representations of sentence or discourse structure. Possibly, some of the apparently truly language-internal elements will turn out to receive more parsimonious explanations as instances of broader phenomena; if so, it will be even more interesting to see what things we cannot explain by invoking general-purpose cognitive processes. The research presented here represents an effort to understand some empirical observations about a linguistic phenomenon from this perspective.

References

Arnold, J. E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4 (4): 187–203.

Arnold, J. E., and Z. M. Griffin. 2007. The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language* 56:521–536.

Asher, N., and A. Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.

Atlas, J. D. 1991 Topic/comment, presupposition, logical form and focus stress implicatures: The case of focal particles *only* and *also*. *Journal of Semantics* 8:127–147.

Aylett, M., and A. Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech* 47 (1): 31–56.

Ballard, D., and M. Hayhoe. 2009. Modeling the role of task in the control of gaze. *Visual Cognition* 17:1185–1204.

Barr, D. J. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* 59:457–474.

de Beaugrande, R., and B. N. Colby. 1979. Narrative models of action and interaction. *Cognitive Science* 3:43–66.

van Berkum J. J., C. M. Brown, and P. Hagoort. 1999. Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language* 41:147–182.

Bicknell, K., and H. Rohde. 2009. Dynamic integration of pragmatic expectations and real-world event knowledge in syntactic ambiguity resolution. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, pages 1216–1221.

Brown-Schmidt, S., C. Gunlogson, and M. K. Tanenhaus. 2008. Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition* 107:1122–1134.

Büring, D. 2003. On d-trees, beans, and b-accents. *Linguistics and Philosophy* 26:511–545.

Clark, H.H. 1992. *Arenas of language use*. Chicago: University of Chicago Press.

Clark, H.H. 1996. *Using language*. Cambridge: Cambridge University Press.

Cooper, R., and J. Ginzburg. 2002. *Using dependent record types in clarification ellipsis*, volume Proceedings of the sixth workshop on the semantics and pragmatics of dialogue, pages 45–52. Edinburgh University.

Farkas, D., and K. Bruce. 2009 On reacting to assertions and polar questions. *Journal of Semantics* 27:81–118.

Frank, A., A. P. Salverda, T. F. Jaeger, and M. K. Tanenhaus. 2009 Analyzing multinomial time-series with "state" dependencies. Poster presented at the 22nd CUNY Sentence Processing Conference, Davis, CA.

Frazier, L., and C. Clifton, Jr. 1996. *Construal*. Cambridge: MIT Press.

Frazier, L., and C. Clifton Jr. 2006. Ellipsis and discourse coherence. *Linguistics and Philosophy* 29 (3): 316–346.

Gee, J. P., and F. Grosjean. 1984. Empirical evidence for narrative structure. *Cognitive Science* 8:59–85.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.

Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In *Image, language, brain,* ed. Y. Miyashita, A. Marantz, and W. O'Neill, 95–126. Cambridge: MIT Press.

Ginzburg, J. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.

Ginzburg, J., and I. A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogative constructions*. CSLI.

Glenberg, A. M., A. C. Wilkinson, and W. Epstein. 1982. The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition* 10:597–602.
Graesser, A. C., and C. L. McMahen. 1993. The influence of connectives on sentence comprehension. *Journal of Memory and Language* 33:128–147.
Graesser, A. C., M. Singer, and T. Trabasso 1994. Constructing inferences during narrative text comprehension. *Psychological Review* 101:371–395.
Groenendijk, J., and M. Stokhof. 1984 *Studies in the semantics of questions and the pragmatics of answers*. PhD thesis, University of Amsterdam, The Netherlands.
Grosz, B., and C. Sidner. 1986. Attention, intention, and the structure of discourse. *Computational Linguistics* 12 (3): 175–204.
Grosz, B., A. Joshi, and S. Weinstein. 1995. A framework for modeling the local coherence of discourse. *Computational Linguistics* 2 (21): 203–225.
Hanna, J. E., M. K. Tanenhaus, and J. C. Trueswell. 2003. The eects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language* 49: 43–61.
Hawkins., J.A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
Hawkins, J.A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
Heim, I. 1982 *The semantics of definite and indefinite noun phrases*. PhD thesis, University of Massachusetts, Amherst.
Heller, D., D. Grodner, and M. K. Tanenhaus. 2008. The role of perspective in identifying domains of reference. *Cognition* 108 (831–836): 831–836.
Hemforth, B., L. Konieczny, and C. Scheepers. 2000. Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In *Architectures and mechanisms for language processing,* ed. M. Crocker and M. Pickering, and C. Clifton Jr, 259–282. Cambridge: Cambridge University Press.
Hintikka, J. 1976. The semantics of questions and the questions of semantics. *Acta Philosophica Fennica* 28(4). Amsterdam: North-Holland Publishing Company.
Hobbs, J. R. 1979. Coherence and coreference. *Cogntive Science* 3:67–90.
Horn, L. 1969. A presuppositional analysis of *only* and *even*. *Proceedings of CLS 5*, pages 97–108, Columbus, OH.
Jaeger, T. F. 2008. Categorical Data Analysis. Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language* 59:434–446.
Jaeger, T. F. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61 (1): 23–62.
Kamp, H., and U. Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer
Karttunen, L. and S. Peters. 1979. Conventional implicature. In *Syntax and semantics, volume 11: Presupposition*, ed. C.-K. Oh and D. A. Dinneen, 1–56. New York: Academic Press.
Kehler, A. 1995. *Interpreting cohesive forms inthe context of discourse inference*. PhD thesis, Harvard University.
Kehler, A. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy* 23 (6): 533–575.
Kehler A. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
Kehler, A., L. Kertz, H. Rohde, and J. L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics: Special Issue on Processing Meaning* 25:1–44.
Kim., C.S. 2012. *Generating alternatives: Interpreting focus in discourse*. PhD thesis, University of Rochester.
Kim, C. S., and J. T. Runner. 2009. Discourse structure and parallelism in VP ellipsis. *Proceedings of Semantics and Linguistic Theory*, Columbus, OH.
King, J., and M. A. Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30:580–602.
Kintsch, W., and T. van Dijk. 1978. Towards a model of text comprehension and production. *Psychological Review* 85:363–394.

Knott, A., and T. J. Sanders. 1998. The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics* 30:135–175.

van Kuppevelt J. 1995. Discourse structure, topicality and questioning. *Journal of Linguistics* 19:393–443

Levin, J. A., and J. A. Moore. 1977. Dialogue-games: Metacommunication structures for natural language interaction. *Cognitive Science* 1 (4): 395–420.

Levy, R., and T. F. Jaeger. 2007. Speakers optimize information density through syntactic reduction. *Proceedings of the twentieth annual conference on neural information processing systems*.

Lewis, D. K. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8:339–359.

Lewis, R. L., and S. Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29:375–419.

Lewis, R. L., S. Vasishth, and J. A. Van Dyke. 2006. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science* 10 (10): 447–454.

Mann, W. C., and S. A. Thompson. 1988. Rhetorical structure theory: Towards a theory of text organization. *Text* 8:243–281.

Mc-Keown, R.K. 1985. *Text generation: Using discourse strategies and focus constraints to generate natural language text*. Cambridge: Cambridge University Press.

Millis, K. K., and M. A. Just. 1994. The influence of connectives on sentence comprehension. *Journal of Memory and Language* 33:128–147.

Morrow, D. G. 1985. Prominent characters and events organize narrative understanding. *Journal of Memory and Language* 24:304–319.

Moulton K. 2006 Small antecedents: Syntax or pragmatics? In *Proceedings of NELS 37* ed. E. Elfner and M. Walkow. GLSA Publications.

Murphy, G. L., and A. M. Shapiro. 1994. Forgetting of verbatim information in discourse. *Memory and Cognition* 22:85–94.

Otero, J., and W. Kintsch. 1992. Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science* 3:229–235.

Polanyi, L., and M. van den Berg. 1996. Discourse structure and discourse contexts. In *Proceedings of the Tenth Amsterdam Colloquium*, ed. P. Dekker and M. Stokhof. Amsterdam: ILLC/Department of Philosophy.

Roberts, C.. 1996. Towards an integrated formal theory of pragmatics. In *OSU Working Papers in Linguistics*, ed. Jae Hak Yoon and Andreas Kathol, vol. 49. Columbus, OH: The OSU Department of Linguistics.

Rohde. H. 2008. *Coherence-Driven Effects in Sentence and Discourse Processing*. PhD thesis, University of California San Diego.

Rohde, H., A. Kehler, and J. L. Elman. 2007. Pronoun interpretation as a side effect of discourse coherence. *Proceedings of the 29th Annual Conference of the Cognitive Science Society,* pages 617–622.

Rohde, H., R. Levy, and A. Kehler. 2011. Anticipating explanations in relative clause processing. *Cognition* 118:339–358.

Roland, D., C. O'Meara, H. Yun, and G. Mauner. 2008. Discourse and object relative clauses: The effect of topic versus mention. Paper presented at the CUNY Conference on Human Sentence Processing, University of North Carolina at Chapel Hill.

Rooth, M. 1985. *Association with Focus*. PhD thesis, University of Massachusetts at Amherst,

Sanders, T. J., W. P. M. Spooren, and L. G. M. Noordman 1992. Toward a taxonomy of coherence relations. *Discourse Processes* 15:1–35.

Sanders, T. J., W. P. M. Spooren, and L. G. M. Noordman. 1993. Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics* 4 (2): 93–133.

van der Sandt R. A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9 (4): 333–377.

Schank, R. C., G. C. Collins, E. Davis, P. N. Johnson, S. Lytinen, and B. J. Reiser. 1982. What's the point? *Cognitive Science* 6 (3): 255–275.

Schwarz, F. 2007. Processing presupposed content. *Journal of Semantics* 24 (4): 373–416.

Simner, J., and M. Pickering. 2005. Planning causes and consequences in discourse. *Journal of Memory and Language* 52:226–239.

Stevenson, R., A. Knott, J. Oberlander, and S. McDonald. 2000. Interpreting pronouns and connectives: Interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes* 15:225–262.

Tanenhaus, M. K., M. J. Spivey-Knowlton, M.K. Eberhard, and J. C. Sedivy. 1995. Intergration of visual and linguistic information in spoken language comprehension. *Science* 268:1632–1634.

Temperley, D. 2007. Minimization of dependency length in written english. *Cognition* 105:300–333.

Traxler M. J., M. J. Pickering, and C. Clifton Jr. 1998. Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language* 39:558–592.

Vasishth S., and R. Lewis. 2006. Argumrnt-head distance and processing complexity: Explaining both locality and antilocality effects. *Language* 82:767–794.

Webber, B., and A. Joshi. 1998. Anchoring a lexicalized tree-adjoining grammar for discourse. *Proceedings of the ACL/COLING workshop on discourse relations and discourse markers*, Montreal.

Wolf, F., and E. Gibson. 2005. Representing discourse coherence: A corpus-based study. *Computational Linguistics* 31 (2): 249–288, ISSN 0891-2017. 10.1162/0891201054223977. http://dx.doi.org/10.1162/0891201054223977.

A Cross-Linguistic Study of the Non-at-issueness of Exhaustive Inferences

Emilie Destruel, Daniel Velleman, Edgar Onea, Dylan Bumford, Jingyang Xue and David Beaver

Abstract Several constructions have been noted to associate with an exhaustive inference, notably the English *it*-cleft, the French *c'est*-cleft, the preverbal focus in Hungarian and the German *es*-cleft. This inference has long been recognized to differ from exhaustiveness associated with exclusives like English *only*. While previous literature has attempted to capture this difference by debating whether the exhaustiveness of clefts is semantic or a pragmatic phenomenon, recent studies such as (Velleman et al. 2012, Proceedings of Semantics and Linguistics Theory (SALT) 22, pages 441–460) supplement the debate by proposing that the notion of *at-issueness* is the culprit of those differences. In light of this notion, this paper reconsiders the results from previous experimental data on Hungarian and German (Onea and Beaver 2011, Proceedings of Semantics and Linguistic Theory (SALT) 19, pages 342–359; Xue and Onea 2011, Proceedings of the ESSLLI 2011 Workshop on Projective Meaning, Ljubljana, Slovenia) and presents new data on English and French, showing that the "Yes, but" test used in these four languages to diagnose the source

E. Destruel (✉)
Department of French and Italian & FLARE, University of Iowa,
111 Phillips Hall, Iowa City, IA, 52242, USA
e-mail: e-destruel-johnson@uiowa.edu

D. Velleman · D. Beaver
Department of Linguistics, University of Texas at Austin,
CLA 4.304, Mailcode B5100, Austin, TX, 78712, USA
e-mail: dvelleman@gmail.com

D. Beaver
e-mail: dib@mail.utexas.edu

E. Onea · J. Xue
Courant Research Centre "Text structures", University of Góttingen,
Nikolausberger Weg 23, Góttingen D-37073, Germany
e-mail: edgar.onea@zentr.uni-goettingen.de

J. Xue
e-mail: jingyang.xue@zentr.uni-goettingen.de

D. Bumford
Department of Linguistics, New York University,
10 Washington Place, 10003 New York, NY, USA
e-mail: bumford@nyu.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_6,

of the exhaustive inference (semantics vs. pragmatics), in fact diagnoses its status (at-issue vs. non-at-issue). We conclude that the exhaustiveness associated with clefts and cleft-like constructions is not at-issue, or in other words, exhaustiveness it is not the main point of the utterance.

Keywords Exhaustivity · Cleft structure · Focus · (Non)-at-issueness · Information Structure

1 Introduction

One of the most important jobs of natural language utterances is to provide information, or, put slightly differently, to provide answers to questions. Questions may be overt, as in dialogue situations, or implicit.

Answers, depending on the circumstances, may be *maximal* or not, as the speaker may or may not have the possibility or desire to give a complete answer to a question. We take an answer p to be *maximal* if no true answer to the question under discussion (hereafter QUD) is strictly stronger than p. In other words, maximal answers are exhaustive. Cooperative speakers will often try to give answers that they believe to be maximal. But then, sometimes a speaker will fall short of this. There are several reasons why a speaker might do so. Perhaps she has limited information—and so give what she knows to be a partial answer, or give an answer without being certain whether it is maximal or not. Perhaps she is not being fully cooperative. Perhaps, given the interlocutor's goals, a partial answer is just as useful as a maximal one.

Given this, one very natural thing that a speaker may want to do is to *indicate* when he believes that his utterance represents a maximal answer to the QUD. In doing so, he signals that the line of inquiry represented by the QUD has been fully explored and can now be closed.

To give a concrete example, consider the question in (1). It may be answered as in (1-a), which may or may not be a maximal answer. Depending on the context, a continuation as in (2) may be justified. This shows that, at least in principle, an answer like (1-a) does not mark the termination of the inquiry started by the question in (1). By contrast, if the speaker chooses to answer (1), a continuation with (2) is completely nonsensical, because the exclusive *only* marks that the answer is maximal, exhaustive; cf. Beaver and Clark (2008). By giving an exhaustive answer, the speaker of (1-b) settles the question for his part, and thereby terminates the line of inquiry and gives rise to the inference in (3).

(1) Who smiled?
   a. MARY smiled.
   b. Only MARY smiled.
(2) Who else smiled?
(3) *Noone else than Mary smiled.*

Clefts seem to be very similar to exclusives in this respect, as illustrated in the examples below for the English *it*-cleft (4-b), the French *c'est*-cleft (5-b), the preverbal focus in Hungarian (6-b) and the German *es*-cleft (7-b). Although the continuation questions in (c) are not as infelicitous as after (1), they are odd precisely because the hearer deliberately ignores the exhaustiveness inference triggered by the construction in the (b) sentences, implying that the speaker of the (b) sentences is not being truthful/cooperative.

(4) a. Who smiled?
b. It is Mary that smiled.
c. # Who else smiled?

(5) a. Qui a ri?
b. C' est Marie qui a ri.
c. # Qui d' autre a ri?

(6) a. Ki mosolygott?
b. MARI mosolygott.
c. # És ki más mosolygott?

(7) a. Wer hat gelächelt?
b. Es war Maria, die gelächelt hat.
c. # Und wer hat noch gelächelt?

Velleman et al. (2012) term constructions that mark utterances as maximally answering the QUD, such as exclusives, *Inquiry-Terminating* constructions (IT-con-struc-tions). Accordint to Velleman et al. (2012), there are two kinds of IT-con-struc-tions: those which mark the maximality of the answer as maximal as part of their at-issue meaning contribution and those which mark the maximality of the answer as part of their not-at issue meaning contribution. The notion of at-issueness is further explained in Roberts et al. (2009), Simons et al. (2011) and Tonhauser (2012). Velleman et al. (2012) argue that exclusives are at-issue IT-constructions while it-clefts are not-at-issue IT-constructions.

One important difference between clefts and exclusives regarding exhaustiveness is then the status of the exhaustiveness inference being at-issue in the case of exclusives and not-at-issue in the case of clefts. Crucially, in both cases, the inferences are conventional, i.e. semantic. More precisely the essential component leading to exhaustiveness is part of the assertion for *only* and part of the presupposition for clefts. Yet, while in the recent theoretical literature there is a tendency to analyze these inferences as semantic (cf. Percus (1997) or Büring and Križ (2013) contra Horn (1981)), experimental evidence has mostly been taken to suggest that exhaustiveness is pragmatic, cf. Onea and Beaver (2011); Drenhaus et al. (2011) and Byram Washburn et al. (2013).

The main goal of this paper is to support an analysis of exhaustiveness inferences of clefts in terms of not-at-issueness Velleman et al. (2012) by giving additional experimental evidence by re-interpreting existing data on Hungarian from Onea and Beaver (2011) using insights from data partly reported in Xue and Onea (2011),

and by providing new data on English and French. In essence, we will claim that the "Yes, but" test used in Onea and Beaver (2011) only shows a lack of at-issue exhaustiveness and cannot be taken as evidence for a pragmatic exhaustification.

The paper is structured as follow: The first section reviews the debate on the source of exhaustiveness—semantics vs. pragmatics—and discusses existing experimental data from the "Yes, but" test applied to Hungarian (Onea and Beaver 2011). Section 3 examines the shift in perspective from the *source* to the *status* of exhaustiveness by (i) re-interpreting the results in Onea and Beaver (2011) and (ii) discussing existing data from the "Yes, but" test for German from Xue and Onea (2011). Section 4 discusses previous related work on implicatures and their cancellation. Section 5 presents new data collected for English and French and shows that these two languages pattern similarly to Hungarian and German. Finally, Sect. 6 concludes that across these four different languages, and maybe more generally, the exhaustiveness in clefts and related constructions must be understood as being not-at-issue.

2 Semantic vs. Pragmatic Exhaustiveness

2.1 Theoretical Discussion

In the literature, it is a well-known fact that there is an exhaustiveness inference associated with *it*-clefts, hence (8-a) licenses the inference (8-b). But at the same time, it is a hotly disputed question as to whether this inference is a semantic one, i.e. a necessary one, or a pragmatic one, in particular an implicature.

(8) a. It is Michael who defeated Jackson.
b. Nobody other than Michael had defeated Jackson.

The semantic position has been advocated by a large number of scholars, though, there have been significant differences regarding the status of the exhaustiveness inference. Essentially, there are three types of prominent approaches. According to the first one, exhaustiveness of clefts is semantic in the narrowest possible sense, i.e. it is part of the proffered content, of the truth conditional contribution of the sentence. For instance, Bolinger (1972) argues that clefts express an equality between two predicates, as suggested in (9-a) for the cleft in (8-a). Atlas and Levinson (1981) take a similar line of attack suggesting an analysis like (9-b). Finally, Szabolcsi (1981) and Kiss (1998) argue that pre-verbal focus in Hungarian, which shares most decisive properties with English *it*-clefts[1], should be exhaustified using an operator like the one in (9-c), which would take Michael as the first argument and the property of having defeated Jackson as the second argument.

[1] Hungarian pre-verbal focus is generally translated as a cleft as observed in Kiss (1998). Eg.

(i) MIHÁLY győzte le Jánost.
Michael defeated PRT Jackson
'It is Michael who defeated Jackson'.

(9) a. $\lambda x.x$ =**Michael** = $\lambda x.defeat$ (x,**Jackson**)
b. The group of individuals that have defeated Jackson is identical to Michael.
c. $\lambda x.\lambda P.P(x) \wedge \forall y.P(y) \rightarrow y = x$

The second type of semantic approach takes a more indirect path. The claim here is that exhaustiveness of clefts is due to the interaction of a presupposition and the asserted meaning. The most prominent proponent of this theory is Percus (1997) for *it*-clefts and Szabolcsi (1994) for pre-verbal Hungarian focus, but, we find this position also in Delin (1990), Delin and Oberlander (1995), Hedberg (2000), Reeve (2010) and some other papers. The essential element of this kind of analysis is a maximality presupposition and an identificational semantic contribution. For (8-a), this boils down to (10). While this time the effect is less direct than in the previous type of approaches, the exhaustiveness inference is not predicted to be by any means weaker or less general (except, of course, in cases in which presuppositions are blocked or filtered).

(10) a. (8-a) presupposes: There is a maximal sum individual X such that X defeated Jackson.
b. (8-a) asserts: X is Michael.

Finally, the third semantic approach is another version of the presupposition analysis, recently advocated in Büring and Križ (2013).[2] The reason why we mention it as a separate type of approach is that it is much more direct than the above. In (10), there is only a general maximality presupposition which does not in itself entail anything about the exhaustiveness inference (8-b), i.e. the maximal sum individual in (10-a) could consist of Michael, John and Brian. The exhaustiveness effect only comes about once the maximal sum-individual is equated to the pivot, in this case, Michael. Büring and Križ (2013) propose a more direct variant in which the presupposition is the exhaustiveness inference itself, as shown in (11).

(11) a. (8-a) presupposes: Michael is not a proper part of the maximal sum-individual who defeated Jackson.
b. (8-a) asserts: Michael has defeated Jackson.

As opposed to this, the pragmatic position assumes that the exhaustiveness inference is a conversational implicature. In particular Horn (1981, 2013) argues that the exhaustiveness inference is not part of the semantic content. Instead it arises as a generalized conversational implicature as a result of the following principle: Whenever a speaker employs an expression which presupposes $\exists x.P(x)$ and asserts $P(a)$, he implicates $\forall y.y \neq a \rightarrow \neg P(y)$. In Horn (1981), one of the main argument is to show differences between sentences with the exclusive *only* and *it*-clefts, as suggested in (12).[3]

[2] Technically, Velleman et al. (2012) propose a very similar analysis essentially involving the same presupposition. The conceptual gist of their argument and some predictions, however, are different.
[3] Büring and Kriz (2013) argue that this is no counter-example against their analysis, since the attitude verb *realise* will only allow the exhaustiveness presupposition of clefts to project, not interfering with the attitude verb.

(12) a. I knew that he ate a pizza, but I just realised that he only ate a pizza.
b. # I knew that he ate a pizza, but I just realised that it is a pizza he ate.

One of the interesting problems of a pragmatic approach is that exhaustiveness of clefts does not seem to be easily cancellable, as readily admitted even by Horn (1981), whence the oddity of (13). Crucially, however, Horn argues that the non-cancellability of the exhaustiveness inference may be related to the fact that the speaker has 'gone out of her way' to use an expression with an existential presupposition.[4] As opposed to this, in other cases of generalized implicatures, the speaker rather seems to use standard, economical expressions. Hence, there seems to be a manner component involved as well in the interpretation of clefts.[5]

(13) # It was a pizza, John ate, indeed, it was a pizza and a calzone.

The main dichotomy in the discussion sketched above is whether clefts are semantically or pragmatically exhaustive. A second line of discussion concerns the question how to model exhaustiveness if clefts are semantically exhaustive, a question which naturally fails to arise if Horn is right.

2.2 Experimental Discussion

A puzzling fact mainly ignored in the theoretical literature is that the exhaustiveness of clefts does not seem to be so strong as suggested. Both for Hungarian focus (Wedgwood et al. 2006) and for clefts (Horn 1981, 2013) a number of examples have been found which seem incompatible with an exhaustive interpretation. See for example the poem in (14), from (Horn 2013):

(14) As we go marching, marching unnumbered women dead,
Go crying though our singing their ancient call for bread.
Small art and love and beauty their drudging spirits knew.
Yes, **it is bread we fight for, but we fight for roses too**!

One natural consideration arising is then the following: if clefts are semantically exhaustive (regardless of the way in which this derivation is performed), one should expect that exhaustiveness effects are strong and consistent. As opposed to this, if clefts are pragmatically exhaustive, one would expect exhaustiveness effects to be significantly weaker.

Onea and Beaver (2011) have conducted a first experimental study to this extent. In particular, they have studied the way participants chose to react in situations in which the exhaustiveness inference of pre-verbal focus in Hungarian was violated given some pictorial stimuli. Given a Hungarian stimulus with pre-verbal focus

[4] Horn calls this a conventional implicature, to be precise.
[5] This argument does not seem to hold for Hungarian focus, however, since Hungarian focus seems to be a fairly economical, unmarked, standard construction.

constructed as (15) and a picture which contradicts the exhaustiveness inference, the possible answers were the ones in (16):

(15) It is John who has a hat.
Hungarian: JNOSNAK van egy kalapja.

(16) a. Yes, and Mary also has a hat.
b. Yes, but Mary also has a hat.
c. No, Mary also has a hat.

Onea and Beaver (2011) assumed that if there is a strong exhaustiveness effect associated with Hungarian focus, in particular a semantic effect, participants would consistently choose the c. answer, to mark that there is a serious violation. As opposed to this, if the exhaustiveness effect was weak or pragmatic, speakers would rather choose the a. or b. answers. The results of the experiment clearly showed that for *only*-sentences as (17), participants consistently picked the most confrontative, c. answer. As opposed to this for Hungarian focus, the number of c. answers was much lower.

(17) Only John has a hat.

From this, Onea and Beaver (2011) concluded that the exhaustiveness of Hungarian focus must be pragmatic. In particular, they proposed an analysis in which pre-verbal focus in Hungarian leads to exhaustification as a pragmatic implicature arising from the fact that the pre-verbal focus in Hungarian marks that the corresponding sentence is to be interpreted as an answer to a *wh*-question, associated with an additional pragmatic principle that answers to a question are interpreted as complete (which mostly means exhaustive). Notice that under the assumption that *wh*-questions have an existential presupposition, this analysis is very similar to the one proposed in Horn (1981).[6]

In a more recent experimental study, Washburn et al. (2013) argue that the fact that Onea and Beaver (2011) compare the exhaustiveness of Hungarian focus, which they also take to be equivalent to English clefts, with the exhaustiveness of *only* leads to a problem, since any difference found may not only be related to the fact that clefts/pre-verbal focus in Hungarian are not exhaustive but also to standard differences between presupposition and assertion. Therefore, they propose a different setup in which they compare the acceptability of the exhaustiveness violation of clefts with the violation of some other presupposition of clefts. In particular, they claim that English *it*-clefts come with a contrastiveness presupposition. This contrast is exemplified in (18).

(18) Contrastive: Jane and Tom painted furniture. Tom painted a chair. Later, Kevin remarks: 'I bet Tom painted only lamps again, didn't he?' Jane responds: 'He doesn't always paint lamps. Yesterday, it was a chair that he painted'.

[6] A similar conclusion is reached by Drenhaus et al. (2011) using an ERP experiment we do not discuss here in detail.

(19) Non-contrastive: Jane and Tom painted furniture. Tom painted a chair. Later, Kevin remarks: 'I bet Tom painted only a chair again, didn't he?' Jane responds: 'Yes. Yesterday, it was a chair that he painted.'

Byram Washburn et al. (2013) observe that the exhaustiveness violation for English *it*-clefts is hardly significant whereas the violation of the contrastiveness presupposition leads to a serious decay in acceptability. They conclude that exhaustiveness must be a Hornian implicature. Notice, however, that while the conclusion seems to support the results in Onea and Beaver (2011), the design is problematic for two reasons. It is a well-known fact that exhaustiveness is always relative to some domain. The comparison with *only* is used as a proof that domain restriction is not interfering, since for all we know, the domain restriction for *only* and for the exhaustiveness of a cleft in the same context should be exactly the same. Moreover, in the design used by Byram Washburn et al. (2013), exhaustiveness was never simply crossed with contrastiveness, hence, the lack of exhaustiveness effects can be due to any other interfering factor.

While we contend that the conclusion drawn by Byram Washburn et al. (2013) is too strong given their design, we accept their critique of Onea and Beaver (2011) that the observed difference between the Hungarian preverbal focus and the only-condition is not enough to support a radical pragmatic approach. Instead we will use the German data from Xue and Onea (2011) to propose a shift in perspective in the interpretation of their result at the same time using additional cross linguistic data showing that the pattern found there seems generally persistent for any cleft-construction.

3 A Shift in Perspective: At-Issue Versus Not-at-issue

The discussion so far has focused on the *source* of the exhaustive inference. The question has been whether this inference forms part of the conventional meaning of the *it*-cleft construction. Onea and Beaver's data was taken to support a negative answer to that question—that is, it was taken as evidence that the exhaustive inference arises instead out of some sort of pragmatic process.

But in fact, we now believe that this is a misinterpretation of the data. In this section, we will argue that the "Yes, but" task—which Onea and Beaver took to diagnose the *source* of the exhaustive inference—actually diagnoses a different property entirely: its *status* as at-issue or not-at-issue. We will make this argument on two grounds. First, we will present new experimental evidence showing that the "Yes, but" task should be reinterpreted as diagnosing status. And second, we will cite recent formal and experimental work by others which supports this reinterpretation of the task.

This reinterpretation has consequences that go beyond the language-specific results reported in Onea and Beaver (2011). First of all, similar diagnostics to the "Yes, but" task have been used in a number of other languages to address the question of whether or not a particular construction has exhaustive semantics. (For instance,

Dyakonova (2009) has argued that Russian "left-edge focus" is not semantically exhaustive, based on the fact that exhaustivity can be denied by a second speaker without using negation.) And more generally, it has consequences for our understanding of confirmation and rejection, and for our use of semantic methodologies involving confirmation and rejection as diagnostics.

Finally, we will clarify the difference between the "Yes, but" test discussed here and Grice's classic cancellation test (Grice 1967)—which has long been used as a diagnostic for the source of an inference. The arguments which we give her for reinterpreting the "Yes, but" test do not apply to the cancellation test; and in many cases, the two tests give different results. Thus, it is consistent to use the "Yes, but" test as a diagnostic for status while continuing to use the cancellation test as a diagnostic for source.

3.1 Reinterpreting the "Yes, But" Test: Evidence From German

In this experiment, the "Yes, but" task was applied to a wider range of triggers, including *es*-clefts as before—but also a number of others. The experiment was designed to distinguish between two hypotheses.

(20) *The source hypothesis*: The "Yes, but" test diagnoses the *source* of the inference being contradicted: "yes" answers indicate that it arises through pragmatic inference.

(21) *The status hypothesis*: The "Yes, but" test diagnoses the *status* of the inference being contradicted: "yes" answers indicate that it is not at issue.

In order to do this, we added a number of items for which there is widespread consensus among linguists concerning the source and status of the inference being contradicted—summarized in the table below.

(22)

	Pragmatic?	Not at issue?
Relevance implicature	Y	Y
Scalar implicature	Y	N
Appositive	N	Y
Nonrestrictive relative clause	N	Y

If the source hypothesis is correct, we predict "Yes, but" answers for the implicatures and "No" answers for the appositive and NRRCs. If the status hypothesis is correct, we predict a different pattern of replies: "Yes, but" answers for the relevance implicatures, the appositives and the NRRCs, and "No" answers for the scalar implicatures.

3.1.1 Methods

In this study, as in Onea and Beaver (2011), German-speaking participants completed a forced-choice task in which they were asked to choose the most natural of three possible conversational responses. A total of 29 German speakers, between the ages of 19 and 61, participated in our web-based experiment. As before, the speakers were presented with a sentence in German which triggered some inference p, and were asked to choose between three possible responses which denied that p—one beginning *Ja, und* "Yes, and," one beginning *Ja, aber* "Yes, but," and one beginning *Nein* "No."

Unlike the experiment in Onea and Beaver (2011), this study covered a wide range of triggers—not only clefts, but a number of other constructions which have been argued to trigger presuppositions, implicatures, or other non-entailment inferences. (Also included were filler items in which the inference to be denied was a simple entailment.) This paper only discusses results on a subset of these triggers; the remainder were presented in Xue and Onea (2011).

We are concerned here with two sets of items in particular. First, there was a set of items in which the inference to be contradicted is a conversational implicature—either a relevance implicature, as in (23), or a scalar implicature, as in (24). These implicatures are not part of the conventional truth conditions of their triggers. The sentence in (23) clearly does not conventionally entail that the bank in question is open, or that it has a working ATM, or anything else of this sort. These inferences arise based on pragmatic reasoning in certain contexts. And the sentence in (24) is standardly taken not to conventionally entail that the soup is not hot—this too is taken to be a pragmatic inference only.

(23) Relevance implicature
   a. Ein paar Schritte weiter ist gleich eine Sparkasse.
      a few steps further is just a bank
      "There' s a bank just a few steps further."
   b. Inference to be contradicted: The bank will let you carry out whatever sort of business you need.
   c. Ja, und/Ja, aber/Nein, der Geldautomat ist kaputt.
      Yes, and/Yes, but/No, the ATM is broken
      "Yes, and/Yes, but/No, the ATM is broken."

(24) Scalar implicature
   a. Die Suppe ist warm.
      the soup is warm
      "The soup is warm"
   b. Inference to be contradicted: The soup is not hot.
   c. Ja, und/Ja, aber/Nein, die Suppe ist heiß.
      yes and/yes but/no the soup is hot
      "Yes, and/Yes, but/No, the soup is hot."

But conversational implicatures *can* be at issue. In fact, the details appear to depend on the type of the implicature. It has been argued that scalar implicatures are *always* at issue: van Kuppevelt (1996) holds that scalar implicatures only arise when the scalar term is the answer to the QUD, and Zondervan (2010) gives experimental evidence that partly supports this claim. Relevance implicatures, on the other hand, may or may not be at issue. For our example (23), the QUD addressed by the (a) utterance may be where a bank is, however, the implicature addresses another question (also relevant in the discourse), namely whether one can withdraw money there (or do some other relevant business). Crucially, at the discourse state at which the (a) inference is interpreted, the latter question is not necessarily the QUD. This suffices to show that relevance implicatures need not always be at issue.

Second, there was a set of items involving appositives and nonrestrictive relative clauses. The properties of this second set of items are exactly opposite to those of the first. On the one hand, the meaning contributed by an appositive or NRRC is clearly part of the *conventional* meaning of the utterance, and not the result of a merely pragmatic process. On the other hand, the meanings of appositives and NRRCs are consistently not at issue.

(25) Appositive
   a. Paula, Peters Schwester, hat ein Kind bekommen.
      Paula Peter' s sister has a child gotten
      "Paula, Peter' s sister, had a baby"
   b. Inference: Paula is Peter' s sister.
   c. Ja, und/Ja, aber/Nein, Paula ist gar nicht Peters Schwester.
      yes and/yes but/no Paula is really not Peter' s sister
      "Yes, and/Yes, but/No, Paula is not really

(26) Nonrestrictive relative clause
   a. Paul, der sehr fleißig ist, sitzt den ganzen Tag am Schreibtisch.
      Paul that very diligent is sits the whole day at.the desk
      "Paul, who is very diligent, sits at his desk all day long."
   b. Inference: Paul is very diligent.
   c. Ja, und/Ja, aber/Nein, Paul ist gar nicht fleißig.
      yes and/yes but/no Paul is really not fleißig.
      "Yes, and/Yes, but/No, Paul is not all that

Finally, there were a number of other triggers which we will not discuss closely here—including items containing factive verbs, items containing "strong" presupposition triggers such as *auch* 'also,' and filler items in which the inference to be contradicted was a simple at-issue entailment. (Results from some of these items are discussed in Xue and Onea 2011.)

The total number of items, including fillers, was 39. As in previous experiments, participants were asked to choose which of the three variant replies was "most natural"—the variant beginning *ja, und* 'yes, and,' the one beginning *ja, aber* 'yes, but,' or the one beginning *nein* 'no.'

3.1.2 Results

As discussed above, there are two hypotheses of interest here. The first hypothesis is that the "Yes, but" test diagnoses the *source* of an inference: semantic or pragmatic. If this is true, it predicts that semantic inferences (consisting here of appositives and NRRCs) will be contradicted with "No" replies, and that pragmatic ones (here, conversational implicatures) will be contradicted with "Yes" replies. The second hypothesis is that the "Yes, but" test diagnoses the *status* of an inference: at-issue or not-at-issue. If this is true, it predicts an outcome which is almost (but not quite) the reverse of this: appositives and nonrestrictive relatives will consistently receive "Yes" replies, while at least some conversational implicatures will receive "No" replies.

Our results were inconsistent with the source hypothesis, and consistent with the status hypothesis. Across the items involving appositives and NRRCs, "Yes, but" consistently receives an overwhelming majority of the votes. The same is true for the items involving relevance implicatures; while in the items involving scalar implicatures, the majority response is "No."

	Relevance (%)	Scalar (%)	NRRC (%)	Appositive (%)
Yes, and	1	2	0	0
Yes, but	97	6	89	90
No	1	92	11	10

This is plainly incompatible with the predictions of the source hypothesis. It is compatible with the predictions of the status hypothesis, so long as we assume that the relevance implicatures in this experiment were interpreted by participants as being not at issue. As we discussed above, conversational implicatures may be either at issue or not at issue, depending on the context.

4 Related Work

One important antecedent for this interpretation is found in work on "attachability"—a concept most recently used by Jayez but originating in work by Ducrot. Ducrot (1972) noted that in general, subsequent discourse moves cannot be "attached" to a presupposition. Jayez (2005, 2010) and Jayez and Tovena (2008) note that conventional implicatures are subject to the same restriction, and cannot serve as attachment sites either.

In this line of work, "attachment" is taken to include many different discourse relations—including, for instance, *Cause*. Consider the biclausal sentence in (27).

(27) a. Unfortunately, Paul has failed his exam . . .
b. . . . because he cannot register for next term.

The clause in (27-a) has two implications—one of which is standardly analyzed as an entailment, and the other of which is standardly analyzed as a CI.

(28) Paul has failed his exam.
(entailment)

(29) The fact that Paul has failed his exam is unfortunate.
(conventional implicature)

But the "because" clause in (27-b) can only attach to the entailment, even though attaching it to the CI would result in an intuitively more plausible meaning.

(30) Paul has failed his exam because he cannot register for next term.

(31) The fact that Paul has failed his exam is unfortunate because [as a result] he cannot register for next term.

In Jayez (2010), experimental evidence is given that French clauses subordinated by *donc* 'so,' *alors* 'so,' *parce que* 'because' or *puisque* 'since' can felicitously attach to at-issue implications, but cannot felicitously attach to not-at-issue implications. Jayez ultimately rejects the idea that at-issueness is the crucial factor, based on objections to the QUD approach as a whole, and to the usefulness of the concept of at-issueness in particular. We believe that his objections can be overcome;[7] but this is not the place to address them and so for now, let us simply note that Jayez's data is consistent on the face of things with our conclusions in this section (and for that matter, that our data here are consistent with his conclusions).

As Jayez himself points out, rejection and confirmation—that is, "yes" and "no" answers—count as a kind of discourse relation. If discourse moves in general attach to at-issue content, then rejection and confirmation in particular should be expected to do so. Our results in this section confirm that that is the case.

(So if Jayez's "Because" test and our "Yes, but" test diagnose the same property, is there any reason to prefer one to the other? Our experience in pilot experiments has been that the "Yes, but" test is less difficult for participants—possibly because it uses syntactically simpler sentences, while the "Because" test requires an additional subordinate clause to be added to every item. Jayez's particular implementation of the "Because" test also depends on speakers having consistent assumptions about what sorts of cause-and-effect relations are plausible. This introduces another possible source of noise which is not present in the "Yes, but" test, though it is possible that this problem can be overcome. For these reasons, we have continued using the methodology from Onea and Beaver rather than adopting the one from Jayez (2010). Still, setting aside practical issues—such as the participants' attention spans and their assumptions about plausibility—we predict that the two methodologies will be interchangeable.)

[7] See Onea (2013) for one framework in which these objections are addressed.

Another point of conceptual support for this interpretation of the "Yes, but" test comes from the connection between at-issueness and projection. It has been observed that in general, not-at-issue inferences project (Simons et al. 2011). Xue and Onea (2011) have shown that responses to the "Yes, but" test are correlated with responses to a standard test of projection. Inferences which are more likely to receive a "Yes" response are also more likely to be treated as projecting, and vice versa. This lends additional plausibility to the idea that these "Yes" responses are indicative of not-at-issue status.

4.1 Comparing the Cancellability Test

So we have seen that the "Yes, but" test does not diagnose the source of an inference, but should be reinterpreted as a diagnostic of status. At this point, one might wonder whether the *cancellability test*—with a much longer history of use as a diagnostic for source—should be reinterpreted as a diagnostic of status as well. Here we argue that it should not be. If a careful methodological distinction is maintained between the two tests, then there is a corresponding difference is their results. This shows that the two tests diagnose different properties. If we take the "Yes, but" test as a diagnostic of status, the cancellability test must be diagnosing something else—plausibly source, as has long been believed.[8]

When Grice drew a distinction between "what is said" and "what is implicated" (Grice 1967) he argued that all conversational implicatures are *cancellable*. This has become a standard test for the source of an inference: inferences are taken to be cancellable if they arise through pragmatic reasoning, and to be uncancellable if they are part of the conventional semantics of an utterance.[9] The classic example here is the conversational implicature in (32), which can be explicitly cancelled by the speaker as in (33).

(32) [In a letter of recommendation] He has excellent penmanship.
→ He has no qualifications beyond his penmanship.

(33) He has excellent penmanship—though of course he has many further qualifications.

It is important to distinguish genuine cancellation, in which the speaker makes it clear that he never intended the inference to arise, from *self-correction*, in which the speaker mistakenly says something that would license the inference and then takes it back.

[8] We should mention, though, that this interpretation of the test is not entirely uncontroversial. In particular, Horn (1981) argues that there are pragmatic implicatures which are nevertheless uncancellable.

[9] Some researchers have argued that there are types of pragmatic inference other than implicatures. In general, these researchers have maintained that *all* pragmatic inferences are cancellable: see for instance Carston (1988) and Recanati (1989).

(34) He has excellent penmanship. (Wait a minute! I' m thinking of the wrong guy!) No, actually, he has a lot of qualifications.

(35) He has excellent penmanship. (Wait a minute! I'm looking at the wrong writing sample!) No, actually, his penmanship is terrible.

Here is one way to distinguish cancellation from self-correction: cancellation can be done *in advance*, by issuing a disclaimer against the unwanted interpretation. Self-correction cannot be done in advance.

(36) Please don't take this to imply that he has no other qualifications. But you have to admit, he really does have excellent penmanship.

(37) Please don't take this to imply that his penmanship isn' t terrible. But you have to admit, he really does have excellent penmanship.

And for that matter, it is important to distinguish cancellation from correction by *another* speaker—which is what we find in the "Yes, but" test.

(38) A: He has excellent penmanship.
B: Yes, and/Yes, but/?? No, he has a lot of other qualifications.

In this particular example, the cancellation test and the "Yes, but" test give parallel results. That is, the relevant inference can be cancelled by the speaker who uttered it *and* denied without negation by another speaker. But in general they often give different results. For instance, scalar implicatures are cancellable—but, as the results above showed, they cannot generally be denied without negation. And the meaning contributed by appositives and NRRCs, while not cancellable, can be denied without negation.

There are two conclusions we can draw from this. The first is a general point: not all ways of denying an inference are the same. Genuine cancellation, self-correction and other-correction all have a superficially similar structure—first p is asserted, then q is denied, where p can be taken to imply q. But they are not all permitted under the same conditions. Cancellation is assumed to be sensitive to the source of an inference, and we see no reason to contradict that assumption; but, as we have shown in this paper, the form of a self-correction or an other-correction is determined by the inference's status.

The second point is more narrowly methodological: it is important, in applying the cancellation and "Yes, but" tests, to maintain a clear distinction between them. It is worth being especially careful with ambiguous examples like the following:

(39) It's John who laughed. Oh, and other people laughed too.

Here it may not be immediately clear whether the second sentence represents a clarification of the first (in which case it counts as a case of genuine cancellation), or whether it should be taken as a self-correction. As a result, it may not be clear what the felicity of (39) is telling us. To make it clear, we need to use an unambiguous test. On the one hand, if we are interested in the status of the inference, we can use the "Yes, but" test—an unambiguous case of other-correction.

(40) A: It's John who laughed.
B: Yes, and other people laughed too.

Or on the other hand, if we are interested in the source of the inference, we might use an example which unambiguously involves genuine cancellation.

(41) # Please don't take this as implying that nobody else laughed. But it's John who laughed.

In hindsight, these unambiguous examples show that (39) has to be interpreted as self-correction, for genuine cancellation here is not possible. But until we have considered the unambiguous examples, the mere fact that (39) is felicitous does not tell us anything.

4.2 Onea and Beaver Revisited: Cleft Exhaustivity is Not at Issue

Everything we have seen in this section suggests a reinterpretation of the data in Onea and Beaver (2011). They concluded that neither Hungarian preverbal focus nor German *es*-clefting has exhaustivity as part of its conventional meaning. This conclusion was based on the assumption—shared with other researchers—that the "Yes, but" test could be used to diagnose the source of an inference. We have now seen that the "Yes, but" test has nothing to do with source; rather, it diagnoses the status of the inference, whether at-issue or not-at-issue. The correct interpretation of Onea and Beaver's data, then, is that these constructions do not have exhaustivity as part of their *at-issue* meaning.

On the other hand, we've argued that the classic cancellation test can still consistently be used as a diagnostic for source. Now, it is not completely uncontroversial that this is the correct interpretation. In particular, Horn (1981) has argued that some pragmatic inferences are nevertheless uncancellable. We disagree; but we will not settle that question in this paper. All we are saying for now is that the standard interpretation of the cancellation test, on which is does diagnose source, is *consistent* with the evidence presented in this section.

5 New Evidence From Two Other Languages: English and French

We now turn to presenting experimental data for two other languages–English and French–that support an analysis of exhaustiveness in cleft constructions as being not-at-issue.

The basic idea behind the design of the experiment, which follows the methodology in Onea and Beaver (2011), is to assume that the most natural way for a speaker to contradict an at-issue inference triggered by some utterance is to use the explicit contradiction marker "No". On the other hand, if an inference is not at-issue, we

assume that it will be most naturally contradicted in some other, weaker way, such as using sequences like "Yes, but" or even the much politer "Yes, and". Of course, depending on circumstances "Yes, but" can be a more polite version of "No", "No" can even be used to contradict inferences that are not even suggested by the utterance, and there may be other strategies of contradicting inferences that we did not consider. However, for now, we assume that in most cases, "No" is predominantly used to contradict at-issue inferences and "Yes, but" and "Yes, and" are used to contradict inferences that are not at-issue. For illustration, consider (42). We assume that the most natural contradiction for (42-a) is (43-a), for (42-b) the most natural contradiction is (44-c), and for (42-c) the most natural contradiction is (45-c).

(42) Mary smiled again.
   a. At-issue meaning: Mary smiled.
   b. Not at-issue meaning: Mary had smiled before.
   c. Not necessarily triggered: Mary is happy.

(43) Mary smiled again.
   a. No, Mary didn' t smile.
   b. Yes, and Mary didn' t smile.
   c. Yes, but Mary didn' t smile.

(44) Mary smiled again.
   a. No, Mary didn' t smile before.
   b. Yes, and Mary didn' t smile before.
   c. Yes, but Mary didn' t smile before.

(45) No, Mary isn' t happy.
   a. No, Mary didn' t smile before.
   b. Yes, and Mary isn' t happy.
   c. Yes, but Mary isn' t happy.

Building on the intuitive appeal of the assumption and the results from (Onea and Beaver 2011; Xue and Onea 2011), we can now experimentally test the idea that, as opposed to exclusives, clefts contribute non-at-issue exhaustiveness. In other words, whatever clefts do, conveying exhaustiveness is not the main point for their usage. In the experimental design, if our hypothesis is correct, the exhaustiveness triggered by *only* will be contradicted using "No" type of answers and the exhaustiveness triggered by clefts will rather be contradicted with "Yes, but" or even "Yes, and" kinds of answers.

5.1 English It-Clefts

Forty five native English speakers participated in the experimental task, a forced-choice task presented on-line using WebExp.[10] Participants were instructed that on

[10] http://www.webexp.info/.

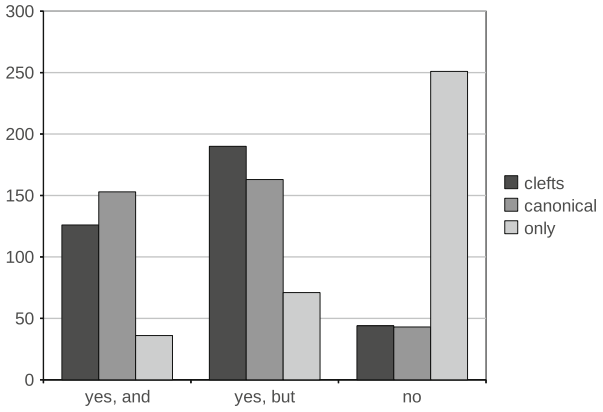


Fig. 1 English results

each slide they would hear a short discourse between two friends Jason and Sarah, and then would be asked to continue or otherwise update the conversation by selecting one of three possible continuation sentences.

Specifically, on each slide, participants clicked a button to hear an audio clip of Sarah posing a question to Jason, such as the one in (46). The question appeared in two conditions: i) question about the grammatical subject of the sentence and ii) question about the grammatical object. Then, participants clicked a second button to hear Jason's response, which appeared in one of the three forms: i) a cleft as in (46-b), ii) an exclusive as in (46-c), or iii) a canonical answer as in (46-a). The task was then for them to choose which of the following continuations in (46-c-i)-(46-c-iii) they considered the most natural way to indicate that Jason had only given a partial (or incorrect) answer to Sarah's question:

(46) What did Phillip buy his sister?
  a. Philipp bought his sister A NECKLACE.
  b. It was a necklace that Phillip bought his sister.
  c. Phillip only bought his sister a necklace.
    (i) Yes, and Phillip also bought his sister a bracelet.
    (ii) Yes, but Phillip also bought his sister a bracelet.
    (iii) No, Phillip also bought his sister a bracelet.

For the experimental stimuli, discourses were built around 8 transitive verbs and 6 experimental conditions (2 question types $\times$ 3 answer types), along with 8 filler question-answer pairs. Following the hypothesis presented earlier, we predicted that participants will opt for the (46-c-i) answer upon hearing a canonical sentence, the (46-c-ii) when hearing a cleft sentence and (46-c-iii) when hearing an exclusive sentence. The results are presented in Fig. 1 in absolute numbers, collapsed for grammatical function of the focused element.

Our predictions are confirmed by the results: participants were far more likely to overtly contradict Jason's answer (by updating the conversation with a *No*-continuation) if it was in the exclusive form. The distribution of continuation

sentences chosen after exclusive answers was significantly different from the distribution of continuations chosen after it-cleft answers ($\chi^2(2) = 249.5$, $p < .0001$), with far more *No*-continuations selected after exclusive sentences than after it-cleft sentences. In addition, it turns out that canonical answers receive a similar answer distribution as it-clefts. In fact there is only a very slight difference between canonical sentences and it-clefts, which turns out to be statistically relevant only at a 10 % error rate, even though the data sample is fairly large ($\chi^2(2) = 4.68$, $p < .1$).

We take these results to be further evidence that the exhaustiveness is triggered by the not at-issue content of clefts.

5.2 French Clefts

Twenty four undergraduates from the Université Toulouse Le Mirail in France participated in the forced-choice task. All participants were native monolingual speakers of French. The experiment was also conducted remotely over the internet via WebExp. On each trial, participants were presented with written stimuli containing a question-answer pair in the upper half of the screen and three continuation sentences in the bottom half of the screen. The instructions emphasized that participants needed to understand each item as being uttered by three different people, thus reading a conversation between three French speakers: Anne asking the question, Paul answering and Nicolas supplementing. The question asked by Anne was included to ensure that subjects correctly identified the focus element. The answer given by Paul appeared in either one of three forms: exclusive (47-a), canonical (47-b) and cleft (47-c). The continuation supplemented by Nicolas was introduced either by *Non* (No), *Oui, mais* (Yes, but) or *Oui, et* (Yes, and), in (47-c-i)-(47-c-iii). The instructions then presented the same task to participants: select the most appropriate continuation to the preceding question-answer pair. Instructions emphasized that there was no correct answer and that participants should base their judgments on their first impressions.

(46) Qui est-ce-que le directeur a grondé?
Who is-it-that the director has scolded?
'Who did the director scold?'

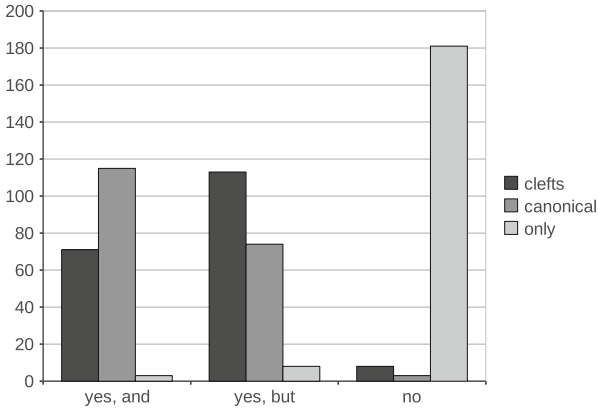
a. Le directeur n' a grondé que la secrétaire.
The director not has scolded only the secretary.
'The director scolded only the secretary.'

b. Le directeur a grondé la secrétaire.
The director has scolded the secretary.
'The director scolded the secretary.'

c. C' est la secrétaire que le directeur a grondé.
It-is the secretary that the director has scolded.
'It's the secretary that the director scolded.'

(i) Non, le directeur a aussi grondé le cadre.
No, the director has also scolded the executive.
'No, the director also scolded the executive.'

Fig. 2 French results



(ii) Oui, mais le directeur a aussi grondé le cadre.
Yes, but the director has also scolded the executive.
'Yes, but the director also scolded the executive.'

(iii) Oui, et le directeur a aussi grondé le cadre.
Yes, and the director has also scolded the executive.
'Yes, and the director also scolded the executive.'

Two variables were controlled for in the experimental stimuli: the form of the answer (exclusive, canonical or cleft-sentence), and the grammatical function of the focused element (subject or object), which yielded a total of six conditions. Within the experiment, each participant judged exactly four items per condition. So, each participant judged a total of twenty four experimental items, as well as twelve fillers which were pseudo-randomized with the experimental items.

Our predictions for French were exactly the same as for English: we will see an effect of the form of the stimuli on the continuation chosen by participants, so that *Non* is selected after exclusive sentences, *Oui, mais* after clefts, and *Oui, et* after canonicals. Results are given in absolute numbers in Fig. 2.

Here again, as predicted, participants did not choose a continuation randomly, the form of the answer did affect their choice. We observe the following: (i) in the cleft and canonical something weaker than *No* is much preferred, (ii) in the canonical *Yes, but* is a bit too strong, (iii) in the cleft something stronger than *Yes, and* is preferred, (iv) in the exclusive, something stronger than *Yes, and* and *Yes, but* is preferred.

A goodness-of-fit chi-square statistic was applied to the data and showed that the difference in distribution of responses across the three answer forms was highly significant ($\chi^2(4) = 100$, $p < 0.001$). The distribution of sentences chosen after exclusives was statistically different from the distribution of continuations chosen after clefts ($\chi^2(2) = 311.9$, $p < 0.001$). The difference in the distribution of continuation between canonical and cleft sentences was also found to be statistically significant, although obviously much smaller ($\chi^2(2) = 20.81$, $p < 0.001$).

We conclude that the predictions we made are confirmed by the experiment for French, and, together with the assumption that the exhaustiveness inference exists, we conclude that the inference is not-at-issue in *c'est*-clefts.

6 Conclusion

In this paper, we were concerned with the analysis of the exhaustiveness inference in clefts and related constructions such as the Hungarian pre-verbal focus construction. While the past literature has concentrated on the *source* of the inference, debating whether it must be considered a semantic or a pragmatic phenomenon, we followed Velleman et al. (2012) in shifting the focus to the *status* of the inference, arguing that the notion of at-issueness is key. Under this view, the differences observed between clefts and exclusives arise from the status of exhaustivity: exclusives make it at-issue whereas clefts make it non-at-issue.

For us, this shift of perspective is crucial and can better explain the results from previously collected data using the "Yes, but" test in Hungarian and German, and applies to newly collected data on English and French. Indeed, we argue that the "Yes, but" test itself should not be understood as a diagnosis for the source of the inference but in fact as a diagnosis for its status.

References

Atlas, J. D., and S. C. Levinson. 1981. It-clefts, informativeness and logical form: Radical pragmatics. In *Radical pragmatics,* ed. P. Cole, 1–62. New York: Academic Press.

Beaver, D., and B. Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Oxford: Blackwell.

Bolinger, D. 1972. A look at equations and cleft sentences. In *Studies for einar haugen,* ed. E. Firchow, 96–114. Den Haag: Mouton de Gruyter.

Büring, D. and Križ, M. 2013. It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics and Pragmatics* 6:1–29.

Byram Washburn, M., E. Kaiser, and M. Zubizarreta. 2013. The English it-cleft: No need to get exhausted. Poster presented at The Linguistic Society of America, Boston.

Carston, R. 1988. Implicature, explicature and truth-theoretic semantics. In *Mental representations: The Interface between language and reality*, 155–81. Cambridge: Cambridge University Press. (Reprinted in Steven Davis (ed.) (1991), Pragmatics: A Reader, 33–51. Oxford: Oxford University Press).

Delin, J. L. 1990. A multi-level account of cleft constructions in discourse. *Coling 90: Proceedings of the 13th International Conference on Computational Linguistics* 2:83–88.

Delin, J., and J. Oberlander. 1995. Syntactic constraints on discourse structure: The case of it-clefts. *Linguistics* 33 (3): 465–500.

Drenhaus, H., M. Zimmermann, and S. Vasishth. 2011. Exhaustiveness effects in clefts are not truth-functional. *Journal of Neurolinguistics* 24:320–337.

Ducrot, O. 1972. *Dire et ne pas dire. Principes de sémantique linguistique*. Paris: Hermann.

Dyakonova, M. 2009. *A phase-based approach to Russian free word order*. Ph.D. thesis, University of Amsterdam, Amsterdam.

Grice, H. P. 1967. *Studies in the way of words*. Cambridge: Harvard University Press.
Hedberg, N. 2000. On the referential status of clefts. *Language* 76:891–920.
Horn, L. R. 1981. Exhaustiveness and the semantics of clefts. In *Proceedings of the Eleventh Annual Meeting of the North Eastern Linguistics Society*, ed. V. Burke and J. Pustejovsky, 125–142. Amherst: University of Massachusetts.
Horn, L. R. 2013. Exhaustivity and focus revisited. Talk given at DGFS. Berlin. Task given at 35th Annual Conference of the German Linguistic Society (DGPS). Potsdam, Germany.
Jayez, J. 2005. How many are "several"? Argumentation, meaning and layers. *Belgian Journal of Linguistics* 19:187–209.
Jayez, J. 2010. Projective meaning and attachment. *Logic, Language and Meaning. Lecture notes in Computer Science* 6042:325–334. (Springer).
Jayez, J., and L. Tovena. 2008. *Presque* and almost: How argumentation derives from comparative meaning. *Empirical Issues in Syntax and Semantics* 7:217–240.
Kiss, K. E. 1998. Identificational focus versus information focus. *Language* 74:245–273.
Onea, E. 2013. *Potential questions in discourse and grammar*. MS. (Grant research in the Semantics Pragmatics Inter face (RiSPI)).
Onea, E., and D. Beaver. 2011. Hungarian focus is not exhausted. In *Proceedings of Semantics and Linguistic Theory (SALT),* 19, ed. E. Cormany, S. Ito, and D. Lutz, 342–359. The Ohio State University.
Percus, O. 1997. Prying open the cleft. In *Proceedings of the 27th Annual Meeting of the North-East Linguistics Society (NELS)*, ed. K. Kusumoto, 337–351. Glsa: Mc Gill University.
Recanati, F. 1989. The pragmatics of what is said. *Mind and Language* 4:295–329.
Reeve, M. 2010. *Clefts*. Ph.D. thesis, UCL.
Roberts, C., M. Simons, D. Beaver, and J. Tonhauser. 2009. Presupposition, conventional implicature and beyond: A unified account of projection. In *New Directions in the Theory of Presupposition, ESSLLI workshop*, ed. N. Klinedinst and D. Rothschild. Toulouse, France.
Simons, M., J. Tonhauser, D. Beaver, and C. Roberts. 2011. What projects and why. *Proceedings of Semantics and Linguistic Theory* XX:309–327.
Szabolcsi, A. 1981. The semantics of topic-focus articulation. In *Formal methods in the study of language*, ed. T. J. Jan Groenendijk and M. Stokhof, 513–541. Amsterdam: Matematisch Centrum.
Szabolcsi, A. 1994. All quantifiers are not equal: The case of focus. *Acta Linguistica Hungarica* 42:171–187.
Tonhauser, J. 2012. Diagnosing (not-)at-issue content. *Proceedings of Semantics of Underrepresented Languages of the Americas (SULA)* 6:239–254. (UMass, Amherst. GLSA).
van Kuppevelt, J. 1996. Inferring from topics. scalar implicatures as topic-dependent inferences. *Linguistics and Philosophy* 19:393–443.
Velleman, D., D. Beaver, E. Destruel, D. Bumford, E. Onea, and E. Coppock. 2012. It-clefts are it (inquiry terminating) construction. *Proceedings of Semantics and Linguistics Theory (SALT)* 22:441–460.
Wedgwood, D., G. Pethő, and R. Cann. 2006. Hungarian 'Focus Position' and english it-clefts: The semantic underspecification of 'Focus' readings. Manuscript.
Xue, J., and E. Onea. 2011. Correlation between presupposition projection and at-issueness: An empirical study. *Proceedings of the ESSLLI 2011 workshop on projective meaning*, 171–184. Ljubljana, Slovenia.
Zondervan, A. 2010. *Scalar implicatures or focus: An experimental approach*. Ph.D. thesis, University of Utrecht, Utrecht.

A Cross-Linguistic Study on Information Backgrounding and Presupposition Projection

Patrícia Amaral and Chris Cummins

Abstract This chapter builds on previous work on the diversity of English presupposition triggers with respect to their projection behavior in an experimental setting (Amaral et al., Proceedings of ESSLLI 2011 Workshop on Projective Content, pp. 1–7, 2011; Cummins et al., Humana Mente 23:1–15, 2012, Proceedings of Sinn und Bedeutung 17, pp. 201–218, 2013). Using the same methodology and similar materials, but in Spanish, we investigate the empirical validity of the distinction between two classes of presupposition triggers posited in the theoretical literature, namely that between lexical and resolution triggers (Zeevat, Journal of Semantics 9:379–412, 1992). The results of this study replicate our previous findings with English data. First, native speakers exhibit the same tendencies with respect to the addressability of foregrounded vs backgrounded content in coherent question-answer pairs. Second, the results point to native speakers' sensitivity to the distinction between lexical and resolution triggers, while further suggesting that distinctions within classes of triggers should be understood as gradient rather than categorical.

Keywords Presupposition triggers · Projection · Information backgrounding · Accommodation · Meaning typologies · Spanish

1 Introduction

Presuppositions are traditionally distinguished from other forms of non-asserted content on the basis of their ability to project to the discourse level from under the scope of operators such as negation. This projection behavior is responsible for the characteristic difficulty associated with responding to polar questions that contain presuppositions. To a question such as (1), a response of either "yes" or "no" appears to endorse the presupposed content (namely that the addressee was an alcoholic at

P. Amaral (✉)
Department of Romance Languages, University of North Carolina at Chapel Hill, 320 Dey Hall, CB 3170, Chapel Hill, NC 27599, USA
e-mail: pamaral@unc.edu

C. Cummins
Department of Linguistics and English Language, The University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD, UK

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_7,

some earlier point in time). In order to deny that presupposition, it is necessary to resort to a circumlocution and, in effect, object to the question at a metalinguistic level.

1. Are you still an alcoholic?

Although this pattern of observations appears broadly consistent for all expressions that trigger presuppositions, recent experimental research has begun to explore the diversity of presuppositions with respect to their projection behavior (Amaral et al. 2011; Cummins et al. 2012; Cummins et al. 2013). This work builds upon some well-established theoretical accounts of differences among presupposition triggers. In particular, the distinction between lexical and resolution triggers, posited by Zeevat (1992), appears to give rise to systematic differences in projection behavior. The empirical data on this point suggests that there may also be gradience among presupposition triggers. If correct, this would potentially support more fine-grained distinctions among triggers, but might also accord with a view on which projection is regarded as closely related to information backgrounding. In either case, these experimental findings could potentially have far-reaching implications for competing accounts of the semantic and pragmatic mechanisms underpinning presupposition projection.

In this paper, we further develop the empirical coverage of presupposition by exploring the behavior of presupposition triggers in another language, namely Spanish. So far, both the theoretical and the empirical work on presupposition has mostly been conducted on English and other Germanic languages (although see Castroviejo-Miró et al. 2013 and Jayez et al. 2014 in this volume for research on the Romance languages). However, the claims that have been put forward about the classification of presupposition triggers derive from rational considerations and are therefore presumed to have cross-linguistic validity. Our own prior work has shown that the distinction between lexical and resolution triggers has measurable correlates in a simple experimental task, while also suggesting that there may be more or less prototypical triggers within these two categories. This line of research has revealed the heterogeneity of presupposition triggers and shown the need for a more fine-grained understanding of strategies of information packaging in discourse. With respect to Spanish, we expect a similar pattern to emerge, i.e. that different categories of presupposition triggers can be distinguished within our experimental paradigm, in which participants are asked to rate the felicity of simple question-answer pairs.

The rationale for this work is twofold. First, the replication of the general pattern across languages would lend support to the notion that the differences between presupposition triggers are rooted in general logical or cognitive principles, rather than arising as a language-specific feature of English. A result of this kind would strengthen the empirical foundations of our proposal regarding the connection between information backgrounding and presupposition projection. Secondly, the possibility of gradience among presupposition triggers, for instance in their information packaging or backgrounding behavior, is relevant for understanding the degree to which certain expressions are cross-linguistic counterparts. For instance, we might judge that the English verb *continue* and the Spanish verb *seguir* are, broadly, translation equivalents, and the intuitions of native speakers might license us to assume

that both are lexical triggers, in Zeevat's (1992) terms. However, if there are potential differences between lexical triggers in the strength of their backgrounding effects, or in the degree to which their presuppositions are inaccessible to further discourse, we cannot assume that the substitution of one for the other can proceed without subtle implications for information structure. Conversely, if they are true translation equivalents, they should behave essentially the same way in respect of how they background their arguments.

In undertaking a cross-linguistic comparison of this kind, we must confront the problem of choosing which presupposition triggers to study in a given language. In this paper, we have identified a representative sample of triggers that will allow us simultaneously to study the possible existence of cross-linguistic and language-specific patterns. In this way, we hope to contribute to the current discussion on the cross-linguistic validity of taxonomies of projective content and on the methodological problems raised by such an endeavor (see e.g. Tonhauser et al. 2013; Matthewson 2004, 2006).

We start out by reviewing some of the properties of resolution and lexical triggers in connection with their projection and accommodation behavior. We then motivate the choice of triggers in our new experiment, and present the results of this study. We conclude by discussing the implications of these findings, both within and across languages.

2 Resolution and Lexical Triggers

As suggested by Zeevat's terminology, *resolution* triggers are fundamentally anaphoric in nature, i.e. their presuppositional content is a matter of retrieval of an entity or eventuality from the context. For this reason, failure to retrieve the presupposed entity or eventuality amounts to a discourse failure rather than to a logical failure. Like other anaphoric elements, their interpretation requires the hearer to identify the entity or eventuality in the discourse context that is being talked about further. We might assume such triggers to be crucial for discourse coherence and updating of the conversational record.

Zeevat identifies definite descriptions and factive *when*- and *after*-clauses as resolutional presupposition triggers. However, here we will include in this class a further category of triggers which Zeevat identifies but does not name, comprising elements such as English *too*, *also*, *another*, and *again*. They share the property of picking up a previously-introduced entity or eventuality from the discourse context and saying something new about it: characteristically, by relating it to a new piece of declarative content. Thus, such triggers participate, in Zeevat's (1992: 22) terms, in "the bookkeeping involved in storing information by humans".

The role of resolution triggers of this type in bookkeeping is underscored by two aspects of their behavior, also discussed by Zeevat (ibid.): they can access antecedents in parts of the context that other triggers cannot reach, and they are less easily dealt with by accommodation. In an example such as (2), the presupposition (that Mary

won an award before) need not have been introduced in the preceding context; it suffices for it to be common ground between speaker and hearer. However, if the presupposition is not common ground, it seems intuitively dispreferred to use the trigger, as in (3). This seems to represent a premature attempt to affirm a connection between two distinct pieces of information when both of them are new to the recipient.

2. Mary won another award.
3. A: Mary has never had any success with her novels.
   B: Didn't you hear? She just won another award.

It should be noted that, partly because of the wide-ranging access that resolution triggers have to the existing context, there is a danger of picking out the wrong antecedent. This is apparent with *too*, and especially in the absence of prosodic information: "Mary won an award too" might mean that someone else also won an award, that Mary did something other than winning an award, or—at a pinch—that someone other than Mary also did something other than winning an award. The purpose, in terms of bookkeeping, of using such a trigger correspondingly varies along with the identity of the anaphor. So the effective use of a trigger such as *too* should require that the context enables the recovery of the correct anaphor; however, the felicitous use of *too* requires a weaker condition, namely that it is possible to recover some antecedent.

On the other hand, in the case of *lexical* triggers, such as *stop* or *regret*, the presuppositional component is a requirement of the asserted content of the lexical item. As Zeevat puts it, "the application of a concept is only an option if certain conditions are already met. The conditions that must be met are the lexical presuppositions of the concept" (1992: 397). In other words, the presupposition is a logical precondition for the implication conveyed by the trigger. For example, one cannot *stop* an activity that has not started before; nor can one *regret* something that is not the case. In effect, the presupposition must hold in order for the asserted content to be meaningful (which accords with the analysis of Heim 1983).

The stronger requirement that lexical triggers thus impose, in order for the declarative content to make sense, seems to lead to different accommodation behavior. Speakers are confidently able to use such triggers on the understanding that the hearer will recover the presupposition and add it to her discourse model if necessary. The presupposition in question is always unambiguous here, unlike in the case of *too*. As a consequence, lexical triggers can be used to convey their presuppositional content to the hearer, as in (4)–(6), i.e. they may introduce informative presuppositions.

4. John recently quit smoking.
5. I just found out that Elaine got the professorship.
6. Miss Otis regrets she's unable to lunch today.

In some cases, it becomes difficult to tease apart the two types of triggers. For lexical triggers, the logical precondition is pre-established knowledge, which may have been established in the previous text. When this happens, the line between anaphorically retrieved content and a pre-condition for a certain concept seems to be blurred. However, as noted by Zeevat, the function of each class of trigger (i.e. the

relation between its asserted and presupposed content) is different, as reflected for example in the ability for presupposed material to be accommodated.

We predict that this difference in accommodation behavior will be manifest in the outcomes of an experimental task in which subjects have to evaluate the acceptability of question-answer pairs that explicitly address either the asserted or the presupposed content of a trigger. We discuss the justification for this prediction in the following section.

3 Presupposition Triggers and Typologies of Meaning Across Languages

One of the main challenges for studies of types of meaning is to identify an appropriate set of tests that can be applied across languages to target different meaning components without overlooking language-specific differences. Traditionally, the "family of sentences test" (Chierchia and McConnell-Ginet 1990) is assumed to provide a diagnostic of presupposed content (but see Levinson 1983; Kadmon 2001, among others, for discussion of the limitations of this test). For example, the implication that Cristina used to attend a Portuguese class, i.e. the presupposition of *dejar de* 'to stop' in (7), seems to project from the corresponding negated sentence in (8a), the interrogative in (8b) and the sentence prefaced by the modal operator of possibility in (8c).

7. Cristina ha dejado de ir a clase de portugués.

   'Cristina has stopped attending the Portuguese class.'

8. a. Cristina no ha dejado de ir a clase de portugués.

   'Cristina hasn't stopped attending the Portuguese class.'

   b. ¿Cristina ha dejado de ir a clase de portugués?

   'Has Cristina stopped attending the Portuguese class?'

   c. Es posible que Cristina haya dejado de ir a clase de portugués.

   'It is possible that Cristina has stopped attending the Portuguese class.'

Under normal circumstances, the speaker who utters (7) is committed to the truth of the proposition that Cristina attended a Portuguese class before. In fact, it is customarily judged as infelicitous, if not contradictory, to continue (7) in such a way as to negate that proposition:

9. Cristina ha dejado de ir a clase de portugués,?pero antes no iba a clase de portugués.

   'Cristina has stopped attending the Portuguese class, but before she didn't go to the Portuguese class.'

Interestingly, even though the implication that Cristina used to attend a Portuguese class projects from the sentence in which the trigger is negated (as in (8a)), it is

possible to provide a continuation of that sentence in which the presupposition is explicitly denied, as shown in (10):

10. Cristina no ha dejado de ir a clase de portugués, porque en realidad antes no iba a clase de portugués.

    'Cristina hasn't stopped attending the Portuguese class, because actually before she didn't go to the Portuguese class.'

One of the possible accounts of the acceptability of examples like (10) is that the presupposition of the trigger *dejar de* 'to stop' is locally accommodated and does not project further—in particular, it does not reach the discourse level. However, the rationale for this behavior remains unaccounted for. In this paper, we will pursue an approach to presupposition in which we regard it as a means of packaging information. In (10), the proposition that Cristina attended a Portuguese class before is backgrounded in the first conjunct and then directly addressed in the second conjunct. The presupposition failure is causally connected to the negation of the foregrounded content. In other words, a speaker who utters (7) makes a contribution to the discourse record (in the sense of Lewis 1979, or the QUD, in the sense of Roberts 1996), namely that Cristina does not attend a Portuguese class anymore, and commits herself to the belief that Cristina used to attend a Portuguese class before.

Native speakers share the intuition that while the former implication, i.e. the "at-issue" content of the sentence, can be directly refuted or questioned, it is less natural or straightforward to call the latter into discussion. This can be shown by constructing dialogues that instantiate a version of the "Hey, wait a minute test!" (Shanon 1976; Von Fintel 2004), as in B's replies to A's statement in (11).

11. A: Cristina ha dejado de ir a clase de portugués.

    'A: Cristina has stopped attending the Portuguese class.'

    B: ¡Cómo! Un momento:– ¿Cristina iba a clase de portugués?

    'B: Hey, wait a minute – Did Cristina attend a Portuguese class?'

    B: #¡Cómo! Un momento:– ¿Cristina ya no va a clase de portugués?

    'B: Hey, wait a minute – Doesn't Cristina attend the Portuguese class anymore?'

This test is argued to measure the addressability of each meaning component: the information which is foregrounded in a sentence is more readily amenable to discussion, while backgrounded information, which does not directly answer the QUD, requires an extra step if it is to be brought to bear on the current information state of the discourse. "Hey, wait a minute" achieves this, in the case of backgrounded information, but is less felicitous when it is used to bring information that is already foregrounded to bear on the current discourse state. The proposition that Cristina doesn't attend a Portuguese class at present could be disputed by just saying "No, that's not true".

Foregrounded content can also be identified by other diagnostics. For example, (7) could be a felicitous continuation to the first sentence in (12):

12. Cristina está ocupadísima este año. (Hasta) ha dejado de ir a clase de portugués.

    'Cristina is very busy this year. She has (even) stopped attending the Portuguese class.'

The relevant conversational implicature that makes the sequence coherent is that Cristina doesn't go to the Portuguese class because she doesn't have time. Crucially, the implication that she attended a Portuguese class before does not directly contribute to discourse update; it is the implication that she doesn't go to that class at present that participates in the calculation of the implicature. On this view, the implication that she attended a Portuguese class before is backgrounded as it is a precondition for the main assertion but it is not available to the main discourse record.

We can use the same diagnostics to compare the behavior of *dejar de* 'to stop' with that of *también* 'too' with respect to local accommodation. In (13), the first conjunct introduces the main proposition that Paco has moved and the presupposition that somebody (presumably known to speaker and hearer at this point in discourse) has moved.

13. Paco también se ha mudado, pero nadie más se ha mudado.

    'Paco has moved too, but nobody else has moved.'

Notice that in (13), although from the perspective of discourse coherence one may ask why the speaker has chosen to use *también* in this context, there is no contradiction; the truth of the foregrounded proposition, i.e. that Paco has moved, is independent of the content of the second conjunct. For this reason, any attempt to establish a causal connection between the negation of the two propositions—that is, to deny the foregrounded proposition on the basis of the falsity of the backgrounded content—results in a potential infelicity:[1]

14. #No es verdad que Paco también se ha mudado, porque nadie más se ha mudado.

    'It is not true that Paco has moved too, because nobody else has moved.'

Once again, in (14) information is introduced and backgrounded in the first conjunct, and then the backgrounded information is denied in the second clause. However, in (14) the presupposition-denying continuation does not entail an answer to the QUD "Why isn't it the case that Paco has moved too?": although the presupposition fails to hold, its falsity does not invalidate the proposition that "Paco se ha mudado".

In this paper, we pursue the hypothesis that the contrast between (9) and (10), and that between (13) and (14), stems in both cases from the relation between the asserted and the presupposed content of the triggers (in these examples, *dejar de* and *también*), adopting Zeevat's distinction between lexical and resolution triggers. Since this theoretical distinction relies upon different possible relations between two types of content, we expect it to have cross-linguistic validity. One of the main goals of

[1] For some native speakers of Spanish, the sentence becomes more acceptable if a pitch accent is placed on *también*. In such a case, we would be dealing with metalinguistic negation, the second clause providing the explanation for the infelicitous use of *también*.

this study is to test whether such a distinction will be measurable in an experimental setting.

Our predictions are the following. First, we predict that the overall preference for directly addressing foregrounded rather than backgrounded content should hold across languages, and hence that subjects should assign higher ratings to question-answer pairs that either affirm or deny the main proposition (i.e. conditions (i) and (iii) of our experiment; see the following section). Second, we predict that the two classes of items should behave differently with respect to presupposition failure (in conditions (ii) and (iv) of our experiment), as the two classes typically display different accommodation patterns, as discussed above. Third, we expect the data from Spanish to reveal a certain degree of gradience among items, just like the English data, on the basis that this is unlikely to be a language-specific property, but rather one that applies for presupposition triggers in general (for details of a proposal of a "continuum" of triggers, see for example Kadmon 2001).

We see this experimental work as exploratory, as it represents a first attempt to test our hypothesis that Zeevat's distinction between types of triggers may be valid across languages and amenable to empirical testing. However, at the same time, it is important to be aware of the possible pitfalls of this endeavor. First of all, we do not want to claim that the items in question have the exact same distribution across the different languages. It seems to us that the chosen items display similar behavior with respect to the categories under consideration and hence are comparable to the English items for the purposes of our study, on the basis of the tests presented. Furthermore, we have not collected information regarding their frequency, and our constructed examples may not be representative of the most common contexts of use of these items. For both these reasons, any claims that we can make, based on the data elicited here, about the general behavior of lexical triggers or resolution triggers as a class are necessarily tentative and subject to future investigation.

4 Our Experiment: Acceptability Judgments in Sentences with Presupposition Failure

The structure of this experiment is parallel to the structure of our previous study using English data. In our experiment, we investigated acceptability judgments for responses to polar questions that contained a range of presupposition triggers. We used four kinds of responses, following a 2×2 design crossing acceptance or rejection ("yes" versus "no") of the main proposition or of the presupposition. These responses involved either (i) responding "yes" and not denying the presupposition, (ii) responding "yes" but denying the presupposition, (iii) responding "no" but not denying the presupposition, or (iv) responding "no" and denying the presupposition. These response types are exemplified for a sample question in (15) and schematically represented in Table 1.

Table 1 Description of how conditions (i)–(iv) manipulate foreground and presupposition. The symbol + denotes positive response, − denotes negative response

	Condition (i)	Condition (ii)	Condition (iii)	Condition (iv)
Main proposition	+	+	−	−
Presupposition	+	−	+	−

15. ¿Sigue siendo Victoria la directora del departamento?

    'Does Victoria continue to be the director of the department?'

    (i) Sí, Victoria sigue siendo la directora del departamento.

    'Yes, Victoria continues to be the director of the department.'

    (ii) Sí, aunque antes Victoria no era la directora.

    'Yes, although Victoria was not the director before.'

    (iii) No, Victoria ya no es la directora.

    'No, Victoria isn't the director anymore.'

    (iv) No, porque Victoria no era la directora.

    'No, because Victoria was not the director before.'

4.1 Materials

We used eight presupposition triggers in Spanish: *lamentar* 'to regret', *seguir* 'to continue', *dejar de* 'to stop', *todavía* 'still', *otra vez* 'again', *también* 'too', *tampoco* (translatable as NPI *either*), and the comparative *mejor que*, which can be translated by a presuppositional comparative (*better . . . than*) in English, e.g. ¿Es Cristina mejor profesora que María? 'Is Cristina a better teacher than María?'. Of these, *lamentar*, *seguir*, *dejar de* and *todavía* correspond to the English words that we have tested in our previous experiment and are lexical presupposition triggers. *Otra vez*, *también* and *tampoco* can be considered as resolution triggers. In our previous study, the behavior of the comparative *better. . . than* suggested that its classification is complex; it seems to fall between both classes of triggers.

For each trigger, two questions were constructed (for a total of 16 questions), and for each question, four response conditions were constructed, as described above.

Four versions of the experiment were constructed. In each version, the 16 questions appeared in turn in a pseudo-randomized order. The design was balanced in such a way that each of the four response conditions appeared four times in each version of the experiment. Within each version, the two questions for each trigger were presented in different response conditions.

Table 2 Mean naturalness ratings for responses in each condition (SD in parentheses)

	Condition (i)	Condition (ii)	Condition (iii)	Condition (iv)
lamentar	3.64 (1.44)	1.60 (0.99)	4.05 (0.76)	3.44 (1.26)
seguir	4.15 (1.35)	2.32 (1.41)	4.56 (1.04)	3.35 (1.42)
dejar de	3.92 (1.47)	1.20 (0.52)	3.85 (1.42)	2.96 (1.34)
todavía	4.15 (1.31)	2.36 (1.29)	4.44 (1.08)	3.60 (1.35)
otra vez	3.68 (1.38)	2.30 (1.38)	3.95 (1.15)	2.48 (1.48)
también	3.90 (1.17)	2.56 (1.26)	3.80 (1.53)	2.50 (1.50)
tampoco	2.12 (1.39)	2.15 (1.09)	2.55 (1.43)	1.88 (1.27)
mejor que	4.10 (1.02)	2.64 (1.66)	3.52 (1.33)	2.95 (1.28)
Overall	3.70 (0.67)	2.14 (0.49)	3.84 (0.62)	2.90 (0.58)

4.2 Participants

45 native adult Spanish speakers, from Spain and Mexico, were recruited via word of mouth. The experiment was administered through an online survey using Ibex, hosted at Ibex farm (http://spellout.net/ibexfarm/). Each participant was randomly assigned to one version of the experiment. 10 participants were assigned to version 1, 15 to version 2, 11 to version 3 and 9 to version 4. The participant group had mean age 24.8 years (SD 7.4 years).

4.3 Procedure

Each version of the experiment comprised 16 trials. Each trial commenced with a fixation cross presented for 500 ms. Following this, a question-answer pair was presented. Participants were instructed to press a key after having read the question and answer, and then to rate the answer according to how natural it sounded ("*según lo natural que te parezca*"), on a five-point Likert scale ranging from 1 (least natural, labelled *menos natural*) to 5 (most natural, labelled *más natural*). The ratings were recorded and analyzed.

4.4 Results

Mean ratings across the four conditions were as shown in Table 2 (SD are in parentheses) and in Fig. 1.

With the exception of items with *tampoco* '(NPI) either',[2] each item attracted relatively high ratings in conditions (i) and (iii), suggesting that these are generally

[2] We suspect that the low ratings obtained in the *tampoco* dialogues in all conditions have to do with two facts. First, the template of answers beginning with "Yes"/"No" that were used in our materials is rather artificial for questions including *tampoco*. Second, our constructed dialogues provided limited contextual background (they consisted of short question-answer pairs). It has been shown

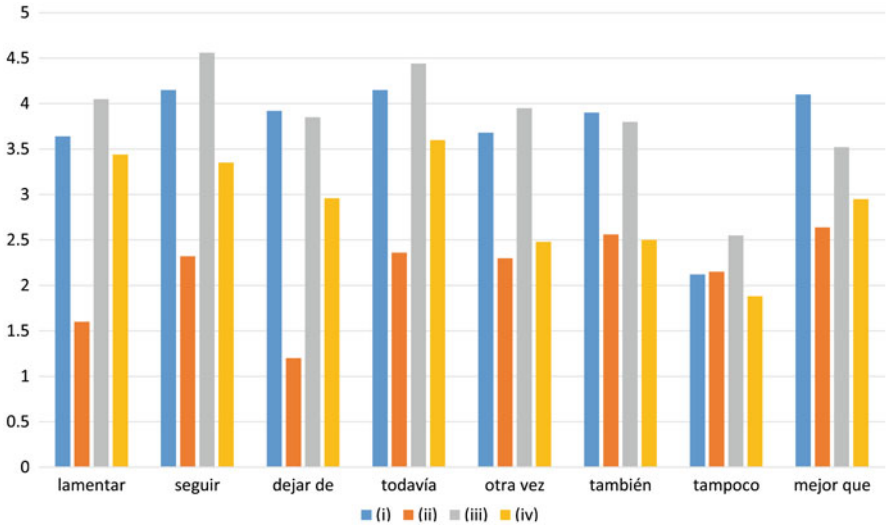


Fig. 1 Naturalness ratings for responses in each condition

regarded as admissible responses to the questions under test. In each case, the ratings for both conditions (i) and (iii) (those conditions in which the presupposition of the question is endorsed) numerically exceeded those for both conditions (ii) and (iv) (in which the presupposition of the question is denied).

Conditions (ii) and (iv) were generally accorded low ratings, although condition (iv) was rated in the upper half of the scale for *lamentar* 'to regret', *seguir* 'to continue' and *todavía* 'still'. The relative acceptability of these two conditions appears to vary substantially between triggers. In the case of *lamentar*, *seguir*, *dejar de* and *todavía*, there is a strong preference for condition (iv) over condition (ii). Planned comparisons revealed that this difference is significant for each of these items (unpaired t-test with Welch correction, all $p < 0.05$). *Mejor que* 'better. . . than' and *otra vez* 'again' exhibit a weaker preference in this direction, and there is a slight preference for type (ii) over type (iv) responses to the questions with *también* 'too' and *tampoco* '(NPI) either', but none of these numerical differences were significant when considered separately (unpaired t-test, all $p > 0.05$). The extent of the preference for condition (iv) over condition (ii) is shown in Fig. 2.

that *tampoco* is associated with complex discourse pragmatic licensing constraints (see Schwenter and Zulaica-Hernández, 2003). In our study, we limited ourselves to brief question-answer pairs with virtually no contextual information for practical reasons of experimental design and in order to keep the task constant across critical items.

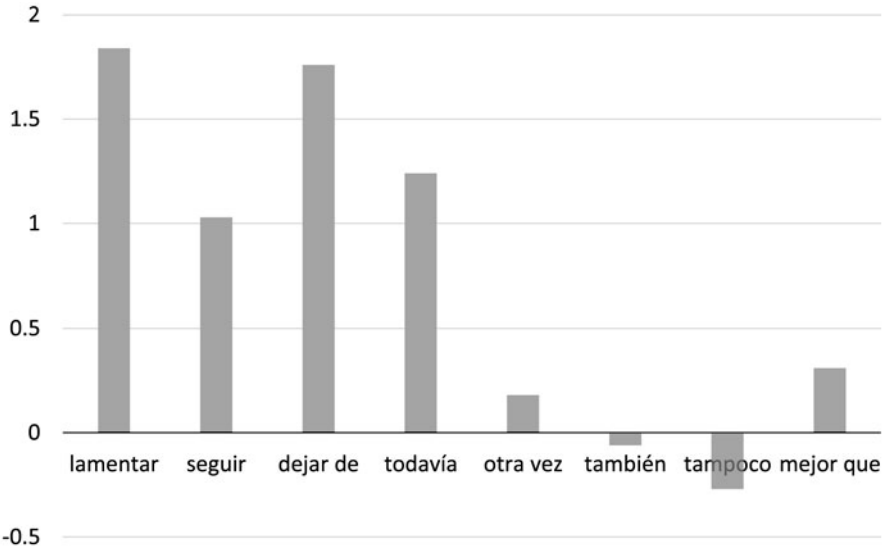


Fig. 2 Preference for condition (iv) over condition (ii) for each trigger

4.5 Discussion

The results show a robust pattern of responses. First, overall, conditions (i) and (iii) receive higher ratings than conditions (ii) and (iv) (with the exception of responses to *tampoco* '(NPI) either'). This pattern indicates that, as predicted, subjects prefer responses that address the foregrounded content rather than the presupposed content of a trigger in a coherent dialogue. This replicates existing findings for foregrounded vs. backgrounded content in English (see Cummins et al. 2012), and also conforms with prior theoretical claims that presupposed content is generally less amenable to being addressed in the following conversational turn than asserted content. This can be construed in terms of higher-level information structure, for instance by presupposition triggers making their corresponding content "not at-issue".

Secondly, in responses to lexical triggers (*lamentar* 'to regret', *dejar de* 'to stop', *todavía* 'still', *seguir* 'to continue'), subjects strongly prefer condition (iv) over condition (ii). For resolution triggers (*otra vez* 'again', *también* 'too', *tampoco* '(NPI) either'), considered separately or as a group, there is no significant preference for either condition (ii) or (iv). In other words, for lexical triggers, subjects consider responses that affirm the foregrounded content while negating the presupposed content to be less acceptable than responses that are negative in their declarative force ("no, because. . . ") as a consequence of presupposition failure.

Thirdly, although there are tendencies pointing to a principled theoretical differentiation, the results suggest that there is gradience within classes of triggers, with some triggers potentially falling between categories. For instance, within the category of lexical triggers, the magnitude of the preference for condition (iv) over

condition (ii) appears greater for *lamentar* 'to regret', *dejar de* 'to stop' and *todavía* 'still' than for *seguir* 'to continue'. Moreover, judgments for the comparative lie between those of the putatively lexical and resolutional classes of triggers. We will discuss the implications of this in more detail in the next section. There is also appreciable variation between triggers in condition (i), but this may be attributable to the varying acceptability of the repetitive formulations used in the answers in this condition, which is not central to the objectives of this piece of research but may be relevant to the use of this methodology in subsequent work. We acknowledge the fact that the repetitive formulations in the answers are artificial from a communicative point of view and may therefore have impacted native speakers' judgments in the task.

5 General Discussion

The results of this experiment by and large replicate our previous findings with English data (Cummins et al. 2012: 210–211). In both cases, the naturalness ratings show a general preference for answers that address the foregrounded content rather than the backgrounded content in coherent question-answer pairs.

The findings from both languages demonstrate that native speakers are sensitive to the distinction between the two classes of triggers, lexical and resolution triggers. If we compare the difference in ratings between condition (ii) and condition (iv), we find a very similar picture in the two languages. In the case of *continue*, *regret, still* and *stop*, there was a significant preference (all $p < 0.01$) for condition (iv) (Cummins et al. 2012: 211). Likewise, for their Spanish counterparts *seguir*, *lamentar*, *todavía* and *dejar de*, we observe that native speakers also prefer negative answers on the basis of presupposition failure rather than positive, presupposition-denying answers (all $p < 0.05$). This is to be expected given the nature of this class of triggers, if we adopt Zeevat's approach. For this class of items, the responses in condition (ii) appear self-contradictory, if we assume that the presupposition is a logical prerequisite for the at-issue content of the trigger. We should also point out that within the lexical triggers, the preference for condition (iv) over (ii) in the Spanish data closely matches the results in the English data, i.e. cross-linguistic counterparts can be ordered in the same way with respect to their preference for one condition over the other. The only exception is *dejar de*, which shows a stronger preference for condition (iv) over (ii) than *stop*.

Note, however, that we do not find a preference for either of these conditions with resolution triggers. In our first experiment, in the cases of *again, too* and the comparative construction, there was no significant preference between conditions (ii) and (iv). There was a slight numerical preference for condition (iv) in the case of *again* and a slight numerical preference for condition (ii) in the case of *too* and the comparative. In the Spanish data, *otra vez* 'again' paired with the comparative in the preference for condition (iv), while *también* 'too'/*tampoco* '(NPI) either'

showed the opposite pattern, but again the differences did not reach statistical significance. In other words, while for lexical triggers subjects showed relatively polarized judgments of the acceptability of negative versus positive answers that introduce a presupposition failure, this was not the case for resolution triggers. This result is highly compatible with a view in which presupposition failure for resolution triggers bears mostly upon discourse coherence, rather than presenting a challenge for the logical coherence of the present declarative utterance, as is putatively the case for lexical triggers.

On a broader theoretical level, these experimental data invite some reflections regarding the use of presupposition triggers in discourse. There seem to be two distinct things that we can do with a presupposition trigger: we can introduce a presupposition that was not previously part of the common ground, and exploit accommodation in order to have it admitted to the common ground; or we can reintroduce some content that was already in the common ground but which we wish to make salient again for some reason. Zeevat's distinction appears to correlate closely with this distinction: hearers seem to expect resolution triggers to reintroduce (and build on) old material, and lexical triggers to allow for the introduction of new material. For instance, *too* seems to be used primarily to reintroduce old material, as in a case like "I was there too" (which typically presupposes e.g. that the addressee or some other salient individual was there, something that had probably been asserted in the preceding discourse), whereas a trigger like *regret* can readily be used to affirm two propositions at once, even when neither is currently in the common ground.

Our perspective is therefore slightly different from that of Jayez et al. (this volume). In our interpretation of Zeevat's distinction, different classes of items may correspond to different strategies of partitioning information in discourse. We would hypothesize that these different strategies give rise to different relations between asserted and presupposed content. This in turn could be expected to influence the way in which contextual information interacts with the presuppositional utterance. Although we would not wish to downplay the importance of context in determining whether or not presuppositions are projected in a given case, we would predict that the interaction between lexical meaning and contextual information might differ systematically between the two classes of triggers. Hence, we are not so much interested in "projection strength" as a property of triggers, as Jayez et al. put it, but rather on the different possibilities for this interaction. Our prediction was that items of each class should display somewhat consistent behavior cross-linguistically, and this prediction—at least with respect to the kinds of dialogic exchanges under test—is borne out by our experiments so far.

We can conjecture that there is a trade-off between the flexibility and economy of the resolution triggers, and their inability to introduce new material. It seems inevitable that there is a trade-off between form and content, inasmuch as words like *too* and *again* cannot have much informational content in their own right. They can each be appended to a declarative sentence to convey something like, respectively, "a similar proposition to the one expressed in this sentence can also be affirmed" and "the event or state of affairs described in this sentence has occurred or arisen before". However, they are essentially parasitic upon the sentence for their propositional

meaning; they are anaphoric both in the sense that they collect entities in order to say new things about them and in the sense that their denotation depends upon that of some other constituent in the sentence in which they occur. By contrast, lexical triggers tend to be less economical in their expression and can be used only in relatively circumscribed syntactic environments, but can convey correspondingly more, and more precise, information. For instance, the construction "N is Adj that p" conveys at once that N is Adj, that p is the case, and that the fact of p being the case is causative of N's status as Adj.

Our results may have implications for currently accepted theories of presupposition projection. Projection theories either tend to rely on a dynamic semantic approach (à la Heim) or on a pragmatic approach (along the lines of Stalnaker 1976 and more recent developments). A related issue has been a putative preference for global accommodation, and whether such a preference relies on semantic or pragmatic mechanisms. Our results in both experiments on English and Spanish suggest that there may be a split in types of presupposition triggers, or at least a tendency for triggers to cluster around different sets of properties, that cannot be accounted for by a unified theoretical approach. Crucially, the experimental data reveal a diversity of behavior of presupposition triggers that may require more nuanced proposals (or a combination of different approaches). If classes of triggers like the ones proposed by Zeevat have some psychological reality, then the constraints (e.g. logical, discourse-based, etc.) that play a role in their accommodation behavior may also vary in their application between triggers, for instance by having different relative weights.

6 Conclusion

Our study has aimed to show that experimental work can contribute to current theoretical debates in the literature on presupposition projection and accommodation. Specifically, we argue that at least some of the broad theoretical distinctions assumed in the field, such as that between foregrounded and backgrounded content, have psychological reality for native speakers of different languages and are amenable to experimental testing. We believe that the questions raised in this work, which is still exploratory and incipient with respect both to the breadth of languages considered and to the coverage within each language, are important to cross-linguistic research on the nature of presupposition triggers. Many of these methodological questions are also relevant to other empirical approaches, e.g. corpus-based and fieldwork studies. With similar work that tries to integrate different strands of research and enlarge our knowledge base, we may be able to create more robustly grounded models of presupposition projection and ultimately arrive at better informed taxonomies of meaning.

References

Amaral, Patrícia, Chris Cummins, and Napoleon Katsos. 2011. Experimental evidence on the distinction between foregrounded and backgrounded meaning. Proceedings of ESSLLI 2011 Workshop on Projective Content, 1–7. (eds. C. Roberts, J. Tonhauser, and G. Kierstead).

Castroviejo-Miró, Elena, Laia Mayol, and E. Allyn Smith. 2013. A cross-linguistic study of direct denial as a diagnostic for meaning types. XPRAG-2013.Utrecht, September 2013.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and grammar: An introduction to semantics*. Cambridge: Cambridge University Press.

Cummins, Chris, Patrícia Amaral, and Napoleon Katsos. 2012. Experimental investigations of the typology of presupposition triggers. *Humana Mente* 23:1–15. (http://www.humanamente.eu/).

Cummins, Chris, Patrícia Amaral, and Napoleon Katsos. 2013. Backgrounding and accommodation of presuppositions: An experimental approach. *Proceedings of Sinn und Bedeutung* 17:201–218.

Heim, Irene. 1983. On the projection problem for presuppositions. In *Proceedings of the second West Coast Conference on Formal Linguistics*, eds. D. Flickinger et al., 114–125. Stanford: Stanford University Press.

Jayez, Jacques, Valeria Mongelli, Anne Reboul, and Jean-Baptiste van der Henst. 2014. Weak and Strong Triggers. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

Kadmon, Nirit. 2001. *Formal pragmatics*: *Semantics, pragmatics, presupposition, and focus*. Malden: Blackwell.

Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8:339–359.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70 (4): 369–415.

Matthewson, Lisa. 2006. Presupposition and cross-linguistic variation. *North East Linguistic Society (NELS)* 26:63–76.

Roberts, Craige. 1996. Information structure: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics* 49:91–136. (Papers in Semantics, eds. J. H. Yoon and A. Kathol. OSU Department of Linguistics).

Shanon, Barry. 1976. On the two kinds of presuppositions in natural language. *Foundations of Language* 14:247–249.

Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. Toward a taxonomy of projective content. *Language* 89 (1): 66–109.

Schwenter, Scott, and Iker Zulaica-Hernández. 2003. On the contextual licensing of *tampoco*. In *Linguistic Theory and Language Development in Hispanic Languages: Papers from the 5th Hispanic Linguistics Symposium and the 4th Conference on the Acquisition of Spanish and Portuguese*, eds. S. Montrul and Ordóñez, 62–80. Somerville: Cascadilla.

Stalnaker, Robert. 1976. Pragmatic presuppositions. In *Semantics and Philosophy*, ed. M. Munitz and P. Unger, 197–213. New York: New York University Press.

Von Fintel, Kai. 2004. Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. In *Descriptions and Beyond*, ed. M. Reimer and A. Bezuidenhout, 315–341. Oxford: Oxford University Press.

Zeevat, Henk. 1992. Presupposition and accommodation in update semantics. *Journal of Semantics* 9:379–412.

Weak and Strong Triggers

Jacques Jayez, Valeria Mongelli, Anne Reboul and Jean-Baptiste van der Henst

Abstract The idea that presupposition triggers have different intrinsic properties has gradually made its way into the literature on presuppositions and become a current assumption in most approaches. The distinctions mentioned in the different works have been based on introspective data, which seem, indeed, very suggestive. In this paper, we take a different look at some of these distinctions by using a simple experimental approach based on judgment of naturalness about sentences in various contexts. We show that the alleged difference between *weak* (or *soft*) and *strong* (or *hard*) triggers is not as clear as one may wish and that the claim that they belong to different lexical classes of triggers is probably much too strong.

Keywords Presupposition projection · Weak/strong triggers · Accommodation · Context abduction · Discourse planing

1 Introduction

Presupposition triggers are often felt to presuppose more or less 'strongly'. For instance, whereas (1b) strongly gives rise to the inference that Paul has missed the point, (1a) does not. Verbs like *discover*, *realize* or *know* are generally considered to be 'weaker' than *regret*, in that they do not (always) give rise to an intuition of entailment.

J. Jayez (✉)
ENS de Lyon and L2C2, CNRS, Lyon, France
e-mail: jjayez@isc.cnrs.fr

V. Mongelli
Institut du Cerveau et de la Moelle Épinière Research Center,
Institut National de la Santé et de la Recherche Médicale, U975 Paris, France
e-mail: valeria.mongelli@ens.fr

A. Reboul · Jean-Baptiste van der Henst
L2C2, CNRS, Lyon, France
e-mail: reboul@isc.cnrs.fr

Jean-Baptiste van der Henst
e-mail: vanderhenst@isc.cnrs.fr

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_8,

(1) a. If Paul realizes he has missed the point, he will probably reformulate his objection.
b. If Paul regrets he has missed the point, he will probably reformulate his objection.

This kind of observation, originating with (Karttunen 1971), suggests that presupposition triggers might differ in their presuppositional strength. Parallel observations for adverbials like *too* or *again* or clefts, as opposed to aspectual verbs like *stop*, might be taken to reinforce this impression.

(2) a. If Paul missed the point again, this proves that he does not really understand the subject.
$\rightsquigarrow$ Paul has missed the point before
b. Maybe Paul is quitting smoking or something. That would explain why he is so edgy.
$\not\rightsquigarrow$ Paul smokes

More recently, Abusch (2002, 2010) has argued for a distinction between *hard* and *soft* triggers on the basis of a simple test. The presupposition of soft triggers like *win* can be accommodated in the antecedent of a conditional although their truth is explicitly suspended, whereas that of hard triggers like *too* cannot.

(3) a. I don't know if Paul participated in the race, but, if he won, he must be very proud.
b. I don't know if Paul participated in the race, but if Mary participated ??too they probably had a drink together just after.

In this paper we examine the distinction between weak (aka soft) and strong (aka hard) triggers from an experimental point of view. In Sect. 2 we examine the major recent works relevant to the weak/soft distinction. In Sect. 3, we present two experiments in French and discuss the results in Sect. 4. Although we do not reject the weak/strong distinction altogether, our findings lead us to adopt a more nuanced view. In view of the experimental results, the most plausible conclusion is that the distinction is not purely lexical but results from the combination of a number of independent factors.

2 Weak and Strong Triggers

It is a truism that the literature on presuppositions is potentially confusing. The interesting question is why. Why should presuppositions be more difficult to describe and categorize than other semantic or pragmatic phenomena? It seems that the main reason is an unresolved tension between two types of evidence.

On the one hand, there are the so-called *projection* tests. Putting a trigger in the scope of certain operators preserves the presupposition, although the very same operators cancel or suspend the main content. Negation, interrogation, *if*-clauses and modal verbs are well-known cases, see (Beaver and Geurts 2013; Chierchia and McConnell-Ginet 1990; Geurts 1999) for different surveys. For instance,

(4) suspends the truth of the main content (the event of Paul forgetting to lock his car) but keeps the presupposition that Paul has a car alive. In such cases, the presupposition is said to *project*.[1]

(4) Maybe Paul forgot to lock his car.

On the other hand, it has been observed repeatedly that projection is not a stable property. From browsing through the literature, one might get the impression that non-projection concerns only a small subset of carefully chosen examples. However, it is not difficult to find less well-known cases. In (5a), where 'F' marks new information, it is possible that no one guessed the secret word, although *only* A normally presupposes that A. In (5b), it is possible that Paul arrived just before running out of power and had time to load his battery, thus preventing any breakdown.

(5) a. If only [Wilma]^{F} guessed the secret word, she has won € 100. (Geurts and Van der Sandt 2004, example 33b)
b. Paul arrived just before his car broke down.

So there is a potential conflict between two sets of observations. A straightforward way out would be to assume that (i) the projection behavior of presuppositions is only a default property and (ii) context can obviate projection more or less easily. Under that perspective, non-projection in a given context would be a matter of plausibility and, when sentences are given in isolation, the final perception would depend on the possibility of abducing reasonable contexts for precluding projection. In essence, this is Stalnaker's project. Stalnaker notes that "the constraints imposed by a statement on what is presupposed seem to be a matter of degree, and this is hard to explain on the semantic account." (Stalnaker 1974, p. 54). By 'semantic account' Stalnaker means any theory that attributes projection to a lexical item or linguistic construction, irrespective of contextual cues. He favors instead a pragmatic account, in which context and inference play a central role.

For instance, Stalnaker would explain the difference between (6a) and (6b) as follows. If the speaker presupposes that she has missed the point, she necessarily realizes that she did so. So, by contraposition, if she does not realize she has missed the point, she cannot presuppose that she did, hence the lack of projection with (6a). In contrast, in (6b), the speaker can perfectly well presuppose that she missed the point and consider the possibility of regretting that later.

(6) a. If I realize I have missed the point, I will probably reformulate my objection. $\not\rightsquigarrow$ I have missed the point.
b. If I regret I have missed the point, I will probably reformulate my objection. $\rightsquigarrow$ I have missed the point.

However, a simple pragmatic strategy does not work as smoothly in each and every case.

[1] Although we are not fond of this terminology, which is heavily syntax-oriented, we will stick to it because it is established usage.

First, Stalnaker's analysis does not seem to extend to (1) very easily. The speaker of (1a) *can* presuppose that Paul missed the point without any contradiction (see Beaver (2004)). Yet, projection is not automatic. Second, when the presupposition does not project, it has to be accommodated in the local context of the trigger. For instance, in a pattern like *If* p, q, where p presupposes some p', if it happens that the truth of p' can be questioned, we have to reconstruct an interpretation corresponding to *If* $p'\&p$, q, where p' is accommodated in the local context, that is, the antecedent of the conditional. If we observe a variation in accommodation, the pragmatic strategy predicts that it is due either to a difference in contexts or to a difference in the possibility of abducing a favorable context. But consider (3) again.

(3) a. I don't know if Paul participated in the race, but, if he won, he must be very proud.
b. I don't know if Paul participated in the race, but if Mary participated ??too they probably had a drink together just after.

We do not need any extra context to tell us that, in *both* cases, we have to accommodate the proposition that Paul participated in the race. Should we argue that this proposition is more accessible, salient, plausible, etc., in the presence of *win* than in the presence of *too*? This is not unreasonable, given that *win* entails or implies very strongly participating, whereas *Mary participated too* does not entail that Paul participated. So, we might say that, in a sense, participating is more strongly associated with winning than *x did y* with *z did y too*. However, two new problems come up. First, one might argue that *someone did y* is strongly associated with *z did y too*.[2] Yet, substituting *someone* for *Paul* in (3b) does not produce a more felicitous text, see (7).

(7) I don't know if someone participated, but, if Mary participated ??too ...

Second, Abusch (2002, 2010) claims that clefts are strong triggers. Clefts have the property that their main content entails their presupposition. For instance, *It's Paul who solved the problem* asserts that Paul solved the problem (main content) and presupposes that someone did (presupposition), an obvious consequence of the main content. If Abusch is right about the status of clefts, we have a case where an entailed presupposition cannot be suspended, in contrast to *participate* with respect to *win*.

Third, the location of verbs like *regret* in the landscape of triggers is not clear. The experimental results reported in (Cummins et al. 2013; Cummins and Amaral 2014)[3] suggest that *regret* patterns with *weak* triggers, which apparently clashes with other taxonomies (Abbott 2006; Abusch 2010).

In view of the variation exhibited by the literature, we decided to investigate the difference between French weak and strong triggers in an experimental perspective, starting with basic experiments based on native speakers intuitions about French sentences. With respect to the introspective observations reported in the present

[2] This is the standard hypothesis on the semantics of *too*, see van der Sandt and Geurts (2001) on *too*.

[3] We discuss some aspects of their work in Sect. 4.2.

section, there are three main differences: (i) We adopted Abusch's basic frame in order to keep the standard of comparison constant, (ii) we used only contextualized sentences in order to reduce as far as possible the effect of context abduction by the subjects, (iii) we systematically compared targets and controls, that is, sentences with and without suspension of the presupposition or sentences with and without the trigger. We did not compare triggers directly, for instance *regret* and clefts, because this amounted to comparing very different sentences. We focused on three French elements that are considered as very close to their English counterparts, *aussi* corresponding to *too*, *regretter* corresponding to *regret*, *qui*-clefts corresponding to subject clefts. In the rest of the paper we refer to these elements using their English counterparts for simplicity.

3 Experimental Findings

3.1 Experiment 1: Too (aussi)

The first experiment is based on an introspective difference illustrated in (8) vs. (9), which are English translations of the original French stimuli (see Appendix). It seems that (9b) is much better that (8b). The intuitive status of (8a) and (9a) is less clear.

(8) **Context**: Paul has been invited to a party. He is on very bad terms with Mary and they would prefer not to come across each other. Paul thinks that Mary might have been invited too.

   a. I don't know whether Paul will go to the party but, if Mary goes, it will be embarrassing.
   b. I don't know whether Paul will go to the party but, if Mary goes *too*, it will be embarrassing.

(9) Context: Paul has been invited to a party. He is on very bad terms with Mary and they would prefer not to come across each other. Paul thinks that Mary might have been invited too.

   a. I don't know whether Paul will go to the party because, if Mary goes, it will be embarrassing.
   b. I don't know whether Paul will go to the party because, if Mary goes *too*, it will be embarrassing.

The goal of the experiment was to take into account the possible effect of discourse structure on the strong trigger *too*. We used a contrast between *mais* (*but*) and *parce que* (*because*), in order to have two very different discourse structures, contrast/opposition and explanation/justification. Moreover, we compared texts with and without the trigger *too* in order to take into account the possibility that accommodating the proposition corresponding to the presupposition could be partly independent from the presence of *too*. If, for instance, the discourse structure is in

Table 1 Design of experiment 1

	because	but
Group 1 (19)	with *too*: 1–3	without *too*: 4–5
Group 2 (19)	without *too*: 1–3	with *too*: 4–5
Group 3 (18)	without *too*: 4–5	with *too*: 1–3
Group 4 (21)	with *too*: 4–5	without *too*: 1–3

itself a sufficient cue to help subjects accommodate, the effect of *too* in inducing an accommodation should be at best marginal.

The difference between *but*/*because* control stimuli where the presupposition was not suspended and target versions with the suspension in effect had already been tested in (Jayez and Mongelli 2013). *But* targets were significantly worse than their control. *Because* targets were not significantly different. We focus here on the comparison between *but* and *because* structures, as illustrated in examples (8) and (9).

3.1.1 Description of the Experiment

Participants Eighty-two subjects were recruited. Five were taken out because they were not native speakers or were bilingual. The remaining subjects (77) were undergraduate students and native speakers of French between 18 and 51 (mean 23.5).

Material and design We used five basic sentences in five contexts (see Appendix). Each final sentence had the form: *I don't know whether p because/but, if p′ too*/ø, *q*. where *p′ too*, but not *p′* alone, presupposes *p* and ø denotes the null string, alternating with *too*. So, there were 2 × 2 conditions: *because*/but × with *too* vs. without *too*. We also had eight fillers, common to all the participants.

Subjects were divided into four groups in a between-subject design shown in Table 1.

Each subject saw five experimental stimuli and the eight common fillers. No subject saw the *but* and *because* versions of the same text or the with and without *too* version of the same text.

Procedure The stimuli were pseudo-randomized and presented on a sheet. The participants were instructed to follow strictly the order of presentation and not to modify a previous answer. They had no time limit or speed indication. The task consisted in reading the context and the sentence and evaluating its comprehensibility on a seven point scale (1–7), where one was the worst and seven the best mark.

3.1.2 Results and Analysis

The basic results are shown in Table 2.

Table 2 Net results for the *aussi* experiment

	Sentence type	Nb. Obs.	Nb. subjects	Mean	Variance
1	bec.with.too	99	40	6.25	3.33
2	but.wo.too	100	40	5.08	3.85
3	bec.wo.too	93	37	5.73	4.24
4	but.with.too	92	37	5.68	3.62

Table 3 Contrasts for the *aussi* experiment

	Contrast	*p*-value	Null hypothesis
1	*because* with/without *too*	0.055	without-score $\geq$ with-score
2	*but* with/without *too*	0.021	without-score $\geq$ with-score
3	*because*/*but* with *too*	0.016	*but* $\geq$ *because*
4	*because*/*but* without *too*	0.0086	*but* $\geq$ *because*

Since the response is ordinal, we analyzed the data with a standard Mann–Whitney test for independent samples.[4] Table 3 reports the *p*-values. The last column indicates under which null hypothesis the *p*-value was estimated. We considered that it was *a priori* more likely to have high scores for a *because* sentence with *too* than with the same sentence without *too*, because the adverb was supposed to help derive the correct inference that the *simultaneous* presence of Paul and Mary may lead to tensions. We made a similar assumption for *but*. For the *but*/*because* contrast, we assumed that the *because*-versions were likely to get higher scores than the *but*-versions.

The results, shown in Table 3, show that the *because* and with *too* levels improve the scores, a fact which can be visualized through density plots (Fig. 1). The *p*-values suggest that:

1. *Because*-sentences are better with *too* (line 1).
2. *But*-sentences are better with *too* (line 2).
3. With *too*, *because*-sentences are better than *but*-sentences (line 3).
4. Without *too*, *because*-sentences are better than *but*-sentences (line 4).

The fact that *but* sentences with *too* are better than without *too* (line 2 of Table 3) is not surprising. In the *but* version without *too*, the conclusion of the *if* sentence bears no clear relation to its antecedent. For instance, in (8a), the dominant interpretation is that if Mary goes to the party, it will be embarrassing, no matter whether Paul goes to the party or not, which clashes with the context.

The main conclusion one can draw from these results is that there is no absolute or even strong prohibition of accommodation with *too* in the antecedent of a conditional. Clearly, the context can help subjects accommodate the missing presupposition, up to the point where the difference with a control where the presupposition is not suspended is no longer discernible (Jayez and Mongelli 2013). The processing of

[4] The results were controlled with the *npar.t.test* function from the R package *nparcomp* (Konietschke 2012), in order to ensure that there was no parasitic effect of difference of variance in the comparisons.

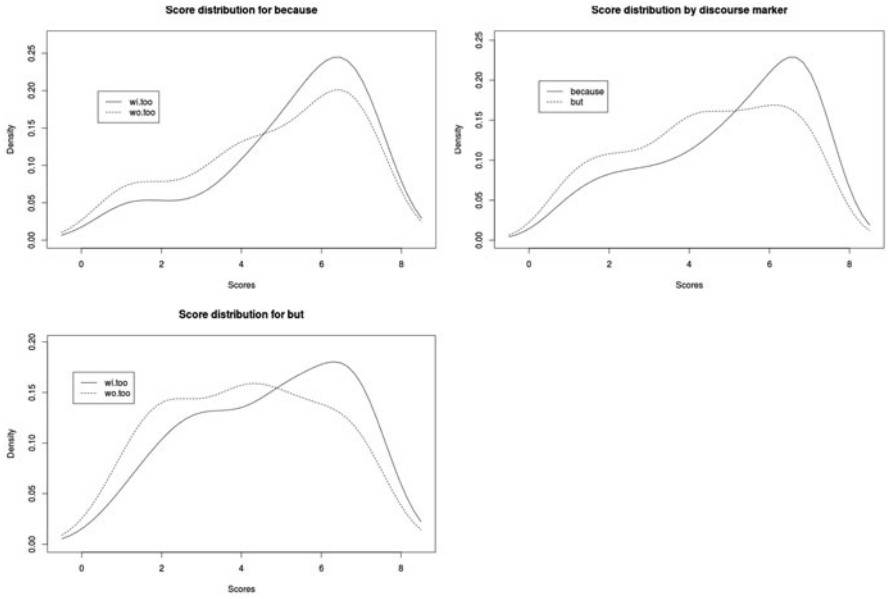


Fig. 1 Density plots for *because* and *but*

accommodation remains, however, unclear. We don't know at which point in the reading sequence subjects accommodate the missing presupposition and it is possible that the accommodation schedule is different across subjects or categories of subjects.

Given that there is no difference between the with/without *too* versions for *because*, it is tempting to assume that, at the moment *too* is read, the missing presupposition has already been made salient. We return to this point in Sect. 4.

3.2 Experiment 2: Regret and Clefts

Regret-sentences and clefts are considered to be strong triggers by Abusch. *Regret* is mentioned as a full factive by Karttunen, in contrast with semi-factives such as *discover*. We tested *regret*-sentences and clefts along the same lines as *too*. Subjects had to evaluate contextualized sentences like the following ones.

(10) **Context**: Véronique is wondering whether she will change her current car for a bigger one.

a. I think that Véronique bought a bigger model. If she regrets it later, it will be difficult to change again. [**control**]
b. I wonder whether Véronique bought a bigger model, but, if she regrets it later, it will be difficult to change again. [**target**]

Table 4 Net results for the *regret*/cleft experiment

	Sentence type	Nb. Obs.	Nb. subjects	Mean	Variance
2	regret.target	52	33	4.87	3.45
3	cleft.target	52	33	5.92	2.23
4	cleft.control	47	33	6.11	2.05
5	regret.control	47	33	5.13	4.03

(11) **Context**: An employee cannot log in on his computer.

a. I think that someone changed the password. If it was my colleague, I just have to wait to ask him. [**control**]
b. I don't know whether someone changed the password but, if it was my colleague, I just have to wait to ask him. [**target**]

3.2.1 Description of the Experiment

Participants Forty undergraduate students were recruited. Seven were taken out because they were not native speakers of French or were bilingual. The 33 remaining subjects ranged over 18–26 years with a mean of 21. They had neither participated in the first experiment nor heard about it.

Material and design Six *regret*-sentences and six clefts were created together with eight fillers. Each subject saw three *regret* items and three clefts, one control and two targets or one target and two controls for each category (*regret*/cleft). No subject saw a target and its control. The design was between-subject in the sense that no two subjects evaluated exactly the same set of stimuli.

Procedure Subjects had to read 14 stimuli (6 experimental items and 8 fillers) and to evaluate the stimuli as in the first experiment.

3.2.2 Results and Analysis

The net results are shown in Table 4.

Unlike in the case of *too*, it is not advisable to run non-parametric comparisons directly because the groups of subjects that would be compared are not independent. The first lines of the design matrix illustrate the problem.

	Subject	Group	Type 1	Cat 1	Type 2	Cat 2	Type 3	Cat 3	Type 4	Cat 4	Type 5	Cat 5	Type 6	Cat 6
1	1	1	regret1	C	regret2	T	regret3	T	cleft1	C	cleft2	T	cleft3	T
2	2	1	regret1	C	regret2	T	regret4	T	cleft1	C	cleft2	T	cleft4	T
3	3	1	regret1	C	regret2	T	regret5	T	cleft1	C	cleft2	T	cleft5	T
4	4	1	regret1	C	regret2	T	regret6	T	cleft1	C	cleft2	T	cleft6	T
5	5	1	regret1	C	regret3	T	regret4	T	cleft1	C	cleft3	T	cleft4	T
6	6	1	regret1	C	regret3	T	regret5	T	cleft1	C	cleft3	T	cleft5	T

Table 5 Distribution of contrasts between controls and targets by trigger

	Sentence type	Mean	≤ 0.01	$0.01 \leq 0.05$	$0.05 \leq 0.1$	$0.1 \leq 0.5$
1	*regret*	0.326	0.041	0.10	0.10	0.50
2	cleft	0.388	0.015	0.065	0.076	0.51

Suppose we want to study the contrast between controls and targets for *regret*. We have then to gather the results for controls and targets and run a non-parametric test. However, the experiment is such that subjects evaluate controls *and* targets. For instance subject 1 sees two targets—stimuli number 2 and 3—for *regret*, and one control—stimulus number 1—for *regret*, subject 2 sees stimuli number 2 and 4 as targets for *regret* and stimulus number 1 as control for *regret*, etc. We might consider a paired non-parametric test, typically a Wilcoxon test for paired samples. Unfortunately, this solution is far from optimal because it requires that we have enough data points to calculate two means—for controls and targets—for each subject and subjects see only *three* stimuli in the *regret* (or cleft) category.

A better option is to divide the set of subjects into two complementary sets and to compare the scores of controls for one of the sets with the scores of targets for the other. By doing this, we get more data points. E.g., we have 32 observations for *regret*-target and 16 observations for *regret*-control for the first 16 subjects. Of course, we have to repeat the non-parametric comparison many times to make it independent from any particular choice of subsets. We randomly divided the 33 subjects of the experiment into two groups of 16 and 17 subjects and compared their scores over 10,000 such samplings. The end result is a distribution of contrasts between controls and targets for 10,000 independent samples of 16 and 17 subjects. Table 5 reports the means and the proportions of *p*-values for various intervals. All the tests assumed that controls are better than targets as the null hypothesis.[5]

As can be expected from the net results in Table 4, the non-parametric comparisons revealed no difference between targets and controls, see Table 5. The distribution of *p*-values after sampling suggests that, overall, clefts are easier to interpret than *regret*, see Fig. 2.

The question arises whether clefts and *regret*-sentences are significantly different. Since the stimuli for the two categories are themselves extremely different, in order to avoid self-adjustment effects on the subjects, it is difficult to answer in a really convincing way. Two points are worth noting, however. First, comparing clefts and *regret*-sentences aggregating the target and control conditions suggests that there is a difference. The *p*-value mean is practically zero with a 10,000-sample test and the distribution of the *p*-values is Gaussian around the mean, which excludes a possible effect of extreme values. Second, comparing clefts and *regret*-sentences under each condition (control/target) with a 10,000 sample test suggests the existence of a difference. The *p*-value is 0.08 for controls and 0.04 for targets, which is compatible

[5] The tests were carried out using the *npar.t.test* function of the *nparcomp* package with the method parameter set to *t.app*. The Mann–Whitney test is not reliable due the presence of ex-aequo.

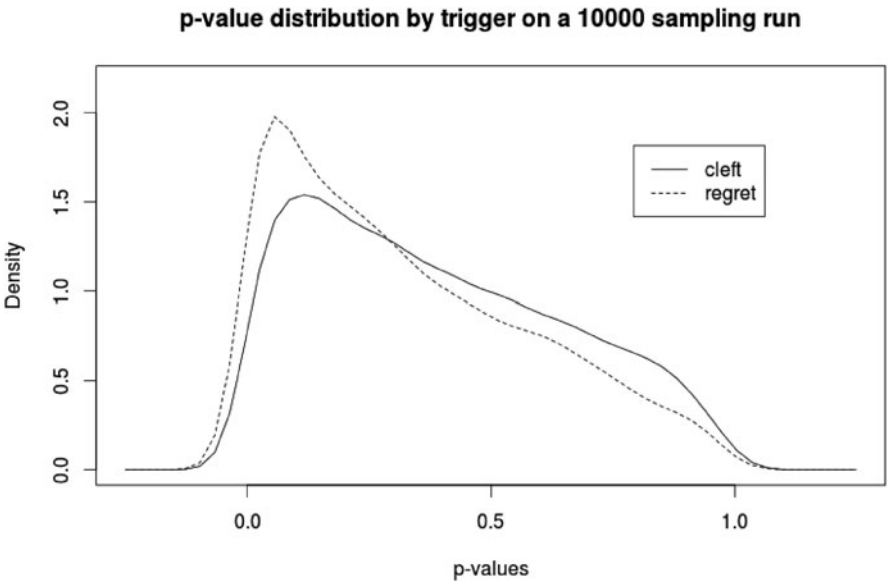


Fig. 2 *p*-value distribution

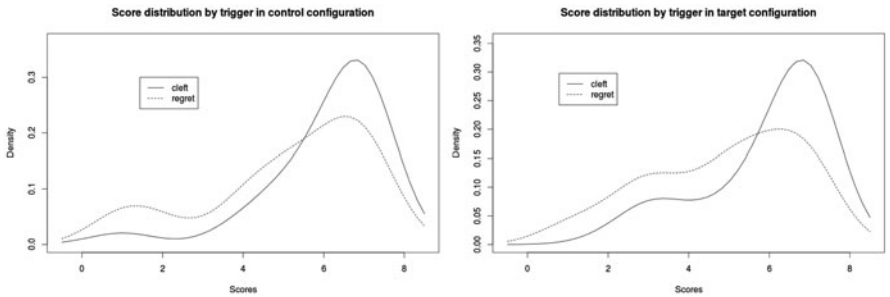


Fig. 3 Density plot by trigger under control and target conditions

with a global difference between clefts and *regret* and a slightly larger difference under the target condition.

It must be borne in mind that the stimuli for clefts and *regret*-sentences are different. In the absence of a comparison based on similar stimuli, there are at least two (mutually compatible) hypotheses: (i) clefts are easier to process because of their presuppositional profile, (ii) clefts are easier to process because the stimuli used in the experiment were more natural, simple, etc. If (i) was the main factor, we would expect to see the difference between clefts and *regret*-sentences decrease in the control condition. Figure 3 does not suggest that there is any difference.

Table 6 Logistic regression for condition and trigger

R pattern: *glm(Score∼Condition*trigger, family = "binomial", data = D)*	
binarization 1	binarization 2
interaction: 0.59	interaction: 0.78

Moreover, if we assimilate the ordinal response to a continuous one, it is useful to note that a linear model with scores as the response variable failed to detect any interaction between control and target conditions.[6] More convincingly, two binarizations of the response produced a similar result. We divided the scores into TRUE (FALSE) according to whether they were superior to 4 (≤ 4) (binarization 1) or whether they were superior ($\leq$) to 5 (binarization 2). A logistic regression model was fitted on the two binarizations with *glm*. The results are shown in Table 6. The results obtained by using *lmer* with subjects as random effect with respect to the intercept are practically identical.

Summarizing, the hypothesis that the difference between *regret*-sentences and clefts is due to the presuppositional profile of the two triggers is at best dubious. The alternative hypothesis that the difference is due to other factors is more plausible. Moreover, the two triggers give rise to accommodation in suitable contexts.

4 Discussion

4.1 General Discussion

The goal of the present work was to assess the robustness of the lexical weak/strong distinction based on Abusch's suspension test. The experimental results show that the distinction is not as robust as one may wish, since what is taken to be a characteristic of strong triggers, namely their resistance to accommodation under Abusch's configuration seems to evaporate in the presence of particular contexts. As noted by a reviewer, we did not compare directly strong and weak triggers, so we cannot conclude that the distinction is illusory, since there is always the possibility that some 'weak' triggers behave as strong triggers are predicted to behave. Our conclusion is, accordingly, more modest. We claim that there is at the moment no clear empirical evidence supporting the existence of a separate class of *lexically* or *conventionally* strong triggers and that it is not clear that triggers in general *encode* directly the strength/persistence/likelihood of projection. In contrast, it is possible that lexical information and context interact in certain ways that may produce the illusion of a

[6] The *lm* function in R with scores as dependent variable and interaction between condition and stimulus type as independent object provides a *p*-value of 0.87 for the interaction. A linear model calculated by *lmer*, with subjects as random effect with respect to the intercept, gave a low *t*-value of −0.187. So, the two models are consistent.

purely lexical distinction. This makes room for various distinctions between triggers, including prototypical weak and strong profiles.

The present results support a more precise claim in line with (Jayez 2014). There are in fact three categories of triggers. With triggers like *discover* or clefts, the presupposition and the main content are not independent of each other. The main content of clefts entails their presupposition: *It's Mary who solved the problem* ⇒ Mary solved the problem (main content) ⇒ someone solved the problem. The situation is the same for *win*: winning ⇒ participating. Concerning *discover* and similar terms, the main content makes the presupposition more probable. It has been noted by Simons (2007) that verbs like *discover* or *realize* imply accessing some evidence that the presupposition is true. In this respect, the examples used by Karttunen and others are similar to examples like (12). In (12), we observe exactly the same behavior as with *discover* and its mates, namely: (i) in positive assertions (12a), the conveyed information strongly implies that Paul missed the point and it is not possible to cancel the latter proposition (12b) and (ii) in suspension environments, there is no longer an implication (12c). With a negation, we have a choice between two interpretations: either we deny the existence of a proof, which amounts to suspending the presupposition in the case of *discover* or we deny that Paul is aware of the truth, which amounts to preserving the presupposition.

(12) a. Paul has a proof/conclusive evidence that he missed the point.
b. Paul has a proof/conclusive evidence that he missed the point but?? he didn't miss it.
c. If Paul has a proof/conclusive evidence that he missed the point, . . .

When the main content entails the presupposition or makes it much probable, the very mention of the trigger is sufficient to activate the presupposition, which predicts that the Abusch suspension test will not cause an impression of anomaly, a prediction which was borne out by the results about clefts. It is also expected that triggers for which the main content does not entail the presupposition or make it significantly more probable could be harder to process *in general*. This is compatible with the fact that *regret* had overall a lower score than clefts in our experiment, although, at this stage, it is impossible to exclude a confound with other factors, for instance the semantic content of the stimuli.

Verbs like *regret* or *stop* can be parceled together. The main content is independent from the presupposition. The differences between the triggers come from the relative degree of difficulty they present when a plausible context has to be constructed. For example, one might feel that it is impossible to find a situation similar to that of (2b) for *regret*, but it is only a matter of degree in context abduction, as shown by (13), where no presupposition emerges and the interpretation is that of a vague speculation, in the manner of (2b).

(13) I don't know what is wrong with the guy. Maybe he regrets having the job he has, or the wife he has, or whatever.

Lastly, there is a class of triggers like *too* or *again* that carry only the presupposition, as noted by Abbott (2006). Jayez (2014) argues that what we observe with such

triggers is a side-effect of a very general discourse constraint, studied in (Ducrot 1972; Jayez 2010; Simons et al. 2011). Non-main content information, including presupposition and implicatures is in general not involved in the normal, that is, non-metalinguistic, flow of discourse. For instance, non-main content information is not 'seen' by discourse markers and cannot be used naturally to answer a question. Operators like negation, *if*, question, etc., contradict or suspend the main content. They tend to ignore the non-main content and this what has been called 'projection'. Projection is stronger when the part of the message that carries the main content is distinct from the part that carries the non-main content. This is clearly the case with conventional implicatures, wich are, in general, separate from the main content (Gutzmann 2013; Potts 2005). For instance, expressives like *the stupid* N do not allow for suspension (14). Examples like (14) exhibit the same mechanism as Karttunen's example (6a). Since expressives are endorsed by the speaker, it is impossible for the same speaker to suspend in the ignorance sentence what he endorses in the rest of the discourse.

(14) I don't know whether Paul is stupid, but, if ??the stupid Paul . . .

Should we then expect that strong triggers behave basically like expressives or similar conventional implicature triggers? There are two sides to the answer. Since the suspension operators target primarily those parts of the message that convey main content information, they tend to ignore other parts, including pure presupposition triggers. This accounts for the fact that we have an impression of automatic ('strong') projection. However, because they are anaphoric,[7] triggers like *too* or *again*, are open to accommodation processes whenever the context provides enough 'independent' cues to make this accommodation plausible, where 'independent' means 'independent from the presence of the trigger itself'. Let us return to the *too* experiment to gain an intuitive understanding of how that would be possible.

When presented with the segment *I don't know whether Paul will go the party because* —, one can expect to find a reason why the speaker is ignorant about Paul's decision or why Paul might not go to the party. Concerning the latter interpretation, it is unlikely that one finds a reason why Paul might go to the party, as evidenced by the contrast in (15).

(15) a. I don't know whether Paul will go to the party because he does not like parties.
b. ?? I don't know whether Paul will go to the party because he likes parties.

As to the reason why Paul might no go to the party, there are two possibilities. First, Paul might run into some unexpected objective obstacle. For instance, he has missed his train or is sick, etc. Second, Paul might plan not to go although he has the possibility to go. In the first case, there is an independent cause that prevents Paul from going. In the second case, Paul has a reason not to go. Summarizing, we have the following possibilities (Fig. 4).

[7] This remark does not commit us to the view that every presupposition is anaphoric.

I don't know whether Paul will go to the party because —
- CAUSE OF IGNORANCE of the speaker
- CAUSE of Paul not going
- REASON for Paul not to go

Fig. 4 Possible families of interpretations after *because*

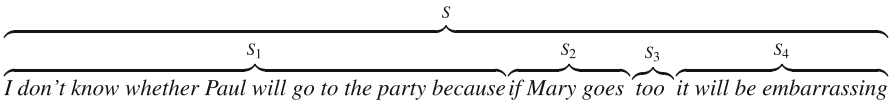


Fig. 5 Possible regions of interest in the sentence

Suppose that we divide the sentence into the regions of interest of Fig. 5. In terms of conditional probability, the three possibilities of Fig. 4 are mutually exclusive and can be represented as the probability of a certain interpretation given the region of interest (linguistic segment) that has just been processed. The interpretation proceeds by capitalizing on a growing chain of successive regions of interest (S_1, $S_1 + S_2$, etc.)

At $S_1 + S_2$, the CAUSE OF IGNORANCE interpretation becomes unlikely, since there is no clear relationship between Mary possibly going to the party and the speaker's state of ignorance. So, the CAUSE and REASON interpretations (Paul might hesitate because ...) are both better candidates. Given that the explicit part of the antecedent in the conditional is the hypothetical proposition that Mary goes to the party, we have two possibilities. Either the conclusion of the antecedent (the situation will be embarrassing) plays a role in allowing subjects to accommodate the missing hypothetical presupposition (Paul goes to the party), or this accommodation is already set at the moment the explicit antecedent (Mary goes to the party) has been processed. Both strategies are *a priori* possible and it is also possible that different subjects apply different strategies. The evaluation experiment does not allow one to tease apart the different processing scenarios. However, it is clear that subjects develop a sort of counterfactual reasoning, which amounts to inferring a reason for Paul for not going to the party from the consequence of two simultaneous events of Paul going and Mary going. How is that possible? Is it a property of the specific sentences we used in the experiment or something more general? Providing a reason for a non-action—here, not going to the party—involves in general a counterfactual reasoning about the possible negative consequences of the suspended action. Suppose for instance that Paul deliberately refrained from signing a document in order to block a project. Paul's non-action does not make much sense if we do not assume that signing the document would have increased the probability that the project is accepted, that is, if we do not take into consideration the effect of the contrary proposition. In the case of (8) and (9), the context makes it clear that one of Paul's goals is to avoid Mary as far as possible. Therefore, one can expect that, for some subjects at least, a counterfactual reasoning is on its way at $S_1 + S_2$ or $S_1 + S_2 + S_3$. So, although it is not possible to be more specific on processing issues (see Jayez and Mongelli (2013) for perspectives on this point), it is very likely that subjects engage at some point in a counterfactual reasoning which helps them abduce the missing presupposition.

4.2 Comparison to Other Work

Recent work by Cummins and colleagues (Cummins et al. 2013; Amaral and Cummins 2014) suggest a partly different picture than the findings reported above. (Jayez 2014) discusses (Cummins et al. 2013) at some length and the remarks made there extend to (Amaral and Cummins 2014) to a large extent. The main goal of these two papers is to investigate experimentally the difference posited by Zeevat (1992) between *resolution* and *lexical* presupposition triggers. Resolution triggers behave like anaphors. They look for a referent already introduced in the context or in the discourse. Lexical triggers refer to a concept whose applicability conditions correspond in part to the presupposition. For instance, the presence of *too*—a resolution trigger—invites one to find a suitable antecedent (an individual or an event), whereas *stop* refers to a transition between two states. The main intuition of Cummins and colleagues is that the negation of the presupposition does not affect a sentence containing a resolution trigger and a lexical trigger in the same way. For resolution triggers, the presupposition is not easily accommodated because it remains 'ambiguous', in Amaral and Cummins 2014 terminology. In contrast, lexical triggers provide more information as to the identity of the presupposition, which makes it more recoverable. Excluding a presupposition with a lexical trigger will affect (negatively) the asserted content, whereas, with a resolution trigger, the relation to the asserted content will remain weak. Accordingly, Cummins and colleagues predict that there are differences in acceptability between the sentence types illustrated in (16) and (17). B's answer in (16) amounts to endorsing the asserted content and refuting the presupposition, whereas C's answer amounts to refuting both. (17) has a similar structure, with a resolution trigger. Since negating the presupposition has more effect with lexical triggers than with resolution triggers, it is expected that C's answer in (16) will be perceived as better than in (17).

(16) A – Did Mary stop smoking?
B – Yes, although she never smoked before.
C – No, because she never smoked before.

(17) A – Did Mary watch the movie again?
B – Yes, although she never watched it before.
C – No, because she never watched it before.

This is indeed what Cummins and colleagues observe on a set of English triggers (Cummins et al. 2013) and their Spanish counterparts (Amaral and Cummins 2014). By and large, their observations and those reported in the present paper are consonant. For instance, in both approaches, *regret* is classified as lexical/weak. It seems that there are three main differences. The first is empirical. Since the experimental settings and the triggers under consideration are different, it is difficult to compare the observations.

The second one has to do with the pre-classification of triggers. the reason why *again* and *still* would be different remains unclear. Intuitively it seems that the presupposition can be very easily accommodated in both cases.[8] For instance, in (17), we accommodate immediately the proposition that Mary has watched the movie before. One might argue that this presupposition is too vague and that we need to know *when* Mary watched the movie (before). But, this would be asking for too much, since a sentence like *I know Mary watched the movie before and watched it again yesterday* is perfectly natural, although the temporal antecedent of *again* is very vague. The results provided by Amaral and Cummins (2014) are of two types. There are means on a five point Likert scale and there are differences between ratings for *Yes, although* and *No, because* patterns. The differences are significant for *lamentar* (*regret*), *seguir* (*continue*), *dejar de* (*stop*) and *todavía* (*still*), whereas they are not for *otra vez* (*again*) and *también* (*too*). However, when we focus on the *Yes, although* pattern, we see that (i) *todavía* gets practically the same score than *otra vez* and *también* and that *lamentar* and *dejar de* do not pattern with *seguir* and *todavía*. In our opinion, this calls for clarification and makes it perhaps difficult to interpret what is going on.

The third difference is a variant of the second. In the Abusch pattern, the content to be accommodated is in general transparent. For instance, in a sentence like*? I don't know whether Mary hit the target before, but if she hit it again, she is really gifted*, it is rather obvious that the proposition to accommodate in the antecedent of the conditional is 'Mary hit the target before'. In this respect, it does not seem possible to explain the difficulties of accommodation with strong triggers in terms of ambiguity or vagueness. The problem is rather with the possibility of the accommodation process itself, given its interaction with discourse coherence.

5 Conclusion

In this paper, we have shown that context plays a crucial role in the perception of the 'projection strength' of triggers, and that, as a result, projection strength can hardly be interpreted as an intrinsic lexical property of the various triggers mentioned in the literature. This is clear with allegedly strong triggers like *regret* or clefts. The situation of *too* is more complex since it is most probably a consequence of several properties, including the fact that it does not carry any main content information, as noted by Abbott, and the facilitation of accommodation under a counterfactual interpretation. Obviously, much work remains to be done, in particular extending the behavioral experiments to other triggers (e.g. factives), exploring the time course of accommodation, and extending the experimental coverage to other languages, as done in (Amaral and Cummins 2014) for Spanish.

[8] *Too* is a different case, see Jayez (2014).

Appendix

Experimental Items with too for Experiment 1

The context is in boldface

because
Paul et Marie se sont disputés et ne tiennent pasà se retrouver ensemble. Paul est invitéà une soirée, où il pense qu'il est possible que Marie ait été invitée. Je ne sais pas si Paul ira à la soirée parce que, si Marie y va aussi, ce sera très embarrassant.
Julien veut offrir un tee shirt à son neveu pour son anniversaire, mais il a peur que quelqu'un ait eu la même idée. Je ne sais pas si Julien offrira un tee shirt parce que, si quelqu'un en offre un aussi, ça sera décevant pour son neveu.
Nadine envisage de faire couper le sapin de son jardin, mais il ne resterait plus que celui de son voisin. Je ne sais pas si Nadine fera couper son arbre, parce que, si son voisin fait aussi couper le sien, il y aura trop de soleil l'été.
La Chine et la Russie veulent éviter une opération militaire contre la Syrie. Il faut que les deux pays donnent leur accord pour que l'intervention ait lieu mais aucun des deux n'est trop sûr du choix que ferait l'autre en cas de crise vraiment grave. Je ne sais pas si la Chine acceptera une intervention proposée par l'ONU, parce que, si la Russie l'accepte aussi, une opération sera lancée contre la Syrie.
Après la tempête de la veille, Louise est inquiète pour son petit bateau et descend au port avec son chien. Le bateau a l'air plein d'eau et Louise aimerait bien le vider. Le bateau supporterait certainement que Louise ou son chien montent, mais pas les deux. Je ne sais pas si Louise va monter dans le bateau parce que, si le chien monte aussi, le bateau risque de chavirer.

For *but*, the stimuli are the same, except that *mais* replaces *parce que*. The contexts are strictly identical. For instance, the first stimulus becomes: *Je ne sais pas si Paul ira à la soirée mais, si Marie y va aussi, ce sera très embarrassant*.

Experimental Items for Experiment 2

The context is in boldface

regret
Paul hésite à resigner dans son club parce que ça ne se passe pas très bien avec son président. Je crois que Paul a resigné dans son club. S'il le regrette plus tard, il ne pourra s'en prendre qu'à lui-même. [control] Je ne sais pas si Paul a resigné dans son club mais, s'il le regrette plus tard, il ne pourra s'en prendre qu'à lui-même. [target]
Véronique se demande si elle va acheter une voiture plus grosse que celle qu'elle a. Je crois que Véronique a acheté un modèle plus gros. Si elle le regrette ensuite, ce sera difficile de changer de nouveau. [control] J'ignore si Véronique a acheté un modèle plus gros mais, si elle le regrette ensuite, ce sera difficile de changer de nouveau. [target]
Céline se demande si elle va signer la pétition pour défendre une collègue de bureau. Je suis sûr que Céline a signé la pétition. Si elle le regrette après coup, ce sera trop tard. [control] Je ne sais pas si Céline a signé la pétition mais, si elle le regrette après coup, ce sera trop tard. [target]
Lucien ne sait pas s'il va voter pour le maire en place ou pour sa concurrente. Je pense que Lucien a voté pour le maire actuel. S'il le regrette par la suite, il ne pourra plus rien y faire. [control] Je ne sais pas si Lucien a voté pour le maire actuel mais, s'il le regrette par la suite, il ne pourra plus rien y faire. [target]
Ariane s'interroge sur l'utilité de faire couper la haie de son jardin. Je suis certain qu'Ariane a fait couper la haie. Si elle le regrette dans quelque temps, elle ne pourra pas revenir en arrière. [control] J'ignore si Ariane a fait couper la haie mais, si elle le regrette dans quelque temps, elle ne pourra pas revenir en arrière. [target]
Nathan ignore s'il va remplacer son gros ordinateur par un portable. Je pense que Nathan a acheté un portable. S'il le regrette, ce sera trop tard et il devra garder son portable. [control] Je ne sais pas si Nathan a acheté un portable mais, s'il le regrette, ce sera trop tard et il devra garder son portable. [target]
clefts
Jules est très vaniteux et ne peut s'empêcher de se vanter dès qu'il réussit quelque chose. Je pense que quelqu'un a trouvé la solution du problème. Si c'est Jules, on n'a pas fini d'en entendre parler. [control] Je ne sais pas si quelqu'un a trouvé la solution du problème mais si c'est Jules, on n'a pas fini d'en entendre parler. [target]
Une société craint d'avoir été espionnée par ses concurrents. Marie est particulièrement compétente sur les nouveaux projets secrets. Je suis sûr que quelqu'un nous a trahi. Si c'est Marie, nos concurrents seront très bien informés. [control] J'ignore si quelqu'un nous a trahi mais, si c'est Marie, nos concurrents seront très bien informés. [target]

Le village a été victime d'un incendie, peut-être dû à un simple court-circuit, après que le village voisin ait été victime d'un pyromane quelques semaines auparavant.
Je crois que quelqu'un a mis le feu. Si c'est le même pyromane que le village voisin, il faudra faire une enquête approfondie. [control]
Je ne sais pas si quelqu'un a mis le feu mais, si c'est le même pyromane que le village voisin, il faudra faire une enquête approfondie. [target]
Un père de famille ne retrouve plus son portable, qu'il doit mettre à jour, et se demande si un membre de la famille l'a emprunté.
Je suis certain que quelqu'un a pris mon portable. Si c'est Emma, il faudra attendre qu'elle rentre du lycée pour que je le mette à jour. [control]
J'ignore si quelqu'un a pris mon portable mais, si c'est Emma, il faudra attendre qu'elle rentre du lycée pour que je le mette à jour. [target]
Un policier se demande quelles sont ses chances de trouver des témoins après une agression.
Je suis convaincu que quelqu'un a été témoin de l'agression. Si c'est la voisine, on va pouvoir l'interroger tout de suite. [control]
Je ne sais pas si quelqu'un a été témoin de l'agression mais, si c'est la voisine, on va pouvoir l'interroger tout de suite. [target]
Un employé n'arrive plus à se connecter à son compte informatique.
Je pense que quelqu'un a changé le mot de passe. Si c'est mon collègue, il n'y a plus qu'à attendre qu'il arrive pour lui demander. [control]
J'ignore si quelqu'un a changé le mot de passe mais, si c'est mon collègue, il n'y a plus qu'à attendre qu'il arrive pour lui demander. [target]

References

Abbott, Barbara. 2006. Where are some of the presuppositions gone? In *Drawing the boundaries of Meaning,* ed. Betty J. Birner and Gregory Ward, 1–20. Amsterdam: John Benjamins.

Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presuppositions. In *Proceedings of Semantics and Linguistic theory XII,* ed. B. Jackson. Ithaca: CLC Publications, Cornell University.

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27:37–80.

Amaral, Patricia, and Chris Cummins. 2014. A cross-linguistic study on information backgrounding and presupposition projection. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

Beaver, David. 2004. Have you noticed that your belly button lint colour is related to the colour of your clothing? In *Presuppositions and Discourse. Essays Offered to Hans Kamp,* ed. R. Bauerle, U. Reyle, and T. E. Zimmerman, 65–100. Bingley: Emerald Group Publishing Limited.

Beaver, David I., and Bart Geurts. 2013. Presupposition. In *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), ed. Edward N. Zalta. Berlin: Springer. http://plato.stanford.edu/archives/fall2013/entries/presupposition/.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and Grammar: An Introduction to Semantics*. Cambridge: MIT Press.

Cummins, Chris, Patrícia Amaral, and Napoleon Katsos. 2013. Backgrounding and accommodation of presuppositions: An experimental approach. In *Proceedings of Sinn und Bedeutung 17,* ed. Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein, 201–218. Paris: ENS. http://semanticsarchive.net/sub2012/.

Ducrot, Oswald. 1972. *Dire et ne pas dire*. Paris: Hermann.

Geurts, Bart. 1999. *Presuppositions and Pronouns*. Amsterdam: Elsevier.

Geurts, Bart, and Rob Van der Sandt. 2004. Interpreting focus. *Theoretical Linguistics* 30:1–44.

Gutzmann, Daniel. 2013. Expressives and beyond. An introduction to varieties of use-conditional meaning. In *Beyond Expressives: Explorations in Use-Conditional Meaning,* ed. Daniel Gutzmann and Hans-Martin Gärtner, 1–58. Leiden: Brill.

Jayez, Jacques. 2010. Projective meaning and attachment. In *Logic, language and meaning. Revised Selected Papers of the 17th Amsterdam Colloquium, Amsterdam 2009, LNAI 6042,* ed. Maria Aloni, Harald Bastiaanse, Tikitu de Jager, and Katrin Schulz, 325–334. Berlin: Springer.

Jayez, Jacques. 2014. Orthogonality and presuppositions. A Bayesian Perspective. To appear in the proceedings of the ESSLLI 2013 workshop *Bayesian Natural Language Semantics and Pragmatics*. http://perso.ens-lyon.fr/jacques.jayez/doc/orthogonality-and-PP.pdf

Jayez, Jacques, and Valeria Mongelli. 2013. How hard are hard triggers? In *Proceedings of Sinn und Bedeutung 17,* ed. Emmanuel Chemla, Vincent Homer and Grégoire Winterstein, 307–324. Paris: ENS. http://semanticsarchive.net/sub2012/,

Karttunen, Laurie. 1971. Some observations on factivity. *Papers in Linguistics* 4:55–69.

Konietschke, Frank. 2012. nparcomp: Perform multiple comparisons and compute simultaneous confidence intervals for the nonparametric relative contrast effects. http://CRAN.R-project.org/package=nparcomp.

Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.

Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presuppositions. *Lingua* 117:1034–1056.

Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts. 2011. What projects and why. In *Proceedings of Semantics and Linguistic Theory XX,* ed. Nan Li and David Lutz, 309–327. Ithaca: CLC Publications (eLanguage).

Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and Philosophy*, ed. Milton K. Munitz and Peter Unger, 197–214. New York: New York University Press.

van der Sandt, Rob, and Bart Geurts. 2001. Too. In *Proceedings of the 13th Amsterdam Colloquium*, ed. Robert van Roy, Martin Stokhof, ILLC, 180–185.

Zeevat, Henk. 1992. Presupposition and accommodation in update semantics. *Journal of Semantics* 9:379–412.

Symmetry and Incrementality in Conditionals

Florian Schwarz

Abstract A central debate in presupposition theory concerns the nature of presuppositions introduced by triggers in conditional sentences. While it is commonly assumed that triggers in the consequent of conditionals give rise (or at least can give rise) to conditional inferences, most traditional accounts assume that triggers in antecedents introduce a non-conditional presupposition. This view has been challenged by recent modular accounts, which argue that the basic projection pattern involves conditional inferences across the board, but that non-conditional inferences can come about due to a processing bias towards at al incrementality. This paper presents an experimental investigation using the covered box paradigm that further assesses the availability of conditional presuppositions for conditional sentences containing a trigger in their antecedent. The results are in line with symmetric account, but are challenging for classic dynamic accounts. However, it may be possible to reconcile the latter with the data by tying together linear order and incremental context update.

Keywords Presuppositions · Presupposition projection · Conditionals · Incremental processing · Experimental pragmatics · Conditional strengthening

1 Introduction

One of the key properties of presuppositions is that they remain unaffected by various embedding operators, i.e., some of their embedded occurrences make contributions at the global level where other types of content would only have a local effect relative to the embedding operator (Karttunen 1973). Conditional sentences constitute a core case where this phenomenon, commonly labelled *presupposition projection*, occurs. Beyond the mere observation that presuppositions project out of conditionals, there are additional issues concerning the exact nature of the projected content. In particular, theories differ with regards to whether they take the presupposition itself to be conditional or not. Recent experimental work by Chemla and Schlenker (2012) argues for uniformly conditional presuppositions, which they analyze in terms of a

F. Schwarz (✉)
Department of Linguistics, University of Pennsylvania,
619 Williams Hall, 255 S. 36th Street, Philadelphia, PA 19104, USA
e-mail: florians@ling.upenn.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_9,

symmetric approach to projection. Asymmetries based on left-to-right ordering can be captured when such an approach integrates an additional incremental component based on left-to-right processing.

The present paper reports new experimental work investigating the nature of presuppositions in conditionals further, using a different methodology and extending the range of cases looked at to conditionals with a final (as opposed to initial) *if*-clause. I begin by reviewing the basic theoretical issues and the details of the experimental work by Chemla and Schlenker (2012). Next I introduce the experimental paradigm used here and the design of the experiments. Taken together, the results provide further evidence for a symmetric view, while also showing that the effects of the incremental processing component are rather strong. An additional aspect of the results is a surprising difference between *if*-initial and *if*-final conditionals, which I suggest is due to a variation in strength of conditional perfection inferences, i.e., inferences that strengthen conditionals to bi-conditionals. Beyond the theoretical implications of the results, the paper also makes a methodological contribution by demonstrating the viability of a picture matching task including a covered box (Huang et al. 2013; Romoli et al. 2011) for the phenomena at hand.

1.1 Theoretical Background

As a detailed introduction to an analysis of presuppositions and their projection behavior goes beyond the scope of the present paper, my presentation of the basic facts will be tailored towards the expressions and constructions at play in the experiment below. (For a recent survey of the state of the art, see Beaver and Geurts (2012) among many others).

The presupposition trigger *again* introduces, roughly speaking, a presupposition that an event of the sort characterized by the clause that *again* adjoins to has already occurred on a previous occasion, which may need to be sufficiently salient (or at least accessible) in the discourse context. For example, *again* in (1a) introduces the presupposition in (1b).

(1) a. John went to the movies again on Thursday.
    b. There is a (sufficiently salient) occasion prior to Thursday where John went to the movies.

As is characteristic of presuppositions, this part of the overall conveyed meaning of (1a) remains constant when the relevant clause is embedded under various operators:

(2) a. It is not true that John went to the movies again on Thursday.
    b. Did John go to the movies again on Thursday?
    c. If John went to the movies again on Thursday, he went to a concert on Friday.

Note that the asserted content of the clause—that John went to the movies on Thursday—is no longer conveyed by any of these examples, in contrast with the presupposition, which is. Homing in on the case of conditionals, the construction

of central concern to us, there is disagreement in the literature, however, as to what exactly is presupposed. In introducing this debate, it is helpful to first step back and consider cases where a presupposition trigger occurs in the consequent of a conditional:

(3) a. If John had company this week, he went to the movies again on Thursday.
b. If John went to the movies last weekend, he went to the movies again on Thursday

Notably, these sentences seem to differ in terms of their presuppositions, as only the first version presupposes a previous trip of John's to the movies. This difference is standardly attributed to the fact that only the latter explicitly introduces information in the antecedent that supports the presupposition of the consequent. One common approach to accounting for this, starting with early work by Karttunen (1974) and Stalnaker (1974), and later developed into dynamic theories of semantics (Heim 1983, and following work), is to assume that presuppositions (and clauses more generally) are evaluated relative to the preceding context, which includes earlier parts of complex sentences. The explanation, in a nutshell, for the absence of the presupposition in (3b) is that the consequent (including its presupposition) is interpreted relative to a context that already has incorporated the antecedent. If the presupposition of the consequent follows from the antecedent, then the sentence as a whole therefore has no (substantive) presupposition. Another way of characterizing this line of analysis is that presupposition triggers in the consequent of a conditional give rise to conditional presuppositions (i.e., *if p then* $\underline{q}q'$, with a presupposition q, gives rise to the presupposition *if p then q*).

For cases such as (3a), these analyses predict a conditional presupposition as well, which may undergo some form of strengthening to account for the perceived unconditional presupposition. Note that this treatment has been subject to extensive debates in the literature, typically discussed under the label of the 'proviso problem' (see, e.g., Geurts 1999, Beaver 2001, Schlenker 2008, 2009). We will put this debate aside, as our main focus is on presupposition triggers in the antecedent of conditionals, to which we turn now.[1]

Given the dynamic view just sketched, there is a crucial difference between conditionals (in canonical order) with presuppositions in the consequent and those in the antecedent. The latter have no preceding clauses within the same sentence, and their presupposition should thus be evaluated unconditionally relative to the discourse in place prior to the overall conditional sentence. In particular, as noted above,

(2c) If John went to the movies again on Thursday, he went to a concert on Friday.

[1] For recent experimental work on the proviso problem, see Romoli et al. (2011), whose experimental evidence is argued to support a conditional presupposition.

is standardly taken to introduce the non-conditional presupposition that John had gone to the movies prior to Thursday. Intuitively, this seems to be the correct result (at least at first sight; see below for more detailed considerations), and therefore, the asymmetric perspective on presupposition triggers in the consequent and antecedent of conditionals has often been considered part of the success story of a dynamic view on presupposition projection.

However, recent years have seen a flurry of renewed theoretical interest in presuppositions, which has given rise to work questioning some basic tenets of (broadly speaking) dynamic accounts. One major challenge is based on the insight that there is a lack of explanatory adequacy, as the properties of the lexical entries of connectives utilized by dynamic semantics that are crucial in accounting for presupposition projection by no means follow from the basic setup of the theory. In other words, variants of the dynamic entries for connectives can be construed which predict unobserved projection behavior. Of particular interest to us in the body of work that has emerged from this observation is that various authors have challenged the notion that incrementality should indeed be hard-wired into the definitions of connectives, as is done on a Heim-style proposal.

In addition to the problem of explanatory adequacy, there has been substantial disagreement over the years about how exactly incrementality should be implemented for various operators and connectives. For disjunction, for example, several possibilities have been proposed. Beaver (2001), for example, argues for a non-conditional presupposition for triggers introduced in the first disjunct and a conditional one for triggers introduced in the second. In contrast, Geurts (1999) argues that triggers introduced on either side of *or* yield non-conditional presuppositions for the overall sentence.

One line of response to these issues has been to posit what is sometimes called a modular approach to presuppositions. The basic idea is that while incrementality has an obvious role to play in presupposition projection, it is not a hard-wired constraint that cannot be overcome, but rather a processing bias. In a nutshell, presuppositions prefer to be supported by information present in the preceding context, but if support is introduced later, that is better than if there is no support at all. As far as projection is determined by the semantics of the various connectives, then, we end up with a symmetric account. Any asymmetry is introduced by the processing bias, which can be overcome at a cost.

To make this picture somewhat more concrete, let me briefly recapitulate Chemla and Schlenker's (2012) characterization of such a modular account. Following Fox (2008), they assume a supervaluationist framework, where the requirement for presuppositional acceptability is that a given sentence receive a classical truth value (true or false; as opposed to #) in every possible world (in the context). Specifically, for conditionals this line of thought goes as follows (assuming a material implication analysis for convenience): if evaluating a case with a presupposition trigger in the consequent and looking at a world where the presupposition is not met and in which the antecedent is true, we cannot determine whether the overall sentence is true or false, because of the undetermined truth value of the consequent (A true antecedent combined with a false consequent would make the conditional false, whereas with

a true consequent, it would be true). If looking at a world where the presupposition is not met and the antecedent is false, however, the truth of the entire sentence is independent of the truth-value of the consequent, since any conditional with a false antecedent is true (on a material implication analysis). This yields the equivalent of the conditional presupposition posited by dynamic semantic accounts, namely that *if p, then qq′* (with *q* as a presupposition of the consequent) presupposes *if p, then q*. For presupposition triggers in the antecedent, we derive a conditional presupposition as well, however, in contrast with dynamic approaches. The reasoning is similar: if evaluating such a conditional in a world where the consequent is true, then the truth value of the antecedent does not matter—whether it is true or false, the entire conditional is true. But when the consequent is false, the truth-value of the antecedent is crucial, and if it cannot be determined due to its presupposition not being met, the truth-value of the entire conditional cannot be determined either (if the antecedent is true and the consequent is false, the entire conditional is false, but if the antecedent is false, the entire conditional is true). Thus, the type of alternative account we are considering predicts that *if pp′, then q* presupposes that *if not q, p*. We end up with conditional presuppositions both for conditionals with a presupposition trigger in the consequent and for ones with a trigger in the antecedent. Summing up the crucial contrast between dynamic and symmetric account schematically, we have the following predictions for the presuppositions of a sentence with a presupposition trigger in the antecedent:

(4) Sentence Type: If pp′, q (with presupposition p of the antecedent)
Presupposition predicted by
a. dynamic account: p
b. modular account: if not q, p

As already mentioned, this is not the full story on the modular account, however. There is no denying that incrementality plays a role for presupposition evaluation. But rather than hard-coding it in the definition of connectives, a modular account can simply assume that incrementality arises as a processing bias. It sure is easier, the story goes, to have a presupposition supported in the preceding context, but if all else fails, support that is introduced later is better than no support at all. Dynamic accounts, which have a strict incremental component cannot allow for such later support (but see discussion for possible reconsiderations in this regard). For modular accounts, some processing cost is expected for such cases, but they are not ruled out categorically. Chemla and Schlenker (2012) present experimental results that they argue favor the violable view of incrementality that modular accounts afford. I turn to a discussion of their results and their interpretation in the following section.

1.2 Experimental Background

The seminal work by Chemla and Schlenker (2012) constituted a first attempt at getting experimental data to bear on the question of which of the two types of theories

outlined above is more appropriate empirically. In addition to conditionals, they also considered disjunctions and sentences with *unless*. Their presupposition trigger of choice is *too* (or rather, it's french counterpart, *aussi*), and they lay out both a carefully worked out theoretical analysis of *too* and a host of cautious choices in the exact construction of their stimuli. For reasons of space, I have to refer the reader to their paper for further details, and will simply present the conditional versions of the sentences they investigated:

(5) a. ***too* in consequent** If Anne decides to study abroad, her brother too will make a stupid decision. (*literally:* ... her brother will-make him too a decision stupid)
b. ***too* in antecedent** If Anne's brother too does not make a reasonable decision, Anne will not decide to study abroad. (*literally:* If the brother of Anne NE make not him too a decision reasonable, ...)[2]

Their design is set up to compare presupposition triggers in the antecedent to ones in the consequent. To prevent any differences in asserted content from coming into play, one of the sentences is the contraposition of the other (*if p, q* is logically equivalent to *if not q, not p*). Based on the reasoning laid out above, both dynamic and modular accounts predict a conditional presupposition for (5a), namely *if Anne decides to study abroad, Anne will make a stupid decision.*[3] However, for (5b), the predictions come apart: a dynamic account predicts a non-conditional presupposition (that Anne will make a reasonable decision), whereas a modular account predicts a conditional one (that if Anne decides to study abroad, she will be making a reasonable decision).

In the main part of their study, they asked subjects to rate the robustness of paraphrases of these conditional and unconditional inferences based on the presupposition. For (5), the inferences were as follows:

(6) a. Unconditional inference: Ann will make a reasonable/stupid decision.
b. Conditional inference: Studying abroad would be reasonable/stupid of Anne.

These inference correspond to the question of whether the presupposition trigger in the one clause is seen as being supported by the other. Note that both the use of a paraphrased version of the conditional presupposition and the fact that the predicates in the two clauses are not identical ensures that there is no entailment relationship between the two inferences.

The results that Chemla and Schlenker (2012) report for the inference task seem to align rather well with the modular, symmetric account. For both versions of the conditionals (and also for the other conditions involving disjunctions in varying orders), subjects judged the conditional inference as more robust then the the unconditional

[2] Recall that the actual experimental materials were in French, so any awkwardness of the English wording here should not be of any concern.
[3] This is under the assumption that *too* requires an antecedent of some sort, and that the link between the distinct predicates in the two clauses can be pragmatically inferred.

one. This is unexpected for a dynamic account, as the second clause (5b) should not be able to play any role in supporting the presupposition in the first.

While much more could be said about the details of the experimental task and the materials utilized by Chemla and Schlenker (2012), two things are clear: first, the results constitute an impressive first instance of experimental support for a symmetric theory of presupposition projection, in contrast with the traditional role assumed for incrementality. But secondly, given the theoretical significance of the matter and the complexity of the issues and judgments involved, further experimental exploration of the general issue is warranted. In the following, I report two experiments that aim at further illuminating the extent to which conditional presuppositions are indeed what we are dealing with in conditionals containing a presupposition trigger in the antecedent.

2 Experiments on *Again* in Conditionals

While our approach also utilized an inferencing task, we aimed to test the relevance of the truth or falsity of the consequent of the conditional more directly. To do so, we employed visual contexts that allowed to control for precisely this. The overall task was a picture matching task, specifically one using the covered box paradigm. This paradigm has proven extremely useful for investigating subtle aspects of meaning (Huang et al. 2013; Romoli et al. 2011; Romoli and Schwarz 2014), as it allows designs where subjects have to decide whether a given picture is compatible with any remotely possible interpretation of a sentence.

In the task we used for the present studies, participants were told that they have to identify suspects based on random bits of intercepted communication that provide limited information about the suspect's activities in the past week. They were shown three pictures, one of them 'covered', and told that they were to match the sentence they saw with a picture. The instructions explicitly stated that only one of the three pictures could be a match, so that they only should choose the covered box if they considered neither one of the overt pictures a match for the sentence. The crucial (overt) picture will be referred to as the 'target' below. The other overt picture was a distractor that never was compatible with the asserted content (in particular, it constituted a case where the antecedent of the conditional would be true and the consequent false). Thus, we were interested in the extent to which subjects resorted to choosing the covered box over the target picture.

The literal, truth-conditional content of the antecedent of the conditional always was false of the target picture. Assuming a material implication analysis, this means that the entire conditional was always true of it, at least as far as it is literal, truth-conditional meaning was concerned. We included the presupposition trigger *again* in the antecedent of the conditional, and varied whether the presupposition was met or not in the target picture. Therefore, whether or not subjects choose the target picture as a match will only hinge on whether they see it as compatible with the presupposition(s) introduced by the sentence. The truth and falsity of the consequent

was also varied across experimental conditions, as we were interested in the impact of this factor on the sentence's presuppositionality.

Given the difference between symmetric and dynamic accounts discussed above, we get distinct predictions for the case where the consequent is false. In particular, a dynamic account predicts that the presupposition of the antecedent should be globally present independently from the truth or falsity of the consequent. Symmetric approaches, on the other hand, take the overall presupposition to be a conditional one, of the form *if not q, p* (based on the conditional *if* pp'*, q*). This predicts that it should only matter whether the target is consistent with $\overline{p}$ if the consequent is false. When it is true, it should not matter whether the target matches p or not.

In addition to looking at conditionals in canonical order (*if . . . , [consequent]*), we also looked at the reverse order (*[consequent], if . . .*). This allows for a more direct investigation of the role of incrementality. The symmetric account of Chemla and Schlenker (2012) does not deny that incrementality plays a role in presupposition projection. However it sees it as a mere processing constraint that can be violated. In conditionals in canonical order, they would plausibly still predict a difference between cases where the presupposition is met and ones where it is not, even if the consequent is true, since the truth of the consequent only enters the picture late in the game, as it were—in particular, after the presupposition in the antecedent has been processed. Conditionals in reverse order remove this processing obstacle, as the truth of the consequent is fully established prior to encountering the presupposition trigger. Note that this manipulation is different from the one used by Chemla and Schlenker, who employed contrapositions to vary whether the presupposition trigger appeared in the first or second clause. Their conditionals were always *if*-initial and varied whether the presupposition trigger appeared in the antecedent or the consequent, while we focus on cases with a trigger in the antecedent and vary whether the *if*-clause comes first or last.

2.1 Materials and Design

Turning to the specific nature of our materials, we used figures of people together with a small calendar strip that included iconic representations of activities, destinations, and food items, which were explained to stand for the relevant activities (e.g., trips to, or consumption of, in the cases of destinations and food items) taking place on that day. Illustrations of the images used are provided in Fig. 1. The sentence in (7a) illustrates the corresponding sentence in the canonical order, and (7b) the reversed variant.

(7) a. If John ate a banana again on Wednesday, he ate a strawberry on Friday.
b. John ate a strawberry on Friday, if he ate a banana again on Wednesday.
c. **Presuppositions:**
   i. ↪ John ate a banana before Wednesday.
   ii. ↪ If John did not eat a strawberry on Friday, then he ate a banana before Wednesday.

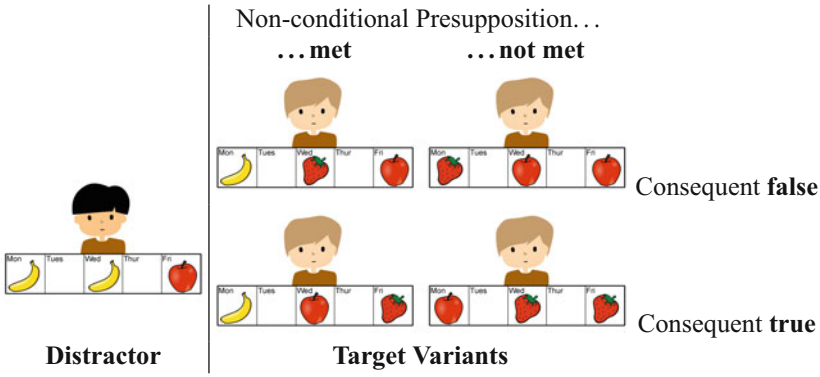


Fig. 1 Examples of distractor and target pictures by condition (relative to the sentence in (7). (Illustrations courtesy of Dorothy Ahn)

On both types of accounts that we are considering here, the basic presupposition should be the same regardless of the order of sentences. For dynamic semantics, the dynamic update formula that gives rise to presupposition projection presumably should not be altered based on order.[4] On a symmetric account, order of course also does not affect the basic logical structure of the sentence, based on which the prediction about what conditionals on the whole presuppose remains constant. However, given the role of incrementality in processing, there is room for acknowledging a difference, which will come into play specifically when a consequent in the initial clause position is true. In that case, any presuppositions in the antecedent can be ignored, and given the clause-order, this is already known at the time the relevant presupposition triggers are encountered. Thus, there should not be any processing cost associated with such a case.

In terms of the concrete predictions of each account, we thus end up with the following picture: a dynamic account predicts that regardless of order and truth or falsity of the consequent, the target should be chosen more frequently when the presupposition in the antecedent is met in the target picture, i.e., a main effect of the presupposition factor. It does not predict an interaction, nor a main effect of the truth-value of the consequent. A symmetric account with an incremental processing component, on the other hand, predicts that both the order of the clauses and the truth-value of the consequent should affect frequency of target choices, in addition to the presupposition factor. While the predicted effect of the latter is the same as on the dynamic account when the consequent is false, this is expected to disappear, or at least to decrease, when the consequent is true. Moreover, the case where consequent

[4] In fact, as Schlenker (2010) notes (in footnote 6, pp. 388–389), dynamic accounts make correct predictions for reverse cases with a presupposition trigger in the consequent, such as *The bathroom is well-hidden, if there is a bathroom.* But this is so, of course, only to the extent that the standard update formula is used for both canonical and reverse orders.

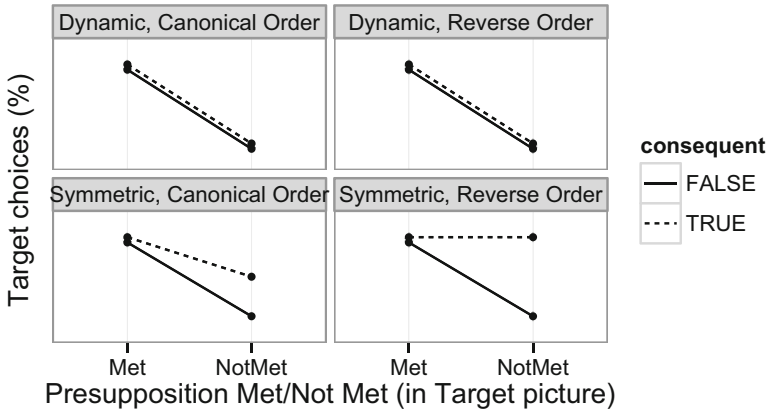


Fig. 2 Schematized predictions by condition and type of account

is true and the presupposition is not met (on the bottom right of Fig. 1) is expected to lead to more frequent target choices than the one where the consequent is false and the presupposition is not met (as on the top right version of the target in Fig. 1). Finally, based on the assumed processing cost of having a presupposition depend on material introduced later in the sentence, there still may be an advantage for the condition where the presupposition is met and the consequent is true over the one where it is not met (with a true consequent). But this effect should disappear once the clause-order is reversed. The symmetric account thus predicts an interaction between the presupposition factor and the truth-value of the consequent for both clause orders. Based on the effect of incrementality, this is expected to be more pronounced in the reverse order, which would furthermore be reflected in an interaction between clause order and the presupposition factor for the true-consequent conditions from both sub-experiments. These predictions are summarized schematically in Fig. 2.

A total of 24 items with target picture versions corresponding to the four variants in Fig. 1 were created, and for each item, there were sentence versions in canonical and reverse order. In addition to the experimental materials, there also were two types of fillers, as well as items from another unrelated sub-experiment. The fillers were conditionals similar to the experimental items. A first group of 12 fillers consisted of cases where both the antecedent and the consequent were true with respect to the target picture. Half of these contained the presupposition trigger *again*, whose presupposition was satisfied in the target picture. A second set of 12 fillers consisted of conditionals (without *again*) where both the antecedent and the consequent were false with respect to the target picture. These provided a useful check on subjects' willingness to choose targets in case the antecedent was false (more on this in the data treatment section below). Finally, there were 24 sentences that were experimental items (from an unrelated sub-experiments) and fillers containing the presupposition trigger *stop* (half of which contained negation).

2.2 Procedure and Participants

Given the necessary setup of having conditionals whose antecedent was false (as far as literal, truth-conditional content is concerned), we provided rather detailed instructions to ensure that participants are clear on the fact that conditionals can strictly speaking be seen as consistent with situations in which the antecedent is false. To further motivate subjects and give them a concrete sense of what their task is, we couched the experiment in a guessing game, where the subjects played the role of a detective. Thus, after signing a consent form, subjects were seated in front of a computer screen and read the following instructions.

(8) You are going to play a guessing game, in which you take on the role of a detective that is trying to identify suspects based on very partial information about what their activities are during a certain week. You will see three pictures showing different people and the things they did throughout the week. You have intercepted one sentence of communication by other people talking about the suspect's activities, which is from some time during the week in question. You assume that the source is reliable and that the sentence is true.
Based on that sentence, your task is to identify the one picture that is consistent with what the sentence says. Note that there will always be only one such picture.
One complication of the game is that one of the pictures will be blocked from your view, so that you can only guess what it depicts. But since there always will be only one picture that is consistent with the sentence, if none of the pictures that are visible are consistent with the sentence, then the hidden picture has to be the one, and you should choose it.
A further complication is that while sometimes it is straightforward to match one of the pictures with the sentence, other times this will involve a more indirect relationship between the two.
[Illustrations of true–true and false–true conditional-picture pairs]
Throughout the experiment, remember to evaluate the sentence and your options very carefully, so that you can be sure to identify the right person. Otherwise, you might lose your detective badge!
At the same time, though, you also should trust your gut feeling and go with what seems right to you, without over-thinking it for too long.
Let us do a couple of practice trials so that you can try out how this works!

After the initial instructions, the types of pictures used in the experimental stimuli where used to illustrate two cases of conditionals as paired with candidate pictures. The first case illustrated a basic, simple match, where both the antecedent and the consequent were true in the target picture. The second illustrated a case where the antecedent was false but the consequent true in the target picture, and it was explained in some detail why such a picture is not really inconsistent with the conditional at hand. Distractor pictures, where the antecedent was true and the consequent was false, were also discussed in both cases to further ensure they properly understood the nature of the task. There was a total of three practice trials. The first involved a simple

match, i.e., a case where both the antecedent and the consequent were true of the target picture. The second and third practice trials involved conditionals where both the antecedent and the consequent were false of the target picture. Subjects received feedback on the correctness of their picture choice after the second practice trial (with a 'correct' indication when they chose the target, and an 'incorrect' indication otherwise). Subjects were free to ask any general questions during the instructions, but did not receive any feedback during the experiment.

The order of trials during the experiment was randomized, constrained in such a way that no more than two subsequent items would come from the same sub-experiment (or filler group). Positioning of the target and distractor pictures as well as the covered box was counter-balanced across items. Participants made their selection via mouse-click.

The canonical and reverse versions of the conditional sentences were treated as a between subjects factor. In other words, we effectively ran two sub-experiment, each with 24 critical items. The clause-order of the conditional fillers was adjusted to match that of the experimental items. A total of 65 undergraduate students from the University of Pennsylvania, all native speakers of English, participated in the study for course credit. 34 of them saw the experimental items with conditionals in canonical order, and 31 in reverse order.

2.3 Results

2.3.1 Data Treatment

Responses were coded as to whether they corresponded to the target picture, the distractor picture, or the covered box. For statistical purposes, we created a binary response variable that was set to 1 whenever subjects chose the target, and to 0 when they did not.

As laid out above, a crucial element of the design employed was that subjects understood that when evaluating a picture with regard to its consistency with a stated conditional, the antecedent of the conditional being false strictly speaking provides no grounds for judging the two to be inconsistent. This was highlighted in the introduction to the experiment, but we were also able to test the extent to which subjects were able to follow these instructions by looking at the fillers where both the antecedent and the consequent of the conditional were false with respect to the target picture. As it turned out, there was a bi-modal distribution amongst subjects in this regard. While about two thirds of the subjects (23 in the canonical order group and 24 in the reverse order group) generally chose the target picture for these 12 items (92 and 93 % of the time respectively)—in line with the provided instructions—, roughly a third of the subjects (9 and 10 in the respective groups) did not in general do so (only 12 and 11 % of the time in the respective groups). Naturally, the subjects in the latter group also did not select the target picture in the experimental conditions, even when the presupposition was met. In the analyses that follow, we only report

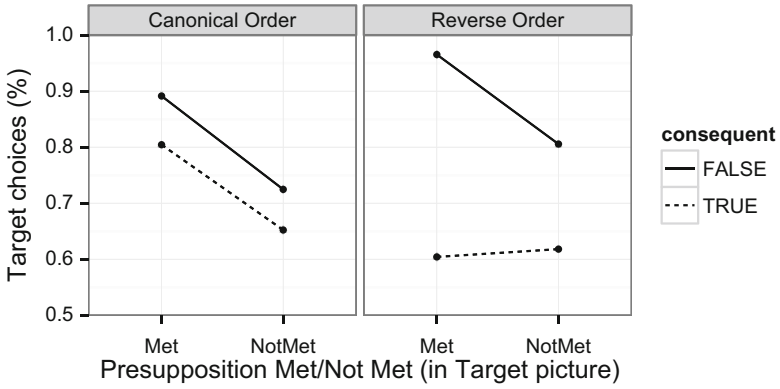


Fig. 3 Proportions of target choices by condition

results from subjects that selected the target picture for at least 7 out of the 12 items in this filler group.

2.3.2 Statistical Analysis

The mean target choice proportions by condition and sentence type are presented in Fig. 3. As is immediately apparent, the results are quite different in the two cases. I first present statistical analyses for the different clause-orders separately, and then proceed to some comparisons between the two. We conducted mixed effect model logistic regression analyses using the *lmer*-package in R (Bates 2005). Interaction analyses used centered values for the factor levels, whereas simple effects were calculated with the appropriate treatment coding. I provide estimates, standard errors, and p-values for each analysis.

Starting with the conditionals in canonical order, a 2×2 interaction analysis did not find a significant interaction but revealed main effects of both presupposition ($\beta = 1.28$, SE $= 0.27$, $p < 0.001$) and truth-value of the consequent ($\beta = 0.75$, SE $= 0.26$, $p < 0.01$), with higher target choice proportions when the presupposition was met and when the consequent was false. Planned comparisons between the relevant individual conditions furthermore confirmed that these effects were present at each level of the respective other factor (all p's < 0.05, with the exception of the simple effect of the truth-value of the consequent when the presupposition was not met ($p = 0.07$)).

At first sight, this result seems very much in line with a dynamic account. There is a uniform effect of whether or not the presupposition is met, regardless of whether the consequent is true or false. The main effect of the truth-value of the consequent could plausibly be seen as being due to the effect of pragmatic strengthening of the conditional, also known as conditional perfection (more on this below). Note in particular, that a symmetric account would lead us to expect the opposite result

for the conditions where the presupposition is not met in the target picture. If the consequent is true, this should be no obstacle for choosing the target picture on the basic symmetric account. And while there may be a processing bias against that, target choices in this condition still would be expected to be more frequent than when the consequent is false and the presupposition is not met, which is inconsistent with the basic presupposition assumed by the symmetric account. However, this initial description of the results for the canonical order conditionals needs to be revised in light of the outcome for the reverse order conditionals, to which we turn next.

In the data from the reverse conditionals, there is a significant interaction ($\beta = 2.41$, SE $= 0.64$, $p < 0.001$), such that while the presupposition factor has the same effect as in the canonical order conditionals when the consequent is false, it does not seem to have any effect when it is true. There also were significant main effects of the truth-value of the consequent ($\beta = 2.69$, SE $= 0.34$, $p < 0.001$) and the presupposition factor ($\beta = 1.10$, SE $= 0.32$, $p < 0.001$), though the latter clearly is dominated by the interaction and thus not generally interpretable. This picture is furthermore confirmed by planned comparisons, which revealed significant simple effects for all pairwise comparisons (p's < 0.001) except for the two conditions where the consequent was true.

In contrast with the finding for the canonical order conditionals, the presence of an interaction is very much in line with the predictions of symmetric accounts that integrate an incremental processing component. In particular, the processing cost that comes with having to wait for the truth value of the consequent in the canonical order is no longer present in the reverse order. Thus, we no longer predict a difference for the two conditions where the consequent is true, since the presupposition should not matter either way on this type of account. For a dynamic account, on the other hand, the interaction is entirely unexpected, as it predicts exactly the same outcome for both clause-orders.

A final statistical analysis looked at the results from the two sub-experiments together, focusing on the conditions where the consequent is true. Here we again find a significant interaction ($\beta = 1.62$, SE $= 0.80$, $p < 0.05$), as well as a significant main effect of presupposition ($\beta = 0.925$, SE $= 0.40$, $p < 0.05$ (which is dominated by the interaction). The presence of this interaction is again entirely in line with the predictions of symmetric accounts, which assume a processing cost to be present in the canonical order, but not in the reverse order. It is entirely unexpected, however, from the point of view of a dynamic account, which assumes a global presupposition to be present regardless of clause order and truth or falsity of the consequent.

2.3.3 Discussion

At first sight, the results from the two sub-experiments seem to be at odds with one another. The first experiment, with initial *if*-clauses, seems to neatly fit with dynamic accounts in that presuppositionality has an effect regardless of truth or falsity of the consequent. Furthermore, the fact that the condition with a true consequent and an unmet presupposition displayed fewer target choices than the one with a

false consequent and an unmet presupposition is inconsistent with the prediction of symmetric accounts, which assume that effectively no presupposition enters the picture if the consequent is false.

The second sub-experiment, on the other hand, exhibits an interaction, just as predicted by a symmetric account. This result is not captured by a dynamic account, as the order of the clauses should not matter. However, the exact nature of the outcome for the second experiment seems at least partly at odds with the predictions of a symmetric account as well. In particular, while we expected an increase in target choices when the consequent was true and the presupposition not met, as compared to the canonical order, we actually see a decrease in target choices when the consequent is true and the presupposition IS met.

But once we consider an additional factor that might be coming into play to affect the results, the aspects of the results that are problematic for symmetric accounts may receive an alternative explanation, thus leaving the symmetric account unchallenged (or at least fully consistent with the present data). The additional factor that I want to suggest is at play concerns the strength of conditional perfection. I spell this out in the following subsection, and briefly discuss results from a follow-up experiment that at least provides tentative support in this direction.

2.4 Follow-Up Study on Conditional Perfection

Starting from work by Geis and Zwicky (1971), it is commonly assumed that conditionals often end up being pragmatically strengthened to bi-conditionals. This (at least in part) accounts for the intuitive oddity of judging conditionals with a false antecedent and a true consequent to be true. This is, of course, highly relevant for our data, as the conditions where the presupposition is supposed to matter on all accounts are precisely instances of this distribution of truth-values (relative to the target image). The hypothesis I want to explore is that a) the presence of conditional strengthening will lower the frequency of Target choices in the conditions where the antecedent is false and the consequent is true; and b), that conditional strengthening has a stronger presence in conditionals where the *if*-clause is final than when it is in the initial position.

This hypothesis would help to explain the initially inconsistent seeming experimental results discussed above, in the following way: In the *if*-initial results, the fact that the true consequent condition where the presupposition is not met yields lower target choice frequencies than the corresponding false-consequent condition could be due to the presence of conditional strengthening. This, together with the assumption that the effect of incrementality is rather strong, could make the results from this study fully consistent with a symmetric account. The predicted interaction apparently has no room to show up, as it were, due to the presence and strength of the other factors. As for the results for the *if*-clause-final sentences, the interaction is already as expected on a symmetric account. The fact that the conditions with a

Table 1 Proportion of target choices (in %) by condition and reaction times (in ms) for target choices by condition

	Target choice %		Reaction times	
	False–False	False–True	False–False	False–True
if-initial	97.2	83.3	7178	8010
if-final	93.8	71.8	7364	9498

true consequent exhibit lower Target-choice frequencies could be attributed to the hypothesized increased strength of conditional perfection with final *if*-clauses.

To test this hypothesis, a follow-up study was conducted with items almost entirely parallel to the experiments above, but without *again*. The only other change was that we simplified the task by leaving out the distractor picture, thus having subjects choose between the target and the covered box. The comparison thus is between conditionals where both the antecedent and the consequent are false, which is consistent with conditional strengthening, and ones where the antecedent is true and the consequent is false. Relative to the pictures illustrated above, this schema would give the following sentences:

(9) a. If John ate a banana on Wednesday, he ate a strawberry on Friday.
b. John ate a strawberry on Friday, if he ate a banana on Wednesday.

Data from 40 subjects was collected. As in the first set of experiments, the condition where both the antecedent and the consequent were false served as a test of whether or not subjects were accepting target pictures in accordance with the instructions. There were eight subjects in the *if*-initial group and four in the *if*-final group that did not choose the target picture over half of the time in this condition, and which thus were excluded from the analysis.

As can be seen from the summary of results in Table 1, there were fewer target choices in the *if*-final condition with a false antecedent and a true consequent than in the the *if*-initial one, but this difference did not reach statistical significance. However, reaction times also yielded an interesting pattern of results that point in the direction of the hypothesis. In particular, target choices were significantly slower in the False–True condition than in the False–False condition for *if*-clause final sentences ($p < 0.05$), but not for *if*-clause initial ones.[5]

To relate the follow-up results directly to the initial experimental data, we also conducted an interaction analysis for target choice proportions comparing the conditions where the presupposition of *again* was met and the follow-up conditions, with the presence of *again* and clause order as factors. If the results for the sentences without *again* in terms of the frequency of target choices for both clause orders were

[5] But note that the corresponding interaction did not reach significance, though there was a significant main effect of condition. As a side note, it's interesting that we here seem to have a case where a literal interpretation—i.e., one without conditional strengthening taking place—yields slower response times than a pragmatically enriched one. This is, of course, in contrast with results on scalar implicatures, where responses based on pragmatically strengthened interpretations have generally been found to be slower. We leave further exploration of this for future research.

different, this would undermine the usefulness of the hypothesis that conditional strengthening has a stronger presence in *if*-clause final conditionals for explaining the original data. However, there was no significant interaction to that effect, which leaves the hypothesis as a viable option for explaining the relatively low frequency of target choices for the *if*-clause final sentences with true consequents from the original study.

In sum, the follow-up study provides some tentative evidence in favor of the suggested hypothesis, and does not yield any indications to the contrary. While further work is needed to establish this possibility more firmly, it does seem that the symmetric approach has a viable option for explaining all of the data from the original study, while dynamic accounts would seem to be at a loss, as things stand, to account for the presence of an interaction for the *if*-clause final sentences.

3 General Discussion and Conclusion

Following up on Chemla and Schlenker (2012), we investigated the effect of clause-order on the interpretation of presuppositions in the antecedent of conditionals. More specifically, we tested the prediction of symmetric accounts that the truth or falsity of a consequent of such sentences would affect the presuppositionality of the sentence as a whole. We utilized an inferencing task quite different from that in Chemla and Schlenker (2012), based on a version of the covered box picture matching task. Furthermore, we directly manipulated the linear order of the antecedent and the consequent in the conditional, which had not been previously done.

The main finding is that the position of the *if*-clause indeed has an effect on the presuppositionality of the entire sentence, in combination with the truth of the consequent. When the consequent is true, symmetric accounts indeed predict there to be no presupposition based on the core projection mechanism alone. However, given the role of incrementality in processing, the presupposition may still exhibit some presence when the true consequent follows the *if*-clause, as the presupposition is evaluated at that point relative to the preceding context. But when the true consequent precedes the *if*-clause, we are factoring out this processing effect. This is indeed what we find: whether or not the presupposition is met in the *if*-clause final conditionals with a true consequent does not affect the frequency of target choices, in contrast with all the other conditions.

While the relevant interaction for the *if*-clause final sentences was exactly in line with the predictions of a symmetric account, the low level of target choices in the relevant conditions presented a new puzzle. I suggest that this can be attributed to an increased strength of conditional perfection for such conditionals, and presented a first bit of tentative evidence along these lines. If this hypothesis indeed can hold its ground, the symmetric account ends up being entirely consistent with the results from the main studies reported here. Classic dynamic accounts, on the other hand, cannot account for the interaction for the *if*-clause final sentences. Note, however, that it may be perfectly possible to consider a variant of a dynamic account that takes

linear order into consideration. For example, Beaver and Geurts (2012) spell out an implementation along these lines (albeit with no reference to processing issues), by suggesting that the context change potential of logical operators should be represented in terms of negation and conjunction, where the order of the conjuncts reflects linear order. While such a proposal would have to be evaluated more generally in light of both the present results and more general projection data (see also Footnote 4 above for a potential problem for such an account), it seems like a welcome possibility for tying together incrementality and linear order on a dynamic account. From such a perspective, the present results would then not necessarily be a reflex of symmetry, but could simply be seen as resulting from the ever-present effect of incrementality. Note, however, that the results from Chemla and Schlenker (2012) are still not straightforwardly captured by such an account, as they find evidence for subjects seeing conditional inferences to be more robust even when the presuppositional support is introduced later. To do so, the adherence to linear order in updating would have to be seen as non-obligatory, as on the relevant variant of the symmetric account. In light of these issues, it is all the more important to assess to what extent Chemla and Schlenker's results generalize to other tasks and methodologies, and ongoing work in my lab is pursuing precisely this point.

In closing, it is worth noting that given the novel application of the experimental paradigm used here, further questions arise that will need to be looked at more closely in future work. In particular, it is noteworthy that even in the conditions where the presupposition is not met (and where all accounts predict a presupposition to be present for the entire sentence), subjects chose the target around 60 % of the time, even though it was inconsistent with the presupposition. In light of the present data, we can only speculate what exactly this hight acceptance rate is due to. It could, in principle, reflect cases of global accommodation, local accommodation, or be the result of subjects ignoring the presupposition altogether in their response behavior. Experimental techniques such as the ones used here should be able to settle the issue of which of these possibilities are indeed behind the present results, but we have to leave this issue for future work. For present purposes, the fact that we get significant variation in proportions of target choices between different conditions suffices entirely to evaluate the predictions of the accounts under consideration.

Acknowledgements Thanks to Rachel Stults, Jamie Fisher, and Robert Wilder for assistance with data collection and to Dorothy Ahn for the visual stimuli. The work reported here was in part supported by a grant from the University Research Foundation of the University of Pennsylvania.

References

Bates, Douglas M. 2005. Fitting linear mixed models in r. *R News* 5:27–30.

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

Beaver, David, and Bart Geurts. 2012. Presuppositions. In *Semantics: An international handbook of natural language meaning*. vol. 3. ed. Claudia Maienborn, Klaus, von Heusinger, and Paul Portner. Berlin: Mouton de Gruyter.

Chemla, Emmanuel, and Philippe Schlenker. 2012. Incremental vs. symmetric accounts of presupposition projection: An experimental approach. *Natural Language Semantics* 20 (2): 177–226.

Fox, Danny. 2008. Two short notes on Schlenker's theory of presupposition projection. *Theoretical Linguistics* 34 (3): 237–252.

Geis, Michael, and Arnold Zwicky. 1971. On invited inferences. *Linguistic Inquiry* 2:561–566.

Geurts, Bart. 1999. *Presuppositions and pronouns*. Amsterdam: Elsevier.

Heim, Irene. 1983. On the projection problem for presuppositions. In *Proceedings of the second West Coast conference on formal linguistics,* ed. D. Flickinger, et al., 114–125. Stanford: Stanford University Press.

Huang, Y., E. Spelke, and J. Snedeker. 2013. What exactly do number words mean? *Language Learning and Development* 9 (2): 105–129.

Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4 (2): 169–193.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1: 181–194.

Romoli, Jacopo, and Florian Schwarz. 2014. An experimental comparison between presuppositions and indirect scalar implicatures. In *Experimental Perspectives on Presuppositions*. ed. Florian Schwarz. Springer International Publishing.

Romoli, Jacopo, Yasutada Sudo, and Jesse Snedeker. 2011. An experimental investigation of presupposition projection in conditional sentences. *Proceedings of SALT* 21: 592–608.

Schlenker, Philipe. 2008. Presupposition projection: The new debate. In *Proceedings of SALT 18,* ed. T. Friedman and S. Ito. Ithaca: CLC Publications.

Schlenker, Philipe. 2009. Local contexts. *Semantics and Pragmatics* 3: 1–78.

Schlenker, Philipe. 2010. Presuppositions and local contexts. *Mind* 119 (474): 377–391.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and philosophy,* ed. Milton K. Milton and Peter K. Unger. New York: New York University Press.

An Experimental Comparison Between Presuppositions and Indirect Scalar Implicatures

Jacopo Romoli and Florian Schwarz

Abstract We compare two aspects of meaning, namely the presupposition of *stop* in the scope of negation (John didn't stop going to the movies. ↪ *John used to go to the movies.*) and the scalar implicature associated with the strong scalar item *always* under negation (John didn't always go to the movies. ↪ *John sometimes went to the movies.*) ('Indirect Scalar Implicatures' (ISIs); Chierchia, Structures and Beyond, 2004). In our results, ISIs are found to pattern with presuppositions in that responses reflecting an interpretation without an inference (corresponding to a 'literal' interpretation) are slower than ones based on the relevant inference (Chemla and Bott, Lang Cogn Process, 38:241–260, 2013), contrary to what has been found for direct scalar implicatures (Bott and Noveck, J Mem Lang, 51:437–457, 2004, among others). These results are puzzling from the traditional perspective that ISIs are generated in the same way as direct implicatures. We explore two possible interpretations: first, strong scalar terms could receive a presuppositional analysis as well and presuppose that their domain is non-empty. Alternatively, we could group stop and ISIs together from another angle and see them as obligatory scalar implicatures, in contrast to the non-obligatory direct ones.

Keywords Presuppositions · Implicatures · Indirect scalar implicatures · Processing · Experimental pragmatics · Covered box task · Reaction time · Presupposition projection · Local accommodation · Negation

1 Introduction

As participants of conversations, we draw a variety of inferences from the sentences we hear. Some of these inferences are thought to be associated directly with the linguistic forms uttered, others instead are considered to go beyond the basic or

J. Romoli (✉)
Room 17E10, School of Communication, University of Ulster,
Jordanstown Campus, Shore Road, Newtownabbey, Co. Antrim BT37 0QB, UK
e-mail: j.romoli@ulster.ac.uk

F. Schwarz
Department of Linguistics, University of Pennsylvania,
619 Williams Hall, 255 S. 36th Street, Philadelphia, PA 19104, USA
e-mail: florians@ling.upenn.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_10,

literal meanings of the sentences they arise from. For instance, from a sentence like (1a) we typically conclude that the information in (1b) is true. Analogously, from a sentence like (2a) we tend to infer that John used to show up late for class, (2b).

(1) a. John sometimes went to the movies.
b. ↪ John didn't always go

(2) a. John didn't stop showing up late for class.
b. ↪ John used to show up late for class

The type of inference in (1b) is generally called a 'scalar implicature', while that in (2b) is a instance of a 'presupposition.'[1] While it is controversial what the place of these inferences should be with respect to the semantics/pragmatics divide, they are not considered to be part of the 'literal' truth-conditional meaning of (1a) and (2b). The main reason for this is that while these inferences are typically drawn, there are cases in which they appear to be absent, as we will see below.[2] Moreover, how these inferences should be derived is far from being a settled matter in the literature. On the contrary, many theories have been proposed, some of which very different from each other. What is most relevant for us in connection to (1b) and (2b), however, is that the majority of these theories agree that these inferences are different in kind.[3]

In the psycholinguistic literature, inferences such as (1a) have been extensively studied, in particular from the perspective of how these inferences are processed (Bott and Noveck 2004; Bott et al. 2012; Breheny et al. 2006; Huang and Snedeker 2009; Chemla and Bott (In press)). More recently, work has also been conducted on presuppositions (Chemla and Bott 2013; Schwarz 2007; Schwarz and Tiemann 2012, 2013). One of the main findings of these studies on the processing of inferences exemplified in (1b) and (2b) is that they appear to have very different processing profiles—a difference that nicely reflects the theoretical distinction standardly posited in the literature. In the case of scalar implicatures such as (1a), there is evidence that they are associated with a processing cost when compared to the corresponding literal meanings. In response times studies, in particular, the delay associated with evaluating a sentence in a situation consistent with its scalar implicature has been consistently found to be longer than that of a situation inconsistent with the scalar implicature (but consistent with the sentence's literal meaning) (Bott and Noveck 2004 and much subsequent work). Similar results have been found with other methodologies (Breheny et al. 2006; Huang and Snedeker 2009; but see Grodner et al. 2010). In other

[1] The literatures on both topics are vast, so let us just point to recent overviews for further background reading: for scalar implicatures, see Chierchia et al. (2012) and Geurts (2010) and references therein; for presuppositions, see Beaver and Geurts (2012) and references therein.

[2] Notice that, strictly speaking, this fact alone does not imply that these inferences could not be part of the literal meaning of one of the possible interpretations of (1a) and (2a), if the latter are considered ambiguous. In fact, some accounts of scalar implicatures and presuppositions do posit some form of ambiguity to account for the optionality just mentioned. As the issue is not relevant for our discussion, we will ignore it here. See Chierchia et al. (2012) for discussion.

[3] There are some exceptions, in particular Chemla (2009b) and Romoli (2012), Romoli (In press). We will come back to these alternative approaches in the discussion section below.

words, there appears to be solid evidence in the literature that responding based on meanings enriched with scalar implicatures is more costly than responding based on literal meaning. While presuppositions have been studied much less than direct scalar implicatures, one common result of the studies conducted is that, unlike the case of scalar implicatures, verifying or deriving presuppositions is not more costly than verifying or deriving the corresponding literal meaning, on the contrary, it appears to be less costly. Chemla and Bott (2013), in particular, have found that sentence verification in contexts consistent with the sentence's presupposition was faster than in contexts inconsistent with it (i.e., in situations only consistent with the sentence's literal meaning).

Seeing the processing results for the two types of inferences in comparison, the effects seem to go in opposite directions. In short, we could characterize this as follows: while the presence of direct scalar implicatures appears to be more costly than their absence, for presuppositions it is their absence that appears more costly than their presence.

This difference in processing between presuppositions and scalar implicatures provides a useful diagnostic for evaluating the nature of other inferences (i.e., evaluating whether they are more like presuppositions or more like scalar implicatures). This is because one can investigate the processing profile of an inference and check where it stands with respect to this distinction (i.e., presence-more-costly-than-absence vs. absence-more-costly-than-presence).

In the experiment reported in this paper, this is precisely what we have done. We have investigated inferences exemplified by (3b), based on sentences such as (3a), against the background of this distinction. Following Chierchia (2004), we will label these 'indirect scalar implicatures', in contrast to the previously considered cases, which we will call 'direct scalar implicatures.'

(3) a. John didn't always go to the movies.
b. ↪ John sometimes went

To our best knowledge, the processing of indirect scalar implicatures has not been systematically investigated.[4] However, this is an important gap to be filled as it could tell us much about the relationship between them and direct scalar implicatures, on the one hand, and presuppositions on the other. Traditionally the inference in (3b) is generally regarded to be a scalar implicature of the very same kind as (1b) and different from the inference in (2b). This traditional view makes two predictions. First, it predicts uniform processing profiles for direct and indirect scalar implicatures.[5] Second, it predicts their processing and that of presuppositions to be different. In

[4] But see Chemla (2009c) for an offline study involving inferences of the type of (3a). For relevant work on the acquisition of indirect scalar implicatures see Musolino and Lidz (2006) and Katsos et al. (2011). Finally, after finishing this paper, the recent manuscript by Cremers and Chemla (2013) came to our attention. We'll have to leave a more detailed comparison with this work to another occasion.

[5] A possible complication in this regard involves the presence of negation and its potential processing implications. We will return to this briefly below.

light of what we know from the previous studies on direct scalar implicatures and presuppositions sketched above, we thus expect that the presence of indirect scalar implicatures should be associated with a higher cost than that of their absence, as with direct scalar implicatures, but in contrast with presuppositions.

We tested these predictions by looking at the processing of indirect scalar implicatures like (3b) in direct comparison with that of presuppositions and against the background of the results in the literature on direct scalar implicatures. As we will see in detail below, the results of our study challenge the prediction that direct and indirect scalar implicatures should pattern uniformly (and support the previous finding on presuppositions). We find the processing of indirect scalar implicatures to actually be more similar to that of presuppositions than that of direct scalar implicatures.

The paper is organized as follows: in Sect. 2, we briefly introduce direct and indirect scalar implicatures as well as presuppositions in more detail, and sketch how they are traditionally derived. In Sect. 3 we present our experiment and its results. In Sect. 4, we discuss the implications of the experimental findings and explore two alternative ways of looking at indirect scalar implicatures in theoretical terms in light of our results.

2 Background

2.1 Basic Properties of Implicatures and Presuppositions

Our example in (1), repeated below, provided a first illustration of direct scalar implicatures. Descriptively speaking, inferences of this sort arise when a weak scalar terms like *sometimes* appears in an upward entailing context. Further examples are provided in (4) and (5), where the sentences in (4a) and (5a), which contain the scalar terms *some* and *or*, respectively, give rise to the implicatures in (4b) and (5b).

(1) a. John sometimes went to the movies.
b. ↪ John didn' t always go

(4) a. Some of the students went to the movies.
b. ↪ Not all of them went

(5) a. John went to the movies or to the beach.
b. ↪ John didn't go both to the movies and to the beach

Analogously, and in a symmetric fashion, indirect scalar implicatures arise from strong scalar terms embedded in downward entailing environments. We already saw the example in (3); other cases include (6) and (7), where the sentences (6a) and (7a) containing the strong scalar terms *all* and *and* give rise to the implicatures in (6b) and (7b).

(3) a. John didn't always go to the movies.
b. ↪ John sometimes went

(6) a. Not all of the students went to the movies.
b. ↪ Some of the students did

(7) a. John didn't both go to the movies and to the beach.
b. ↪ John went to one or the other

One main characteristic of scalar implicatures is that they can easily be suspended. More neutrally, we will say that they appear to be 'absent' in certain cases. For instance, in (8a) and in (8b), the direct scalar implicature that John didn't always go to the movies, for (8a), and the indirect one that John went sometimes to the movies, for (8b), are not present (at least at the end of the continuations which directly contradict them).

(8) a. John sometimes went to the movies . . .
In fact, he always did!
b. John didn't always go to the movies . . .
In fact, he never went!

In brief, direct and indirect scalar implicatures are suspendable inferences of sentences such as (1a) and (3a) and, at least superficially, they are symmetrical and very similar. Indeed, as we will soon see in more detail, they have been treated in a completely unified way. Before sketching the traditional way of deriving these inferences, we turn to presuppositions in a bit more detail.

The example in (2), repeated below, provided a first illustration of a presuppositional inference. Other inferences of this sort include (9b) and (10b), which are associated with sentences like (9a) and (10a), containing an it-cleft and an achievement verb like *win*, respectively.

(2) a. John didn't stop showing up late for class.
b. ↪ John used to show up late for class

(9) a. It is John who showed up late for class.
b. ↪ Somebody showed up late for class

(10) a. John didn't win the marathon.
b. ↪ John participated in the marathon

The main characteristic property of presuppositions is their behavior in complex sentences (Karttunen 1973). This is their so-called 'projection' behavior, which can be described as follows: if we consider a sentence like (9a) and we embed it under negation (11a), in the antecedent of a conditional (11b), under a possibility modal (11c) or in a question (11d), we still draw the inference in (9b). In traditional terminology, the inference in (9b) 'projects' through the embeddings in (11a)–(11d). This projection behavior is generally taken to be a characteristic feature of presuppositions and it is used as a diagnostic for presuppositionality (Chierchia and McConnell-Ginet 1990; Beaver 2001).

(11) a. It isn't John who showed up late for class
b. If it is John who showed up late for class, he should apologize.
c. It's possible that it is John who showed up late for class
d. Is it John who showed up late for class?

Similarly to scalar implicatures, we can easily construct cases in which presuppositions appear to be suspended, or 'absent,' in the terminology we used above. In (12), at least after hearing the continuation, we certainly do not conclude that the presuppositional inference that John used to show up late for class is true.

(12) John didn't stop showing up late . . .
because he never did!

As noticed above, this is not a property of all types of inferences. Compare, for instance, the behavior of entailments, as exemplified by (13): the attempt of suspending/contradicting the inference that John sometimes went to the movies, in parallel to what was done above, sounds contradictory.

(13) John sometimes went to the movies last week . . .
#In fact he never went!

Summing up, presuppositions are inferences of sentences like (8a) that are not strictly speaking obligatory, and which display a characteristic projection behavior in complex sentences. The theoretical goals of a theory of presuppositions and of scalar implicatures is to explain how these inferences arise and to predict precisely in what circumstances they are arise. In the next section, we turn to sketch the traditional ways in which this is done.

2.1.1 Traditional Derivation of Implicatures and Presuppositions

In light of the theoretical goals just stated, we now proceed to briefly sketch what could be described as the 'traditional' take on each of these inferences. This will be a highly simplified and somewhat idealized description, not the least because many of the theories that our description loosely encompasses are very different from each other. Nonetheless, it will suffice for our purpose of illustrating something like the standard theoretical treatment of these inferences.

Starting with scalar implicatures, we can simply use an idealized Gricean algorithm, as represented in (14) (Grice 1975). As (14) indicates, the basic idea is that when we hear an utterance, we reason about what the speaker might have said instead (among a restricted set of competitors). We then conclude that some of these competitors are false. More precisely, the competitors that we deem false are those that are stronger than the speaker's utterance.

(14) a. The speaker said A.
b. The speaker might have said B.
c. It' s false that B.

One question that arises at this point is how to determine competitors—that is how do we determine B in (14). A simple response is the following: certain words or morphemes, sometimes called 'scalar terms,' are associated with others in the lexicon.

For example, *sometimes* is associated with *always*, *some* with *all*, and *or* with *and*.[6] When we have a sentence containing one or more scalar terms, we can obtain sentential alternatives by replacing the scalar terms in question with their associates. For instance, if the assertion is (15a), we can obtain its sentential competitor in (15b) by replacing *sometimes* with *always*. Once we have (15b), we can apply our algorithm in (14) and derive the inference in (16).

(15) a. John **sometimes** went to the movies.
b. John **always** went to the movies.

(16) John didn't always go to the movies.

While we are glossing over many details here, what is important for us is that the very same ingredients that we used above for deriving the direct scalar implicature in (16) can automatically account for indirect scalar implicatures like (17b) as well.

(17) a. John didn't always go to the movies.
b. John sometimes went to the movies.

To illustrate, consider the assertion in (17a): we can obtain the competitor in (18) by replacing *sometimes* with *always*.[7] (18) is stronger than the assertion in (17a), and therefore winds up being negated by the algorithm in (14). This yields the inference that it's not true that John didn't sometimes go to the movies, which is equivalent to (17b).

(18) John didn't sometimes went to the movies.

In sum, on the traditional view—and essentially in all theories of scalar implicatures that we are aware of—direct and indirect scalar implicatures are derived in the very same way. Before discussing how presuppositions are derived, let us sketch how cases where the implicature winds up being absent can be derived, e.g., in cases like (8a) and (8b), repeated from above.

(8a) John sometimes went to the movies. . .
In fact, he always went!

(8b) John didn't always go to the movies. . .
In fact, he never went!

As already mentioned, on the traditional view implicatures arise as reasoning about what the speaker could have said instead of what she actually said. This perspective can, therefore, easily account for the absence of scalar implicatures. This is because if the speaker makes it clear, as in the continuation in (8a) and (8b) above, that the

[6] See Horn (1972); Rooth (1992); Sauerland (2004) among many others. For a more articulate theory of alternatives see Katzir (2007); Fox and Katzir (2011).

[7] Notice that the positive polarity nature of *sometimes* makes (18), if asserted, marginal if not completely infelicitous in its interpretation in which *sometimes* takes narrow scope with respect to negation. As it is generally implicitly done in the literature, we will assume that this is not a feature that applies to alternatives.

competitor is true, then as hearers we will obviously not reason that she must think it to be false.[8]

Turning our attention to the derivation of presuppositions, there are many different proposals in the literature on how to derive them, just as with in the case of scalar implicatures. One approach, stemming from work by Stalnaker (1974, 1978), Karttunen (1974), and Heim (1983), is to consider presuppositions as definedness conditions on the 'update' of the context by the information associated with the assertion. So for instance, a sentence like (19) can only update a context that already entails the information that John used to show up late for class. If the context doesn't entail this information, either the sentence sounds infelicitous, or this information is understood as an inference of the sentence.[9]

(19) John stopped showing up late for class.

Beyond basic cases such as (19), a theory of presuppositions needs to explain the behavior of presuppositions in complex sentences as well. In the approach above, in particular in Heim (1983), this is done by redefining the semantics of connectives and quantifiers in such a way as to derive the projection properties of presuppositions. In particular, this is done by identifying the meanings of sentences with the ways in which they change the contexts they update (also called 'context change potentials'). Given its identification of the meanings of sentences with the way they change the context they occur in, this approach is called 'dynamic.' It can be shown that this way of proceeding ensures that not only is (19) predicted to presuppose that John used to show up late for class, but also (20a)–(20d) are.

(20) a. John didn't stop showing up late for class.
b. If John stopped showing up late for class, Mary will be pleased.
c. It's possible that John stopped showing up late for class
d. Did John stop showing up late for class?

Many other theories of presuppositions have been proposed as well, of course, from partial or trivalent logics to DRT-style approaches Kamp (1981), to more recent pragmatic accounts (for a recent overview, see Schlenker (2008)). But what is most important for our purposes is that these mechanisms are all different from those assumed for deriving scalar implicatures.

As discussed in Sect. 2.1, both implicature- and presupposition-based inferences can be absent in certain circumstances. Given the different theoretical perspectives on the two types of inferences, the absence of presuppositions in cases like (21) is generally derived in a way that's very different from the way absent inferences are derived in the case of scalar implicatures. Theoretically speaking, presuppositions are generally seen as non-negotiable inferences (at least within the accounts we're considering here), and thus always have to be taken into account. How can this,

[8] In other cases, other considerations enter into the picture, in particular a notion of 'relevance' is used, so that if it is clear in the context that the competitor would have not been relevant for the goals of the conversation, the scalar implicature is not derived.

[9] In this perspective, the presupposition is said to be 'accommodated' in the context (Lewis, 1979); see von Fintel (2008); Beaver and Zeevat (2012) and references therein for further discussion.

then, be reconciled with their apparent absence in cases like (21)? The key idea is that presuppositions can be interpreted at different levels: when they are computed globally, we expect an inference, but we do not if they are computed in more local positions.

(21) John didn't stop showing up late for class . . .
because he never did!

For instance, imagine—for the sake of presentation—that your favorite theory of presuppositions is a function **PS**, which, given sentences of any complexity, returns their presuppositions: for any p, $\mathbf{PS}(p) = p$ and p's presuppositions. Imagine also that **PS** can recursively be applied at any scope site of a sentence. For instance, for a sentence like (22), repeated from above, we have two possible scope sites at which **PS** could apply. First, it could apply at a global level, as in (23a) and this would give rise to the inference that John used to show up late for class. Second, however, it could also apply vacuously at a local level below negation, as in (23b). This latter option is what corresponds to the 'suspension' or the 'absence' of presuppositions in the traditional view, as this gives rise to an interpretation that is compatible with John never having showed up late for class.

(22) John didn't stop showing up late for class.

(23) a. **PS** [not[John stopped showing up late for class]]
b. not [**PS** [John stopped showing up late for class]]

In a case like (21), given that the continuation is incompatible with the global derivation of the presupposition, it is assumed that the presupposition is instead derived locally, as in (23b). In other words, in this approach the presence vs. absence of presuppositions lines up with their global vs. local derivation.

This concludes our rough sketch of possible derivations of scalar implicatures and presuppositions, as well as of accounts of cases where the respective inferences are absent. While there are many different alternative implementations, what is most relevant here is that the on essentially all accounts, the stories are quite different for implicatures and presuppositions. In the next section, we turn to a brief summary of what has emerged from studies on the processing of direct scalar implicatures and presuppositions in the literature and their implications for an investigation of indirect scalar implicatures.

2.2 Processing Implicatures and Presuppositions

Among the three inferences that we are looking at in this paper, direct scalar implicatures are the ones that have been studied most extensively in the psycholinguistic literature (Bott and Noveck 2004; Bott et al. 2012; Breheny et al. 2006; Huang and Snedeker 2009; Chemla and Bott 2013). A central result that has emerged from these studies is that the processing of direct scalar implicatures appears to be costly. More precisely, evaluating a sentence in a situation consistent with its direct scalar

implicature appears to be associated with a higher processing cost than evaluating it in a situation incompatible with its direct scalar implicature (Bott and Noveck 2004; Breheny et al. 2006; Huang and Snedeker 2009; but see Grodner et al. (2010)). More specifically, in reaction times studies, the reaction time associated with evaluating a sentence in a situation compatible with its scalar implicature was consistently longer than evaluating it in a situation only compatible with its literal meaning. We can summarize this result as in (24) (Bott and Noveck (2004) among others).

(24) **Result on direct scalar implicatures**
presence of direct scalar implicatures > absence of direct scalar implicature

The processing of presuppositions has been studied less than direct scalar implicatures, but the number of studies investigating it has been growing recently (Chemla and Bott 2013; Schwarz 2007; Schwarz and Tiemann 2012, 2013). One common result that has emerged from these recent studies is that, contrary to direct scalar implicatures, it is the absence of presuppositions that gives rise to a higher processing cost. In other words, evaluating a sentence in a situation compatible with its presupposition is associated with a longer reaction time than evaluating it in a situation incompatible with its presupposition.[10] We can formulate this result as in (25).

(25) **Result on presuppositions**
presence of presupposition < absence of presuppositions

As already mentioned, the processing of indirect scalar implicatures has not been investigated in the literature. Given what we know from the studies just described and the unified treatment of direct and indirect scalar implicatures sketched above and repeated in (26), we expect indirect scalar implicatures to behave like direct ones in terms of processing. In other words, everything being equal we expect that evaluating a sentence in a situation consistent with its indirect scalar implicature should be slower than evaluating it in a situation which is inconsistent with its indirect scalar implicature, (27).

(26) **Traditional grouping:**
direct scalar implicatures = indirect scalar implicatures ≠ Presuppositions

(27) **Prediction for Indirect scalar implicatures**
Presence of indirect scalar implicatures > absence of indirect scalar implicatures

In the next section, we report on an experiment in which we tested the prediction in (27). In particular, we tested (27) in direct comparison to the case of presuppositions and against the background of the results in the literature about direct scalar implicatures.

[10] Notice that this does not mean that presuppositions *per se* are not associated with a cost. See Schwarz and Tiemann (2013) for discussion and relevant results on the processing of presupposition projection.

3 Experiment: *Stop* vs. *Not Always* in the Covered Box Paradigm

In our experiment, we compared indirect scalar implicatures to presuppositions by testing for the availability of different interpretations for strong scalar items and presupposition triggers under negation using the covered box paradigm. This paradigm is particularly well-suited for investigating the existence of non-dominant interpretations and has already been fruitfully utilized in the study of implicatures (Huang et al. 2013) and presuppositions in conditionals (Romoli et al. 2011). It is a variant of a picture-matching task, where participants have to judge the fit of a picture with a given linguistic stimulus. The covered box version adds a layer to the ask by including a (representation of) a covered box, which participants are told hides another picture. In our variant, they are also told that only one of the pictures matches the sentence. If they find none of the overtly shown pictures to match the sentence, the match must be the hidden one and they should select the covered box. The methodology is particularly useful for testing for the availability of non-dominant interpretations, because participants are forced to consider whether an overtly shown picture matching such an interpretation could possibly be seen as corresponding to any available interpretation of the sentence in question. If they choose the covered box instead, this is a clear indication that the relevant interpretation is not available to them. Dominant interpretations can be included as well, of course, and we also included controls to ensure that participants did indeed understand the task and choose the covered box in cases where none of the overt pictures matched any interpretation of the presented sentence.

Our implementation of this paradigm compared cases corresponding to overall interpretations that either did or did not include an inference of interest. In the case of indirect scalar implicatures, this was the inference that some activity took place some of the time when the sentence said that it did not always take place. In the case of presuppositions, it was the global interpretation of the presupposition of *stop* under negation that some activity had been going on prior to the time mentioned in the sentence. The comparison cases involved no such inference, i.e., they involved target pictures where the activity did not go on at all (in the case of indirect scalars) or didn't go on prior to the mentioned time (in the case of *stop*. Illustrations of the actual stimuli used are introduced in the following subsection.

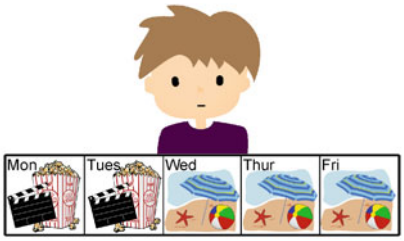
Based on the traditional picture sketched above, which assumes that indirect scalar implicatures are equivalent to direct scalar implicatures in both their derivation and in processing, we'd expect response behavior results parallel to the well-studied case of the latter. In particular, we'd expect that the generation of the scalar inference comes with a processing cost that is reflected in an increase in (when comparing trials where the inference is drawn to ones where it is not), and possibly also affects the frequency with which it is chosen as a match. With presuppositions, we expect a different outcome, both based on the standard assumption in the theoretical literature that global interpretations of presuppositions are the default and the initial processing results by Chemla and Bott (2013), where global interpretations are argued to be faster than local ones. Based on these expectations, we thus predict a cross-over interaction, with slower for inference trials with *always* and faster ones for inference trials with *stop*.

3.1 Materials and Design

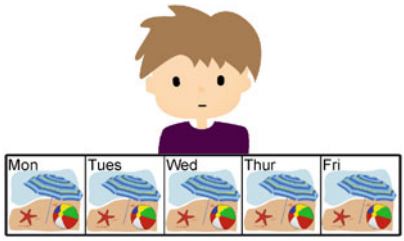
Our materials utilized pictures depicting individuals together with a 5-day calendar strip that contained icons representing various activities or destinations. Subjects were told that these represented what the individuals did during the past week, and they had to identify which of them match the provided linguistic description (see below for procedural details). The critical manipulation varied whether the target picture corresponded to an interpretation that included the inference of interest or not. Illustrations are provided in (28):

(28) a. Benjamin didn't always go to the movies last week.
↪ Benjamin sometimes went to the movies last week
b. Benjamin didn't stop going to the movies on Wednesday
↪ Benjamin went to the movies prior to Wednesday

i. Inference

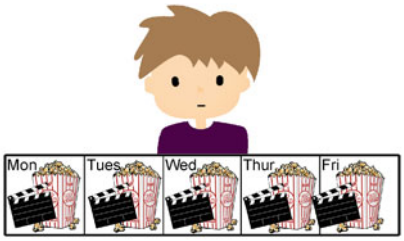


ii. NoInference

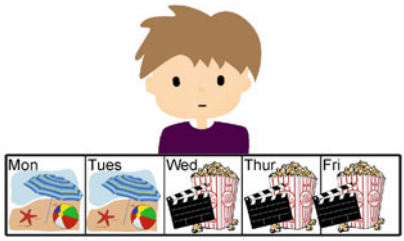


(Illustrations Courtesy of Dorothy Ahn)

i. Inference



ii. NoInference



(Illustrations Courtesy of Dorothy Ahn)

The inference target picture in the *always*-condition matched the inference at stake for the sentence in (28a), as this person did go to the movies some but not all of the time. The no-inference picture, on the other hand, does not match this inference and thus should only be selected if the inference is not present. Similarly, the inference *stop*-picture matches the inference in that this person did go to the movies prior to Wednesday (and continued to do so from then on, to match the asserted content as well). But the no-inference picture did not match the inference, as this person did not go to the movies prior to Wednesday. Note that the asserted content - that they did go

to the movies from Wednesday on - is still matched by the no-inference picture, so that the participants' decision about whether this picture matches the sentence is a clear indicator of whether or not the inference based on the global presupposition is present or not. Note furthermore that under a local interpretation of the presupposition, this picture does match the sentence, as the negation of the conjunction *Benjamin went to the movies before Wednesday and he went to the movies on Wednesday and thereafter* is indeed true of it.

In addition to the target and the covered box, a third distractor item was included as well. Distractors included a picture of another individual and were constructed so as not to match either interpretation of the respective sentences (e.g., the distractor for the *always*-conditions had a boy that went to the movies every day, and the one for the *stop*-condition had a boy that went to the beach from Monday through Thursday and to the movies on Friday).

A total of 12 items were created with variants in all four conditions described above. Furthermore, an additional experimental manipulation involving the interpretation of the verb *think* under negation was included, which we will not discuss here in detail for reasons of space. In brief, a sentence such as *Benjamin doesn't think he will go to the movies on Thursday* can, but does not necessarily, give rise to the inference that Benjamin is certain that he won't go to the movies on Thursday. We varied pictures here by either including another activity on Thursday or placing a question mark there. Given the future-orientedness of these stimuli, they were presented in a separate block of the experiment where participants were told that the calendar strips represented the corresponding individuals' plans for the coming week. Order of blocks was counter-balanced between participants.[11]

In addition to the experimental stimuli, several types of filler items without negation were included. First, there was a set of 12 items with sentences containing *stop* without negation. For half of these, the target picture matched the sentence, and for the other half it did not and thus required selection of the covered box. Similarly, there were 12 items using *always* without negation, again split in half with respect to whether or not the target matched the sentence. In the block on *think*, parallel control variants without negation were included.

3.2 Procedure and Participants

After signing a consent form, participants were seated in front of a computer, where the experimental program prepared in SR Research's Experiment Builder was started. They were then shown instructions that told them that they were taking on the role of a detective in a guessing game, where they would see three pictures, one of them blocked from view, and a sentence, which was a piece of intercepted communication

[11] For those curious about the results, we found no differences between inference and no-inference trials for this sentence type, in contrast with the other two, as discussed below. We have to leave the interpretation of this result for another occasion.

from a suspect about their activities during a certain week. Based on that intercepted information, they were to decide who was the suspect, under the assumption that only one of the three pictures was compatible with the sentence, and that the covered box would therefore need to be chosen if neither one of the overt pictures matched the sentence. The choice was carried out by using a mouse to click on the selected picture. Prior to displaying the images and the sentence, a fixation cross was presented in the center of the screen for 1000 ms.

After the instructions, the first block started with two practice trials to familiarize them with the task. One block consistent of the *stop* and *not always* items and fillers, and the other of the *think* items. Each block had its own practice trials, as they differed in that one was about the past week and the other about the upcoming one.

The expression factor was used as a within subject factor, i.e., participants saw three out of the six variations of each item, one with each expression (*stop*, *not always*, *think*). Whenever they saw the *stop* version of an item in the inference condition, they saw they *always* version in the no-inference condition and vice versa. Each subject completed a total of 36 experimental trials, (24 and 12 in the respective blocks), as well as 36 filler trials (again 24 and 12 in the different blocks).

Twenty-five undergraduate students from the University of Pennsylvania, all native speakers of English, participated in the experiment for course credit. They were randomly assigned to four groups for counterbalancing both the inference vs. no-inference conditions and the block order. Each subject saw six items in each condition.

3.3 Results

3.3.1 Data Treatment

For purposes of analysis, responses were coded with respect to whether participants selected the target picture or the covered box. Trials where the distractor was chosen were considered as errors and were removed (2.5 % of trials fell into this category, leaving 585 out of 600 data points for analysis). Reaction times were calculated as the time that passed from initial display of the images and the sentence until the mouse click occurred.

3.3.2 Statistical Analysis

The average proportions of target choices by condition is depicted in Fig. 1. As can be seen from the graph, target choices were much less frequent in the No Inference condition, both for *stop* and *always*: While the Inference conditions were at ceiling with close to 100 % target choices, there were only around 27 % (*always*) and 29 % (*stop*) of trials with target choices. Note that there also was a rather striking divide between participants with slightly over half of them (14) never making any target choices in the No Inference condition, and slightly less then half of them (11) making

Fig. 1 Mean percentages of target choices proportion by condition

target choices around two thirds of the time (for both *stop* and *always*). Furthermore, inspecting target choices relative to trial order, it appeared to be the case for most of the participants belonging to the latter group that they did not initially select the target on No Inference trials, but then switched and stuck with target choices from then on. There were two participants which only had target choices for *stop* in the No Inference condition, and one that only had target choices for *always* in that condition. All others had target choices for both expressions.

To analyze the pattern in responses statistically, we carried out mixed effect model logistic regression analyses using the *lmer* package in *R* (Bates 2005). While we tried to include random slopes for both factors and their interaction, following the recommendation of Barr et al. (2013), the corresponding models did not converge, and we thus only included random intercepts (for both participants and items). We report estimates, standard errors, z-values, and p-values from the lmer output. A 2×2 interaction analysis (*stop* vs. *always* and Inference vs. No Inference) using centered factors yielded a main effect of the Inference factor ($\beta = 6.73, SE = 0.62, z = 10.8, p < 0.001$). There was no main effect of expression and no significant interaction, but the effect of the Inference manipulation was slightly more pronounced in the *always* conditions in numerical terms. Planned comparisons for the effect of the Inference factor within each type of expression were carried out by using the appropriate treatment coding. There was a simple effect of the Inference factor for both *stop* ($\beta = 5.90, SE = 0.61, z = 9.66, p < 0.001$) and *always* ($\beta = 7.53, SE = 0.95, z = 7.91, p < 0.001$).[12]

In summary, targets were selected much less frequently in the No Inference condition. For the indirect scalar implicature with *not always*, this is somewhat surprising given previous results for direct implicatures. Using essentially the same task, involving picture selection with a covered box, Huang et al. (2013) found 87 % target choices in the equivalent of a No Inference (≈ literal) condition with the quantifier *some*. Furthermore, several studies conducted in our lab using a truth-value judgment task rendered similarly high acceptance rates in (the equivalent of) No Inference

[12] These effects were also present when looking only at data from the participants with target choices in the No Inference condition.

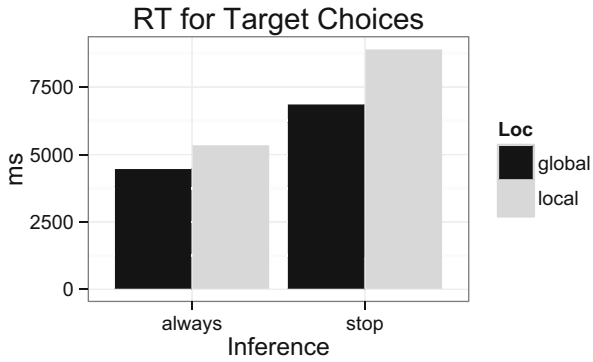


Fig. 2 Mean for target choices by condition

conditions. For presuppositions, there is much less precedent, though Chemla and Bott (2013) found about 50 % of (the equivalent of) No Inference responses in a truth-value judgment task with factive verbs such as *realize* and *know*. Looking at both expressions together, what stands out that they seem to behave quite similarly as far as participants' response behavior is concerned, even though we traditionally would not group them together, as discussed above.

Turning to reaction times, we again find similar patterns of results for the two expressions we investigated. The mean for target choices in the Inference and No Inference conditions are graphed in Fig. 2. For both *always* and *stop*, target choices in the No Inference conditions were slower than in the Inference conditions. To evaluate these differences statistically, we ran mixed effect model analyses, again using the *lmer* function in *R*. Statistical analyses used mixed-effect models with participants and items as random effects, using the *lmer* function of the *lme4* package in *R* Bates (2005). Following Barr et al. (2013), the maximal random effect structure that was suitable for the design and that would converge was used, with a random intercept as well as random slopes. To assess whether inclusion of a given factor significantly improved the fit of the overall model, likelihood-ratio tests were performed that compared two minimally different models, one with the fixed effects factor in question and one without, while keeping the random effects structure identical (Barr et al. 2013). We report estimates, standard errors, and t-values for all models, as well as the χ^2 and p-value from the likelihood-ratio test for individual factors.

A 2×2 interaction analysis, with random intercepts and random slopes for type of expression for participants and items and a random slope for the Inference factor for items, yielded significant a main effect of Inference ($\beta = 1517, SE = 437, t = 3.47$; $\chi^2 = 7.73$, $p < 0.01$), with slower in the No Inference condition. There also was a main effect of expression ($\beta = 2124, SE = 558, t = 3.81$; $\chi^2 = 11.40, p < 0.001$) with slower for *stop*, but no significant interaction. The main effect of expression includes reading times for the sentence, and thus may at least in part be due to differences in sentence structure. Planned comparisons between the Inference and No Inference conditions for *stop* and *always* suggest that the difference is present for both expressions, reaching significance for *stop* ($\beta = 2110$, $SE = 619$, $t = 3.41$; $\chi^2 = 4.16$, $p < 0.05$), and approaching it for *always* ($\beta = 948$, $SE = 618$,

$t = 1.53$; $\chi^2 = 2.37$, $p < 0.13$).[13] Taken together, the upshot of the reaction time results is that target choices in the No Inference condition are slower than in the Inference condition.

Given the split in participants with respect to whether or not they had any target choices in the No Inference condition at all, we conducted some post-hoc analyses as well to further investigate differences between these groups. The No Inference participants seemed to exhibit slower overall. Comparing for covered box choices in the local conditions revealed a significant difference between them and the group without any target choices in the No Inference condition ($\beta = 4365$, $SE = 1637$, $t = 2.67$; $\chi^2 = 6.70$, $p < .01$). Similarly, target choices in the global conditions were marginally slower for them as well ($\beta = 1162$, $SE = 659$, $t = 1.76$; $\chi^2 = 3.09$, $p < 0.1$)

There are some interesting methodological points to note as well, concerning the use of the covered box design for investigating response times. Note that the central comparisons of interest above are all comparisons between equivalent choices, all involving trials where the target picture was chosen. This allows for a more straightforward comparison between different interpretations, as it avoids issues such as the well-known response bias effect, which commonly yields slower response times for *false* answers in truth-value judgment tasks. Interestingly, covered box choices were not significantly slower than target choices (looking at the local conditions for the No Inference participants). Another interesting point to note in this regard is that comparing target vs. covered box choices for the No Inference participants yields a rather different pattern from the results discussed above. While the No Inference responses were slower than the Inference responses, rejections of the target picture in the No Inference conditions was even slower than their acceptance (though only numerically). A result along these lines alone might have suggested a very different line of interpretation, though it might face some of the same issues as existing reaction time studies (e.g., with respect to response bias). Of further interest is the fact that covered box choices of those participants that never had any target choices in the No Inference condition were (numerically) faster even than the target choices of the participants from the other group. While we cannot pursue these issues in further detail here for reasons of space, these methodological points seem well worth investigating in greater depth, and we are currently extending the paradigm to direct scalar implicatures as well.

4 Discussion

We saw above how direct and indirect scalar implicatures are typically considered part of the same phenomenon and different from presuppositions. Our results challenge this standard picture. In particular, we found that the proportion of No inference

[13] As with the response data, largely parallel results are also obtained when just looking at the data from participants with target choices in the No Inference condition.

choices with the strong scalar term *always* was much lower than in other studies investigating direct scalar implicatures as arising from weak scalar terms like *some* (Huang et al. (2013) in particular). More importantly, the reaction time associated with the No inference choice was slower than that associated with the Inference choice, not only with *stop* but also with *always*, in contrast with findings in the literature for direct scalar implicatures. The result for *stop* is consistent with previous findings (Chemla and Bott 2013), while that associated with *always* is a novel result. The most informative part of the results, however, is the direct comparison between the two inferences.

To illustrate the implication of our results more schematically, recall that the traditional grouping of these inferences is that sketched in (29). Given what we know about the processing of direct scalar implicatures and that of presuppositions, the prediction of the traditional approach was that the presence of indirect scalar implicatures should give rise to a longer response time than its absence.

(29) **Traditional grouping**
direct scalar implicatures = indirect scalar implicatures ≠ presuppositions

(30) **Predictions for indirect scalar implicatures**
a. presence of indirect scalar implicatures
> absence of indirect scalar implicatures
b. presence/absence of indirect scalar implicatures
≠ presence/absence of presuppositions

Our results thus do not support the prediction in (30): the processing of indirect scalar implicatures appears more similar to that of presuppositions than that of direct scalar implicatures. In other words, the reaction time associated with the presence of indirect scalar implicatures was shorter than that associated with their absence.

(31) **Main results:**
a. presence of indirect scalar implicatures
< absence of indirect scalar implicatures
b. presence/absence of indirect scalar implicatures
≈ presence/absence of presuppositions

The outcome in (31) is a challenge for the traditional view, and in the next section we briefly consider two responses that we are exploring in light of this challenge. Before doing so, let us briefly touch on two relevant points. First, let us emphasize that at this point we did not test direct scalar implicatures in our design. Therefore our discussion below relies on previous results in the literature. While the methodologies of other response time studies are comparable to ours, we are currently in the process of testing direct scalar implicatures in a within-subject experiment with direct ones. Until then, our discussion relies on the assumption that direct scalar implicatures will give rise to a delay in the same way they did in previous studies. Secondly, we want to discuss the relevance of our results for some more recent accounts of presuppositions that treat them more like indirect scalar implicatures (Chemla 2009a; Romoli 2012; Romoli (In press); see also Simons (2001); Abusch (2010)). In these approaches, (at least

some) presuppositions are essentially considered to be indirect scalar implicatures. This can in principle account very well for the similarity in processing that we found. However, there are two issues that are also not straightforwardly solved in these approaches. First, as we will see below, there are other differences between indirect scalar implicatures and presuppositions that need to be accounted for. Once these differences are explained in these recent approaches, however, it is not clear that the similarity in processing is still expected.[14] Secondly, these accounts still propose that direct and indirect scalar implicatures should be handled in a unified way and this appears problematic given our results.

In sum, the results presented above are challenging for a traditional view of indirect scalar implicatures, and they are also not straightforwardly accounted for in more recent approaches treating presuppositions and indirect scalar implicatures in a more or less parallel way. In the next section, we turn to two hypotheses in response to this challenge.

4.1 Presuppositions or Obligatory Implicatures?

Our results challenge the idea that indirect scalar implicatures are simply regular scalar implicatures, because they appear to have different processing profiles. What could indirect scalar implicature be then if not simply regular scalar implicatures? We explore two hypotheses. First, we focus on the idea that indirect scalar implicatures could actually be presuppositions.[15] Second, we consider a different option, one that takes indirect scalar implicatures to be a different type of scalar implicature.

4.1.1 Indirect Scalar Implicatures as Presuppositions

The first hypothesis could be formulated as in (32). According to (32), a sentence like (33a) not only entails but also presupposes, in the traditional sense, its sentential alternative (33b).

(32) **Indirect scalar implicatures as presuppositions**
strong scalar terms presuppose their weakest competitor.[16]

(33) a. John always went to the movies.
b. John sometimes went to the movies.

[14] For instance, Romoli (2012); Romoli (In press) proposes that some presuppositions are obligatory indirect scalar implicatures and this can account for certain differences with 'regular' indirect scalar implicatures. But then, once we distinguish between presuppositions and indirect scalar implicatures in this way, it is not clear anymore that the similarity in processing is expected.

[15] Thanks to Danny Fox (p.c.) for suggesting this way of looking at indirect scalar implicatures.

[16] We only consider here the weakest competitor as a presupposition. This immediately raises the question as to whether also other intermediate members of the scale, like 'often' in this case, should be considered presuppositions. We leave this issue for future research.

endequation Now, of course, if (33a) presupposes (33b), it follows from any mechanism for deriving presupposition projection, that (33b) is also a presupposition of (34).

(34) John didn't always go to the movies.

Moreover, if (33b) is a presuppositions of sentences like (33a) or (34), then it is indeed expected that its processing should be similar to that of presuppositions such as that associated with *stop*. The hypothesis in (32) can, therefore, account for our processing results.

There are, however, several issues with the hypothesis in (32) that remain open at this point. These issues are related to (at least potential) differences between presuppositions and indirect scalar implicatures. In particular, in relation to the projection behavior of presuppositions, e.g., in the context of connectives, modals and questions, as well as to the persistence of the projection of the inference in quantificational environments.

Starting with the first point concerning connectives, modals, and questions, we saw above that we want our theory of presuppositions to predict that all of (35a)–(35e) should give rise to the inference that somebody showed up late for class.

(35) a. It is John who showed up late for class
b. It isn't John who showed up late for class
c. It's possible that it is John who showed up late for class
d. Is it John who showed up late for class?
e. If it is John who showed up late for class, he should apologize.

If indirect scalar implicatures are presuppositions, we expect the same projection behavior. In other words, we expect all of (36a)–(36e) to presuppose that John sometimes went to the movies.

(36) a. John always went to the movies.
b. John didn't always go to the movies.
c. It's possible that John always went to the movies.
d. Did John always go to the movie?
e. If John always went to the movies, he will be able to suggest a good movie.

Based on our own intuitions, it is not clear that the inference that John sometimes went to the movies from (36b)–(36e) is as robust as the one that someone showed up late for class arising from (35b) to (35e), or even whether it is present at all. Such intuitions are rather subtle, however, so a more systematic investigation is needed to establish this data point.

Another relevant case, investigated by Chemla (2009a), is that of negative quantifiers like (37a) and (38a). When presented with a sentence like (37a) and asked whether the universal inference in (37c) followed, the participants of the experiments by Chemla (2009a) were largely willing to give an affirmative answer.

(37) a. None of these ten students won the marathon.
b. Some of these ten students participated in the marathon.
c. Each of these ten students participated in the marathon.

More precisely, the inference in (37c) was accepted more often (≈ 80 %) than the analogous universal inference with a scalar implicature in (38c) from (38a) (≈ 25 %).

(38) a. None of my professors failed all of their students.
b. Some of my professors failed some of their students.
c. All of my professors failed some of their students.

This result tells us that in the scope of negative quantifiers there is a difference in the degree to which presuppositional triggers and strong scalar terms give rise to universal inferences.[17] And this difference is unexpected based on the hypothesis in (32) above, which treats indirect scalar implicatures and presuppositions uniformly.

Finally, the hypothesis in (32) does not say anything about whether the inferences that we have labeled indirect scalar implicatures arise only because they are presuppositions, or whether the same inference is additionally supported by standard implicature reasoning. In fact, for most theories of scalar implicatures, it seems hard to even block deriving them as scalar implicatures as well, based on an account of direct ones in terms of reasoning about scalar alternatives in terms of logical strength. If indirect scalar implicatures are derived both based on scalar reasoning and because they are presupposed, this raises obvious questions about how these two components interact. We have to leave the exploration of these open issues for a future occasion.

4.1.2 Indirect Scalar Implicatures as Obligatory Scalar Implicatures

Another idea possibility to explore in order to account for our results is that indirect scalar implicatures could be a type of scalar implicatures distinct from direct ones, with corresponding differences in their processing properties. In particular, we focus on the notion of 'obligatory' scalar implicatures, recently proposed by Spector (2007); Chierchia et al. (2012). This second hypothesis can be formulated as in (39).

(39) **Indirect scalar implicatures as obligatory scalar implicatures**:
Indirect scalar implicatures are obligatory.

The obligatoriness in (39) can be implemented in different ways and we will remain neutral on how this should be done. What is relevant for us is that if indirect scalar implicatures are obligatory scalar implicatures, we can account for our results about their processing in the following way. First, consider how we would explain cases in which indirect scalar implicatures are absent in this approach. This is not straightforward, because if they are obligatory, we cannot simply say that they are not derived in the first place, based on contextual information about the speaker's beliefs. However, in the same way as what we described is standardly done for presuppositions, we can resort to a local derivation of scalar implicatures. To illustrate, imagine your theory of scalar implicatures to be a function **SI**, which applied to a sentence returns its meaning strengthened with a scalar implicature. For instance, for any p,

[17] Notice that, importantly, these results by no means imply that sentences like (38a) do not give rise at all to universal inferences like that in (38c). See Romoli (2012); Romoli (In press) for arguments in favor of the existence of tg inference.

SI(p) = p and p's scalar implicatures. Moreover, imagine that **SI** can be recursively applied at any scope site of a complex sentence. If you have a sentence like (40), repeated from above, there are at least two scope sites at which you can apply your **SI**. First, you can apply it globally, as in (41a), which gives rise to the indirect scalar implicature that John sometimes went to the movies. Second, however, you could apply it locally, under the scope negation (41b). This local derivation is vacuous, therefore this is equivalent to the absence of the indirect scalar implicatures (i.e., in this interpretation, (40) is compatible with a situation in which John never went to the movies).

(40) John didn't always go to the movies.

(41) a. **SI**[not[John always went to the movies]]
b. not[**SI**[John always went to the movies]]

In sum, if we take this approach, we can account for the cases in which indirect scalar implicatures are suspended as cases of (vacuous) local scalar implicatures. Having local interpretations available is useful independently. We know from the literature that the distribution of scalar implicatures generally appears to be sensitive to the polarity of the context in which the scalar term is embedded. In other words, scalar implicatures tend not to arise when the corresponding scalar term is embedded in downward entailing contexts.[18] However, it is possible to force the strengthened interpretation of a scalar term in a downward entailing context, though such an interpretation appears to be marked. For instance, one way to account for how (43) is compatible with its continuation is to assume that it should be interpreted with a local scalar implicature under negation, as in (44). This interpretation could be paraphrased as either John met neither Paul nor Mary or he met both of them and it is, therefore, compatible with the continuation that he met both of them.

(43) John didn' t meet Paul or Mary . . .
he met both of them!

(44) not[**SI**[John met Paul or Mary]]

Now, if we make the plausible assumption that the intuitive markedness of local scalar implicatures in downward entailing contexts is reflected in processing and assume that our No Inference condition involves precisely that, we can account for the processing cost associated with that condition. This is because, on this approach, the only way the participants of our experiment could chose a picture incompatible with the indirect scalar implicature for a case like (45) is if they computed it locally

[18] To illustrate, consider the following minimal pair from Chierchia (in press): while (42a) is easily interpreted with an exclusive reading of disjunction (i.e., everyone likes Mary or Sue but not both), this is not the case for (42b). In other words, (42b) is generally interpreted as not suggesting that if someone likes both Mary and Sue, she won't write to the dean (see Panizza et al. 2009 for discussion and experimental data that show the sensitivity of scalar implicatures to polarity).

(42) a. Everyone either likes Mary or likes Sue and will write to the dean.
b. Everyone who either likes Mary or likes Sue will write to the dean.

under negation. The observed markedness plausibly can be seen as corresponding to the observed delay in processing.

(45) John didn't always go to the movies.

In light of the differences between indirect scalar implicatures and presuppositions discussed above, which are potentially problematic for an account of the former in terms of the latter, it is worth considering how the present hypothesis fares in this regard. Given that this hypothesis does not assume that indirect scalar implicatures are presuppositions, it could, in principle, explain the differences with presuppositions better. To illustrate, consider first the case of connectives, modals, and questions. Here we do not predict the same projection pattern for both types of inferences.[19] For illustration, take the case of possibility modals and the antecedent of conditionals. In the former case, the alternative that we obtain in (b) for (a) is entailed by the assertion so no inference is predicted—specifically the inference that John sometimes went to the movies is not predicted.

(46) a. It's possible that John always went to the movies.
b. It's possible that John sometimes went to the movies.

Similarly, in the case of antecedents of conditionals, the alternative that we have for (47a) is (47b), which is stronger than the assertion, so it winds up being negated. The inference that we obtain, however—it's false that if John sometimes went to the movies, he skipped the cooking classes—is not the projection-like inference that a theory of presuppositions would obtain (namely that John sometimes went to the movies).

(47) a. If John always went to the movies, he skipped the cooking classes.
b. If John sometimes went to the movies, he skipped the cooking classes.

As far as the difference between universal and negative quantifiers in (48b) and (b) is concerned, while (b) might as well be an inference of (a), given that in this approach presuppositions and indirect scalar implicatures are different in nature, it is not expected that the rate of acceptance of (b) should be the same as that of (b).[20] In other words, this approach is compatible with the results by Chemla (2009a).

(48) a. None of these students stopped showing up late for class.
b. All of these students used to show up late for class.

(49) a. None of these students did all of the readings
b. All of these students did some of the readings.

In sum, we have sketched two alternative theoretical perspectives on indirect scalar implicatures that differentiate them from direct ones, and discussed how this could account for our results. At this point both hypotheses require fleshing out and further

[19] We do not predict projection inferences, unless we make additional assumptions. See Chemla (2009a); Romoli (2012); Romoli (In press) for scalar implicature-based account of presuppositions, which do make additional assumptions for deriving their projection behavior.
[20] See Romoli (2012); Romoli (In press) for arguments in favor of having the inference in (49b) from (49a) and a proposal on how to derive it as a scalar implicature,

exploration, but we have to leave a more detailed investigation for future research. Relatedly, we note that, as part of the general project of comparing direct and indirect scalar implicatures and presuppositions, we are currently investigating the case of *always* under negation in direct within-participants comparison with that of *sometimes* in positive sentences. Additionally, we are also in the process of comparing children's knowledge of direct and indirect scalar implicatures, as well as presuppositions, to provide yet another angle on the comparison of these inferences. We hope that taken together, this series of studies will provide the basis for a more comprehensive understanding of the types of inferences discussed here.

Acknowledgements For discussion and suggestions, thanks to Cory Bill, Emmanuel Chemla, Gennaro Chierchia, Stephen Crain, Alexandre Cremers, Raj Singh, Benjamin Spector, Yasutada Sudo, and the audience at Sinn und Bedeutung 18 in Vitoria. Thanks to Rachel Stults, Jamie Fisher, and Robert Wilder for assistance with data collection and to Dorothy Ahn for images used in the experimental stimuli. The work reported here was in part supported by a grant from the University Research Foundation of the University of Pennsylvania.

References

Abusch, D. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27 (1): 37–80.

Barr, D. J., R. Levy, C. Scheepers, and H. J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68 (3): 255–278.

Bates, D. M. 2005. Fitting linear mixed models in R. *R News* 5:27–30.

Beaver, D. 2001. *Presupposition and assertion in dynamic semantics*. Stanford: CSLI.

Beaver, D., and B. Geurts. 2012. Presupposition. In *Semantics: An international handbook of natural language meaning, vol. 3*, ed. C. Maienborn, K. von Heusinger, and P. Portner. Berlin: Mouton de Gruyter.

Beaver, D., and H. Zeevat. 2012. Accommodation. In *Oxford handbook of linguistic interfaces*, ed. G. Ramchand and C. Reiss. Oxford: Oxford University Press.

Bott, L., and I. A. Noveck. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* 51 (3): 437–457.

Bott, L., T. M. Bailey, and D. Grodner. 2012. Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language* 66 (1): 123–142.

Breheny, R., N. Katsos, and J. Williams. 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.

Chemla, E. 2009a. Presuppositions of quantified sentences: Experimental data. *Natural Language Semantics* 17 (4): 299–340.

Chemla, E. 2009b. Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. Under revision for Semantics and Pragmatics.

Chemla, E. 2009c. Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics* 2 (2): 1–33.

Chemla, E., and L. Bott 2013. Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and Cognitive Processes* 38 (3): 241–260.

Chemla, E., and L. Bott. In press. Processing inferences at the semantics/pragmatics frontier: Disjunctions and free choice. Manuscript, Ecole Normal Superieure, Paris.

Chierchia, G. 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In *Structures and beyond*, ed. A. Belletti. Oxford: Oxford University Press.

Chierchia, G. In press. *Logic in grammar: Polarity, free choice, and intervention*. Oxford: Oxford University Press.
Chierchia, G., and S. McConnell-Ginet. 1990. *Meaning and grammar. An introduction to semantics*. Cambridge: MIT Press.
Chierchia, G., D. Fox, and B. Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In *Semantics: An international handbook of natural language meaning*, vol. 3, ed. C. Maienborn, K. von Heusinger, and P. Portner. Berlin: Mouton de Gruyter.
Cremers, A., and E. Chemla. 2013. Direct and indirect scalar implicatures share the same processing signature. Manuscript, Ecole Normal Superieure, Paris.
Fox, D., and R. Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19 (1): 87–107.
Geurts, B. 2010. *Quantity implicatures*. Cambridge: Cambridge University Press.
Grice, H. 1975. Logic and conversation. In *Syntax and semantics: Speech acts*, ed. P. Cole and J. Morgan. New York: Academic.
Grodner, D. J., N. M. Klein, K. M. Carbary, and M. K. Tanenhaus. 2010. "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition* 116 (1): 42–55.
Heim, I. 1983. On the projection problem for presuppositions. In *WCCFL 2*, ed. M. Barlow, D. Flickinger, and M. Wescoat, 114–125.
Horn, L. 1972. On the semantic properties of logical operators in English. Ph.D. Diss., UCLA.
Huang, Y. T., and J. Snedeker. 2009. Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology* 58 (3): 376–415.
Huang, Y., E. Spelke, and J. Snedeker. 2013. What exactly do number words mean? *Language Learning and Development* 9 (2): 105–129.
Kamp, H. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language: Proceedings of the third Amsterdam colloquium*, vol. I, ed. J. Groenendijk, T. Janssen, and M. Stokhof, 227–321. Amsterdam: Mathematical Center, University of Amsterdam.
Karttunen, L. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4 (2): 169–193.
Karttunen, L. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1:181–194.
Katsos, N., C. Andrés Roqueta, R. A. C. Estevan, and C. Cummins. 2011. Are children with specific language impairment competent with the pragmatic and logic of quantification? *Cognition* 119:43–57.
Katzir, R. 2007. Structurally-defined alternatives. *Linguistic and Philosophy* 30 (6): 669–690.
Lewis, D. 1979. Scorekeeping in a language game. In *Semantics from different points of view*, ed. R. Bäuerle, U. Egli, and A. von Stechow. Berlin: Springer.
Musolino, J., and J. Lidz. 2006. Why aren't children universally successful with quantification? *Linguistics* 44 (4): 817–852.
Panizza, D., G. Chierchia, and Charles Clifton Jr. 2009. On the role of entailment patterns and scalar implicatures in the processing of numerals. *Journal of Memory and Language* 61 (4): 503–518.
Romoli, J. 2012. Soft but strong: Neg-raising, soft triggers, and exhaustification. Ph.D. Diss., Harvard University.
Romoli, J. In press. The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics*.
Romoli, J., Y. Sudo, and J. Snedeker. 2011. An experimental investigation of presupposition projection in conditional sentences. In *Proceedings of SALT 21*, ed. N. Ashton, A. Chereches, and D. Lutz, 592–608.
Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics* 1 (1): 117–121.
Sauerland, U. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27 (3): 367–391.
Schlenker, P. 2008. Presupposition projection: The new debate. In *Proceedings of SALT 18*, ed. T. Friedman and S. Ito. Ithaca: CLC.

Schwarz, F. 2007. Processing presupposed content. *Journal of Semantics* 24 (4): 373–416.
Schwarz, F., and S. Tiemann. 2012. Presupposition Processing—The case of German *wieder*. In *Proceedings from the 18th Amsterdam colloquium*, ed. M. Aloni, V. Kimmelman, F. Roelofsen, G. Weidman Sassoon, K. Schulz, and M. Westera. Berlin: Springer.
Schwarz, F., and S. Tiemann. 2013. The path of presupposition projection in processing—The case of conditionals. In *Proceedings of SuB 17*, ed. E. Chemla, V. Homer, and G. Winterstein, 509–526. Paris: semanticsarchive.net
Simons, M. 2001. On the conversational basis of some presuppositions. In *Proceedings of SALT 11*, ed. R. Hastings, B. Jackson, and Z. Zvolenszky, 431–448.
Spector, B. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In *Presupposition and implicature in compositional semantics*, ed. U. Sauerland and P. Stateva. Basingstove: Palgrave.
Stalnaker, R. 1974. Pragmatic presuppositions. In *Semantics and philosophy*, ed. M. K. Milton, and P. K. Unger. New York: New York University Press.
Stalnaker, R. 1978. Assertion. In *Pragmatics*, ed. P. Cole., vol. 9, 315–322. New York: Academic.
von Fintel, K. 2008. What is presupposition accommodation, again? *Philosophical Perspectives* 22 (1): 137–170.

Three-Year-Olds' Understanding of *Know* and *Think*

Rachel Dudley, Naho Orita, Valentine Hacquard and Jeffrey Lidz

Abstract This study investigates three-year-olds' representations of the verbs *think* and *know*, in attempt to assess their understanding of factivity. *Know*, being factive, is used in contexts where the complement is taken to be true. *Think*, although non-factive, is often used in contexts where the complement also is taken to be true. Despite this, can children recognize the difference between them and understand that the truth of the complement is presupposed in only one case? Acquisition studies find that children do not have an adult-like understanding of these verbs before age four, but the tasks used are often inappropriate for testing preschoolers. We designed an interactive game to implicitly evaluate their knowledge of the verbs in a task that more directly targets factivity. Our results show that some three-year-olds are able to distinguish *think* and *know* in ways indicating they understand *know* presupposes the truth of its complement and *think* does not. The remaining three-year-olds seem to treat both verbs as non-factive. This suggests that early representations of *know* may be non-factive, and raises the question of how children come to distinguish the verbs.

Keywords Factivity · Presupposition · Acquisition · Semantics-pragmatics interface · Attitude ascriptions · Attitude verbs · Factive verbs · Child pragmatics · Theory of mind

1 Introduction

Children's understanding and use of pragmatics poses several observational and theoretical challenges. On the one hand, children seem quite competent in this domain. They make inferences about the goals and desires that drive people's behavior as

R. Dudley (✉) · N. Orita · V. Hacquard · J. Lidz
University of Maryland, College Park, MD, USA
e-mail: redudley@umd.edu

N. Orita
e-mail: naho@umd.edu

V. Hacquard
e-mail: hacquard@umd.edu

J. Lidz
e-mail: jlidz@umd.edu

F. Schwarz (ed.), *Experimental Perspectives on Presuppositions,*
Studies in Theoretical Psycholinguistics 45, DOI 10.1007/978-3-319-07980-6_11,

young as one year of age (e.g., Gergely et al. 1995; Warneken and Tomasello 2006; Woodward 1998). Moreover, early word learners seem to be able to use such inferences in determining the intended meaning of novel nouns (Merriman and Bowman 1989; Bloom 2002; Halberda 2003). Similarly, in the domain of language use, we often find evidence that children are exquisitely sensitive to the felicity conditions of a given grammatical expression (Hamburger and Crain 1982; Thornton and Crain 1999; Gualmini et al. 2008; Musolino and Lidz 2006). On the other hand, children are notoriously susceptible to pragmatic errors in making inferences about a speaker's communicative intent when this goes beyond the literal meaning of the expression (Noveck 2000; Papafragou and Musolino 2002; Huang and Snedeker 2009; Papafragou and Skordos in press). And, they fail to use pragmatic information to guide parsing decisions, even when they appear to use the very same information to shape their own productions (Trueswell et al. 1999; Hurewitz et al. 2000). Together, these literatures highlight a tension between children's abilities to make inferences about the goals and intentions that underlie what people do and between those that underlie what people say.

In this paper, we explore this tension in the domain of attitude verb learning. This class of verbs presents an interesting puzzle from the perspective of the semantics-pragmatics interface. Attitude verbs are used to convey rich content, but that content is sometimes packaged into the conventional meaning of the verb and sometimes not. For example, the verb *know* is factive, in that it presupposes the truth of its complement and is therefore used in contexts where the speaker takes the complement to be true. The verb *think* is non-factive, but it is nonetheless often used in contexts where the speaker is committed to the truth of the complement. How is a learner to recognize that the truth of the complement is encoded as a presupposition in the *know* case but not the *think* case? As a first step towards addressing that question, this paper examines the age at which children distinguish *know* from *think* with respect to factivity.

2 Know and Think

Think and *know* both report the beliefs of a subject. For example, (1) and (2) both convey that John has the belief that Mary is at the office.

(1) John knows that Mary is at the office
(2) John thinks that Mary is at the office

They differ however in that (1), unlike (2), can only be true if Mary is in fact at the office. *Know* is factive. Factive verbs like *know* are typically taken to presuppose, and not merely entail, the truth of their complements (Kiparsky and Kiparsky 1968; Karttunen 1971; Hooper and Thompson 1973; Hooper 1975). Indeed, the truth of the complement seems to project out of p-family contexts like negation (Chierchia and McConnell-Ginet 2000) with *know* sentences like (3), but not with the equivalent *think* sentences like(4).

(3) John doesn't know that Mary is at the office
(4) John doesn't think that Mary is at the office

Because *know* and *think* differ in the presuppositional status of their complements, their uses license different inferences. Taken in isolation, (1) and (3) both indicate that the speaker takes it for granted that Mary is at the office. On the other hand, (2) and (4) are consistent with Mary either being at the office or not: we cannot draw inferences about her actual location without further assumptions about John and the speaker. In contexts where the speaker could have uttered "John knows that Mary is at the office", or simply "Mary is at the office", we might infer, via Gricean reasoning, that by uttering "John *thinks* that Mary is at the office", the speaker is indicating that she doesn't believe that Mary is at the office. However, in contexts where the speaker takes John to be a reliable source of information, and the competing sentences are irrelevant or not accessible, her use of (2) might invite the inference that Mary actually is at the office, and her use of (4) that she isn't[1]. Thus, in contexts where speakers and hearers take John to be a reliable source of information, both (1) and (2) will invite the inference that Mary is at the office. However, we expect divergent inferences for their negated counterparts: (3) still indicates that Mary is at the office (and John simply isn't aware of this fact), while (4) indicates that she isn't.

Do children recognize that sentences with *know* presuppose the truth of their complements, but that *think* sentences do not? As we will see in the next section, previous research shows that young children behave as though *think* sentences report true beliefs: across various tasks, their responses seem to suggest that they take both sentences like (1) and (2) to indicate that Mary is in the office. If speakers use sentences like (2) to implicate that they do not believe that Mary is in the office, children seem to be oblivious to it. This pattern could result from children not differentiating *think* and *know* at all, or from a failure to derive implicatures using these verbs (either due to difficulty with quantity implicatures in general, or in realizing that *think* and *know* can be relevant alternatives to each other). To test whether young children can differentiate *think* and *know*, it is thus important to look at their understanding of negated sentences like (3) and (4): do they realize that (3) presupposes that Mary is at the office, but that (4) does not?

While we cannot directly probe children's representations of verbs like *know* and *think* (or adults' for that matter), we can use behavioral methods to assess the kinds of inferences they make upon hearing such sentences and then try to reconstruct the semantic representations that underlie such inferences. Specifically, we ask whether three-year-olds can demonstrate recognition of the factivity of *know* and the non-factivity of *think*. Once we know the answer to this question, we will be in a better position to address further questions about (a) the nature of the target (adult) representation of these verbs (e.g., how is factivity encoded?), (b) how children eventually

[1] The reason why the inference for (4) is that Mary is not at the office, rather than agnosticism about Mary's actual location is that *think* is a neg-raising predicate: "John doesn't think that Mary is at the office" implies that John thinks that Mary is not at the office. If John is a reliable source, we will infer that Mary isn't at the office.

reach such representations, and (c) how children come to associate a presupposition with certain lexical items, but not others.

3 Past Acquisition Studies

3.1 Studies on Children's Understanding of Think and Their Developing Theory of Mind

Previous studies show that children have difficulty with *think* until at least four years of age. Unlike adults and older children, three-year-olds typically reject a sentence like (2) in contexts where Mary is not at the office, even if John thinks that she is (Johnson and Maratsos 1977; Wellman et al. 2001; de Villiers and Pyers 2002; de Villiers 2005; Sowalsky et al. 2009; Lewis et al. 2012; a.o.). For example, consider the scenario in (5):

(5) False belief scenario: Mary has already made it home for the day, but John wrongly believes that she is at the office.
   a) John thinks that Mary is at the office
   b) Mary is at the office

In this scenario, adults and older children assent to (5a), but three-year-olds reject it. Three-year- olds seem to respond based on the truth of the complement clause (5b) (false in this scenario), instead of the entire sentence. There are at least four possible explanations for three-year-olds' non-adult-like responses in contexts like (5).

1. **Conceptual Hypothesis.** Children's initial difficulty with *think* reflects difficulty with the belief concept that *think* expresses. They reject sentences like (5a) because of their inability to attribute a false belief to John. This failure could either be due to a lack of understanding that people can have false beliefs or because they cannot deal with a belief representation that conflicts with their own (cf. Johnson and Maratsos 1977; Tardif and Wellman 2000; Perner et al. 2003; a.o.).
2. **Complement-only Hypothesis.** Three-year-olds only attend to the embedded clause because they do not understand the matrix clause and thus ignore it altogether. This could be due to a lack of understanding of the verb *think*, or an inability to embed a finite complement clause (cf. de Villiers 1995; Diessel and Tomassello 2001).
3. **Pragmatic Hypothesis.** Children's difficulty with *think* is not semantic, syntactic, nor conceptual, but pragmatic in nature (Lewis et al. 2012, in progress; Lewis 2012). The reason children respond to the truth of the complement clause in (5) is the same reason adults sometimes respond to the truth of the complement of *think*. Verbs like *think* are sometimes used to proffer the content of its complement, in which case the complement clause carries the main point of the utterance, and the matrix clause get demoted to parenthetical status (cf. Urmson 1952; Hooper 1975;

Rooryck 2001; Simons 2007; a.o). Children's failures are due to a tendency to assign such parenthetical readings to *think*, even in situations where adults do not.

4. **Factive-*think* Hypothesis.** Children's non-adult responses are due to a failure to recognize the non-factivity of *think* and instead treat it in essentially the same way adults treat *know*. (Johnson and Maratsos 1977; Abbeduto and Rosenberg 1985).

Lewis et al. (2012, in progress), and Lewis (2013), provide initial evidence against the first two hypotheses. They show that three-year-olds are not attending solely to the complement clause, and argue that children respond to the truth of the complement clause only in contexts in which they assume that it is being proffered by the speaker. In contexts in which parenthetical interpretations are blocked, three-year-olds respond to the truth of the entire clause, in an adult-like way, even in false belief scenarios. Consider the variant of scenario (5) in (6):

(6) False belief scenario 2: Mary *is* at the office, but John wrongly believes that she is at at home.
   a) John thinks that Mary is at the office
   b) Mary is at the office

In this scenario, the sentence (6a) is *false*, even though the complement clause (6b) is true. In such cases, three-year-olds, rejected sentences like (6a), just like adults. Lewis et al. argue that the reason three-year-olds' performance improves in this kind of false belief scenario is that parenthetical interpretations in which the speaker endorses the reported belief are blocked: the speaker cannot endorse a belief of John's that John does not hold. These results argue against the Complement-only Hypothesis. Indeed, children's adult-like responses are unexpected if children merely respond to the truth of the complement: they should accept the sentence, since the complement is true. Furthermore, it shows that three-year-olds are able to provide adult-like responses, even in contexts in which the subject has a false belief, suggesting that their difficulty is not conceptual, contra the Conceptual Hypothesis.

Lewis's results, however, are still consistent with the last two hypotheses: children's difficulty with *think* could either be due to a (cancelable) tendency to assume that speakers typically endorse the truth of the complement, or to a factive interpretation of *think.* Note that in the scenario in (6), we would typically reject the sentence "John knows that Mary is at the office". Perhaps three-year-olds' responses then reflect a factive understanding of *think*.

3.2 Past Studies on Children's Understanding of Know

Previous research suggests that children do not differentiate verbs like *know* and *think* until at least age four (Macnamara et al. 1976; Johnson and Maratsos 1977; Abbeduto and Rosenberg 1985; Moore and Davidge 1989; Moore et al. 1989). Some authors even argue that children might not have a fully adult-like understanding of *know* well into the grade school years (Harris 1975, Hopmann and Maratsos 1978,

Scoville and Gordon 1980, Falmagne et al. 1994; Léger 2007). However, many of these studies involve complex tasks that could be independently difficult, especially for three-year-olds. Some require children to make metalinguistic judgments or to compare the relative acceptability of two sentences. Consequently, these tasks may underestimate children's knowledge. Studies involving more naturalistic use of the verbs could serve as a better probe of children's knowledge.

3.2.1 Negation Tasks

Several past studies have attempted to assess whether children understand that the presupposition associated with *know*, but not *think*, projects out of negation. Harris (1975) tested preschoolers, kindergarteners and older children in first through sixth grade. On one of Harris's tasks, the "truth questioning" task, participants responded to questions of the form "The teacher did not know that Tim was absent. Was Tim absent?" On another, participants made judgments on whether sentences like the following sounded funny "John {knew, didn't know} that {his father was a tree, his sister was not a girl}." Harris concluded that comprehension of factive verbs is a lengthy process that might begin in preschool but definitely extends past sixth grade, although he found the largest improvement in performance was between the ages of four and seven. Harris's measures, while later adopted widely in this literature, might not be appropriate for younger children because they involve activities that are not a common part of everyday life.

Hopmann and Maratsos (1978) tested four-, five- and seven-year-olds on an act-out task with affirmative and negative sentences that contained either factive or non-factive verbs. They found that some children responded as if the complement was true, regardless of the matrix verb, and that this tendency was stronger with the youngest children. Like Harris (1975), these authors concluded that development of an understanding of these verbs as a protracted process (and that acquisition of the factivity of factive verbs was not uniform), with full competence achieved by age seven. However, this apparent delay in understanding might be due to the added pressures of an act-out task, and not the children's comprehension of the verbs.

Scoville and Gordon (1980) tested children aged 6, 8, 12 and 14. In their task, participants watched a lottery-style game show where one character would report on another (semi-omniscient) character's understanding of the outcome. Again, like other authors, Scoville and Gordon concluded that acquiring an understanding of factive verbs was an extended process, where each verb is learned to be factive on a case-by-case basis. However, children's apparently poor performance in this task might be due to Scoville and Gordon's strict criteria for success.

Léger (2007) tested children aged 6, 7, 9 and 11. Participants were presented with the following four attitude reports (7–10):

(7) She knows she has a turtle
(8) She knows she doesn't have a turtle
(9) She doesn't know she has a turtle
(10) She doesn't know she doesn't have a turtle

In Léger's task there were four dolls, each of which was uniquely described by one of the sentences in (7–10) and participants were asked to pick the appropriate doll after hearing the attitude reports. She found that even the youngest children tended to pick the right answer but participants had not attained 100 % accuracy even by age 11. Yet, Léger's conclusion is based children's performance on sentences like (10), which could be hard to process independently of factivity.

3.2.2 Metalinguistic Tasks

Other studies in this literature require explicit comparison of *know* and *think* statements, which could be too metalinguistically difficult for young children. Macnamara et al. (1976) told four-year-olds different stories and then probed their participants about the mental states of the characters, including whether the characters *knew* a certain proposition. They found that four-year-olds performed well on this task. Johnson and Maratsos (1977) tested three- and four-year-olds, also on whether characters in a story *knew* or *thought* a certain proposition. Like Macnamara et al, these authors concluded that four-year-olds could succeed on such a task, but that three-year-olds could not. Abbeduto and Rosenberg (1985) tested three-, four- and seven-year-olds on three different tasks: (i) a modified Harris (1975) "truth questioning" task, (ii) a verb choice task where participants determined whether *know* or *think* was a more accurate description of a character's mental state, and (iii) a definitional task with questions like "What does it mean to know?" These authors concluded that three-year-olds had not achieved mastery of the verbs but that four- and seven-year-olds had. However, three-year-olds' poor performance may be due to the metalinguistic nature of these tasks and to difficulty comparing the relative acceptability of two sentences.

3.2.3 Relative Strength Tasks

Finally, some studies in this literature assessed children's understanding of the relative strengths of these predicates. These studies were concerned with whether *know* indicates more certainty or confidence about the truth of a complement than *think* does. Moore and colleagues (Moore and Davidge 1989; Moore et al. 1989; a.o.) tested children ages 3, 4, 5, 6 and 8. In their tasks, children were presented with two boxes, only one of which contained a toy. The participants' job was to determine which box contained the toy after hearing two puppets utter sentences like "I know it's in the red box" and "I think it's in the blue box." The studies by Moore and colleagues all found that three-year-olds were unable to reliably use the *know* statement over the *think* statement, but that children four and over could. Falmagne

et al. (1994) tested third and sixth graders on four different variations of Harris's (1975) "truth questioning" task. These authors also found that development of these verbs was a process that continued through grade school and even after sixth grade. These studies suggest that children have difficulty computing quantity implicatures with these verbs, but this could be due to a variety of reasons that are independent from their understanding of the two verbs' (non-)factivity, for example: difficulties with computing implicatures in general, or with realizing when *know* should be a relevant alternative to *think*.

Three-year-olds appeared to fail all the above tasks, but their poor performance could be due to extra-linguistic, task-related difficulties rather than to a lack of understanding the factivity of *know*. Most of these tasks require children to make explicit judgments about the truth of a *know* sentence or its complement clause given a context (that was often very sparse). This is arguably difficult for naïve adult participants to do, let alone grade school children or preschoolers. Some of these tasks required children to answer definitional questions. Some required participants to either explicitly or implicitly compare a *think* sentence with a *know* sentence (*e.g.:* Does John know that Mary's at the office or does John think that Mary's at the office?). This could be independently difficult for preschoolers for many reasons, not the least of which is that adult understanding of *know* logically implies *think* as well, and preschoolers might not have the pragmatic competence to choose the more informative *know* statement as the "correct" one in cases where both statements are true (Grice 1975).

Several authors suggest that an understanding of factive verbs continues to develop well into the grade school years. On a certain level, that seems to be an apt description of the developmental trajectory of these verbs; there are intricacies of their use that surely only adults could grasp. However, we should disentangle assessment of three-year-olds' basic understanding of factivity from a more sophisticated holistic understanding of the verbs. Several of these studies failed to assess that basic understanding for a combination of the following reasons: (i) their age ranges did not go as low as three years (Harris 1975, Macnamara et al. 1976; Scoville and Gordon 1980; Falmagne et al. 1994; Léger 2007); (ii) their tasks incorporated extraneous difficulties (Harris 1975; Johnson and Maratsos 1977; Abbeduto and Rosenberg 1985; Moore and Davidge 1989; Moore et al. 1989); or (iii) their measures of success were too strict (Harris 1975; Scoville and Gordon 1980; Falmagne et al. 1994).

3.3 Past studies on Children's Understanding of Presuppositions

While the findings in this literature are mixed, there is some indication that children are aware of some presuppositions quite early. Despite this early awareness, children may not deploy their understanding of presuppositions in the full range of contexts that adults do. Berger and Höhle (2012) show that German children are aware of the presupposition associated with the focus particles *auch* 'also' and *nur* 'only'. Hamburger and Crain (1982) show that preschoolers' performance on relative clause

interpretation is a function of the pragmatic use of relative clauses; children are able to succeed at interpreting object relatives only when relative clauses are used to distinguish two entities that are otherwise similar (e.g., the sheep that the lion bit vs. the sheep that the dog bit). Syrett et al. (2010) find that three-year-olds are aware of the uniqueness presupposition associated with *the*, and that they are able to use that information in an online task. Trueswell and colleagues show that children fail to use the discourse context in concert with the uniqueness presupposition of *the* in order to help them resolve a PP attachment ambiguity (Trueswell et al. 1999), but that they are nonetheless able to use one structure when the discourse demands it (Hurewitz et al. 2000). Together, these findings suggest an initial understanding of some presuppositional phenomena, but one that is emerging earlier than the literature on children's understanding of factive verbs would suggest.

4 Method

In order to assess three-year-olds' understanding of the factive and non-factive verbs *know* and *think* and the inferences that they license, we designed a simple task that allows them to demonstrate their knowledge without being hindered by orthogonal difficulties. We ask participants to find a toy hidden in one of two boxes (much like Moore and Davidge 1989 and Moore et al. 1989) using clues in the form of attitude reports containing *think* and *know* (like Scoville and Gordon 1980). The participant's goal in our task is to uncover the location of the toy.

4.1 Subjects

Child participants were 40 three-year-olds (age range: 3;1–3;11 years;months, mean age: 3;6, 19 boys). All children were reported to be monolingual speakers of English by their parents and all were recruited from the University of Maryland Infant Studies Database.

Ten adults also participated. They were recruited from an undergraduate introductory linguistics course at University of Maryland and participated for course credit.

4.2 Design

Participants were seated in front of two boxes (one red and one blue). They were told that the experimenter would hide one toy in either of the boxes and their task was to find the toy, after the experimenter gave them a clue. Participants were also informed that a puppet (Lambchop) would be joining the game as well, but was too shy to

Verb-type / Negation-type	Think	Know
None	Lambchop **thinks** that it's in the red/blue box (A)	Lambchop **knows** that it's in the red/blue box (B)
Matrix	Lambchop **doesn't think** that it's in the red/blue box (C)	Lambchop **doesn't know** that it's in the red/blue box (D)
Embedded	Lambchop **thinks** that it's **not** in the red/blue box (E)	Lambchop **knows** that it's **not** in the red/blue box (F)

Fig. 1 Test sentences by factor

do anything but whisper to the experimenter. An occluder kept participants from seeing where the experimenter hid toys and there was always a toy hidden in each box, despite what participants were told. This was done in order to avoid participants learning from negative evidence.

On each trial, the occluder would rise and the puppet would whisper in the experimenter's ear before the experimenter delivered a clue in the form of an attitude report (test sentence). Upon hearing the clue, participants were asked to demonstrate which box they thought the toy was in.

We manipulated two factors within subjects: verb-type (*think*, *know*) and negation-type (none, matrix, embedded). Accordingly children heard *think* and *know* sentences, with or without negation, as in Fig. 1.

Participants were given three trials for each of the sentence types in Fig. 1, as well as three control trials with the test sentence *It's not in the red/blue box*. Responses were coded as selections of the box mentioned in the test sentence, or as selections of the other box.

Note that this task requires children to accommodate information that is not in the common ground. We decided that this was a necessary trade-off in order to better assess young children's knowledge, using a natural, non-metalinguistic task. However, this task may still underestimate their knowledge of presupposition. Even if children recognize *know*'s factivity, they may still have difficulty accommodating the presupposition that the complement clause is true in order to pick the mentioned box, as we discuss later.

4.3 Predictions

Based on the above discussion, there seem to be only three logical possibilities for children's understanding of these verbs that are consistent with the literature: (i) children understand the (non-)factivity of these verbs in a fully adult manner but previous tasks have obscured their competence; (ii) children lack the understanding that *know* is factive, thereby treating both verbs as non-factive; or (iii) children understand *know* in an adult manner but also treat *think* as a factive, which is why they tend to assume that its complement is true.

Hypothesis	Trials where mentioned box should be selected					
	No negation		Matrix negation		Embedded negation	
	Think (A)	Know (B)	Think (C)	Know (D)	Think (E)	Know (F)
1) children are adult-like *know* = factive *think* = non-factive	✔	✔	✖	✔	✖	✖
2) children have non-adult-like *know* *know* = non-factive *think* = non-factive	✔	✔	✖	✖	✖	✖
3) children have non-adult like *think* *know* = factive *think* = factive	✔	✔	✔	✔	✖	✖

Fig. 2 Summary of predictions

These possibilities make the following predictions: If children are adult-like, they will only pick the mentioned box when they hear sentences like A, B and D. If children treat *know* as a non-factive, they should only pick the mentioned box when they hear A and B. If children treat *think* as factive, they should pick the mentioned box when they hear A, B, C and D. See Fig. 2 for a summary of these predictions.

As shown in Fig. 2, the matrix negation trials will be the crucial ones for determining participants' understanding of the factivity of the two verbs.

5 Results

5.1 Control Items

Control items were three trials with the following clue:

(11) It's not in the red/blue box

For these trials, participants needed to choose the other box (not the mentioned box) at least two out of three times in order to be included in the analyses. Nine out of the ten adult participants chose the other box on every trial. The tenth participant failed to choose the correct box on these trials, and was excluded from analyses. Out of the 40 three-year-old participants, 9 of them failed the control items (by picking the other box only once or never), and were therefore excluded from analyses. Additionally, three child participants were excluded due to experimenter error, leaving a total of 28 children (age range: 3;1–3;11, mean age: 3;6, 12 boys).

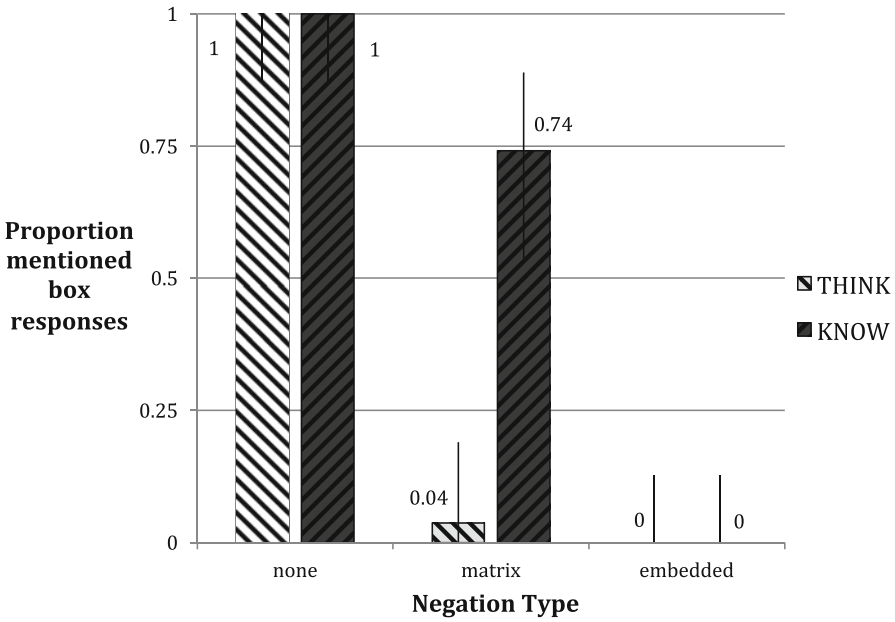


Fig. 3 Adult performance

5.2 Test Items

Adult data ($n = 9$) is given in Fig. 3. Adults always chose the mentioned box for affirmative *think* sentences (A) and affirmative *know* sentences (B). They never chose the mentioned box on *think* sentences with embedded negation (E) and *know* sentences with embedded negation (F). Finally, adults chose the mentioned box on 4 % of *think* trials with matrix negation (C) and 74 % of *know* trials with matrix negation.[2]

Child participants' performance ($n = 28$) is given in Fig. 4. Overall, children picked the mentioned box for affirmative *think* (A) and *know* (B) sentences. They picked the other box for *think* sentences with matrix negation (C) and well as for both *think* and *know* sentences with embedded negation (E, F). On *know* sentences with matrix negation (D), they picked the mentioned box about 40 % of the time.

A 2 × 3 ANOVA revealed a significant main effect of verb-type ($F(1,21) = 28$, $p < 0.017$) and negation- type ($p < 2.0e{-}16$) and a significant interaction between verb-type and negation-type ($p < 0.0072$). Planned comparisons revealed that children treat *think* sentences with matrix negation (C) differently from *know* sentences

[2] The 74 % (instead of the expected 100 %) in this condition comes from two participants who reported being unsure about sentences like (D) because they did not know "if the puppet wasn't aware it was in that box, or if it thought something else," otherwise performance was at 95 % in this condition. We return to this in section 6.

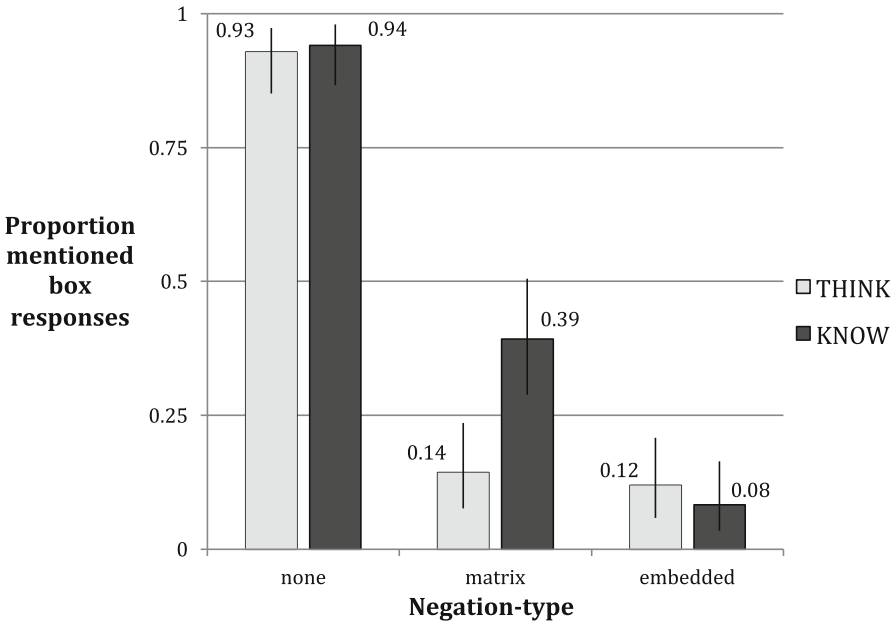


Fig. 4 : Three-year-olds' responses

with matrix negation (D) ($p < 0.017$) and that they treated *know* sentences with matrix negation (D) differently from *know* sentences with embedded negation (E) ($p < 0.0088$).

5.2.1 Think

All child participants performed completely adult-like on *think* trials; both child and adult participants picked the mentioned box for affirmative *think* sentences (A) but they picked the other box for both kinds of negative *think* sentences (C, E). See Fig. 5 for a comparison of adults' and three-year-olds' performance on *think* trials.

Note that the performance of both adults and children in this task is consistent with the assumption that Lambchop was a reliable source of information. Neither adults nor children seemed to compute a quantity implicature from the use of "think p" in the context of "know p" and "p": they always picked the mentioned box with affirmative *think* sentences. Given previous results from the literature, we expect that children would do so, but we had no such expectation for adult participants because an adult-like understanding of sentences like *Lambchop thinks that it's in the blue box* is consistent with the toy being in either the red or the blue box. We take the apparent lack of implicature computation in this task to be due to the "clue" status of the utterance: participants do not necessarily assume that the speaker is going

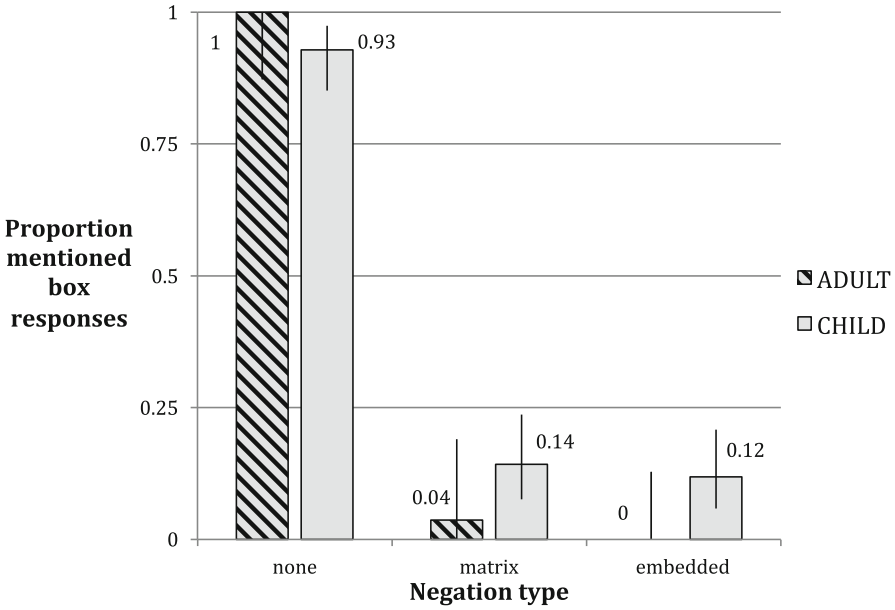


Fig. 5 Comparison of performance on *think* for adults and children

to make her contribution as informative as possible, but that she will provide just enough information to help them guess the correct location of the toy.

5.2.2 Know

Children appear to perform like adults in some *know* conditions, but not others; they pick the same box as adults on *know* trials with no negation (B) or embedded negation (F), but not with matrix negation (D). See Fig. 6 for a comparison of adults' and three-year-olds' performance on *know*.

On matrix negation (D) trials, which is where their behavior differs, adults tend to pick the mentioned box (consistent with a factive interpretation), but three-year-olds only pick the mentioned box about 40 % of the time. An examination of individual performance on this measure suggests that this 40 % performance is not due to chance performance (e.g., if children did not know which box to pick, they would alternate between picking the mentioned box and the other box). If all three-year-olds were guessing on sentences like (D), we would expect to see children distributed normally around a mean accuracy of approximately 50 %. However, children's performance on *know* sentences with matrix negation is distributed *bimodally*, and not normally around the mean, as it is in other conditions. There is a group that seems to perform below chance (consistent with a non-factive interpretation) and a group that seems to perform above chance (consistent with a factive interpretation). See Fig. 7 for an individual measure of performance.

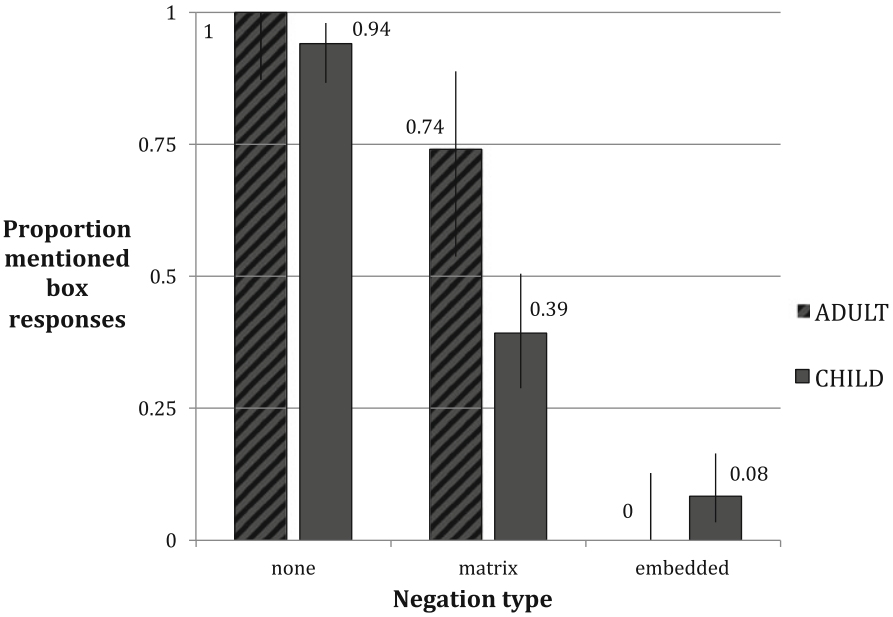


Fig. 6 Comparison of performance on *know* for adults and children

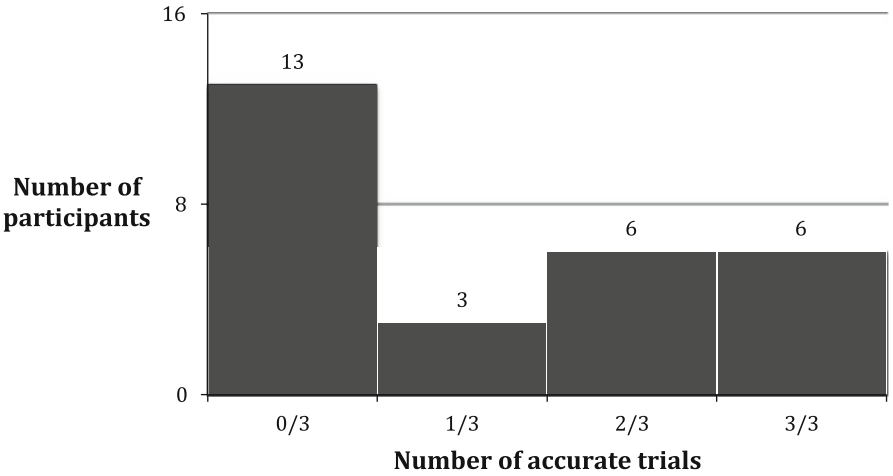


Fig. 7 Individual accuracy on *know* with matrix negation for children

Figure 7 shows that 6 three-year-olds (21 % of the group) always picked the mentioned box (consistent with a factive representation of *know*), 13 three-year-olds (46 %) always picked the other box (inconsistent with an adult-like understanding of *know*), and 9 three-year-olds (32 %) had more variable performance.

6 Discussion

These results demonstrate that three-year-olds, as a group, differentiate the factive verb *know* from the non-factive verb *think*, based on their significantly different responses to *think* and *know* sentences with matrix negation. This finding contrasts with previous findings in the literature where three-year-olds tested on the distinction between these verbs systematically failed (Johnson and Maratsos 1977; Abbeduto and Rosenberg 1985), or were found to be at chance (Moore and Davidge 1989; Moore et al. 1989). Unlike those studies, this one did not require participants to explicitly compare *think* and *know* sentences to decide which was a better description of the events, or to provide definitions of the verbs. Instead, this task required children to make choices in a game based on some linguistic stimuli. We believe that the metalinguistic nature of the previous tasks masked children's understanding of these verbs, and that our task was better able to assess their understanding.

Three-year-olds' high accuracy in all *think* conditions indicates that they have an adult-like understanding of *think*. Given their performance on the *think* sentences with matrix negation, we can conclude that three-year-olds, just like adults, understand *think* to be non-factive. When they hear sentences like (12), they do not infer that the toy is in the red box (which would be the expected outcome for a factive verb), but rather that the toy is in the blue box.

(12) Lambchop doesn't think [that its in the red box]

These results suggest that previous studies in which children failed to differentiate *know* and *think* could be due to extra pragmatic processing associated with the metalinguistic nature of tasks.

These findings also suggest the need for a more sophisticated analysis of the developmental trajectory of verbs like *know* than was previous thought. Our results show that a factive understanding of *know* may emerge earlier than four years of age. Some three-year-olds (about 43 %) consistently behave like they have an adult-like understanding of *know.* However, other three-year-olds (about 57 %) reliably treat *know* exactly like they treat *think*, namely as if it were non-factive. The apparent chance performance of three-year-olds in the matrix negation *know* sentences results from averaging the performance of the adult-like and non-adult- like children together. Therefore, past studies which found three-years-olds to be at chance in *know* conditions, like the studies by Moore and colleagues, might have yielded similar results if individual performance were measured.

To the extent that children's performance on our task is a direct reflection of their semantic representations for *know*, our data suggest that some children understand *know* to be factive by age three, but that others do not. It is, however, possible that even this simplified task is still obscuring three-year-olds' performance and that the failure of some to behave as if they understand *know* to be factive derives from an additional factor masking their knowledge.

Consider the design of the experiment. Recall that the puppet always whispered something to the experimenter, and after listening to the puppet, the experimenter

gave the clue (e.g., Lambchop knows it's in the red box). Participants never heard what the puppet actually said. So, the participant may have made inferences not about the puppet's beliefs about the location of the toy, but instead about what the puppet said. One possibility is that the puppet uttered statements of the form "It's in the red box". The experimenter, who was aware of the actual location of the toy, would then report what the puppet knows or thinks based on what it said. However, it's also possible to imagine that the puppet instead whispered statements such as "I think it's in the red box" or "I know it's not in the blue box", in which case, the experimenter would merely serve as translator for the shy puppet by reporting "She thinks it's in the red box" or "She knows it's not in the blue box." This conception of the interaction between the puppet and experimenter would lead to the expected "adult-like" behavior in every condition but the matrix negation *know* condition. In that case, the experimenter would be perceived to be relating the puppet's message: "I don't know that it's in the red box." But then the expected inference that the toy is in the red box would not be licensed. The only possible continuations of "I don't know that p" are "I don't know that p because p is not true," "I don't know that p because I know that q" or "I don't know that p because I don't have enough evidence." (An experiment in progress controls for this possibility by ensuring that the participants in the experiment know that the puppet's statement was of the form "it's in the. . . ".) While this possible interpretation of the experimental materials may have affected some participants, most adults and at least half of the children however behaved in a way consistent with the experimenter being responsible for giving the clue, and not serving as a mere translator.

Given this discussion, there seem to be three possible explanations for these children's failures to treat *know* as factive under matrix negation in this task:

(13) **Semantic failure:** These children have the wrong semantics for know. They understand know to be non-factive.

(14) **Processing failure:** These children have an adult-like semantics for know. Their failure lies in their difficulty accomodating the presupposition associated with know. Consider the processing demands to succeed on our task Children hear clues in the absence of context. Upon hearing the sentence Lambchop doesn't know that its in the red box, they have to realize that the speaker used a factive verb, which presupposes the truth of its complement. They then have to infer from her use of know that the speaker takes it for granted that the toy is in the red box. If the speaker it for granted that it's in the red box (and is in a good position to be justified in doing so, since she hid the toy), it must be that it is in the red box. Children should then choose the red box. It is possible that for some children, this inference process is too demanding. In effect, they would have all the right pieces but they would be unable put them together in this task.

(15) **Local accomodation failure:** These children have an adult-like semantics for know. They differ from adults in this task by favoring local, rather than global

accomodation of the presupposition. This preference may be driven by an assumption that the experimenter is reporting the puppets utterance of the form "I don't know that it's in the red/blue box".[3]

The first possibility (13) raises interesting questions both about how factivity is encoded in the target (i.e., adult) representation of *know*, and about how children come to acquire it.[4] How do learners determine that that *think* and *know* are different, and specifically that *know* is factive and *think* is not (which half of our three-year-olds seem to have already done)? What gives away the difference between the verbs? What gives away *know*'s factivity?

There are two possible sources of evidence that children might use to infer the meaning of novel words: the conversational context in which these words are used, and the linguistic environment in which they appear. As for the former, perhaps children can glean some meaning differences from the context in which verbs like *think* and *know* are used? Take the sentence '*x thinks that p*'. In contexts in which the speaker could have uttered '*x knows that p*', or simply *p*, but used '*x thinks that p'* instead, we, as Gricean adults, might infer that the speaker does not in fact endorse *p*. Could it be that children pick up on this, and somehow use it figure out the difference between *know* and *think*? They might reason that *think* is used when speakers want to indicate uncertainty, and use *know* when they want to indicate certainty: this, they might reason, must mean that *know* lexically encodes full certainty or 'knowledge'. We believe that this scenario is actually quite unlikely (at least for our successful three-year-olds). The literature on children's understanding of *think* shows that independently of their understanding of *know*, children overwhelmingly tend to assume that *p* is true whenever they hear *'x thinks that p*', even at age four (Lewis et al. 2012). If speakers truly and frequently use *think* to distance themselves from the truth of the complement, and use *know* to endorse it, children are not picking up on this, at least not by age 3 or 4. Moreover, this theory would entail that acquisition of both *think* and *know* is dependent on acquisition of the other, and that learners keep track not just of the interpretations that they assign to sentences but also to the pragmatic conditions that led them to make their interpretive decisions.

Perhaps then what gives away the difference in meaning between *think* and *know* is the syntactic environment in which these verbs appear. Syntactic bootstrapping, or learning about the meaning of a novel verb via its syntactic frames, occurs in conjunction with learning by observation and relies on systematic relationships between syntactic and semantic properties (Landau and Gleitman 1985; Gleitman 1990;

[3] Recent work by Bill et al. (2014), which argues that children do not locally accommodate, may provide support against such a possibility. (Thanks to F. Schwarz, for the reference and for helping us articulate this third option).

[4] Note that the question might partly depend on the representational status of this presupposition in the adult grammar. The standard view is that *know p* asserts that the subject believes p, and presupposes that p is true. Alternatively, it could be that *know* is veridical; it entails that p is true and pragmatically (instead of lexically) presupposes p (cf. Stalnaker 1978; Abusch 2002; Simons 2001).

Pinker 1989; Lidz 2006; a. o.). Developmental work on verb learning shows that syntactic bootstrapping is a viable option for simple verbs or verbs that span across broad semantic classes (Naigles 1990; Fisher et al. 2010; a. o.). Using syntactic cues to learn meaning differences between various attitude verbs that are impossibly difficult to figure out from direct observation should be particularly useful, though syntactic bootstrapping studies for attitude verbs are still inconclusive (cf. Asplin 2002).

The theoretical literature on attitude verbs shows that there are systematic relationships between the syntactic properties of biclausal sentences and the semantic classes of attitude verbs generally (Bolinger 1968; Hooper 1975; Stowell 1981; Pesetsky 1992; Grimshaw 1990; a.o.). As for *think* vs. *know* in particular, some authors have shown that there is a correlation between a predicate's factivity and its ability to take both declarative and interrogative complements (cf. Karttunen 1977; Hintikka 1975; Ginzburg 1995; Egré 2008; a.o.). Assuming that the link is principled, and that question-embedding is a reliable cue to factivity, it could be that hearing *know* sentences with interrogative complements provides evidence that it is factive. Under this view we would then want to ask: Why is it that some children have not arrived at the right meaning for *know* when others have? It's possible that the answer to this question boils down to the quantity and quality of input that children hear with respect to *know* and *think*. Some children may not have heard enough sentences of the right type to decide for sure that *know* is factive and *think* isn't. While we do not yet have a good idea of the quality and distribution of *know* and *think* in the input, a principled investigation is now underway.

7 Conclusion

Our data suggest that some children might begin to understand *know* in an adult-like way at an earlier age than the literature has indicated. The behavior of roughly half of our child participants is consistent with an adult-like understanding of *know*. The other half, however, do not distinguish *think* and *know*, even under negation, effectively treating neither one as factive. Thus some children distinguish *think* and *know* before age 4, even when they still assume (by default) that *think* sentences report true beliefs. Moreover, we find no evidence that children build a factive representation for *think*. Still, our results suggest that children's early representations of *know* may be non-factive and raise the question of how children come to recognize that *know* is factive and *think* is not.

These results also have impacts outside of the literature on children's understanding of factivity. Our results, taken in concert with those of Lewis et al, allow us to rule out the possibilities that children's difficulties with *think* are (i) conceptual, (ii) a result of ignoring the matrix verb, or (iii) due to a factive understanding of *think*. Instead, we can conclude that children's difficulties are pragmatic in nature. Additionally, there are implications for work on children's theory of mind. It is a longstanding puzzle that infants seem to track other people's knowledge states but that explicit measures with preschoolers seem to find no evidence of this capacity.

The fact that roughly half of our three-year-olds successfully treated *know* as factive suggests that there is more continuity between infants and preschoolers than the explicit measures indicate because it shows preschoolers can be sensitive to their interlocutors' knowledge and belief states in a linguistic task. Finally, these results help to refine the boundary between the pragmatic phenomena that children are good at and those that they are bad at.

Acknowledgement The authors would like to thank Kate Harrigan, Shevaun Lewis, Morgan Moyer, Aaron Steven White, and Alexander Williams for fruitful discussion of this work as various stages, as well as the members of the UMD Cognitive Neuroscience of Language Lab, the UMD Language Development and Parenting Lab and audiences at BUCLD 38, PHLINT 1, MACSIM 3rd, and CLS 49. The authors would also like to acknowledge members of the UMD Project on Children's Language Learning, including Morgan Moyer, Anne Ramirez and Susan Ojo for helping to collect data. This work was supported in part by a UMD Baggett Fellowship awarded to R. Dudley, NSF grant BCS-1124338 awarded to V. Hacquard and J. Lidz, and UMD's NSF-IGERT DGE-0801465.

References

Abbeduto, L., and S. Rosenberg. 1985. Children's knowledge of the presuppositions of know and other cognitive verbs. *Journal of Child Language* 12 (3): 621–641.

Abusch, D. 2002. Lexical alternatives as a source of pragmatic presuppositions. In *Proceedings of SALT*, vol. 12, ed. B. Jackson, 1–19. Ithaca: CLC Publications.

Asplin, K. N. 2002. Can complement frames help children learn the meaning of abstract verbs? PhD dissertation, UMass.

Berger, F., and B. Höhle. 2012. Restrictions on addition: Children's interpretation of the focus particles auch ('also') and nur ('only') in German. *Journal of Child Language* 39:383–410.

Bill, C., J. Romoli, F. Schwarz, and S. Crain. 2014. Indirect scalar implicatures are neither scalar implicatures nor presuppositions (or both). Poster presented at the 27th Annual CUNY Conference on Human Sentence Processing, Ohio State University, Ohio, March 2014.

Bolinger, D. 1968. Postposed main phrases: An English rule for the Romance subjunctive. *Canadian Journal of Linguistics* 14 (1): 3–30.

Bloom, P. 2002. *How children learn the meaning of words*. Cambridge: The MIT Press.

Chierchia, G., and S. MacConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics*. Cambridge: The MIT Press.

de Villiers, J. G. 1995. Questioning minds and answering machines. In *Proceedings of the Boston University Conference on Language Development*, eds. D. MacLaughlin and Susan McEwen. Somerville: Cascadilla.

de Villiers, J. G. 2005. Can language acquisition give children a point of view? In *Why language matters for theory of mind*, Apr, 2002.

de Villiers, J. G., and J. E. Pyers. 2002. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development* 17 (1): 1037–1060.

Diessel, H., and M. Tomasello. 2001. The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12 (2): 97–142.

Egré, P. 2008. Question-embedding and factivity. *Grazer Philosophische Studien* 77 (1): 85–125.

Falmagne, R. J., J. Gonsalves, and S. Bennett-Lau. 1994. Children's linguistic intuitions about factive presuppositions. *Cognitive Development* 9 (1): 1–22.

Fisher, C., Y. Gertner, R. M. Scott, and S. Yuan. 2010. Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science* 1 (2): 143–149.

Gergely, G., Z. Nádasdy, G. Csibra, and S. Bíró. 1995. Taking the intentional stance at 12 months of age. *Cognition* 56 (2): 165–193.
Ginzburg, J. 1995. Resolving Questions, I. *Linguistics and Philosophy* 18 (5): 459–527.
Grice, H. P. 1975. *Logic and conversation*, 41-58. Amsterdam: Elsevier.
Gleitman, L. 1990. The structural sources of verb meanings. *Language Acquisition* 1 (1): 3–55.
Grimshaw, J. 1990. *Argument structure*. Cambridge: The MIT Press.
Gualmini, A., S. Hulsey, V. Hacquard, and D. Fox. 2008. The question-answer requirement for scope assignment. *Natural Language Semantics* 16 (3): 205–237.
Halberda, Justin. 2003. The development of a word-learning strategy. *Cognition* 87 (1): B23–B34.
Hamburger, H., and S. Crain. 1982. Relative acquisition. *Language Development* 1:245–274.
Harris, R. J. 1975. Children's comprehension of complex sentences. *Journal of Experimental Child Psychology* 19:420–423.
Hintikka J. 1975. Different constructions in terms of the basic epistemological verbs. In *The intentions of intentionality*, 1–25. Dordrecht: Kluwer.
Hooper, J. B. 1975. On assertive predicates. In *Syntax and semantics*, vol. 4, ed. James Kimball, 91–124. New York: Academic.
Hooper, J. B., and S. A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4 (4): 465–497.
Hopmann, M. R., and M. P. Maratsos. 1978. A developmental study of factivity and negation in complex syntax. *Journal of Child Language* 5 (02): 295–309.
Huang, Y., and J. Snedeker. 2009. Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology* 58 (3): 376–415.
Hurewitz, F., S. Brown-Schmidt, K. Thorpe, L. R. Gleitman, and J. C. Trueswell. 2000. One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research* 29 (6): 597–626.
Johnson, C. N., and M. P. Maratsos. 1977. Early comprehension of mental verbs: Think and know. *Child Development* 8 (4): 1743–1747.
Karttunen, L. 1971. Some observations on factivity. *Research on Language & Social Interaction* 4 (1): 55–69.
Karttunen, L. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1 (1): 3–44.
Kiparsky, P., and C. Kiparsky. 1968. *Fact*. Linguistics Club, Indiana University.
Landau, B., and L. R. Gleitman. 1985. *Language and experience: Evidence from the blind child*, vol. 8. Cambridge: Harvard University Press.
Léger, C. 2007. The acquisition of two types of factive complements. In *Language acquisition and development: Proceedings of GALA*, eds. A. Gavarró and M. J. Freitas, 337–347. Newcastle: Cambridge Scholars.
Lewis, S. 2013. Pragmatic enrichment in language processing and development. PhD dissertation, University of Maryland.
Lewis, S., V. Hacquard, and J. Lidz. 2012. The semantics and pragmatics of belief reports in preschoolers. In *Proceedings of SALT,* ed. A. Chereches, vol. 22, 247–267.
Lidz, J. 2006. Verb learning as a probe into children's grammars. In *Action meets word: How children learn verbs,* eds. Kathy Hirsh-Pasek and Roberta M. Golinkoff, 429–449. Oxford: Oxford University Press.
Macnamara, J., E. Baker, and C. L Olson. 1976. Four-year-olds' understanding of pretend, forget, and know: Evidence for propositional operations. *Child Development* 47 (1): 62–70.
Merriman, W. E., and L. L. Bowman. 1989. The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development* 54 (3–4): 220.
Moore, C., and J. Davidge. 1989. The development of mental terms: Pragmatics or semantics. *Journal of Child Language* 16 (03): 633–641.
Moore, C., D. Bryant, and D. Furrow. 1989. Mental terms and the development of certainty. *Child Development* 60 (1): 167–171.
Musolino, J., and J. Lidz. 2006. Why children aren't universally successful with quantification. *Linguistics* 44 (4): 817–852.

Naigles, L. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17 (02): 357–374.

Noveck, I. A. 2000. When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition* 78 (2): 165–188.

Papafragou, A., and J. Musolino. 2002. The pragmatics of number. In *Proceedings from the 24th annual conference of the cognitive science society,* eds. W. D. Gray and C. Schunn. Austin: Congitive Science Society.

Papafragou, A., and D. Skordos. In press. Scalar implicature. In *Oxford handbook of developmental linguistics,* ed. J. Lidz, W. Snyder, and J. Pater. Oxford: Oxford University Press.

Perner, J., M. Sprung, P. Zauner, and H. Haider. 2003. Want that is understood well before say that, think that, and false belief: A test of de Villiers's linguistic determinism on German-speaking children. *Child Development* 74 (1): 179–188.

Pesetsky, David. 1992. *Zero syntax II: An essay on infinitives*. Cambridge: Manuscript, MIT.

Pinker, S. 1989. *Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge: The MIT Press

Rooryck, J. 2001. Evidentiality, part I. *Glot International* 5 (4): 125–133.

Scoville, R. P., and A. M. Gordon. 1980. Children's understanding of factive presuppositions: An experiment and a review. *Journal of Child Language* 7 (02): 381–399.

Simons, M. 2001. On the conversational basis of some presuppositions. In *Proceedings of SALT*, vol. 11, ed. R. Hastings, B. Jackson, and Z. Zvolensky, 431–448. Ithaca: Cornell University.

Simons, M. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117 (6): 1034–1056.

Sowalsky, E., V. Hacquard, and T. Roeper. 2009. Is PP opacity on the path to false belief? *Generative Approaches to Language Acquisition North America (GALANA)* 3:263–261.

Stalnaker, Robert C. 1978. Assertion. In: *Pragmatics (Syntax and semantics 9)*, ed. Peter Cole, 315–332. New York: Academic Press.

Stowell, T. A. 1981. Origins of phrase structure. Doctoral dissertation, Massachusetts Institute of Technology.

Syrett, K., C. Kennedy, and J. Lidz. 2010. Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics* 27 (1): 1–35.

Tardif, T., and H. M. Wellman. 2000. Acquisition of mental state language in Mandarin-and Cantonese-speaking children. *Developmental Psychology* 36 (1): 25–43.

Thornton, R., and S. Crain. 1999. Levels of representation in child grammar. *Linguistic Review* 16:81–123.

Trueswell, J. C., I. Sekerina, N. M. Hill, and M. L. Logrip. 1999. The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition* 73 (2): 89–134.

Urmson, J. O. 1952. IV.-Parenthetical verbs. *Mind* 61 (244): 480–496.

Warneken, F., and M. Tomasello. 2006. Altruistic helping in human infants and young chimpanzees. *Science* 311 (5765): 1301–1303.

Wellman, H. M., D. Cross, and J. Watson. 2001. Metaanalysis of theory of mind development: The truth about false belief. *Child Development* 72 (3): 655–684.

Woodward, A. L. 1998. Infants selectively encode the goal object of an actor's reach. *Cognition* 69 (1): 1–34.