Remarks on the experimental turn in the study of scalar implicature Part I and Part II

Emmanuel Chemla and Raj Singh

This document includes the pre-final version of a review study that was published in *Language and Linguistics Compass* in two parts:

– Chemla, E. & Singh, R. (2014). Remarks on the experimental turn in the study of scalar implicature, Part I. *Language and Linguistics Compass.*

– Chemla, E. & Singh, R. (2014). Remarks on the experimental turn in the study of scalar implicature, Part II. *Language and Linguistics Compass.*

Remarks on the experimental turn in the study of scalar implicature, Part I

Emmanuel Chemla and Raj Singh

December 4, 2013

Abstract (for Part I and Part II)

There has been a recent 'experimental turn' in the study of scalar implicature, yielding important results concerning online processing and acquisition. This paper highlights some of these results and places them in the current theoretical context. We argue that there is sometimes a mismatch between theoretical and experimental studies, and we point out how some of these mismatches can be resolved. We furthermore highlight ways in which the current theoretical and experimental landscape is richer than is often assumed, and in light of this discussion we offer some suggestions for what seem to us promising directions for the experimental turn to explore.

The article is divided in two parts. Part I first presents the two dominant families of accounts of scalar implicature, the domain-general Gricean account and the domain-specific grammatical account. We try to separate the various components of these theories and connect them to relevant psycholinguistic predictions. Part II examines and reinterprets several prominent experimental results in light of the theoretical presentation proposed in the first part.

1 Background on scalar implicature

1.1 The traditional introduction

Hearing a sentence such as (1), *Some of Mary's students got an A*, a competent speaker of English will most naturally conclude that not all of Mary's students got an A. This inference is called a *scalar implicature*. Uncontroversially, this inference is the result of some competition between two alternatives: the uttered sentence (1) and a minimally different unspoken sentence (2). The competition here resolves in the inference that the alternative in (2) is false, i.e., in the scalar implicature in (3). The conjunction of the literal meaning of the sentence and its scalar implicature is the 'strengthened meaning' of the sentence, (4).

(1) Some of Mary's students got an A.

(2) *Alternative:* All of Mary's students got an A.

(3) *Scalar implicature:* Not all of Mary's students got an A.

(4) *Strengthened Meaning:* Some but not all of Mary's students got an A.

The adjective 'scalar' witnesses the idea that the source of the relevant alternatives lies in lists of words ordered such that replacing one with another would, in many environments, yield sentences ordered by informativity, where 'informativity' is often measured by logical strength (though cf. section 3.2 from Part II). In our example, the relevant scale is ⟨*some*, *all*⟩, and placing one word or the other in ____ *of Mary's students got an A* provides the two alternatives (1) and (2), with (2) logically stronger than (1). The mechanism generalizes: scales are used to explain similar *X* implies *not-Y* inferences, using ⟨*X*, *Y*⟩ pairs such as ⟨*or*, *and*⟩, ⟨*possibly*, *necessarily*⟩, etc.

The word 'implicature' comes from works inspired by Grice's seminal work on other phenomena, following which these types of inference are assumed to be the result of pragmatic reasoning. A goal of the Gricean line of research is to reduce the computation of scalar implicatures to principles of social interaction which apply beyond linguistic communication. These principles are spelled out as a set of conversational maxims in Grice (1967). The driving principle behind a Gricean explanation of scalar implicatures is that, all things being equal, cooperative speakers should convey as much relevant information as possible. The starting point of the explanation is that in contexts in which (1) and (2) are both relevant, a rational and cooperative speaker who has enough evidence to warrant assertion of (2) should assert (2), since (2) is more informative than (1). As a result when the hearer processes (1) they can safely conclude that the speaker was not in position to assert (2), and with some additional assumptions about the speaker's epistemic state (e.g., that the speaker knows *whether* (2)) the hearer can conclude that the speaker knows that (2) is false (see e.g., Grice, 1967; Horn, 1972; Gamut, 1991; Sauerland, 2004; van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Russell, 2006; Spector, 2006; Chemla, 2009a).

1.2 Properties of the Gricean approach

Precise implementations of the Gricean approach make several important predictions, two of which have been particularly central to debates concerning the mental mechanisms governing implicature computation:

(5) `Stronger Alternatives`: Scalar implicatures should only arise when a sentence is asserted in a context in which a more informative (logically stronger) alternative is relevant; recall that the rationale behind implicature computation is that the speaker could have provided more information, but chose not to.

(6) `Global Reasoning`: The computation of scalar implicatures is post-compositional; because the reasoning applies at the level of speech acts (*why did the speaker say what s/he said, rather than saying something else?*), it must involve comparing the logical strength of full sentences.

Motivated by the governing principle of rational cooperation, these predictions provide straightforward explanations for otherwise complex data. For example, they explain why a sentence such as (7-a) does not trigger a *not-all* inference of the kind we saw for (1): the alternative (7-b) is not logically stronger than (7-a) and thus (because of `Stronger Alternatives`) does not trigger the Gricean reasoning sketched above, and since (1) is in an embedded position in (7-a), it cannot (because of `Global Reasoning`) receive a so-called 'embedded implicature,' which reading we could paraphrase as 'if some but not all of Mary's students got an A, she will be happy' (see also note 7).

(7) a. If some of Mary's students got an A, she will be happy.
b. If all of Mary's students got an A, she will be happy.

1.3 Some challenges for the Gricean approach

Despite their considerable empirical success and their grounding in intuitively appealing principles of rational cooperative action, `Stronger Alternatives` and `Global Reasoning` were challenged rather soon after Grice (1967)'s seminal work appeared, and have continued to face resistance in more recent times (see e.g., Cohen, 1971; Horn, 1972; Hirschberg, 1985 for early challenges, and see Chierchia et al., 2008 for a summary of recent challenges). As an example challenge to `Global Reasoning`, consider Hurford (1974)'s generalization that disjunctions whose disjuncts entail one another are infelicitous:

(8) a. #John ate an apple or a fruit.
b. #John drives a car or a Toyota.

Gazdar (1979) noted that Hurford's generalization is systematically obviated by disjunctions whose disjuncts form $\langle X, Y\rangle$ pairs:

(9) Some of Mary's students got an A or all of them did.

Why should there be a contrast between (8) and (9)? Assuming the correctness of Hurford (1974)'s generalization, a natural explanation of the felicity of (9) is that the embedded disjunct *some of Mary's students got an A* (= (1)) is interpreted with its strengthened meaning (4), so that the relevant reading of (9) might be paraphrased as *only some of Mary's students got an A or all of them did.* With this so-called 'embedded implicature' the Hurford configuration is avoided because the disjuncts no longer stand in an entailment relation. Since no such implicature is available in (8) (e.g., *John ate a fruit* does not implicate that John didn't eat an apple), embedded implicatures cannot rescue the sentences in (8) from oddness.[1] The appeal to embedded implicatures is only possible, of course, if `Global Reasoning` is incorrect (see e.g., Chierchia et al., 2008 for extensive discussion).

As an example challenge to `Stronger Alternatives` consider the sentence in (10-a), its scalar alternative in (10-b), and the reading that results from conjoining (10-a) with the negation of (10-b), paraphrased in (10-c):

(10) a. Exactly one girl passed some of her exams.
b. Exactly one girl passed all of her exams.
c. Exactly one girl x passed some of x's exams, x did not pass all of x's exams, and for all other girls y, y did not pass any of y's exams.

[1]See Gazdar (1979), Hirschberg (1985), Krifka (1999) for discussion of the lack of implicature in hyponym-hyperonym pairs like (8), and see Singh (2008b) for an attempt to understand this gap in a formal framework. This is not just a matter of hyponym-hyperonym pairs. Note for example that if we change the first disjunct in (9) to *at least some of Mary's students got an A*, which does not implicate (3) (e.g., Krifka, 1999), the oddness reappears: # *At least some of Mary's students got an A or all of them did.*

The important question is whether (10-c) is a possible strengthened meaning of (10-a). If it is, then the alternatives that get negated when computing implicatures cannot, as demanded by `Stronger Alternatives`, be restricted to *stronger* alternatives because (10-b) is not stronger than (10-a).[2] Indeed, the two are logically independent.[3] Many recent works have argued that (10-c) is indeed a possible strengthened meaning of (10-a) (e.g., van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Spector, 2005, 2006; Chemla, 2009a; Chierchia et al., 2008; Chemla and Spector, 2011).[4]

1.4 The grammatical approach

In response to these and other challenges an alternative view of scalar implicatures has emerged according to which scalar implicatures are the result not of pragmatic reasoning, but rather of a grammatical operation (e.g., Chierchia, 2004, 2006; Fox and Hackl, 2006; Fox, 2007; Chierchia et al., 2008; Magri, 2009, 2011; Gajewski and Sharvit, 2012). A useful metaphorical representation of this 'grammatical theory of scalar implicature' is that the relevant grammatical operation is the application of an invisible operator, call it O, which would to a large extent have the same effect as the word *only*. Under this approach, the scalar implicature for (1) would be explained by (i) the fact that the sentence *Only some of Mary's students got an A* conveys a *not-all* inference and (ii) the assumption that the O operator is real and has a meaning which for current purposes we can identify with the meaning of *only*. These two assumptions lead to the consequence that a sentence like (1) is structurally ambiguous between a parse with O, equivalent to *only some of Mary's students got an A* and therefore responsible for the *not-all* inference in (3), and a parse without O, giving rise to the classical reading which is compatible with all of Mary's students getting an A. Under this view, interpretation of (1) requires access to a disambiguation mechanism. This is one of the ways in which the grammatical theory differs from the Gricean picture – the sentence in (1) generates only one form-meaning pair under the Gricean view.[5]

The grammatical approach raises important questions about the nature of this ambiguity, such as what the factors are that govern its resolution. We will return to this matter shortly. For now, we would like to highlight the ways in which the grammatical view provides a perspective on strengthened meanings from which it is natural to reject `Stronger Alternatives` and `Global Reasoning`. It is natural to reject `Stronger Alternatives` under the gram-

[2]This is because *exactly one* is a non-monotonic operator. A non-monotonic operator OP is an operator that neither preserves nor reverses entailment relations: if X entails Y, and X and Y are appropriate arguments to Op, $Op(X)$ and $Op(Y)$ need not stand in an entailment relation.

[3]To see that (10-b) does not entail (10-a), consider that several girls passed at least one of the exams, and that only one of these girls passed all of them: (10-b) is then true but (10-a) is false. To see that (10-a) does not entail (10-b), consider that only one of the girls passed any exams, and that this girl failed to pass all of them: (10-a) is then true but (10-b) is false.

[4]In fact, the experimental results reported in Chemla and Spector (2011) reveal this violation of `Stronger Alternatives` and *also* suggest that, in violation of `Global Reasoning`, an embedded implicature is available in (10-a). The reading resulting from this possibility can be paraphrased as 'exactly one student of Mary's passed some but not all of her exams.'

[5]With n sentential constituents in the sentence the possibility of adding O to any constituent generates 2^n parses unless restrictions are placed on O-application (see Fox and Spector, 2008; Singh, 2008b,a; Ippolito, 2011; Gajewski and Sharvit, 2012 for evidence that there are restrictions on O-application, and see Fox and Spector, 2008 and Singh, 2008a for attempts to characterize these restrictions). Recursive application of O will increase this number, but under reasonable conditions the number of form-meaning pairs that get generated will remain finite (Fox, 2007).

matical theory because it identifies O as a silent variant of *only*, and since *only* is commonly analyzed as an operator that negates alternatives that are not only stronger than its sentential argument (the so-called 'prejacent'), but also those which are merely non-weaker than it (see e.g., Horn, 1972; Hirschberg, 1985; Rooth, 1992; van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Spector, 2005, 2006; Chierchia et al., 2008; Chemla, 2009a; Chemla and Spector, 2011), it would be natural to interpret O as an operator that does just this. And the grammatical theory makes it natural to reject `Global Reasoning` because O-insertion is a syntactic process, and hence should apply to embedded constituents. This possibility leads to the expectation that embedded implicatures should exist (cf. our discussion of (9)).

1.5 A note about Relevance Theory

One framework that has often been called for in the experimental literature to derive online predictions from offline theories is Relevance Theory (Sperber and Wilson, 1986). However, Relevance Theory does not offer a theory of scalar implicatures. Instead, it provides a general framework and terminology to draw and describe expectations about how theories of scalar implicatures and almost just any other phenomena can lead to processing predictions (in fact, Noveck and Sperber, 2007, who are good representatives of relevance theorists, develop quite the opposite of a theory of scalar implicatures, arguing against the existence of the phenomenon). Hence, Relevance Theory *per se* is not at all a competitor of the Gricean and grammatical theories we described: its purpose is different. One may see it as a good complement of these theories, but it is not a necessary one either as long as an effort is made to be explicit about cognitive aspects of the implementation of these theories.

1.6 The debate: the division of labour between grammar and pragmatics in the computation of scalar implicature

The upshot is that the study of scalar implicature in the last decade has witnessed a lively period of debate and development (see e.g., Horn, 2005 and Chierchia et al., 2008 for surveys), and `Stronger Alternatives` and `Global Reasoning` have played a prominent role in these debates. For example, because the Gricean approach is committed to these predictions it can explain the disappearance of scalar implicatures in (7-a), but it is faced with the challenge of making sense of the contrast between (8) and (9) and of the possibility of the scalar implicature in (10-c). On the other hand, while the grammatical theory can appeal to embedded implicatures to account for the felicity of (9), and it exploits the possibility of negating non-weaker alternatives to account for the scalar implicature in (10-c), it is faced with the challenge of explaining why an embedded implicature in (7-a) seems to be dispreferred (e.g., Chierchia et al., 2008; Geurts and Pouscoulous, 2009; Panizza et al., 2009), and of explaining new puzzles that arise when the negation of non-weaker alternatives becomes possible.[6]

Underlying these empirical questions lies a fiercely disputed question of modularity: are scalar implicatures the ouptut of domain-general reasoning procedures derived from principles of rational cooperation, as in the Gricean picture, or are they the output of domain-specific

[6]For instance, why does *All of Mary's students got an A* (= (2)) not implicate that none of Mary's students got an A ('not (1)'; see Schlenker, 2012)? In textbook treatments of classical logic (2) and the negation of (1) are incompatible only if it assumed that the domain is non-empty. Unless we assume, then, that there are boys, (1) and (2) are logically independent and so we are left with the problem of explaining why assertions of (2) do not implicate that there are no boys. (Note that unless this existence assumption is made `Stronger Alternatives` will not be met, which in turn would leave the Gricean theory without an explanation of the observation that (1) implicates (3)).

computations taking place within the linguistic system, as proposed by the grammatical theory?[7] We hope to impress upon the reader that this way of formulating the question, prominent in much of the literature, is something of a false dichotomy. As we hope to clarify, both approaches involve interactions between grammar (syntax/semantics) and pragmatics at all relevant stages of implicature computation. Given the highly interactive nature of implicature computation, the only legitimate question is what the right division of labour is between the grammatical system and grammar-external cognitive systems in carrying out these computations.[8] This question of cognitive architecture is intimately connected to questions of processing, acquisition, and information flow within the mind (e.g., questions of encapsulation and the like, cf. Fodor, 1983). It is thus no surprise that these developments have led to a burst of work introducing the methods of experimental psychology in service of resolving this debate.

1.7 Experimental turn

The existence of survey articles on experimental studies of scalar implicature (e.g., Noveck and Reboul, 2008; Katsos and Cummins, 2010) can be taken as evidence of what we might call an 'experimental turn' in this domain. The origin of the recent interest from experimentalists can be located in Bott and Noveck (2004), who revived work by Rips (1975) investigating the complexity of processing quantified sentences like (1) and (2), and in Chierchia et al. (2001), Gualmini et al. (2001), and Noveck (2001), who revived work on the acquisition of quantifiers and connectives by Paris (1973), Smith (1980), and Braine and Rumain (1981). This revival provides a welcome example of linguists and psychologists studying a phenomenon in parallel and offering results that inform each other's scientific enterprise. The sociological consequences go beyond scalar implicatures, with the rise of 'experimental pragmatics and semantics' as a field. Linguists are now accustomed to experimental methods which provide access to data that are not available with standard tools. In addition to introspective judgments the theory of implicature now benefits from evidence obtained using a broad range of methodologies including offline judgments about the availability of inferences or picture matching studies (e.g., Geurts and Pouscoulous, 2009; Chemla, 2009b; Clifton and Dube, 2010; Chemla and Spector, 2011; Romoli and Schwarz, 2013), self-paced reading (e.g., Breheny et al., 2006; Chemla et al., 2013), eye-tracking and mouse tracking (e.g., Huang and Snedeker, 2009; Tomlinson et al., 2011, 2012), response times and Speed Accuracy Trade-Off paradigms (e.g., Bott and Noveck, 2004; Degen and Tanenhaus, 2011; Chemla and Bott, 2011; Bott et al., 2012; Cremers and Chemla, 2013), dual-task studies investigating the use of working memory resources (De Neys and Schaeken, 2007; Marty and Chemla, 2011), betting-paradigms (e.g., Goodman and Stuhlmuller, 2013), and acquisition studies using the truth-value judgment task and variations thereof (e.g., Noveck, 2001; Chierchia et al., 2001; Gualmini et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005; Crain, 2008; Crain and Khlentzos, 2010; Barner and

[7]This modern controversy revives a 19th century debate, discussed by Horn (2004, 2005), between those who argued that sentences like (1) were ambiguous (e.g., Hamilton, 1860), and those who argued that the inference to (3) came about through pragmatic reasoning (e.g., Archbishop Whately, 1848; De Morgan, 1847; Mill, 1867).

[8]Our focus here is on empirical issues, but as noted in Fox (2007, 2013) there are conceptual tradeoffs also: the grammatical theory buys a simple pragmatics (a pure 'Gricean' one) at the cost of complicating grammar (by positing O), while the neo-Gricean approach buys a simple grammar (no O) at the cost of complicating pragmatics (with stipulations about ALT in the pragmatics itself, turning 'Griceanism' into 'neo-Griceanism'). We will steer clear of this conceptual stalemate, and will use the terms 'Gricean' and 'neo-Gricean' interchangeably (as is commonly done). However, if Fox is right, this terminological confusion is at best highly misleading. See note 10.

Bachrach, 2010; Barner et al., 2011; Stiller et al., 2011; Foppolo et al., 2012; Zhou et al., 2013; Singh et al., 2013).

What is the source of this experimental interest and success? We submit that scalar implicature is a phenomenon with the following two properties: (i) its most widespread theoretical descriptions and disputes (ambiguity versus pragmatic reasoning) can be sketched in intuitive terms (see above); (ii) it raises questions of fundamental importance for cognitive science, such as modularity; since it is a phenomenon at the interface between semantics and pragmatics, an understanding of scalar implicature is likely to yield insights into the broader organization of language and mind. One consequence of (ii) is that data gathered using the methods of experimental psychology (e.g., derivation time course or stages and paths of acquisition) gain immediate importance. For example, if scalar implicatures are pragmatic in nature, one might expect them to be acquired along with other pragmatic skills (see e.g., de Villiers, 2007 for general discussion of the bidirectional influences of the development of linguistic and extra-linguistic capacities). One problematic consequence of (i), however, is that it creates a risk of over-simplification, a matter to which we will soon return.

2 Integration of theoretical models and experimental methods

The new kinds of data made available by the experimental turn hold the promise to radically advance our understanding of scalar implicature. However, this potential has not been fully realized, partly due to simplifying assumptions about the competing theories and about their relation to processing and development. To discuss these matters precisely it will be useful to identify what the competing theories view as the computational problems the implicature mechanism solves, and to state some auxiliary assumptions that relate these computations to models of performance.

2.1 Computational Problems

This section identifies the tasks the implicature mechanism has to solve under the competing theories. We will see that both the grammatical and the Gricean theory posit a highly interactive mechanism, with communication between grammar and pragmatics throughout the computation.

2.1.1 Gricean view

The Gricean view posits the path from utterance to strengthened meaning outlined in Figure 1. Under this view, we start with an utterance of sentence S, to which the grammar assigns its literal meaning, [[S]]. For example, the literal meaning of (1) is that John ate at least some of the cookies. A set of alternative sentences is then derived in a collaborative effort between grammar and pragmatics: a set of *formal* alternatives $ALT(S)$ is derived by the grammar (see note 7 in Part II), a set of *relevant* alternatives $\mathcal{R}$ is derived by context (often identified with the complete answers to the question under discussion; see Groenendijk and Stokhof, 1984; Lewis, 1988), and the intersection of these two sets, $ALT^{*}(S)$, is the set of *actual* alternatives used in the computation (for details, see e.g., Horn, 1972; Gazdar, 1979; Rooth, 1992; Katzir, 2007; Chemla, 2009a; Fox and Katzir, 2011).[9] For example, returning to (1), if we assume

[9] It is natural to think of this process as the context pruning members of $ALT(S)$ based on considerations of relevance. Such an interpretation would bring this process in line with contextual domain restriction elsewhere. See e.g., Rooth (1992) and Fox and Katzir (2011).

that $ALT((1))$ = {(1), (2)} and that context decides that each element of $ALT((1))$ is relevant, then $ALT^*((1))$ = {(1), (2)}, but if context decides that (2) is irrelevant, then $ALT^*((1))$ = {(1)}.

The sentence S, and its actual alternatives $ALT^*(S)$, serve as input to a defeasible pragmatic strengthening mechanism which considers those members $S^+ \in ALT^*(S)$ that are stronger than S (cf. `Stronger Alternatives`), and concludes that the speaker believes S (speakers assert only those propositions they believe to be true) and does not believe S^+ (for otherwise the speaker would have uttered one of the stronger alternatives S^+ instead). These inferences are sometimes called 'primary implicatures.' In our running example, if $ALT^*((1))$ = {(1), (2)}, at this stage we have $B_S(1)$ and $\neg B_S(2)$, where $B_S(x)$ is shorthand for 'the speaker believes that x'. The inference $\neg B_S(2)$ is the primary implicature of (1). In a subsequent stage, the hearer makes a judgment on whether the speaker can be assumed to have an opinion on whether these stronger alternatives S^+ are true or false, that is, on whether the so-called 'competence assumption' holds (e.g., van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Chemla, 2008). The factors that enter into this judgment are unclear (though see Goodman and Stuhlmuller, 2013 for explicit manipulation), but if the hearer makes the competence assumption they will conclude that these stronger alternatives are false (secondary implicatures about each S^+), and if the hearer does not make the competence assumption they will conclude that the speaker is ignorant about the truth-value of all these stronger alternatives (ignorance implicatures about each S^+). Finally, with the assumption that the speaker has sound evidence for their beliefs, the hearer concludes that A for each conclusion $B(A)$ produced by the reasoning. Thus, an utterance of (1) will lead the hearer to conclude that (1) is true. If the competence assumption holds, they will also conclude that (2) is false (speakers have sound evidence for their beliefs), and if the competence assumption does not hold, they will conclude that the the speaker is ignorant about the truth of (2). And had context pruned (2) to begin with, no inferences about this proposition would be made.

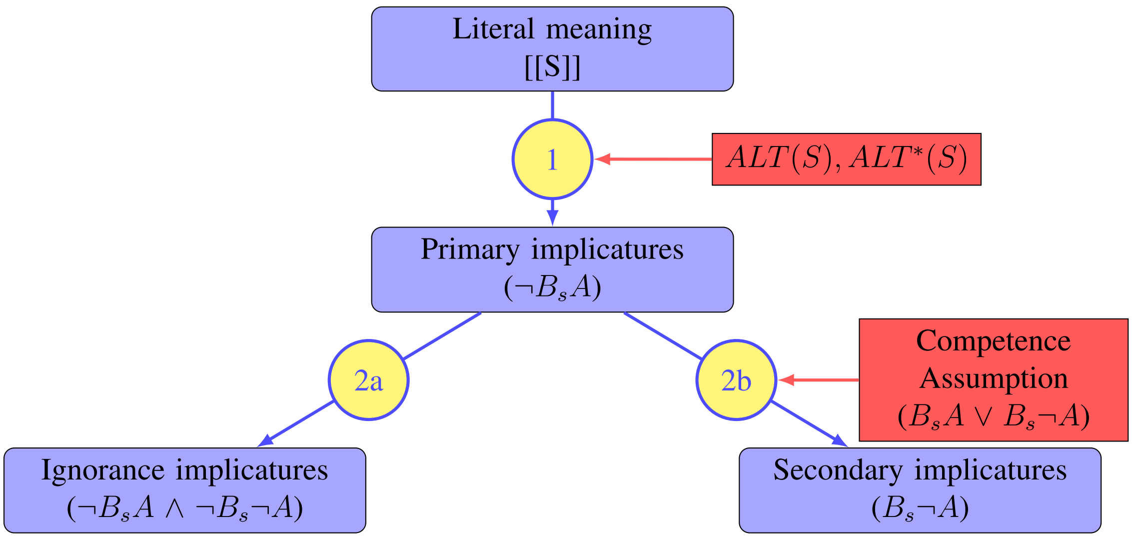

Figure 1: Steps in the computation of the meaning of sentences with potential scalar implicatures according to the Gricean family of accounts

2.1.2 Grammatical view

Figure 2 is the counterpart of Figure 1 for the grammatical view. We start again from an utterance of a sentence S. Analysis of the utterance requires the hearer to decide whether to parse the sentence with or without the O operator, as S or $O(S)$. This decision does not have to be made in the Gricean view. However, as under the Gricean view, each parse is assigned a literal meaning: if the sentence is parsed without O, then S gets assigned its literal meaning [[S]] as before, but if the sentence is parsed with an O, then [[O(S)]] is computed. The computation of this meaning involves several steps:

(11) Steps in the computation of [[O(S)]]

a. Alternatives: O, like *only*, works over a set of alternatives, and the alternatives are assumed to be the same as those involved with implicature computation (e.g., Rooth, 1992; Fox, 2007; Fox and Katzir, 2011); thus, as in the pragmatic approach, the system needs to derive $ALT(S)$ and its contextually modified variant $ALT^*(S)$.

b. Strengthening: Given S and $ALT^*(S)$ as input, O decides by a formal algorithm which members of $ALT^*(S)$ should be negated (the precise mechanics are not relevant for current purposes – see Fox, 2007 for a formal statement).

There are two things to note so far. First, there needs to be a decision about how to disambiguate the sentence. Second, the parse with O brings about additional computations, such as those mentioned in (11-a) and (11-b), in addition to the syntactic operations required to merge O and $ALT^*(S)$ into the structure (to avoid clutter we will continue to write $O(S)$ as shorthand for the more articulated logical form $O(ALT^*(S))(S)$).

Next, no matter how the sentence is parsed, the result is delivered to a pragmatic mechanism that makes various inferences about the beliefs of the speaker. First, if the parse of the sentence is A, the hearer concludes $B_S A$. Second, for all relevant propositions C whose truth-value is left open by the parse of the sentence, the hearer generates ignorance implicatures about C, $\neg B_S C$ and $\neg B_S \neg C$. This follows from assumptions about the nature of quantity reasoning that we will put aside here (see note 10). To see how the reasoning works in our example, suppose that (1) is uttered in a context in which each member of $ALT(1) = \{(1), (2)\}$ is relevant, and nothing else is. If the hearer parses the sentence with an O, we have [[O(S)]] = (4). If the hearer parses the sentence without an O, we have [[S]] = that at least some of Mary's students got an A (possibly all). This leaves open the truth-value of (2), and so the hearer concludes that the speaker is ignorant about whether all of Mary's students got an A.

2.1.3 Summary

The two theories have a different architecture. In particular, they assume that the strengthened meaning of a sentence is produced in different cognitive systems. Each assumes that there is some function, f, responsible for strengthening the meaning of a sentence like (1) into (4). However, the Gricean view places f in the pragmatic system. The grammatical theory identifies f with an unpronounced morpheme, O. At the same time, the theories share various components. Surely, both of them involve pragmatic and grammatical components. For example, the derivation of alternatives is identical in the two systems, involving both grammatical (ALT) and pragmatic considerations (ALT*). Second, the decision to apply f or not, or to

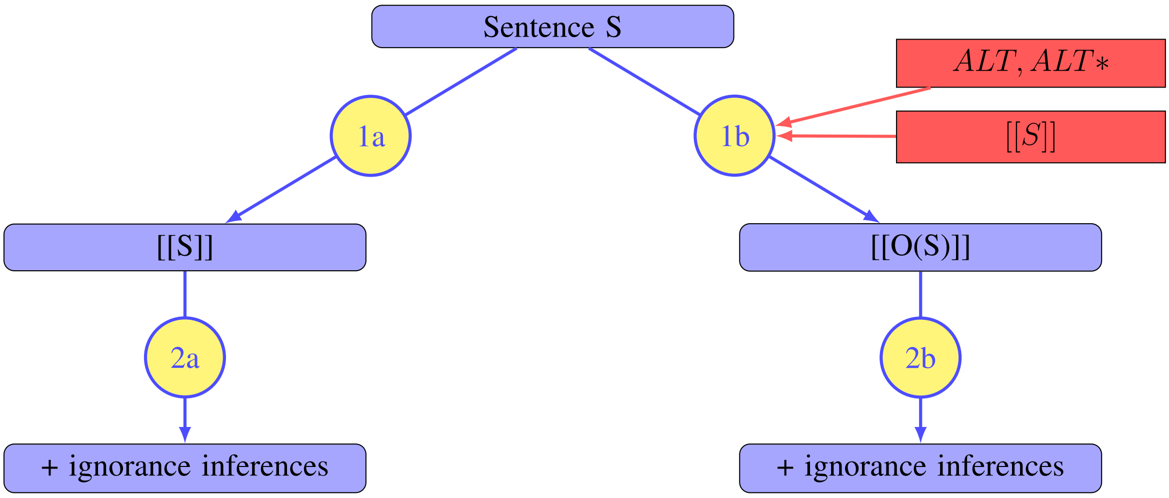

Figure 2: Steps in the computation of meaning for sentences with potential scalar implicatures according to the grammatical family of acccounts

stop at any point in the computation, is a pragmatic one in both cases. These similarities and differences are summarized in Table 1.[10]

	Grammatical theory	Neo-Gricean theory
Grammatical operations	ALT f	ALT
Pragmatic operations	ALT^* decision to apply f (disambiguation)	f ALT^* decision to apply f

Table 1: Grammar vs Pragmatics: all accounts have some grammatical and pragmatic components

The question to which we would like to now turn is how these models should be operationalized as representations of what occurs in the minds of speakers and hearers.

[10]Fox (2007, 2013) points out that the pragmatic demands on speakers are also different in the two approaches: under the grammatical theory the quantity maxim guides speakers to say the strongest relevant proposition that they know to be true, while under the neo-Gricean theory the maxim guides speakers to say the strongest relevant proposition in ALT^* that they know to be true (see note 8 and section 3.1 in Part II).

2.2 Auxiliary assumptions about processing

We have so far been concerned with laying out the principles behind the Gricean and grammatical theories of implicature, and with specifying the computational problems the implicature mechanism is thought to solve under each approach. These theories have been largely motivated by a particular kind of data, namely the range of interpretations that get assigned to various sentences in various contexts of use, typically attained using elicited judgments from informants. The theories have thus been concerned with stating explicit assumptions about grammar (syntax and semantics) and pragmatics that yield the desired interpretations. It would seem fair to conclude, then, that these theories have primarily been concerned with capturing a speaker-hearer's *knowledge* of grammar and pragmatics, sometimes called a speaker-hearer's 'competence' (e.g., Chomsky, 1965).

What do these theories of grammatical and pragmatic competence predict about various questions we might be interested in asking, such as how long it would take a subject to compute the inference in (3) upon hearing or reading (1)? So far these theories make *no* predictions about this. What is needed is a set of auxiliary assumptions about how these knowledge systems are used in behaviour, assumptions that are sometimes said to belong to a theory of 'performance' (e.g., Chomsky, 1965).[11] Note that we are *not* saying that these theories are not responsible for processing data. To the contrary, a complete theory of scalar implicature will aim to capture data of whatever kind that is relevant to characterizing the systems responsible for implicature computation. These data may come from offline elicited informant judgments, complexity measures in comprehension/production tasks, brain imaging studies, corpus studies, stages in acquisition, and much else. However, any theory aiming to describe and explain these data will likely be decomposed into various component parts $T_1, T_2, \ldots, T_k$, the joint description of which constitutes the entire theory.

To illustrate, suppose that T_1 is one of our candidate competence theories, and that we wish to extend T_1 to a theory of processing, say to account for response time data concerning comprehension of (1). We will proceed to make explicit assumptions about processing. These might include assumptions such as that in production/comprehension the processing mechanism executes computations specified by the knowledge systems of syntax/semantics/pragmatics. This so-called 'Competence Hypothesis' (Bresnan and Kaplan, 1982) relates competence systems to their realization in performance. While other assumptions about this relation are imaginable (see e.g., Fodor et al., 1974), our reading of the experimental literature on implicature suggests that the Competence Hypothesis has been implicitly assumed throughout (though see section 2 in Part II). We will also assume this here. Call this assumption T_2.

There are many other choice points in the specification of a processing model. For example, does the processor compute serially or in parallel? What information does the system have access to at each step of computation? What is a 'step' in the computation? Which steps are costly? When is one problem that the system has to solve 'harder' than another?

We will try to clarify these questions, and possible answers to them, when we discuss some experimental results in Section 3. Here, to illustrate some of the issues they raise, consider how we might deal with the fact that each theory needs to say something about how the system

[11] Similar distinctions are also made in other cognitive sciences, such as in the theory of vision, where a distinction is sometimes made between what Marr (1982) calls the 'computational level,' responsible for characterizing the problems the system is designed to solve, and the 'algorithmic level,' responsible for characterizing the processes that are executed in solving the problems at the computational level.

decides whether or not to execute f (i.e., whether or not to strengthen the sentence). Recall that while both systems face this problem, the problem imposes different demands on the competing implicature mechanisms. Under the Gricean view the system has to decide whether or not to carry out the sequence of pragmatic reasoning that leads to implicatures. What are the factors that govern this? Some authors have proposed that the reasoning is triggered by default, with contextual factors sometimes overriding this default (e.g., Levinson, 2000). This assumption would predict that interpreting (1) with its strengthened meaning (4) would be faster than interpreting it with its literal meaning (see section 2.1 in Part II). It goes without saying that other choices are imaginable, and one would hope that any such choice would be motivated by general considerations. So far as we are aware, the factors that govern this decision have not been clarified.

Under the grammatical view, the question of whether to apply f reduces to a question of ambiguity resolution. The immediate question that arises is whether the disambiguation method employed is the same method used to resolve other ambiguities, such as lexical ambiguities, PP-attachment ambiguities, or scope-ambiguities. There are several proposals here. Some authors have followed Dalrymple et al. (1998) in assuming that there is a preference for *stronger meanings* (e.g., Chierchia, 2004; Chierchia et al., 2008 – see e.g., Geurts and Pouscoulous, 2009; Chemla and Spector, 2011 for critical discussion). Others have followed Gualmini et al. (2008) in assuming that the decision is based in part on considerations of how well the parse answers the question under discussion (e.g., Fox, 2007; Singh et al., 2013). Magri (2009, 2011) argues that there is no ambiguity at all, with sentences mandatorily parsed with an O (all apparent optionality is relegated to the contextual considerations that create ALT^* from ALT). And it remains an open question how, if at all, these factors interact with other well-known factors in disambiguation, such as considerations of plausibility (e.g., Crain and Steedman, 1985; Trueswell et al., 1994; Jurafsky, 1996; Hale, 2006), structural complexity (e.g., Miller and Chomsky, 1963; Frazier and Fodor, 1978; Gibson, 2000), tradeoffs between computation and storage (e.g., Johnson et al., 2007; O'Donnell et al., 2011), and so on.

A complete theory of scalar implicature requires answering these questions about processing and their relation to knowledge systems. Such a theory will thus be given by a conjunction $T_1 \wedge T_2 \wedge T_3 \wedge \dots T_k$, with each T_i specifying some set of assumptions about the system. When we evaluate any such theory against data, the evaluation will necessarily be holistic (we are evaluating the entire conjunction of assumptions). If the theory comes across data that it is unable to handle, all that we are licensed to conclude is that one of the conjuncts must be given up (though of course the conclusion has little force in absence of an alternative conjunction of assumptions that better explains the data). It should be clear that negating any particular conjunct, or even worse negating each conjunct, is not a valid inference.

3 Interim conclusion: the theoretical situation and its challenges

In the first part of this paper, we have introduced two of the main approaches to the mental computation of scalar implicatures: the Gricean approach and the grammatical approach. We have argued that although they rely on conceptually different views about the phenomenon, they share various insights. We have shown in particular that both of these approaches rely on some grammatical properties and on some pragmatic, domain general properties. Importantly, we have tried to delineate the empirical challenges for each of these approaches and to determine which ways can and cannot serve to tease them apart. We hope this first part will help reduce some current misunderstandings which often make their way to experimental studies. In the second part of this review paper, we will sample some experimental results and discuss their relevance for this task.

References

Barner, D. and A. Bachrach (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology 60*, 40–62.

Barner, D., N. Brooks, and A. Bale (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition 118*, 87–96.

Bott, L., T. M. Bailey, and D. Grodner (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language 66*, 123–142.

Bott, L. and I. Noveck (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language 53*, 437–457.

Braine, M. D. and B. Rumain (1981). Development of comprehension of 'or:' Evidence for a sequence of competencies. *Journal of Experimental Child Psychology 31*, 46–70.

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

Bresnan, J. and R. M. Kaplan (1982). Introduction. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, pp. xvii–lii. Cambridge, MA: MIT Press.

Chemla, E. (2008). An epistemic step for anti-presuppositions. *Journal of Semantics 25*, 141–173.

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

Chemla, E. (2009b). Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics 2*, 1–33.

Chemla, E. and L. Bott (2011). Processing inferences at the semantics/pragmatics frontier: disjunctions and *free choice*. Ms. LSCP & Cardiff University.

Chemla, E., C. Cummins, and R. Singh (2013). Reading Hurford disjunctions: Processing evidence for embedded implicatures. Poster presented at XPRAG 2013.

Chemla, E. and B. Spector (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics 28*, 359–400.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax-pragmatics interface. In A. Belleti (Ed.), *Structures and Beyond*. Oxford: Oxford University Press.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the 'logicality' of language. *Linguistic Inquiry 37*(4), 535–590.

Chierchia, G., S. Crain, M. T. Guasti, A. Gualmini, and L. Meroni (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *Proceedings of BUCLD25*, pp. 157–168.

Chierchia, G., D. Fox, and B. Spector (2008). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, and K. . Heusinger (Eds.), *Handbook of Semantics*. New York, NY: Mouton de Gruyter.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Clifton, C. J. and C. Dube (2010). Embedded implicatures observed: A comment on Geurts and Pouscoulous. *Semantics and Pragmatics 3*(7), 1–13.

Cohen, L. J. (1971). Some remarks on Grice's views about the logical particles of natural language. In Y. Bar-Hillel (Ed.), *Pragmatics of Natural Languages*, pp. 50–68. Dordrecht: D. Reidel.

Crain, S. (2008). The interpretation of disjunction in Universal Grammar. *Language and Speech 51*, 151–169.

Crain, S. and D. Khlentzos (2010). The logic instinct. *Mind & Language 25*(1), 30–65.

Crain, S. and M. Steedman (1985). On not being led up the garden path: the use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, and A. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, pp. 320–358. Cambridge: Cambridge University Press.

Cremers, A. and E. Chemla (2013). Direct and indirect scalar implicatures share the same processing signature. Manuscript, LSCP-ENS, Paris.

Dalrymple, M., M. Kanazawa, Y. Kim, S. Mchombo, and S. Peters (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy 21*, 159–210.

De Morgan, A. (1847). *Formal Logic*. London: Taylor and Walton.

De Neys, W. and W. Schaeken (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental psychology 54*(2), 128–133.

Degen, J. and M. K. Tanenhaus (2011). Making inferences: The case of scalar implicature processing. In L. Carlson, C. Holscher, and T. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, pp. 3299–3304.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. A., T. G. Bever, and M. F. Garrett (1974). *The Psychology of Language*. New York, NY: MgGraw-Hill.

Foppolo, F., M. T. Guasti, and G. Chierchia (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development 8*, 365–394.

Fox, D. (2007). Free choice disjunction and the theory of scalar implicature. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. New York, NY: Palgrave Macmillan.

Fox, D. (2013). Cancelling the Maxim of Quntity: Another challenge for a Gricean theory of scalar implicature. *Semantics and Pragmatics, in press*.

Fox, D. and M. Hackl (2006). The universal density of measurement. *Linguistics and Philosophy 29*, 537–586.

Fox, D. and R. Katzir (2011). On the characterization of alternatives. *Natural Language Semantics 19*, 87–107.

Fox, D. and B. Spector (2008). Economy and embedded exhaustification. Handout, MIT and Harvard University/ENS.

Frazier, L. and J. D. Fodor (1978). The Sausage Machine: A new two-stage parsing model. *Cognition 6*, 291–325.

Gajewski, J. and Y. Sharvit (2012). In defense of the grammatical approach to scalar implicature. *Natural Language Semantics 20*(1), 31–57.

Gamut, L. T. F. (1991). *Logic, Language, and Meaning*, Volume 1. Chicago, IL: University of Chicago Press.

Gazdar, G. (1979). *Pragmatics*. New York, NY: Academic Press.

Geurts, B. and N. Pouscoulous (2009). Embedded implicatures?!? *Semantics and Pragmatics 2*, 1–34.

Gibson, E. (2000). The dependency locality theory: A distance-based measure of linguistic complexity. In Y. Miyashita, A. Marantz, and W. O'Neil (Eds.), *Image, Language, Brain*, pp. 95–126. Cambridge, MA: MIT Press.

Goodman, N. D. and A. Stuhlmuller (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science 5*, 173–184.

Grice, H. (1967). Logic and conversation. William James Lectures, Harvard University.

Groenendijk, J. and M. Stokhof (1984). Studies on the semantics of questions and the pragmatics of answers. Doctoral Dissertation, University of Amsterdam.

Gualmini, A., S. Crain, L. Meroni, G. Chierchia, and M. T. Guasti (2001). At the semantics/pragmatics interface in child language. In *Proceedings of SALT 11*, pp. 231–247.

Gualmini, A., S. Hulsey, V. Hacquard, and D. Fox (2008). The question-answer requirement for scope assignment. *Natural Language Semantics 16*, 205–237.

Guasti, M. T., G. Chierchia, S. Crain, F. Foppolo, A. Gualmini, and L. Meroni (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes 20*(5), 667–696.

Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science 30*, 643–672.

Hamilton, S. W. (1860). *Lectures on Logic, Volume 1*. Edinburgh: Blackwood.

Hirschberg, J. (1985). A theory of scalar implicature. Doctoral Dissertation, University of Pennsylvania.

Horn, L. R. (1972). On the semantic properties of logical operators in english. Doctoral Dissertation, UCLA.

Horn, L. R. (2004). Implicature. In L. R. Horn and G. Ward (Eds.), *Handbook of Pragmatics*, pp. 3–28. Oxford: Blackwell.

Horn, L. R. (2005). The border wars. In K. von Heusinger and K. Turner (Eds.), *Where Semantics Meets Pragmatics*, pp. 21–48. Amsterdam: Elsevier.

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

Hurford, J. (1974). Exclusive or inclusive disjunction. *Foundations of Language 11*, 409–411.

Ippolito, M. (2011). A note on embedded implicatures and counterfactual presuppositions. *Journal of Semantics 28*, 267–278.

Johnson, M., T. L. Griffiths, and S. Goldwater (2007). Adaptor grammars: a framework for specifying compositional nonparametric bayesian models. In B. Sch"olkopf, J. Platt, and T. Hoffman (Eds.), *Advances in Neural Information Processing Systems 19*, Cambridge, MA, pp. 641–648. MIT Press.

Jurafsky, D. (1996). A probabilistic model of lexical access and disambiguation. *Cognitive science 20*, 137–194.

Katsos, N. and C. Cummins (2010). Pragmatics: From theory to experiment and back again.

Language and Linguistics Compass 4, 282–295.

Katzir, R. (2007). Structurally defined alternatives. *Linguistics and Philosophy 30*, 669–690.

Krifka, M. (1999). At least some determiners aren't determiners. In K. Turner (Ed.), *The semantics/pragmatics interface from different points of view*, pp. 257–291.

Levinson, S. C. (2000). *Presumptive Meanings*. Cambridge, MA: MIT Press.

Lewis, D. (1988). Relevant implication. *Theoria 54*, 161–174.

Magri, G. (2009). A theory of individual level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics 17*, 245–297.

Magri, G. (2011). Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics 4*, 1–51.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman and Company.

Marty, P. and E. Chemla (2011). Scalar implicatures: working memory and a comparison with 'only'. Ms. LSCP.

Mill, J. S. (1867). *An Examination of Sir William Hamilton's Philosophy* (3rd ed.). London: Longman.

Miller, G. and N. Chomsky (1963). Finitary models of language users. In R. Luce, R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Volume 2, pp. 419–491. New York, NY: Wiley.

Noveck, I. (2001). When children are more logical than adults: Investigations of scalar implicature. *Cognition 78*, 165–188.

Noveck, I. and A. Reboul (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences 12*, 425–431.

Noveck, I. and D. Sperber (2007). The why and how of experimental pragmatics: The case of scalar inferences burton-roberts noel (ed.), advances in pragmatics.

O'Donnell, T. J., J. Snedeker, J. B. Tenenbaum, and N. D. Goodman (2011). Productivity and reuse in language. In N. Taatgen and H. van Rijn (Eds.), *Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society*. The Cognitive Science Society.

Panizza, D., G. Chierchia, Y. T. Huang, and J. Snedeker (2009). The relevance of polarity items for the online interpretation of scalar terms. In *Proceedings of SALT 19*.

Papafragou, A. and J. Musolino (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition 80*, 253–282.

Paris, S. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology 16*, 278–291.

Rips, L. (1975). Quantification and semantic memory. *Cognitive Psychology 7*(3), 307–340.

Romoli, J. and F. Schwarz (2013). An experimental comparison between presuppositions and indirect scalar implicatures. Presented at *Sinn und Bedeutung 18*, Vitoria, Spain.

van Rooij, R. and K. Schulz (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language, and Information 13*, 491–519.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics 1*, 75–116.

Russell, B. (2006). Against grammatical computation of scalar implicatures. *Journal of Semantics 23*, 361–382.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy 27*, 367–391.

Schlenker, P. (2012). *Maximize Presupposition* and Gricean reasoning. To appear in *Natural Language Semantics*.

Schulz, K. and R. van Rooij (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy 29*, 205–250.

Singh, R. (2008a). Modularity and locality in interpretation. Doctoral Dissertation, MIT.

Singh, R. (2008b). On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy 31*, 245–260.

Singh, R., K. Wexler, A. Astle, D. Kamawar, and D. Fox (2013). Children interpret disjunctions as conjunctions: Cconsequences for the theory of implicature. Manuscript, Carleton University and MIT and HUJI.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology 30*, 195–205.

Spector, B. (2005). Scalar implicatures: Exhaustivity and Gricean reasoning. In M. Aloni, A. Butler, and P. Dekker (Eds.), *Questions in Dynamic Semantics*. Amsterdam: Elsevier.

Spector, B. (2006). Aspects de la pragmatique des opérateurs logiques. Doctoral Dissertation, Université Paris 7.

Sperber, D. and D. Wilson (1986). *Relevance: Communication and Cognition*. Blackwell.

Stiller, A., N. D. Goodman, and M. C. Frank (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of the Cognitive Science Society*.

Tomlinson, J., L. Bott, and T. Bailey (2011). Understanding literal meanings before pragmatic inference: Mouse-trajectories of scalar implicatures. In *4th Biennial Conference of Experimental Pragmatics, June*, pp. 2–4.

Tomlinson, J. M. J., T. M. Bailey, and L. Bott (2012). Possibly all of that and then some: Scalar implicatures are understood in two steps. Ms. Cardiff University.

Trueswell, J. C., M. K. Tanenhaus, and S. M. Garnsey (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language 33*, 285–318.

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua 117*(11), 1858–1878.

Whately, R. (1848). *Elements of Logic* (9th ed.). London: John W. Parker.

Zhou, P., J. Romoli, and S. Crain (2013). Children's knowledge of free choice inferences. In *Proceedings of SALT 23*, pp. 1–15.

Remarks on the experimental turn in the study of scalar implicature, Part II

Emmanuel Chemla and Raj Singh

December 4, 2013

Interim abstract

In Part I, we have introduced two of the main approaches to scalar implicature: the Gricean approach and the grammatical approach. We have argued that although they rely on conceptually different views about the phenomenon, they share various insights and we argued that their empirical differences were more subtle than what one may have expected. In this second part of this review paper, we will sample some experimental results with two goals in mind. First, we will exemplify some simplifications that are found in the literature and examine the consequences of these simplifications on the interpretability of experimental results. Second, we will suggest future directions that the experimental turn might consider exploring, directions that seem to us to have the potential to illuminate the richness of the competing theories and the potential to dissociate or improve them.

1 Interpreting Experimental Results

In the second part of this paper, we will review some results from experimental studies of scalar implicature (henceforth SI), with the first part serving as the theoretical matrix. We do not offer a comprehensive review of the literature. Rather, we will discuss a handful of results with the following two goals in mind: (i) exemplifying some simplifications that are found in the literature and examining the consequences of these simplifications on the interpretability of experimental results, and (ii) suggesting future directions that the experimental turn might consider exploring, directions that seem to us to have the potential to illuminate the richness of the competing theories and the potential to dissociate or improve them. We will do this in two steps. First, we will argue that the current theoretical terrain is richer than is often assumed (section 2), and in particular we will extract a generalization that relates stages of acquisition to processing complexity in the adult: SIs that are acquired late are processed slower in the steady state (Reinhart, 2006; Pouscoulous et al., 2007; Noveck and Reboul, 2008; see Table 1 in Section 3.3 for a summary). Second, we will suggest some new directions that experimental studies might take in helping adjudicate the debate between the neo-Gricean and grammatical theories (section 3).

2 The challenge from processing complexity and acquisition

2.1 Processing complexity: The automaticity/defaultness debate

A well-known finding is that scalar implicatures are derived with a delay: it takes more time to read or evaluate the truth-value of a sentence like (1) if its scalar implicatures are taken into account. This result has been obtained with various methodologies, such as the truth-value judgment task (Bott and Noveck, 2004; Degen and Tanenhaus, 2011), self-paced reading (Breheny et al., 2006; Chemla et al., 2013), and eye-tracking studies (Huang and Snedeker, 2009a; Storto and Tanenhaus, 2005). This result has been obtained with various subspecies of scalar implicatures as well, including e.g., so-called 'indirect implicatures,' (such as the inference from a sentence like *not all of Mary's students got an A* to *some of Mary's students got an A*, Cremers and Chemla, 2013).[1]

(Same introducing examples as in Part I, repeated here for convenience)

(1) Some of Mary's students got an A.

(2) *Alternative:* All of Mary's students got an A.

(3) *Scalar implicature:* Not all of Mary's students got an A.

(4) *Strengthened Meaning:* Some but not all of Mary's students got an A.

There have been two provisos to the result that SI computation is costly. First, there is a worry that the time difference could be due to the fact that scalar implicatures normally add complexity to the meaning of the sentence (by way of providing a strictly stronger meaning or of removing simple monotonicity properties), independent of whether or not the *derivation* of this scalar implicature is costly (see Bott et al., 2012 for discussion). Second, some may understand eye-tracking results showing that scalar implicatures are derived quickly, before a sentence is fully processed, as evidence that their derivation is not costly. However, such results merely show that even though it may have a cost, the parser is willing to pay that price early on. We would like to stay away from these two questions, and discuss instead the relevance of this cost result for deciding between theories.

Assuming the Competence Hypothesis (Bresnan and Kaplan, 1982), the null hypothesis under both the Gricean approach and the grammatical one is that SI computation should be costly. This is visible already from figures 1 and 2 from Part I, in both of which [[S]] occurs 'above' any implicature-related reading. The execution of the strengthening function f will involve retrieval of alternatives $ALT(S)$ together with their contextual restriction $ALT^*(S)$, and either a sequence of pragmatic reasoning (Gricean theory) or addition of syntactic structure to S to create $O(ALT^*(S))(S)$ and corresponding semantic computations demanded by the O operator (grammatical theory). This work presumably takes time and effort.[2] Why, then, has

[1] Cremers and Chemla (2013) and Romoli and Schwarz (2013) present evidence that indirect SIs are not costly, but Cremers and Chemla (2013) argue that there are confounds in these findings, and they show that once these confounds are overcome the computation of indirect SIs does come with a processing delay.

[2] One difference between the two theories is that even if the meaning of S needs to be derived before scalar implicatures, this derivation is a mere intermediate computational step in the grammatical view, while it should behave like a plain, potential meaning for the whole sentence in the neo-Gricean view. Tomlinson et al. (2011; 2012) present

the cost result been given so much attention in the literature?

The reason is that writers from both the neo-Gricean camp (Levinson, 2000) and from the grammatical perspective (Chierchia, 2004) made the ambitious and surprising claim that SIs are derived by 'default.' A natural interpretation of this defaultness claim is that deriving an SI is the best bet, and hence will be done 'automatically' (despite the costs associated with doing so), incurring the risk that this might have to be undone at a later stage (e.g., if contextual considerations about relevance or the competence assumption suggest that the inference is not warranted). The default assumption predicts, then, that a subject should be faster in understanding a sentence with its SIs than without. This prediction is not supported by experimental results, and authors have used this as an argument against the neo-Gricean view of Levinson and the grammatical view of Chierchia (see e.g., Noveck and Reboul, 2008; Katsos and Cummins, 2010 and references therein).

This conclusion is not warranted, since it is possible that the difficulty lies in the auxiliary assumptions about processing that implement defaultness within a neo-Gricean or grammatical framework.[3] The defaultness assumption was a provocative and sophisticated attempt to suggest that certain preferences in interpretation are worth the computational costs that go with them and thus impose greater costs for reverting to the dispreferred interpretation. The default assumption was clearly and precisely stated, and perhaps for this reason it has proven to be extremely fruitful in pushing our understanding of the time course of implicature computation. Experimental results are inconsistent with the defaultness assumption, but this negative result leaves many questions unanswered about the source of the costs. This is where we think attention should be focused. As already noted, the costs could be linked to the computation of the alternatives ALT, the comparison of alternatives (cf. `Stronger Alternatives`), the negation of alternatives (execution of f), the need to access contextual information in deciding which of the alternatives are relevant (ALT^*) and which of them the speaker is opinionated about (the competence assumption), and the relative complexity of the meanings obtained with and without scalar implicatures (see recent experimental work by De Neys and Schaeken, 2007; Grodner et al., 2010; Degen and Tanenhaus, 2011; Bott et al., 2012; Tomlinson et al., 2011, 2012; Marty and Chemla, 2011). Furthermore, not all alternatives ALT are derived in

mouse-tracking results showing that hearers who eventually respond taking scalar implicatures into account go through a stage at which scalar implicatures are absent and behave as if S were the reading assigned to the sentence. These types of finer-grained results seem to favor a neo-Gricean model in which S systematically appears as a potential meaning for the sentence. However, other results show that intermediate computational stages behave very much like potential readings in that respect. These other results might be taken as evidence for online 'bottom-up' compositional interpretation. For instance, in interpreting negative sentences of the form *not* S, hearers also go through a stage in which they behave as if S were the reading assigned to the sentence *not* S (Dale and Duran, 2011).

[3] There are many ways of formulating a processing model with a defaultness assumption. One possibility is to give up the Competence Hypothesis. For example, a learner might pick up on the fact that (1) often has the strengthened meaning (4), and simply store this and use it as a heuristic instead of running through the computations laid out in Figure 1 or Figure 2 from Part I each time they encounter an occurrence of (1). An alternative that is consistent with the Competence Hypothesis might require – at a first pass at interpretation – that certain options be taken at the relevant choice-points in the derivation, subject to later revision. For example, under both the grammatical and Gricean theory a subject would have to delay the contextual pruning of ALT. In addition, a subject working through Figure Part 1-1 would have to execute all the steps in the reasoning (which are defeasible in principle), and they would assume by default that the competence assumption holds. And a subject working through Figure Part 1-2 would have to parse the sentence with an O. This implementation of the default principle would force the parser to make certain choices with the risk that these choices may need to be revised if contextual considerations decide the result is inappropriate.

the same way: some alternatives are derived by replacing lexical items with other lexical items (e.g., *A and B* from *A or B*), while others are derived by deleting parts of the tree (e.g., *A* from *A or B*) (e.g., Sauerland, 2004; Katzir, 2007; Chemla, 2009a; see note 7), and it is conceivable that these differences matter to processing.

2.2 Processing profiles of a wider range of scalar implicatures

The problem of localizing the source(s) of costs during SI computation is further complicated by the observation that the computation of some other inferences which are often analyzed as scalar implicatures does not exhibit the same cost.[4] Specifically, there is evidence that computing the 'exactly' interpretation of numerals is not costly (e.g., Huang and Snedeker, 2009a; Marty et al., 2013) and there is evidence that computing free-choice inferences is not costly (Chemla and Bott, 2011). These inferences are often assumed to be SIs. Let us illustrate these phenomena: a sentence like *John ate three cookies* is often assumed to have a literal 'at least three' reading (that John ate at least three cookies) which can get strengthened through SI computation to an 'exactly three' reading (see e.g., Spector, 2013 for a recent overview), and a sentence like *John is allowed to eat the cake or the ice-cream* is often assumed to have a literal meaning that John has permission to eat at least one of the cake or ice-cream, which reading can get strengthened by SI computation to the free-choice inference that John is allowed to eat the cake and that he is allowed to eat the ice-cream (e.g., Kratzer and Shimoyama, 2002; Schulz, 2005; Alanso-Ovalle, 2005; Fox, 2007; Chemla, 2009a).[5]

The challenge is to explain all of these results in a general theory of semantics, pragmatics, and processing. There are many reasons for thinking that the exactly-readings of numerals and free-choice inferences should be computed as SIs in the same way as what we will call 'standard SIs' – those involving Horn scales consisting of duals, such as such as $\langle$*some*, *all*$\rangle$ and $\langle$*or*, *and*$\rangle$.[6] For example, they tend to exhibit many of the same properties as standard SIs: they can be described as the output of f on a classical meaning, the classical meanings are preferred in downward-entailing environments (e.g., *no one is allowed to eat the cake or the ice-cream* means no one is permitted to eat either, not that no one has free-choice), and they can be cancelled (e.g., it is consistent to say *John has three sons, and possibly even more*). It is consequently puzzling that the processing signatures of otherwise similar inferences should diverge in this way. At the same time, if numerals and free-choice are to be explained with means other than f, then not only are their similarities to standard SIs left as something of a puzzle, they also give rise to complications in the semantics (cf. note 5). The literature on this is too vast for us to do justice to here; what we would like to contribute is the generalization about processing complexity, and to suggest that with a clearer understanding of which linguistic inferences lead to costs, and why, only then will we be able to use processing complexity to assess and refine current theories of SI computation (for relevant recent discussion, see e.g.,

[4]van Tiel et al. (2013) explores a related question, showing that the rate of derivation of scalar inferences varies in important ways from one scalar item to another. Although these results are not yet associated with processing results, they are certainly relevant to the enterprise of understanding the core of the phenomenon by looking at a broader class of exemplars.

[5]Under common assumptions, the logical form *allowed (A or B)* does not entail *allowed A and allowed B*, though cf. Zimmerman (2000); Geurts (2005).

[6]The alternatives for a sentence like *John ate 3 cookies* consists of sentences of the form 'John ate n cookies' for all natural numbers 'n,' and the alternatives of *John is allowed to eat cake or ice-cream* that are needed to derive free-choice are *John is allowed to eat cake* and *John is allowed to eat ice-cream* (see note 7).

Geurts, 2006; Breheny, 2008; Chemla, 2009a; Huang and Snedeker, 2009a; Panizza et al., 2009; Chemla and Bott, 2011; Singh et al., 2013; Spector, 2013; Cremers and Chemla, 2013; Marty et al., 2013).

2.3 Child development

Acquisition studies have demonstrated that, descriptively, children (roughly between the ages of three and seven) who master much of their native language grammar behave as if they do not derive standard scalar implicatures like the one described in (1)-(4) (e.g., Noveck, 2001; Papafragou and Musolino, 2003; Reinhart, 2006; Crain, 2008; Barner and Bachrach, 2010; Crain and Khlentzos, 2010). More precisely, it has been shown that the observed rate of implicature-compliant behavior can be manipulated by changing the 'pragmatics of the task' (e.g., Chierchia et al., 2001; Gualmini et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005; Miller et al., 2005; Verbuk, 2009; Barner et al., 2011; Stiller et al., 2011; Foppolo et al., 2012). This could be taken as evidence that there is modularity in the system: SIs are acquired at a stage when, say, a pragmatic module develops, while a semantic module could develop earlier.

We do not think the evidence warrants such a strong interpretation. First, children have been shown to have quite sophisticated pragmatic abilities even at the stage at which they do not compute SIs (e.g., Clark, 1993; Clark and Amaral, 2010). Second, as discussed earlier, the computation of SIs involves several components (e.g., the generation of alternatives, negations of them, reasoning about the speaker's epistemic state, etc), and it is possible that these components mature at different stages of development. For example, it has been argued that the reason children behave as if they do not compute scalar implicatures is that they are at a stage of development at which they do not generate alternatives, but otherwise the rest of the system might well be intact (e.g., Chierchia et al., 2001; Gualmini et al., 2001; Barner and Bachrach, 2010; Barner et al., 2011).[7] In a recent argument in support of this, Barner et al. (2011) provided evidence that children judge sentences like *only some of the animals are sleeping* as true in contexts in which all of the animals in a given picture are sleeping. It seems that even the presence of *only*, which like SI computation takes a set of alternatives as input and negates some of them (replace the O in (Part I:11) with *only*), does not compel the child to perform the required substitution. However, if the relevant alternatives are made explicit for the child, they reveal their underlying knowledge of the strengthening function f. For example, Barner et al., 2011 showed that children reject sentences like *only the dog and the cat are sleeping* when they are explicitly given the information that the dog, the cat, and the cow are sleeping (see similar results in Chierchia et al., 2001; Gualmini et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005; Barner and Bachrach, 2010; Minai and Fiorentino, 2010).

As with our discussion of processing complexity, our summary of the acquisition of SIs is complicated by results concerning numerals and free-choice. Specifically, there is evidence that children pass through a stage of development at which they do not compute standard SIs but nevertheless compute 'exactly' interpretations of numerals (e.g., Papafragou and Musolino,

[7]More concretely, assume with Katzir (2007) that alternatives in the steady state are derived by a function ALT which takes the parse of the uttered sentence as input and produces a set of alternatives by replacing nodes in the tree with elements from the following substitution sources: (i) the lexicon (e.g., replacing *some* in (1) with *all* to create (2)), (ii) subconstituents of the target node (e.g., replacing *A or B* with its constituent *A*), (iii) explicitly mentioned linguistic material (e.g., if *the dog and the cat* has been mentioned, then that phrase can be replace *the dog* in *the dog is sleeping*). Singh et al. (2013) localize the child's difficulties with ALT with operation (i): children do not access the lexicon but can access subconstituents and linguistically salient material.

2003; Musolino, 2004; Huang and Snedeker, 2009b) and free-choice inferences (Zhou et al., 2013; Singh et al., 2013).[8] As noted earlier, these early-developing inferences are also processed faster in the adult than are standard SIs. These results raise the challenge of explaining (i) the stages the child passes through as it develops into the steady state, (ii) when and why online processing costs are observed, and (iii) why implicatures available early to the child are processed fast in the steady state.

Here we would like to highlight another question that remains unanswered in studies of the acquisition of SI computation. Let us assume that children first acquire the core 'at least some' meaning of *some* (and similarly for other scalar items) and later acquire the relevant competence to derive scalar implicatures. How do they come to this state of affairs, given that many (conceivably most) occurrences of *some* in the target adult community give rise to a 'some but not all' interpretation? Several types of answer may be entertained: there could be innate constraints on possible lexical meanings, either domain-specific (e.g., Katzir and Singh, 2013) or domain-general (e.g., Piantadosi et al., 2012), or it may be that a few counter-examples suffice to rule out a 'some but not all' entry for *some* (e.g., a few occurrences of *some* in the antecedent of a conditional, cf. (7-a) in Part I). This puzzle offers a line of inquiry that could yield insights into the division of labour between domain-specific and domain-general constraints on the acquisition of the various components involved in the computation of SIs.

2.4 Summary

As noted, studies of the adult reveal many similarities between standard SIs, free-choice, and exact-readings of numerals. Yet, processing results reveal that scalar inferences are slow to process and develop while free-choice and numerals are not. Moreover, the child – despite these inferences – learns the classical meanings for *or, some, allow*, etc. So far as we are aware, there is no unifying theory that predicts and explains these observations.

3 New directions for dissociating theories

Section 2 summarized experimental results and embedded them in the current theoretical context, raising new challenges that any adequate theory of SI computation will have to address. In this section we point to new avenues that might be worth pursuing with the aim of dissociating the neo-Gricean and grammatical theories of SI, and point out some ways in which some experimental paradigms should be avoided in light of the goal of dissociating theories.

3.1 SIs and ignorance inferences

As discussed in section 2 from Part I, both the Gricean approach and the grammatical approach predict that in addition to the *not-all* scalar implicature in (3) associated with *some* (= (1)) it is also possible to derive so-called 'ignorance inferences' about the *all* alternative in (2), which is an inference that the speaker is ignorant about the truth-value of (2).[9] The distinction between

[8]More accurately, Singh et al. (2013) provide evidence that children interpret disjunctive sentences $A \vee B$ as conjunctions (replicating earlier results in Paris (1973) and Braine and Rumain (1981)). Given the observation that children understand *or* as inclusive disjunction in downward-entailing environments (e.g., Chierchia et al., 2001; Gualmini et al., 2001; Crain, 2008; Crain and Khlentzos, 2010), Singh et al. (2013) explain the conjunctive interpretation with the assumption that the child strengthens $A \vee B$ to $A \wedge B$ as an SI. This follows as a consequence of the assumption about child alternatives in note 7 together with theories of free-choice (Fox, 2007; Chemla, 2009a; Franke, 2011).

[9]There are additional inferences about the speaker's epistemic state, such as 'primary implicatures' and 'secondary implicatures.' These are relevant to our broader point, but we focus here on the distinction between SIs and ignorance

these inferences has received little attention in the experimental literature. Current processing results thus remain underspecified: it is unclear whether the costs of SI computation that have been discovered in the experimental literature (cf. section 2.1) are associated with plain SIs or with ignorance inferences (see Breheny et al., 2013 for preliminary results that take the speaker's epistemic state into account). However, information about the processing properties of *both* SIs and ignorance inferences may distinguish the two main families of approach.

This distinction is evidenced even at the the schematic level of figures 1 and 2 from Part 1, but let us make this more concrete with an ideal case study: disjunctions. Sentences of the form 'John talked to Mary or Sue' are associated with the set of alternatives {*John talked to Mary*, *John talked to Sue*, *John talked to Mary and Sue*} (e.g., Sauerland, 2004; Fox, 2007; Katzir, 2007; Chemla, 2009a – see note 7), and they give rise to (a) the inference that John did not talk to both Mary and Sue, as a plain SI, and (b) ignorance inferences about whether John talked to Mary and whether John talked to Sue. Under the Gricean view (see figure Part 1-1), ignorance inferences about each disjunct follow from the conjunction of the literal reading and primary implicatures of the form $\neg B_s(A)$ for each alternative A.[10] In other words, the final leaf of the right branch is irrelevant, and so ignorance inferences are predicted to arise before plain SIs (which follow from secondary SIs). Under the grammatical view (see figure Part 1-2), ignorance inferences always come after plain SIs, which follow from the selected literal meaning for a sentence. Hence, establishing a processing priority between these two types of inferences would provide theoretically important information. Of course, one would have to control for irrelevant factors such as the role of the content of the inferences, which makes this a non-trivial task. Nevertheless, establishing the time course of these inferences could provide valuable information that would help move the discussion beyond the mere existence of a cost, which is unsurprising (though cf. section 2.4), to fine-grained questions about the sources of cost, stages of computation, etc.

3.2 Back to Global Reasoning and Stronger Alternatives

One of the virtues of the original view of scalar implicatures is that it predicts that SIs should arise only in certain contexts, namely contexts that satisfy (Part I:5) `Stronger Alternatives` and (Part I:6) `Global Reasoning`. The grammatical theory and Gricean theory sometimes agree on predictions when both of these assumptions hold. Examining processing costs under the assumption that these principles both hold might give useful data about complexity that cuts across controversial issues. For such investigations, it is reasonable to ensure that the experimental materials indeed satisfy these constraints. Other times, to adjudicate between theories, one needs materials that lead to different expectations depending on whether (Part I:5) and/or (Part I:6) hold. Unfortunately, experimental paradigms come with their own practical constraints, and these practical constraints sometimes bias against using environments that would satisfy (Part I:5) and/or (Part I:6).

As a case in point consider visual world eye-tracking studies, which routinely use imperatives: the eye-gaze of participants is recorded while they listen to *instructions*. In SI studies of this sort (e.g., Huang and Snedeker, 2009a), at some point in the instructions participants will encounter a scalar item such as *some*, and the crucial question is whether anything can

inferences for expository purposes.

[10] This is a special and somewhat accidental property of disjunctions, which follows from the logical fact that $B_s(A \vee B) \wedge \neg B_s(A) \wedge \neg B_s(B)$ entails $\neg B_s(\neg A)$ and $\neg B_s(\neg B)$.

be inferred from participants' behavior after this point. In interpreting such results, it is important to note that the relevant sentence (e.g., *pick the girl who has some of the socks*) and its alternative (e.g., *pick the girl who has all of the socks*) are not informatively ordered, since these sentences do not denote propositions with truth-conditional content. Thus they escape from the scope of a theory that embraces (Part I:5) `Stronger Alternatives`, raising the challenge of reconciling this result with any theory committed to (Part I:5). (The absence of an entailment relation between alternatives also complicates comparison with results from other visual world eye-tracking studies that do not embed scalar terms under imperatives (e.g., Panizza et al., 2009; Degen and Tanenhaus, 2011).)

In fact, studies using simpler methods such as the truth-value judgment task have also failed to ensure that `Stronger Alternatives` was satisfied. For example, the critical items in Bott and Noveck (2004)'s pioneering study included sentences like *some elephants are mammals*. This sentence, given the common knowledge that all elephants are mammals, is no less informative to the hearer than its alternative *all elephants are mammals*. The Gricean intuition about informativity is certainly context-sensitive: the maxim of quantity demands that the speaker give their hearer as much *new* information as possible (*ceteris paribus*). Thus, so far as informativity considerations go, *some elephants are mammals* should be no worse than *all elephants are mammals*. What does this mean, then, for the finding that there is nevertheless a cost to comprehending this sentence when it is strengthened to *some but not all elephants are mammals*? And how, given that `Stronger Alternatives` does not hold, does such strengthening take place anyway?[11]

The difficulty with ensuring that the relevant conditions are satisfied might stem from a tempting simplification of what theories of SIs describe. SIs are not *some but not all* interpretations of *some*. In many cases, a *some but not all* interpretation is predicted to be unavailable, for example because `Stronger Alternatives` is violated, as in the cases discussed above. As a result, the interpretation of any distinctive behavior that would be observed in environments that violate (Part I:5), be it a processing cost or an absence thereof, would have to discuss the consequences for theories that embrace this assumption.

Similar traps exist in experimental investigations of the behavior of scalar items in embedded syntactic environments. The conceptual goal behind this enterprise is to evaluate whether apparent violations of (Part I:6) `Global Reasoning` warrant a grammatical account of SIs. For example, the sentence in (5) has been argued to allow not only the strengthened meaning in (5-a), but also the strengthened meaning in (5-b).

(5) Every girl passed some of her exams
  a. Possible strengthened meaning: Every girl x is such that x passed some of her exams, but it is false that every girl x passed all of x's exams
  b. Possible strengthened meaning: Every girl x is such that x passed some but not all of x's exams

Although the strengthened meaning in (5-a) can be derived under both the Gricean and grammatical theories by negating the alternative *every girl passed all of her exams*, the possibility of the reading in (5-b) on the face of it seems to be allowed only by the grammatical the-

[11]For related discussion, see e.g., Heim (1991); Percus (2006); Fox and Hackl (2006); Chierchia et al. (2008); Magri (2009); Singh (2009); Magri (2011); Schlenker (2012).

ory: the logical form 'every girl $x, O(x$ passed some of x's exams)' entails (5-b). As we noted earlier, allowing f to operate in embedded positions means giving up `Global Reasoning`. By identifying the strengthening function f with a covert operator O, the grammatical theory accommodates the availability of (5-b) without difficulty. However, theories that remain committed to `Global Reasoning` can capture the reading in (5-b) by enriching the set of alternatives of (5) to include *some girl passed all of her exams* (e.g., Chemla, 2009a). This move, of course, requires giving up `Stronger Alternatives` (see also Fox, 2007; Magri, 2009; Sauerland, 2012 for discussion). Thus, while sentences such as (5) have received significant attention in the experimental literature (e.g., Geurts and Pouscoulous, 2009; Clifton and Dube, 2010; Ippolito, 2010; Sauerland, 2010; van Tiel, 2013), these cases may not be informative with respect to questions of theory comparison.

We submit that embedding scalar items under non-monotonic operators, such as in (Part I:10-a), is a better test of the possibility of embedded implicatures because there is no plausible set of alternatives which, together with `Global Reasoning`, allows for the strengthened meaning in (Part I:10c) to be generated (see Chemla and Spector, 2011). Another case of interest concerns the Hurford disjunctions discussed earlier. For example, consider (Part I:9), which we repeat here as (6):

(6) Some of Mary's students got an A or all of them did.

Recall that the data from Hurford's Constraint suggest that the first disjunct *must* be interpreted with an SI (Chierchia et al., 2008), which is only possible under theories that give up (Part I:6) `Global Reasoning`. Under the further assumption that strengthening (1) to (4) is costly, the prediction from the grammatical approach is that there should be a cost to processing the first disjunct if an SI is computed. Some tentative processing evidence in favour of this prediction – and hence for the grammatical approach – is presented in Chemla et al. (2013).

3.3 Toward a theory of preference

Let us return to Figures Part I:1 and Part I:2. Traversing through either figure involves making several choice points. Does the competence assumption hold in the given context? What is relevant in the given context? Should f be applied? Under the grammatical theory, this last question raises a disambiguation problem at each scope site: should the sentence, embedded or not, be parsed with f or without? This permissiveness of the theory needs to be tempered by restrictions on or preferences over readings, while the Gricean theory needs an account of why embedded strengthenings are available at all. Ultimately, we should be able to predict how a sentence is interpreted given full knowledge of the conversational situation and possibly of the resources of the participants. How are these decisions made? Given the many factors that might enter into these decisions, formulating a decision theory is likely to be a rather complex enterprise. In section 2.2 from Part I we mentioned some proposals that have been advanced about factors that could play a role in ranking alternative decisions (e.g., preferences for stronger meanings, or for better answers to the question under discussion, or for simpler structures, or for plausible interpretations, among others).

We submit that in addition to *a priori* motivations for positing certain preferences in communication, a general theory of preference could also benefit from the processing results we mentioned in earlier sections. We summarized relevant results in Table 1. Recall first that 'standard' SIs involving scalar items like *some* are slower to process in the adult state and are acquired later than other SIs such as 'exactly-'interpretations of numerals and free-choice. It turns out that investigations of how accessible these strengthenings are in embedded positions cuts along similar lines. Specifically, while the accessibility of embedded strengthenings of *some* seems to vary with task-demands (e.g., Geurts and Pouscoulous, 2009; Clifton and Dube, 2010; Ippolito, 2010; Sauerland, 2010; Chemla and Spector, 2011; van Tiel and Geurts, 2013), embedded free-choice strengthenings are quite readily observed (e.g., Chemla, 2009b; Singh et al., 2013), as are embedded 'exactly'- readings of numerals, even in downward-entailing environments (e.g., Panizza et al., 2009). Any theory of preference over readings will need to account for this generalization.

What we have, then, is a generalization about acquisition, processing complexity, and embeddability: SIs come in at least two kinds, those that are processed fast, learned fast, and easily embedded, and those that are learned slow, processed slow, and are not easily embedded.

Kinds of SI	**Embeddability**	**Processing results**	**Acquisition**
Standard cases	Weak but possible	slow processing (for direct and indirect SIs)	Rather late
Numerals	Stronger	faster Weaker memory demands	Earlier
Free choice inferences	Stronger	faster (no info about memory)	Earlier

Table 1: Summary of the results about various kinds of scalar implicatures (results for Numerals and Free choice inferences are stated relative to the results for the standard case).

4 Concluding Remarks

In the first part of this paper (Part I) we have described the current theoretical situation in the study of scalar implicatures, and in the second part (Part II) we have discussed some experimental results in light of this background.

The study of scalar implicature over the last decade has benefitted from both theoretical and empirical innovations. The merits of undertaking such a joint approach are uncontroversial. The goal of linguistics is to propose a model of all relevant data, hence it cannot be restricted to offline introspective judgments, nor can it ignore them. The advance of experimental methods offering new processing, acquisition, and imagery data is thus welcome most generally and should be supported. The application of these methods to scalar implicatures is certain to have a representative and long lasting echo. However, perhaps because the relations between theory and experiment in the domain of scalar implicature are in their infancy the benefits of these innovations have not yet realized their full potential. We tried to clarify the assumptions of the major competing theories, and clarify how their differences are or are not reflected in current experimental results. We hope that future experiments will place greater emphasis on using the methods of experimental psychology to tease apart – and possibly refine – the predictions of competing theories. At the same time, we hope that linguists will clarify and test the predictions their theories make concerning performance. We tried to identify issues that are raised by the current state of knowledge and that we think stand to benefit from joint experimental and theoretical scrutiny.

References

Alanso-Ovalle, L. (2005). Distributing the disjuncts over the modal space. In L. Bateman and C. Ussery (Eds.), *Proceedings of the North East Linguistic Society (NELS) 35*. GLSA.

Barner, D. and A. Bachrach (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology 60*, 40–62.

Barner, D., N. Brooks, and A. Bale (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition 118*, 87–96.

Bott, L., T. M. Bailey, and D. Grodner (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language 66*, 123–142.

Bott, L. and I. Noveck (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language 53*, 437–457.

Braine, M. D. and B. Rumain (1981). Development of comprehension of 'or:' Evidence for a sequence of competencies. *Journal of Experimental Child Psychology 31*, 46–70.

Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics 25*, 93–139.

Breheny, R., H. J. Ferguson, and N. Katsos (2013). Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition 126*(3), 423–440.

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

Bresnan, J. and R. M. Kaplan (1982). Introduction. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, pp. xvii–lii. Cambridge, MA: MIT Press.

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

Chemla, E. (2009b). Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics 2*, 1–33.

Chemla, E. and L. Bott (2011). Processing inferences at the semantics/pragmatics frontier: disjunctions and *free choice*. Ms. LSCP & Cardiff University.

Chemla, E., C. Cummins, and R. Singh (2013). Reading Hurford disjunctions: Processing evidence for embedded implicatures. Poster presented at XPRAG 2013.

Chemla, E. and B. Spector (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics 28*, 359–400.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax-pragmatics interface. In A. Belleti (Ed.), *Structures and Beyond*. Oxford: Oxford University Press.

Chierchia, G., S. Crain, M. T. Guasti, A. Gualmini, and L. Meroni (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *Proceedings of BUCLD25*, pp. 157–168.

Chierchia, G., D. Fox, and B. Spector (2008). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, and K. . Heusinger (Eds.), *Handbook of Semantics*. New York, NY: Mouton de Gruyter.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clark, E. V. and P. M. Amaral (2010). Children build on pragmatic information in language acquisition. *Language and Linguistics Compass 4*, 445–457.

Clifton, C. J. and C. Dube (2010). Embedded implicatures observed: A comment on Geurts and Pouscoulous. *Semantics and Pragmatics 3*(7), 1–13.

Crain, S. (2008). The interpretation of disjunction in Universal Grammar. *Language and Speech 51*, 151–169.

Crain, S. and D. Khlentzos (2010). The logic instinct. *Mind & Language 25*(1), 30–65.

Cremers, A. and E. Chemla (2013). Direct and indirect scalar implicatures share the same processing signature. Manuscript, LSCP-ENS, Paris.

Dale, R. and N. D. Duran (2011). The cognitive dynamics of negated sentence verification. *Cognitive science 35*(5), 983–996.

De Neys, W. and W. Schaeken (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental psychology 54*(2), 128–133.

Degen, J. and M. K. Tanenhaus (2011). Making inferences: The case of scalar implicature processing. In L. Carlson, C. Holscher, and T. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, pp. 3299–3304.

Foppolo, F., M. T. Guasti, and G. Chierchia (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development 8*, 365–394.

Fox, D. (2007). Free choice disjunction and the theory of scalar implicature. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. New York, NY: Palgrave Macmillan.

Fox, D. and M. Hackl (2006). The universal density of measurement. *Linguistics and Philosophy 29*, 537–586.

Franke, M. (2011, June). Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics 4*(1), 1–82.

Geurts, B. (2005). Entertaining alternatives: Disjunctions as modals. *Natural Language Semantics 13*, 383–410.

Geurts, B. (2006). Take 'five': the meaning and use of a number word. In *Non-definiteness and plurality*, pp. 311–329. Amsterdam/Philadelphia: John Benjamins.

Geurts, B. and N. Pouscoulous (2009). Embedded implicatures?!? *Semantics and Pragmatics 2*, 1–34.

Grodner, D. J., N. M. Klein, and K. M. C. ad Michael K. Tanenhaus (2010). 'Some,' and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition 116*, 42–55.

Gualmini, A., S. Crain, L. Meroni, G. Chierchia, and M. T. Guasti (2001). At the semantics/pragmatics interface in child language. In *Proceedings of SALT 11*, pp. 231–247.

Guasti, M. T., G. Chierchia, S. Crain, F. Foppolo, A. Gualmini, and L. Meroni (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes 20*(5), 667–696.

Heim, I. (1991). Artikel und definitheit. In *Semantics: An International Handbook of Contemporary Research*. Berlin: de Gruyter.

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

Huang, Y. T. and J. Snedeker (2009b). Semantic and pragmatic interpretation in 5-year olds: Evidence from real-time spoken language comprehension. *Developmental Psychology 45*, 1723–1739.

Ippolito, M. (2010). Embedded implicatures? Remarks on the debate between globalist and localist theories. *Semantics and Pragmatics 3*, 1–15.

Katsos, N. and C. Cummins (2010). Pragmatics: From theory to experiment and back again. *Language and Linguistics Compass 4*, 282–295.

Katzir, R. (2007). Structurally defined alternatives. *Linguistics and Philosophy 30*, 669–690.

Katzir, R. and R. Singh (2013). Constraints on the lexicalization of logical operators. *Linguistics and Philosophy 36*(1), 1–29.

Kratzer, A. and J. Shimoyama (2002). Indeterminate pronouns: The view from japanese. In Y. Otsu (Ed.), *Proceedings of the 3rd Tokyo Conference on Psycholinguistics*, pp. 1–25.

Levinson, S. C. (2000). *Presumptive Meanings*. Cambridge, MA: MIT Press.

Magri, G. (2009). A theory of individual level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics 17*, 245–297.

Magri, G. (2011). Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics 4*, 1–51.

Marty, P. and E. Chemla (2011). Scalar implicatures: working memory and a comparison with 'only'. Ms. LSCP.

Marty, P., E. Chemla, and B. Spector (2013). Interpreting numerals and scalar items under memory load. *Lingua 133*, 152–163.

Miller, K., C. Schmitt, H.-H. Chang, and A. Munn (2005). Young children understand *some* implicatures. In A. Brugos, M. R. Clark-Cotton, and S. Ha (Eds.), *Proceedings of the Twenty-Ninth Annual Coston University Conference in Language Development*, pp. 389–400.

Minai, U. and R. Fiorentino (2010). The role of the focus operator *Only* in children's computation of sentence meaning. *Language Acquisition 17*(3), 183–190.

Musolino, J. (2004). The semantics and acquisition of number words: integrating linguistic and developmental perspectives. *Cognition 93*, 1–41.

Noveck, I. (2001). When children are more logical than adults: Investigations of scalar implicature. *Cognition 78*, 165–188.

Noveck, I. and A. Reboul (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences 12*, 425–431.

Panizza, D., G. Chierchia, Y. T. Huang, and J. Snedeker (2009). The relevance of polarity items for the online interpretation of scalar terms. In *Proceedings of SALT 19*.

Panizza, D., G. Chierchia, and C. C. 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.

Papafragou, A. and J. Musolino (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition 80*, 253–282.

Paris, S. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology 16*, 278–291.

Percus, O. (2006). Antipresuppositions. In A. Uyema (Ed.), *Theoretical and Empirical Studies of Reference and Anaphora: Toward the Establishment of Generative Grammar as an Empiricial Science*, pp. 52–73. Report of the Grant-in-Aid for Scientific Research (B), Project No. 15320052, Japan Society for the Promotion of Science.

Piantadosi, S. T., J. B. Tenenbaum, and N. D. Goodman (2012). Modeling the acquisition of quantifier semantics: a case study in function word learnability. Manuscript, University of Rochester, MIT, and Stanford University.

Pouscoulous, N., I. A. Noveck, G. Politzer, and A. Bastide (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition 14*(4), 347–375.

Reinhart, T. (2006). *Interface Strategies*. Cambridge, MA: MIT Press.

Romoli, J. and F. Schwarz (2013). An experimental comparison between presuppositions and indirect scalar implicatures. Presented at *Sinn und Bedeutung 18*, Vitoria, Spain.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy 27*, 367–391.

Sauerland, U. (2010). Embedded implicatures and experimental constraints: A reply to Geurts & Pouscoulous and Chemla. *Semantics and Pragmatics 3*(2), 1–13.

Sauerland, U. (2012). The computation of scalar implicatures: Pragmatic, lexical or grammatical? *Language and Linguistics Compass*, 36–49.

Schlenker, P. (2012). *Maximize Presupposition* and Gricean reasoning. To appear in *Natural Language Semantics*.

Schulz, K. (2005). A pragmatic solution for the paradox of free choice permission. *Synthese 147*(2), 343–377.

Singh, R. (2009). *Maximize Presupposition!* and informationally encapsulated implicatures. In A. Riester and T. Solstad (Eds.), *Proceedings of* Sinn und Bedeutung 13, pp. 513–526.

Singh, R., K. Wexler, A. Astle, D. Kamawar, and D. Fox (2013). Children interpret disjunctions as conjunctions: Cconsequences for the theory of implicature. Manuscript, Carleton University and MIT and HUJI.

Spector, B. (2013). Bare numerals and scalar implicatures. *Language and Linguistics Compass 7*(5), 273–294.

Stiller, A., N. D. Goodman, and M. C. Frank (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of the Cognitive Science Society*.

Storto, G. and M. Tanenhaus (2005). Are scalar implicatures computed online? In C. B. Emar Maier and J. Huitink (Eds.), *Proceedings of Sinn und Bedeutung 9*, pp. 431–445. Nijmegen: Nijmegen Centre for Semantics.

Tomlinson, J., L. Bott, and T. Bailey (2011). Understanding literal meanings before pragmatic inference: Mouse-trajectories of scalar implicatures. In *4th Biennial Conference of Experimental Pragmatics, June*, pp. 2–4.

Tomlinson, J. M. J., T. M. Bailey, and L. Bott (2012). Possibly all of that and then some: Scalar implicatures are understood in two steps. Ms. Cardiff University.

van Tiel, B. (2013). Embedded scalars and typicality. *Journal of Semantics*.

van Tiel, B. and B. Geurts (2013). Embedded scalars. *Semantics and Pragmatics, in press*.

van Tiel, B., E. van Miltenburg, N. Zevakhina, and B. Geurts (2013). Scalar diversity. Manuscript, University of Nijmegen.

Verbuk, A. (2009). The role of the lexical component in the acquisition of scalar implicatures. In J. Craford, K. Otaki, and M. Takahashi (Eds.), *Proceedings of the Third Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)*, pp. 303–312. Cascadilla Proceedings Project.

Zhou, P., J. Romoli, and S. Crain (2013). Children's knowledge of free choice inferences. In *Proceedings of SALT 23*, pp. 1–15.

Zimmerman, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics 8*, 255–290.