IMPLICATURE PROCESSING

Yasutada Sudo

Name: Yasutada Sudo
Affiliation: University College London
Mailing address: Department of Linguistics, UCL, 2 Wakefield St., London WC1N 1PF, UK
email: y.sudo@ucl.ac.uk

*I would like to thank Daniele Panizza for discussion and for sharing his teaching material, which was very helpful in preparing this paper. My thanks also go to two anonymous reviewers for very helpful suggestions for improving the article, as well as to Hajime Ono for the editorial work. Needless to say, all remaining errors are my own.

Abstract: This article gives a brief overview of the current literature on how a class of inferences called *implicatures* are processed, mostly focusing on so-called *scalar implicatures* triggered by *some* and *or*. This topic is of relevance not only for psycholinguists interested in sentence processing and its relations with other cognitive functions such as reasoning and memory, but also for theoretical linguists, as experimental results in this area have relatively direct implications for semantic and pragmatic theories. It is notable that a number of different experimental methods and paradigms have been deployed to investigate processing of implicatures, and the literature is currently rapidly growing, making this topic one of the most actively studied areas in experimental linguistics today.

Keywords: *implicature, scalar implicature, pragmatics, processing, psycholinguistics*

1. Introduction

Implicature is a type of inference that cannot be attributed to the conventional meaning (or semantics) of linguistic expressions, but instead is considered to arise due to reasoning about how language users (should) communicate.[1] Herbert Paul Grice, who pioneered the study of implicatures in the 1960s, offers the following example that clearly illustrates what an implicature is:

> *A is writing a testimonial about a pupil who is a candidate for a philosophy job, and his letter reads as follows: "Dear Sir, Mr. X's command of English is excellent, and his attendance at tutorials has been regular. Yours, etc".* (Grice, 1975, p. 33)

One can infer from this letter that according to A, X is not excellent as a philosopher. Evidently, this inference cannot be attributed to the semantics of the words and phrases used in this letter. Rather, it is reasonable to think that it arises because the letter does *not* mention anything about X's expertise as a philosopher, which it should have mentioned if A thought Mr X was a good candidate for the job. We will come back to the details of this reasoning below in light of Grice's influential ideas.

Given examples like this, that implicatures exist is undeniable and uncontroversial, and the literature is mostly concerned with the following questions:

- What is the mechanism that is used to generate implicatures?
- Are all implicatures generated by the same mechanism?
- How and when are implicatures processed during comprehension?

This article will focus primarily on the third question and review the results of experimental research on it, but the other two questions cannot be ignored. For this reason, we will first briefly review major views and issues from the theoretical literature on implicatures in

[1] This type of implicatures is more specifically called *conversational implicatures.* Conversational implicatures contrast with *conventional implicatures*, which are considered to be part of the semantic properties of expressions. Since we do not deal with conventional implicatures in this article, we will call conversational implicatures, simply implicatures. See Grice (1989), Potts (2005), and Gutzmann (2015), among others, for more theoretical discussion about conventional implicatures and related ideas.

Section 2, and then in Section 3, we will discuss the experimental literature on processing of implicatures, especially focusing on so-called *scalar implicatures* arising from *some* and *or*. Section 4 contains concluding remarks and further references.

2. Theories of Implicatures

2.1 Gricean Pragmatics

As mentioned above, it is Grice who initiated the study of implicatures (Grice, 1989; see Davis, 2019 for an overview). His ground-breaking idea is that one can draw certain inferences from utterances based on the assumption that one's interlocutors are rational and cooperative, and that being a rational cooperative interlocutor can be understood as knowing and obeying the normative rules of conversation, or the *Conversational Maxims*. Grice postulated the following four Conversational Maxims.[2]

- Maxim of Quantity: Don't say too much or too little.
- Maxim of Quality: Only say what you have enough evidence for.
- Maxim of Manner: Be clear.
- Maxim of Relation: Be relevant.

According to Davis (2019), Grice's motivation behind these maxims comes from general considerations about cooperation among rational agents that go beyond linguistic behavior. The idea is that these maxims are natural rules for rational agents that try to achieve some common goal, e.g. sharing some information among them.

With these Conversational Maxims, we can understand the above example of an implicature arising from a recommendation letter as follows. Assuming that the author is a rational cooperative communicator, we know that if they thought the candidate was a good philosopher, they would have said so in their recommendation letter, given the Maxims of Quantity and Relation. If, on the other hand, they didn't think the candidate was suitable for the job, they might not have said so directly, given the conventions about recommendation letter writing that are shared by the parties involved. Now, upon observing that the letter contains no mention of X's expertise as a philosopher, the reader of the letter reasons that writing so would have violated some maxim or maxims. As mentioned above, it would have satisfied the Maxims of Quantity and Relation, and also, the Maxim of Manner could have been respected easily, so these cannot be the reason. Then it must be because it would have violated the Maxim of Quality, viz. the author doesn't have enough evidence for asserting that X is a good philosopher. One can reason further that as a recommender for X, the author should have enough evidence to judge whether X is a good or bad philosopher, so the reader can conclude that the author thinks that X is *not* a good philosopher, which is the inference we wanted to derive.

2.2 Scalar Implicatures

Later authors, most notably, Relevance Theorists like Sperber & Wilson (1986) and so-called neo-Griceans like Gazdar (1979), Horn (1989; 2004), and Levinson (2000) have proposed to extend the empirical landscape of implicatures, by giving broadly Gricean analyses to further types of pragmatic inferences, and also to refine and improve on Grice's original set of maxims. The details of their theoretical proposals do not concern us

---

[2] The content of the maxims is often stated differently, usually with more details, but since these details are not very important for our purposes here, I will stick to simple formulations.

very much here, but it is of importance for us that they attempt to account for so-called *generalised implicatures* (as opposed to *particularised* or ad hoc *implicatures*). Generalised implicatures are those implicatures that seem to not rely much on contextual factors. The above example with a recommendation letter is not a case of a generalised implicature, as the context plays a crucial role in the derivation of the inference, as we have seen. In other words, when put in other contexts, the same expressions may not suggest that X is not a good philosopher.

Prime examples of generalised implicatures are so-called *scalar implicatures,* which are implicatures that arise from items that are 'on a scale' in some sense, and are generally about higher ranked items on the same scale being false*.* The quantifier *some* is a good example of this, with the scale being <some, all> (or perhaps <some, most, all>). Concretely, one normally understands *Some of them are Americans* as meaning *some but not all* of them are Americans, but the "not all" part of the meaning is considered to be not coming from the lexical semantics of *some* itself*.* There are several reasons to think so. One is that this inference is actually cancellable, at least in certain contexts. For example, *Some of them are Americans, in fact all of them are,* is deemed consistent. Also, the very fact that one can say *some but not all* without a feeling of redundancy also suggests that the semantics of *some* itself does not mean "not all". Similarly, the consistency of expressions like *some and possibly all* points to the same conclusion. In addition, when *some* is put in linguistic contexts called Downward Entailing (DE) contexts, the "not all" inference may disappear altogether. For example, a conditional antecedent is a DE context, and a sentence like *If you read some of these papers, you will pass for sure* does not mean the same thing as *If you read some but not all of these papers, you will pass for sure.* Rather, this example does not seem to be associated with a "not all" inference at all. Furthermore, in a number of experiments that assessed speakers' judgments of statements like *Some of them are Americans*, it is found over and over again that the interpretation without the "not all" scalar implicature is often, if not always, accessed (e.g., Bott & Noveck, 2004; Van Tiel, Van Miltenburg, Zevakhina & Geurts, 2016). We will discuss related experimental findings later in this article.

Under the Gricean approach, the "not all" inference of *some* can be understood as an implicature in the following manner. We assume that the core semantics of *some* is merely lower-bounded, and is similar to "some and possibly all" or "at least some". Upon hearing *Some of them are Americans,* the hearer wonders why the speaker didn't say *All of them are Americans* instead. This hypothetical utterance would have conveyed more information, because it asymmetrically entails what is actually uttered and hence is more informative. Therefore, by not uttering it, the rational and cooperative speaker is at the risk of violating the Maxim of Quantity. Given that this alternative statement is also relevant, the hearer reasons further that it must be because the speaker couldn't say it without violating the Maxim of Quality, which means that the speaker does not have enough evidence for this statement. Assuming that the speaker is knowledgeable about the nationalities of the relevant people, the hearer may furthermore conclude that according to the speaker, this statement is false, which amounts to the "not all" inference.

A nice feature of this analysis is that it accounts for why *some* does not mean "not all", when put in a DE context like a conditional antecedent. The reasoning goes as follows. When one hears an utterance of *If you read some of these papers, you will pass for sure,* one should wonder about an alternative utterance of *If you read all of these papers, you will pass for sure,* but this sentence is already entailed to be true by what was uttered, and a likely reason why the cooperative speaker didn't use it instead is because, although relevant and as succinct, it would actually have conveyed less information and thereby violated the Maxim of Quantity. As a result, one does not draw a "not all" implicature in this case.

Another example of an expression that is often associated with a scalar implicature that is worth mentioning here is *or,* which is considered to form a scale with *and.* For example, a sentence like *Alice speaks Burmese or Cantonese* has a number of inferences. For one, it implies that Alice dose not speak both of these languages. It also implies that the speaker does not know whether or not Alice speaks Burmese or whether or not she speaks Cantonese. These inferences are also considered to be generalised implicatures, rather than arising from the lexical semantics of *or,* in light of evidence parallel to the case of *some* mentioned above*.* For instance, let us put *or* in a DE context, e.g. in the scope of negation, as in *Alice does not speak Burmese or Cantonese.*[3] The meaning of this sentence is not the negation of the positive sentence above augmented with all these extra inferences. Rather, it entails that Alice speaks neither of the languages. This is taken to suggest that the lexical semantics of *or* is inclusive, rather than exclusive, disjunction, which is compatible with both disjuncts being simultaneously true, and therefore the additional inferences we saw above are all implicatures.

Under the Gricean account, we can derive these implicatures as follows. When one hears A*lice speaks Burmese or Cantonese*, one reasons about an alternative utterance of *Alice speaks Burmese and Cantonese*. This hypothetical utterance would have been more informative than what was actually uttered, because the former asymmetrically entails the latter on the assumption that the lexical meaning of *or* is inclusive disjunction. Assuming that the speaker is rational and cooperative, the hearer can then conclude that the speaker didn't say so, because she would have violated the Maxim of Quality with this alternative utterance, meaning that they don't have enough evidence for the truth of *Alice speaks Burmese and Cantonese*. Parallel reasoning about shorter but more informative alternative utterances, *Alice speaks Burmese* and *Alice speaks Cantonese*, also results in the inference that the speaker does not have enough evidence for the truths of these either. Taken together, the speaker must have meant that Alice speaks one of these two languages but she has no evidence for the truths of more specific statements.[4]

There is a number of other cases that are considered to be part of the same phenomenon. For instance, a bare numeral in an argument position is often assumed to have an interpretation that is not upper-bounded and its upper-bounded interpretation is a scalar implicature (see Spector, 2013 for an overview). Concretely, Daniel *has two laptops* is often taken to imply that he does not have three laptops and this is considered to be a scalar implicature. In addition, theories of scalar implicature often account for the exhaustive interpretation of an answer to a *wh*-question by the same mechanism. That is, it is clear that *Ed came* does not semantically entail that Fred didn't come, but when uttered as an answer to the question *Who came to the party?,* it is often used to imply so. The empirical landscape of (scalar) implicatures has been considerably enriched since Grice's work, and the literature is especially rapidly growing in recent years, drawing novel data not only from familiar languages like English but also from languages that have traditionally been less often mentioned in relation to pragmatic inferences.

---

[3] In languages like Japanese, disjunction exhibits properties of Positive Polarity Items and resist narrow scope readings relative to a negation and other negative items in the same clause. For such a language, a different, perhaps syntactically more complex, DE context, for example a conditional antecedent, needs to be used to make the same point.

[4] Note that the inferences we drew here does not imply that the speaker thinks that Alice doesn't speaks both of the languages. To derive this stronger inference, we have to resort to an extra assumption, similar to the case of *some,* that the speaker either believes that Alice speaks both of the languages or that she speaks neither of them. It remains to be explained, however, why the hearer doesn't go through analogous reasoning about one of the shorter sentences, e.g. *Alice speaks Burmese,* and conclude that Alice doesn't speak Burmese. See Sauerland (2004), Spector (2006), and Fox (2007) for specific proposals about this.

2.3 Theoretical Debates about Scalar Implicatures

We have just gone through the derivations of scalar implicatures arising from *some* and *or* under the Gricean theory of implicatures, but it is debated in the literature how exactly this is done. As mentioned above, some neo-Griceans, especially Levinson (2000), claim that generalised implicatures, including scalar implicatures, are computed by default and do not trigger context dependent reasoning like we saw. They are, however, still implicatures, and are cancellable, if one notices that they conflict with contextual knowledge, for example.

More recently, an approach to scalar implicatures that stands on an entirely different conception of what they are has been advocated by Chierchia (2004), Fox (2007) and Chierchia, Fox & Spector (2011), among others. According to this approach, scalar implicatures are not really implicatures in the sense that they actually arise as part of the compositional semantics of the sentence, rather than via pragmatic reasoning. This idea could be formalised in terms of a phonologically empty operator (which is often written as *Exh* or *O*) that introduces relevant inferences (Fox, 2007; Chierchia, Fox & Spector, 2011), or in terms of a special compositional rule in the semantics (Chierchia 2004).

One of the primary motivations for this *grammatical approach* to scalar implicatures comes from so-called embedded scalar implicatures. The observation is that scalar implicatures seem to be able to take scope under certain expressions, at least in some cases. For instance, in a sentence like *If some of us lie, we are screwed, but if all of us lie, we will be fine, some* needs to be understood as meaning *some but not all* within the conditional antecedent, in order for the entire sentence to be coherent.

However, whether or not such observations really require the grammatical view is particularly theoretically contentious in the literature. Since our main interest here is processing of implicatures, we will not delve into this debate (see, e.g., Sauerland, 2004; Russell, 2006; see also Geurts, 2010 for defense of the pragmatic view). Yet, it should be mentioned here that this is one of the areas where theoretical and experimental research has been advancing hand in hand. Major experimental studies include the following journal articles: Geurts & Pouscoulous (2009), Clifton & Dube (2010), Chemla & Spector (2011), Van Tiel & Geurts (2013), Cummins (2014), Van Tiel (2014), Potts, Lassiter, Levy & Frank (2016), Franke, Schlotterbeck & Augurzky (2017), Van Tiel, Noveck & Kissine (2018). At this point, there is still an ongoing controversy regarding the interpretations of these experimental results, but it is certainly the case that more and more data is being collected with a wide range of quite sophisticated experimental methods, which will certainly help us eventually settle the theoretical debate.

2.4 Towards Predictions about Processing

Different theories of scalar implicature make different predictions about the time course of their generation, although the predictions are based upon specific assumptions about processing.

For instance, according to the Gricean view (and other related ones like Relevance Theory), the computation of a scalar implicature involves reasoning on the hearer's part about the speaker's epistemic state in a particular context of utterance, as well as various contextual factors (on a par with non-generalised cases of implicatures like the recommendation letter example mentioned at the outset) and therefore is expected to involve more processing and hence should take additional processing time, relative to when the relevant scalar item receives an interpretation without the scalar implicature

By contrast, according to a theory like Levinson's (2000) where scalar (and other generalised) implicatures are computed by default and cancelled only when they clash with contextual knowledge, it is expected that interpretation with a scalar implicature should be processed faster than interpretation without a scalar implicature, as the latter involves an additional process of cancelling the scalar implicature.

The grammatical view of implicatures does not bear directly on processing on its own, but certain auxiliary assumptions may lead to testable predictions. Under this view, each sentence containing a scalar item like *some* has multiple semantic interpretations, usually one with a scalar implicature and one without.[5] One can assume that when a sentence with a scalar item like *some* is processed, the parser faces an issue of ambiguity resolution. There are several possible assumptions we can make here. If one adopts the operator-based view of scalar implicatures where scalar implicatures are generated by a phonologically silent operator, it might sound reasonable to assume that the parser prefers the interpretation that is generated by the structurally simpler option, namely the one without the operator. Or, one could assume that the parser first tries the stronger interpretation. Then the interpretation with the scalar implicature should be processed faster, unless the scalar item is in a DE context. One could alternatively assume that these two factors are both operative at the same time, in which case the predictions will be more nuanced. See Chemla & Singh (2014a; 2014b) for more detailed discussion on this point.

In sum, there are a number of different theories of how scalar implicatures are generated, and behavioral data of online processing will be useful in reducing the space of such possible theories. In the next section we will review major experimental results, especially focusing on the scalar implicatures associated with *some* and *or,* which are the most widely studied scalar expressions in the current experimental literature (see Section 4 for references for experimental research on other items).

3. Processing Studies on Implicature

Partly thanks to relatively direct implications that processing data can have on theories of scalar implicatures, processing of scalar implicatures has been particularly vigorously investigated using many experimental methods. Among scalar items, *some* (and its analogues in other languages) is by far the most frequently studied one. Some of the experiments we will review below also have tested disjunction *or.* In addition to these two representative scalar items, many experiments also tested numerals, often along with *some,* but there is a separate theoretical question as to whether numerals actually trigger scalar implicatures at all (see Spector, 2013 for an overview), and experimental results show more or less consistently that numerals are processed at least partially differently from other scalar items like *some* (see the references in the survey articles mentioned at the end of this article).

These different methods have so far yielded converging results, suggesting that the interpretation without a scalar implicature is accessed first, and the interpretation with scalar implicatures requires some extra processing time. However, as we will discuss in detail below, the results of visual world eye-tracking experiments are interestingly equivocal, although more recent studies have shed some light on what factors might be behind the conflicting results reported in earlier eye-tracking studies.

---

[5] The grammatical theory of scalar implicature allows for multiple interpretations with a scalar implicature for complex sentences, corresponding to different 'scopes' of the scalar implicature. This is often taken to be an empirical advantage of the theory, although not completely uncontroversially, as mentioned at the end of the previous section.See Chierchia, Fox & Spector (2011), Fox & Spector (2018) for extensive discussion of such cases.

In the rest of this section, we will first review experimental evidence for delayed processing of scalar implicatures, and then summarise representative visual world eye-tracking experiments, and open issues that remain.

3.1 Evidence for Delayed Processing of Scalar Implicatures

One of the most influential studies on processing of implicatures is Bott & Noveck (2004). Bott & Noveck investigated how scalar items are processed by looking at Reaction Times (RTs) of speakers evaluating statements like *Some elephants are mammals.*[6] The idea is that since these statements are false with the "not all" scalar implicature, while they are true without it, we can see which of these interpretations is computed faster by measuring how fast speakers make their response.

Their Experiment 1 is divided into two blocks. In one of them the participants were instructed to treat *some* as meaning "some but not all", and in the other block, they were instructed to treat it as meaning "some and possibly all". With the first type of instructions, therefore, the participants should reject *Some elephants are mammals,* while in the second type of instructions, they should accept it. The results show that it takes on average about 500 ms longer to reject these statements. Furthermore, they observed a similar difference in RTs in their Experiment 3, in which participants were not given explicit instructions about how to understand *some.* In this experiment, statements like *Some elephants are mammals* were judged as true 40.7% of the time, and accepting it was on average about 600 ms faster than rejecting it.

These results have been replicated many times in later studies (e.g. Noveck & Posada, 2003; Bott, Bailey & Grodner, 2012; Chemla & Bott, 2014; Chemla & Marty, 2014). Furthermore, a picture verification task, instead of a sentence verification task, has also been employed to test the same thing, and yielded converging results (Van Tiel & Schaeken, 2017; but Degen & Tanenhaus, 2015 report more nuanced results).

These RT data can be taken to be suggesting that scalar implicature computation indeed does take time. However, several alternative possible explanations have also been suggested for the observed difference in RTs. For example, one might wonder if this difference is due merely to the difference between accepting vs. rejecting statements. Bott & Noveck (2004) themselves raise this question and point out that in their results, rejecting the target statements required longer RTs on average than rejecting the control statements that were straightforwardly false, while accepting the target sentences was as fast as accepting the true control statements. This suggests that there is an additional processing time associated with rejecting the target statements, as opposed to rejecting control statements, which could be seen as due to the scalar implicature computation.

Another conceivable explanation is that the interpretation with the scalar implicature involves a positive ("at least some") and a negative ("not all") component in the meaning, while the interpretation without it is simply positive, and this interpretative difference, rather than online generation of scalar implicatures itself, is what is behind the observed difference in RTs (Grodner, Klein, Carbary & Tanenhaus, 2010; Bott, Bailey & Grodner, 2012; Chemla & Marty, 2014). If this were the case, then, Bott & Noveck's results would not tell us much about processing of implicatures *per se*. In order to test this alternative explanation, Experiment 2 of Bott, Bailey & Grodner's (2012) compared *Some elephants are mammals* with *Only some elephants are mammals,* which has a similar

[6] All the experiments reported in Bott & Noveck (2004) were conducted in French, and the relevant scalar item was *certains* 'some', but later studies replicated their main findings in English, so we will stick to English examples below to facilitate the exposition. However, it is certainly an empirical question whether there is any crosslinguistic variation here, which I believe is still largely an uncharted domain of inquiry at this point.

interpretation as the former statement with its “not all” scalar implicature. Their results indicate that the statement with *only* is rejected faster than the statement without *only*, while the latter was processed as fast when it was accepted, suggesting that the delay observed with the rejection of *Some elephants are mammals* is not entirely due to the interpretation that is simultaneously positive and negative.

In addition to RTs, other experimental measures have also been used to discern processing properties of scalar implicatures. Breheny, Katsos & Williams (2006), for example, conducted a series of self-paced reading experiments. Their Experiment 1 is of particular interest for us here, where they recorded the reading time of disjunctive phrases like *the class notes or the summary* preceded by two different types of discourse, one supporting the exclusive disjunction reading and one supporting the inclusive disjunction reading. Recall that the “not both” part of the exclusive disjunction reading is considered to be a scalar implicature. The results of this experiment indicate that disjunctive phrases are read more slowly when the context puts bias towards the exclusive reading, which suggests, again, that a reading with a scalar implicature requires more processing time than a reading without. Similar results are reported by Bergen & Grodner (2012) for *some.*

Tomlinson, Bailey & Bott (2013) conducted ‘mouse-tracking’ experiments, whose results point to the same conclusion. Specifically, in each trial of their experiments, the participant read a sentence displayed in the center of the computer screen word by word, and reported whether they judge it to be true or false by navigating the cursor with a mouse, which started in the center bottom position of the screen, towards either a box that says True or a box that says False, which were located at the top-left and top-right corner of the screen. They tested sentences similar to Bott & Noveck’s (2004) and found that True answers, which required no scalar implicature, tended to have more or less straight trajectories, while False answers, which requires a “not all” scalar implicature, exhibited curved trajectories where the cursor initially moved towards True and then veered towards No. This indicates that the interpretation with the “not all” scalar implicature is accessed later than the interpretation without it, which is consistent with the findings from the aforementioned experimental studies.

In addition, a number of studies have made use of experimental paradigms that manipulate cognitive resources available to the participants. The logic behind this is that if scalar implicatures require extra computations in online processing, then limiting cognitive resources should make the necessary computations harder to execute, and hence scalar implicatures should be observed less often. Different ways of limiting cognitive resources have been employed but the experiments have generally yielded converging results, indicating that scalar implicature computation uses extra cognitive resources.

For instance, Experiment 4 of Bott & Noveck (2004) is one of the first studies to implement this idea. They manipulated the time that participants are allowed to spend before providing their answer in each trial between 900 ms (Short) and 3000 ms (Long). They used the same stimuli as their Experiment 1 described above, and observed that sentences like *Some elephants are mammals* were judged as true 72% of the time in the Short condition, while only 56% of the time in the Long condition, meaning that more scalar implicatures were observed when more time was available. On the other hand, the control conditions that do not involve scalar implicatures exhibited much smaller discrepancies between the Short and Long conditions and also the direction of the difference was generally the opposite. These results suggest that scalar implicature computation requires extra time, which is in harmony with the other results we have reviewed so far.

Another experimental task that builds on the same idea of manipulating cognitive resources is the so-called dual task, where the participant is first given an irrelevant task and while they are still engaging with it, they also do the target linguistic task. De Neys & Schaeken (2007) is one of the first studies to apply this technique to scalar implicatures, where the participants were asked at the beginning of each trial to remember the placement of three horizontally arranged dots (Low Load) or four randomly arranged dots (High Load) in a 3 x 3 grid, then to judge the truth or falsity of sentences similar to those used in Bott & Noveck (2004), and finally to reproduce the grid with dots they were asked to memorise at the beginning of the trial. In the Hight Load condition, the target sentences like *Some oaks are trees* were judged as true 26.8% of the time, as opposed to 21.1% in the Low Load condition. This means that the less working memory is available, the fewer scalar implicatures are computed, suggesting, again, that scalar implicature computation is cognitively effortful. Similar results are reported by Dieussaert, Verkerk, Gillard & Shaeken (2011) and Marty & Chemla (2013), which address some potential concerns of De Neys & Schaeken's (2007) experimental design. Marty, Chemla & Spector (2013) employed a slightly different version of the task, where in each trial, the participants were asked to memorise a sequence of letters, rather than a placement of dots, and to reproduce it in the reverse order after the linguistic task. Also, they employed a picture verification task, rather than a sentence verification task as their linguistic task. Despite these differences in the task, their results exhibited decreasing rates of scalar implicatures as the memory load increased, pointing to the same conclusions as the other studies mentioned so far.

3.2 Visual World Eye-Tracking Paradigm

All the experimental evidence we have reviewed so far suggests that scalar implicature computation takes time and requires more cognitive resources like working memory. Interestingly, however, experiments with the so-called visual world eye-tracking paradigm have yielded conflicting results. In a visual world eye-tracking experiment, the participant's gaze is tracked, while they are looking at some visual scene displayed on a computer screen and listening to an auditory linguistic stimulus. It has been found in some studies using this method that the "not all" scalar implicature of *some* requires extra processing time relative to a baseline without a scalar implicature, while other similar studies report that it was processed as rapidly as a sentence without a scalar implicature. Below, we will first review the key findings of these studies, and then discuss more recent studies that aim at explaining factors behind the divergent results.

Huang & Snedeker (2009) is one of the first eye-tracking studies that employed the visual world paradigm to investigate processing of scalar implicatures, and their results are in line with the idea that scalar implicature computation is costly. Specifically, they recorded participants' eye-movements as they listened to a sentence like *Point to the girl that has some of the socks* against a visual scene with four characters: a boy with two socks, a girl with two socks, a boy with nothing, and a girl with three soccer balls. The idea is that if *some* is understood as meaning "some but not all", they do not need to wait until the noun to know which girl this sentence is about, namely, the girl with two socks. In their Experiments 1 and 2, they compared sentences like this to sentences containing *all* instead, e.g. *Point to the girl that has all of the soccer balls,* for which it is clear that the quantifier is enough to disambiguate the referent. They also tested sentences with bare numerals like *Point to the girl that has two of the soccer balls,* which turned out to be crucial, as we will discuss later. The results of the experiments indicate that in the *all* condition, participants started to shift their gaze towards the intended referent within 400 ms from the onset of the quantifier *all,* meaning they make use of the meaning of the determiner to predict the noun to come, while a similar shift was observed much later in the case of *some.* If *some* had been read as meaning "some but not all" from the beginning, it should have resulted in an early shift as in the case of *all,* and therefore, these results are taken to be evidence that the "not all" implicature requires extra time.

Huang & Snedeker (2011) used pairs of nouns with longer ambiguous periods and replicated these results.

Other studies, however, pointed out several potential confounds of Huang & Snedeker's experimental design. In particular, Grodner, Klein, Carbary & Tanenhaus (2010) argue that the phrase *some of the NP* used by Huang & Snedeker independently might have delayed the disambiguation effect for the following reason. The quantifier *some* itself has a use that is not reliably associated with a scalar implicature, especially when it is not in a partitive construction. The contrast between a non-partitive and partitive *some* is intuitively robust, and and also consistently verified with quantitative data in studies like Degen (2015) and Degen & Tanenhaus (2011, 2015). Then, it could well be the case that the participants of Huang & Sendeker's experiments had to wait until they heard *of* before realizing that the "not all" scalar implicature should be computed. On the other hand, in the case of *all,* such an ambiguity is not present. This alone, therefore, could explain the relative delay of referent ambiguity resolution for *some.* Grodner et al. (2010) furthermore raised the possibility that the presence of numerals, *two* and *three,* in the same experiment might have affected the processing of *some.* In Huang & Snedeker's (2009) experiments, the participants heard numerals more often than *some,* and as indicated in the results, the numerals were interpreted as 'exactly two' and 'exactly three' very rapidly (which, incidentally, could be taken as evidence that the 'exactly'-readings of numerals might not be due to scalar implicatures; see Spector 2013 for an overview on different theoretical views about the semantics and pragmatics of numerals). Also, the visual scene always contained either two or three relevant items, which are well within the subitising range and hence were easy to count, which led Grodner et al. (2010) to argue that thanks to their precise 'exactly'-readings, numerals were generally better descriptions of the relevant objects than *some*, and for this reason *some* was perceived less natural and might have incurred more processing time just for that reason. Grodner et al. (2010) present the results of an acceptability judgment experiment that support this claim (see also Degen & Tanenhaus 2011, 2015, Sun & Breheny 2020 for related data).

In light of these and other considerations, Grodner et al. (2010) made the following changes to Huang & Snedeker's (2009) experimental design.

1. Use of a reduced pronunciation, *summa,* in place of a more fully articulated *some of,* so that the phonological cue for the partitive construction is heard earlier. Correspondingly, *all of* is also pronounced as *alla.*
2. Exclude numerals from the experiment altogether.
3. Inclusion of a statement at the beginning of each trial, describing the visual scene. This is to make sure that participants don't get drawn to visually salient objects, especially an individual with the most items, which could facilitate the processing of *alla.*
4. Inclusion of *nunna,* a contracted pronunciation of *none of,* as a baseline condition, in addition to *alla.*
5. Inclusion of a baseline condition (late *summa*) where the scalar implicature of *summa* is not enough to uniquely determine the referent, and the following noun must be heard to do so.

In this experiment, the disambiguation effects of *summa* were observed within 200–400 ms from the onset of the quantifier, which, considering the time it takes to execute eye movements, suggests that it was accessed as soon as *summa* was heard. Similar early disambiguation effects were observed for the two baselines, *alla* and *nunna,* and as expected, identification of the referent in the late *summa* condition took place significantly later, namely when the noun was heard.

Grodner et al. (2010) argue that one of the key differences between their experiment and Huang & Sndeker's (2009) was the absence of numerals, but because of the other differences, the results of this experiment alone do not constitute convincing evidence for their claim (see also Breheny, Ferguson & Katsos, 2013; Degen & Tanenhaus, 2016; Foppolo & Morelli, 2017; Huang & Snedeker, 2018) . To test which of the above changes had crucial effects more directly, Huang & Snedeker (2018) conducted two experiments. Their Experiment 1 investigated whether the faster pronunciation *summa* had any significant effect. Its design was essentially identical to Huang & Snedeker's (2009) original experiments except that *some of* was replaced with the reduced form, *summa.* They observed delayed processing of *summa,* replicating Huang & Snedeker's (2009) earlier results, meaning the pronunciation of *some of* is not responsible for the diverging results. Their Experiment 2 tested the effects of the pronunciation as well as the presence of items with bare numerals in the same experiment by crossing two factors, *summa* vs. *some of,* and whether or not numerals were included. While the former manipulation exhibited no discernible difference, as in their Experiment 1, they observed a reliable effect of the latter such that when numerals were present, the processing of *summa/some of* was delayed relative to *all,* but not when numerals were absent.

One thing to be noted at this point about these eye-tracking experiments is that the crucial scalar item *some* occurs inside a definite description, as in *the girl with some of the socks,* and from a theoretical point of view, a scalar inference in such a case might have a different property from *bona fide* scalar implicatures. More specifically, such cases are considered to involve scalar inferences in the presuppositional domain, which are called *anti-presuppositions* or *implicated presuppositions.* It is currently hotly debated whether such scalar inferences are generated by the same mechanism as scalar implicatures (see, for example, Percus, 2006; Sauerland, 2008; Schlenker, 2012; Spector & Sudo, 2017; Marty, 2017; Anvari, 2019). Breheny, Ferguson & Katsos (2013) explicitly raise this concern (see also Sun & Breheny 2020). However, we can find studies that didn't use definite descriptions but nonetheless yielded results that are generally in line with the other eye-tracking experiments reviewed above. For example, Breheny, Ferguson & Katsos (2013) used sentences like *The man has poured some of the water with limes into the bowl on tray A and all of the water with oranges in to the bowl on tray B.* The logic of their experiment is the same as in the above experiments, and crucially, they did not have items with numerals. Their results indicate that the scalar implicature of *some* was accessed as soon as the quantifier was heard, replicating Grodner et al.'s (2010) results. Similarly, Degen & Tanenhaus (2016) used sentences like *You got some of the blue gumballs.* Their Experiment 1 did not contain numerals and their Experiment 2 did. In their results, the effects of the "not all" scalar implicature of *some* were delayed only in Experiment 2, which is consistent with the results of the other experiments above.

There are several ways to interpret the observed effects of numerals. As Grodner et al. (2010) suggested, this might mean that having seen numerals, which are arguably more natural descriptions than *some* in relevant contexts with a small number of relevant items (see Grodner et al., 2010; Degen & Tanenhaus, 2011; 2015; 2016; Sun & Breheny 2020), participants simply slowed down upon encountering the less natural expression *some.* Another conceivable interpretation is put forward by Huang & Snedeker (2018), according to which the absence of numerals in the experiment invited participants to pre-encode the visual scene in terms of *some* and *all,* allowing them to predict that a character with a subset of a particular set of items would be referred to using *some,* and hence facilitated its processing. Thus, according to this view, the true processing effects of the "not all" scalar implicatures are observed when alternative expressions block such verbal encoding.

While these two views are still being discussed in the literature, Degen & Tanenhaus' (2016) observe interesting effects of 'set size' that are not expected under either of these views. Specifically, this study observed that processing of *some* was delayed relative to *all,* when there were four or five relevant 'verifiers' in the scene, while not when there were only two or three. This goes against Grodner et al.'s (2010) idea above based on naturalness, because when the number is big, it is observed that *some* becomes as good a description as numerals, at least in offline judgment studies (Degen & Tananhaus, 2011; 2015). However, such an effect of number is also not expected under Huang & Snedeker's (2018) idea of verbal encoding. More recently, Sun & Breheny (2020) investigated the possibility that this has little to do with scalar implicatures and simply comes from processing of *all,* rather than that of *some.* Their idea is that when the total number is small, the use of *all* is not so natural, as indicated by their offline acceptability judgment experiments, and hence its processing slows down, which makes the relative processing difference with *some* relatively small. The results of their eye-tracking experiments provide support for this idea. Notice, however, that it is not so clear if Sun & Breheny's idea is applicable to *none,* which has also been used as a baseline condition to compare *some* to, but this is of course an empirical question that needs to be investigated.

To summarise, all experimental methods, other than the visual world eye-tracking paradigm, have yielded more or less converging results, showing that processing of the "not all" scalar implicature of *some* requires extra processing time and cognitive effort. It should be mentioned also that a number developmental studies also suggest the same conclusion (e.g. Noveck, 2001; Papafragou & Musolino, 2003; Barner & Bachrach, 2010; see Chemla & Singh, 2014a; 2014b for an overview) The results of the visual world eye-tracking paradigm are interestingly intricate, and subtle difference in experimental design seems to have large effects. We would ultimately like to have a unified view that could explain these divergent results obtained from different experimental methods, but this requires further investigation.

4. Summary and Outlook

In this article, we have reviewed the theoretical and experimental literature on scalar implicatures, mostly focusing on *some* (and *or*). As mentioned above, *some* is still by far the most frequently investigated scalar item in this context, but there is also a sizable body of experimental work on other types of scalar implicatures.

In particular, numerals are very well studied, partly because of the theoretical controversy surrounding their semantics and pragmatics. The current literature on numerals is copious and would require a survey article on its own, so I simply mention here some references: Spector (2013) is a nice overview of different theoretical views and debates, and references to experimental work can be found in survey articles and books such as Noveck & Reboule (2008), Katsos & Cummins (2010), Chemla & Singh (2014a; 2014b), Noveck (2018), and Schumacher (2018).

It should also be mentioned that these survey articles and books also serve similar purposes to the present article as well as complement it, as they review other kinds of experimental research than sentence processing, such as child language development and neuro-imaging.

Lastly, more recent studies have started to expand the empirical landscape by looking at even more types of scalar implicatures. Above all, it is observed in offline judgment studies (Van Tiel, Van Miltenburg, Zevakhina & Geurts, 2016, among others) that different scalar items give rise to scalar implicatures with different degrees of robustness, and it is one of the new challenges for theories of scalar implicatures to explain why this is so. Related to this is the question about how they are processed, and the experimental literature on this question is currently rapidly growing (e.g. Cremers & Chemla, 2014; Romoli & Schwarz, 2015; Van Tiel & Schaeken, 2017; Bill, Romoli & Schwarz, 2018; Van Tiel, Marty, Pankratz & Sun, 2019; Van Tiel, Pankratz & Sun, 2019; Marty, Romoli, Sudo, Van Tiel & Breheny, 2020).

References

Anvari, A. (2019). *Aspects of contextual enrichment.* Unpublished doctoral dissertation, Ecole Normale Supérieure, Paris.

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology,* 60(1), 40–62. https://doi.org/10.1016/j.cogpsych.2009.06.002

Bill, C., Romoli, J., & Schwarz, F. (2018). Processing presuppositions and implicatures: similarities and differences. *Frontiers in Communication,* 3(44), 1–22. https://doi.org/10.3389/fcomm.2018.00044

Bott, L., Bailey, T. M., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language,* 66(1), 123–142. https://doi.org/10.1016/j.jml.2011.09.005

Bott, L., & Novek, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language,* 51(3), 437–457. https://doi.org/10.1016/j.jml.2004.05.006

Breheny, R., Ferguson, H. J., & Katsos, N. (2012). Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes,* 28(4), 443–467. https://doi.org/10.1080/01690965.2011.649040

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition,* 100(3), 434–463. https://doi.org/10.1016/j.cognition.2005.07.003

Chemla, E., & Bott, L. (2014). Processing inferences at the semantics/pragmatics frontier: disjunctions and *free choice. Cognition,* 130(3), 380–396. https://doi.org/10.1016/j.cognition.2013.11.013

Chemla, E., & Spector, B. (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics,* 28(3), 359–400. https://doi.org/10.1093/jos/ffq023

Chemla, E., & Singh, R. (2014a). Remarks on the experimental turn in the study of scalar implicature, Part I. *Language and Linguistics Compass,* 8(9), 373–386. https://doi.org/10.1111/lnc3.12081

Chemla, E., & Singh, R. (2014b). Remarks on the experimental turn in the study of scalar implicature, Part II. *Language and Linguistics Compass,* 8(9), 387–399. https://doi.org/10.1111/lnc3.12080

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond: The Cartography of Syntactic Structures,* pp. 39–103. Oxford: Oxford University Press.

Chierchia, G., Fox, D., & Spector, B. (2011). Scalar implicature as a grammatical phenomenon. In C. Maienborn, K. von Heusinger, & P. Porter (Eds.), *Semantics: An International Handbook of Natural Language Meaning,* pp. 2297–2331. Berlin: de Gruyter.

Clifton, C., Jr., & Dube, C. (2010). Embedded implicatures observed: a comment on Geurts and Pouscoulous (2009). *Semantics and Pragmatics,* 3(7), 1–13. http://dx.doi.org/10.3765/sp.3.7

Cremers, A., & Chemla, E. (2014). Direct and indirect scalar implicatures share the same processing signature. In S. Pistoia Reda (Ed.), *Pragmatics, Semantics and the Case of Scalar Implicatures,* pp. 201–227. London: Palgrave Macmillan.

Cummins, C. (2014). Typicality made familiar: A commentary on Geurts and van Tiel (2013). *Semantics and Pragmatics,* 7(8), 1–15. http://dx.doi.org/10.3765/sp.7.8
Davis, W. (2019). Implicature. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy,* Fall 2019 Edition. https://plato.stanford.edu/archives/fall2019/entries/implicature/
Degen, J. (2015). Investigating the distribution of *some* (but not *all*) implicatures using corpora and web-based methods. *Semantics and Pragmatics,* 8(1), 1–55. http://dx.doi.org/10.3765/sp
Degen, J., & Tanenhaus, M. K. (2011). Making Inferences: The Case of Scalar Implicature Processing. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science,* pp. 3299–3304.
Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: a constraint-based approach. *Cognitive Science,* 39(4), 667–710. https://doi.org/10.1111/cogs.12171
Degen, J., & Tanenhaus, M. K. (2016). Availability of alternatives and the processing of scalar implicatures: a visual world eye-tracking study. *Cognitive Science,* 40(1), 172–201. https://doi.org/10.1111/cogs.12227
Dieussaert, K., Verkerk, S., Gillard, E., & Schaeken, W. (2011). Some effort for some: Further evidence that scalar implicatures are effortful. *Quarterly Journal of Experimental Psychology,* 64(12), 2352–2367. https://doi.org/10.1080/17470218.2011.588799
Foppolo, F., & Marelli, M. (2017). No delay for *some* inferences. *Journal of Semantics,* 34(4), 659–681. https://doi.org/10.1093/jos/ffx013
Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland, & P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics,* pp. 71–112. London: Palgrave Macmillan.
Fox, D. & Spector, B. (2018). Economy and embedded exhaustification. *Natural Language Semantics,* 26(1), 1–50. https://doi.org/10.1007/s11050-017-9139-6
Franke, M., Schlotterbeck, F., & Augurzky, P. (2017). Embedded scalars, preferred readings and prosody: an experimental revisit. *Journal of Semantics,* 34(1), 153–199. https://doi.org/10.1093/jos/ffw007
Geurts, B. (2010). *Quantity implicatures.* Cambridge: Cambridge University Press.
Geurts, B., & Pouscoulous, N. (2009). Embedded implicatures?!? *Semantics and Pragmatics,* 2(4), 1–23. http://dx.doi.org/10.3765/sp.2.4
Geurts, B. & van Tiel, B. (2013). Embedded scalars. *Semantics and Pragmatics,* 6(9), 1–37. http://dx.doi.org/10.3765/sp.6.9
Grice, P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.
Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition,* 116(1), 42–55. https://doi.org/10.1016/j.cognition.2010.03.014
Gutzmann, D. (2015). *Use-conditional meaning: Studies in multidimensional semantics.* Oxford: Oxford University Press.
Horn, L. (2004). Implicature. In L. R. Horn, & G. Ward (Eds.), *The handbook of pragmatics,* pp. 2–48. Oxford: Blackwell.
Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology,* 58(3), 376–415. https://doi.org/10.1016/j.cogpsych.2008.09.001
Huang, Y. T. & Snedeker, J. (2011). *Logic and conversation* revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes,* 26(8), 1161–1172. https://doi.org/10.1080/01690965.2010.508641

Huang, Y. T., & Snedeker, J. (2018). *Some* inferences still take time: Prosody, predictability, and the speed of scalar implicatures. *Cognitive Psychology,* 102, 105–126. https://doi.org/10.1016/j.cogpsych.2018.01.004

Katsos, N., & Cummins, C. (2010). Pragmatics: from theory to experiment and back again. *Language and Linguistics Compass,* 4(5), 282–295. https://doi.org/10.1111/j.1749-818X.2010.00203.x

Levinson, S. C. (2000) *Presumptive meanings: The theory of generalized conversational implicature.* Cambridge, MA: MIT Press.

Marty, P. (2017) *Implicatures in the DP domain.* Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Marty, P., Chemla, E., & Spector, B. (2013). Interpreting numerals and scalar items under memory load. *Lingua,* 133, 152–163. https://doi.org/10.1016/j.lingua.2013.03.006

Marty, P., Romoli, J., Sudo, Y., Van Tiel, B., & Breheny, R. (2020). Processing implicatures: a comparison between direct and indirect SIs [conference presentation]. CUNY Processing Conference 2020, University of Massachusetts, Amherst, United States of America.

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load. *Experimental Psychology,* 54(2), 128–133. https://doi.org/10.1027/1618-3169.54.2.128

Noveck, I. A. (2001). When children are more logical than adults: investigations of scalar implicature. *Cognition,* 78(2), 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Noveck, I. A. (2018). *Experimental pragmatics.* Cambridge: Cambridge University Press.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language,* 85(2), 203–210. https://doi.org/10.1016/S0093-934X(03)00053-1

Noveck, I. A., & Reboul, A. (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences,* 12(11), 425–431. https://doi.org/10.1016/j.tics.2008.07.009

Panizza, D., Chierchia, G., & Clifton, 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. https://doi.org/10.1016/j.jml.2009.07.005

Panizza, D., Huang, Y. T., Chierchia, G., & Snedeker, J. (2009). Relevance of polarity for the online interpretation of scalar terms. In E. Cormany, S. Ito, & D. Lutz (Eds.), *Proceedings of Semantics and Linguistic Theory 19,* pp. 360–378. eLanguage.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition,* 86(3), 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8

Percus, O. (2006). Antipresuppositions. In *Theoretical and Empirical Studies of Reference and Anaphora: Toward the Establishment of Generative Grammar as an Empirical Science,* pp. 52–73. Grant report, Japan Society for the Promotion of Science.

Potts, C. (2005). *Logic of conventional implicatures.* Oxford: Oxford University Press.

Potts, C., Lassiter, D. Levy, R., & Frank, M. C. (2016). Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics,* 33(4), 755–802. https://doi.org/10.1093/jos/ffv012

Romoli, J., & Schwarz, F. (2015). An experimental comparison between presuppositions and indirect scalar implicatures. In Florian Schwarz (ed.), *Experimental Perspectives on Presuppositions,* pp. 215–240. Berlin: Springer.

Russell, B. (2006). Against grammatical computation of scalar implicatures. *Journal of Semantics,* 23(4), 361–382. https://doi.org/10.1093/jos/ffl008

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics & Philosophy,* 27(3), 367–391. https://doi.org/10.1023/B:LING.0000023378.71748.db

Schlenker, P. (2012). *Maximize Presupposition* and Gricean reasoning. *Natural Language Semantics,* 20(4), 391–429. https://doi.org/10.1007/s11050-012-9085-2
Schmacher, P. B. (2018). Semantic-pragmatic processing. In E. M. Fernández, & H. Smith Cairns (eds.), *The handbook of psycholinguistics,* pp. 392–410. New York: Wiley.
Spector, B. (2006). *Aspects de la pragmatique des operateurs logiques.* Unpublished doctoral dissertation, Université Paris 7, Denis Diderot.
Spector, B. (2013). Bare numerals and scalar implicatures. *Language and Linguistic Compass,* 7(5): 273–294.
Spector, B., & Sudo, Y. (2017) Presupposed ignorance and exhaustification: how scalar implicatures and presuppositions interact. *Linguistics and Philosophy,* 40(5), 473–517. https://doi.org/10.1111/lnc3.12018
Sperber, D., & Wilson, D. (1986) *Relevance: Communication and cognition.* Cambridge, MA: Harvard University Press.
Sun, C., & Breheny, R. (2020). Another look at the online processing of scalar inferences: an investigation of conflicting findings from visual-world eye-tracking studies. *Language, Cognition and Neuroscience,* 35(8), 949–979. https://doi.org/10.1080/23273798.2019.1678759
van Tiel, B. (2014). Embedded scalars and typicality. *Journal of Semantics,* 31(2), 147–177. https://doi.org/10.1093/jos/fft002
van Tiel, B., Marty, P., Pankratz, E., & Sun, C. (2019). Scalar inferences and cognitive load. In *Proceedings of Sinn und Bedeutung 23,* pp. 427–442.
van Tiel, B., van Miltenburg, E., Zevakhina, E., & Geurts, B. (2016). Scalar diversity. *Journal of Semantics,* 33(1), 137–175. https://doi.org/10.1093/jos/ffu017
van Tiel, B., Noveck, I., & Kissine, M. (2018). Reasoning with 'some'. *Journal of Semantics,* 35(4), 757-797. https://doi.org/10.1093/jos/ffy012
van Tiel, B., Pankratz, E., & Sun, C. (2019). Scales and scalarity: processing scalar inferences. *Journal of Memory and Language,* 105, 93–107. https://doi.org/10.1016/j.jml.2018.12.002
van Tiel, B., & Schaeken, W. (2017). Processing conversational implicatures: alternatives and counterfactual reasoning. *Cognitive Science,* 41(S5), 1119–1154. https://doi.org/10.1111/cogs.12362
Tomlinson, J. M., Jr., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language,* 69(1), 18–35. https://doi.org/10.1016/j.jml.2013.02.003