# Scalar Implicature

Anna Papafragou and Dimitrios Skordos
University of Delaware

## 1. Introduction

One of the hallmarks of human communication is the ability to go beyond the words uttered to compute what the speaker meant by an utterance. Speakers tacitly exploit inferential principles to bridge the gap between words and meaning and count on hearers to invoke the same principles for the purposes of utterance interpretation. To become fully competent communicators, children need to be able to “read between the lines,” to calculate aspects of speaker meaning that go beyond the literal meaning of the words. The ubiquity of inference in communication complicates the task of word learning: In acquiring language, children need to do more than map word meanings onto speaker intentions: more specifically, they need to discover the intricate division of labor between semantic and pragmatic aspects of word meaning (roughly, what words linguistically encode versus what inferences words can give rise to on the basis of their linguistically encoded meaning).

Despite the tight links between lexical-semantic content and pragmatic inference, we know relatively little about how the ability to infer aspects of speaker meaning develops in children. Early work on a number of phenomena from indirect speech acts (“Can you give me an apple?”) to figurative language (“My sister is an angel”) showed that children are oblivious to their interlocutors’ intentions and fail to compute pragmatic meaning (see review in Shatz 1983). As a result, many researchers thought it safe to assume that theories of lexical and grammatical development can be advanced without considering the development of the pragmatic component. Moreover, until recently there was a lack of methodological tools specifically designed for the study of pragmatic interpretation in children (for exceptions, see Crain and Thornton 1998; Trueswell et al. 1999).

In the last decade, there has been renewed interest in the psychology of pragmatic inference. This interest has been fueled by a resurgence of theoretical investment in the semantics–pragmatics interface in theoretical linguistics and by serious attempts to develop experimental methods for studying pragmatic, as opposed to purely semantic, contributions to meaning by psychologists and psycholinguists. Most of this growing literature has concentrated on a sub-type of pragmatic inference known as *scalar implicature* which presents a sharply defined and easily testable division between encoded and inferred meaning. In this chapter, we take scalar implicature as a paradigm case for children's pragmatic development and use the available developmental evidence to probe questions about the scope and nature of early pragmatic abilities. We begin by offering some background on scalar implicatures and the theories that have been proposed to account for their derivation in adults (section 2). We go on to review the developmental evidence on whether and how young children calculate scalar implicatures and propose an account of early successes and failures with pragmatic inferences (section 3). We next consider the implications of developmental findings for adult (and child) semantic representations of different scalar expressions, paying particular attention to a recent debate about the semantics of numerals and related terms (section 4). We conclude with a discussion of how findings from children's pragmatic development can both inform theories of word learning and constrain theories of semantic representation in children as well as adults.

## 2. Theoretical background on scalar implicatures

*Implicatures* are components of speaker meaning that constitute an aspect of what is *meant* in a speaker's utterance without being part of what is *said*. The utterances in (1a)–(3a) are typically taken to implicate the propositions in (1b)–(3b):

(1) a. Mary ate some of the cakes.
b. Mary did not eat all of the cakes.

(2) a. Elmo will buy a car or a boat.
b. Elmo will not buy both a car and a boat.

(3) a. We saw three bears.
b. We did not see more than three bears.

The standard account of how these implicatures are derived stems from Grice's (1975) theory of inferential communication. As is well known, Grice suggested that communication is essentially a cooperative enterprise governed by certain rational expectations ("Maxims") about how a conversational exchange should be conducted. According to Grice's maxims, interlocutors are normally expected to offer contributions which are truthful, informative, relevant to the goals of the conversation, and appropriately phrased. These expectations about rational conversational conduct constrain the range of inferences which hearers are entitled to entertain when interpreting utterances. Furthermore, these expectations can be violated (or exploited) to create a variety of effects. According to Grice's theory, in producing (1a), the speaker violated the maxim of informativeness (or Quantity):

> *Quantity maxim*
> (i) Make your contribution as informative as is required.
> (ii) Do not make your contribution more informative than is required.

Specifically, the speaker in (1a) violated the submaxim (i) since he/she chose a relatively weak term from among a range of items logically ordered in terms of informational strength (<*all*, . . . , *some*>). Assuming that the speaker is trying to be cooperative and will say as much as he/she truthfully can that is relevant to the exchange, the fact that he/she chose the weaker term (*some*) gives the listener reason to think that he/she is not in a position to offer an informationally stronger statement ("Mary ate all of the cakes."). This leads to the inference that, as far as the speaker knows, the stronger statement does not hold, that is, to (1b).[1] The type of conversational implicature in (1b) is known as a *Quantity* or *scalar* implicature (SI; so called because of the informational scale < *all, some* >; Horn, 1972).

On this account, the quantifier *some* has a lower-bound semantics ("at least some and possibly all") that is upper-bound by the implicature. Evidence for the lower-bound semantics comes from the fact that the implicature can be explicitly canceled (e.g. the speaker in (1a) may continue: "In fact, Mary ate all of the cakes.") without logical

---

[1] We omit reference to the speaker's epistemic state from (1b) for simplicity.

contradiction. The rest of the examples above work in similar ways: (2b) rests on the scale < *and*, *or* > and (3b) on the scale < *four, three* . . .> (see Horn, 1972, 1984 for more examples). In some cases, SIs can be derived from non-logical scales that are based on contextual information (Hirschberg 1985). For instance, the child's response in (4) implicates that she did not complete the action based on a temporal scale:

(4) Mother: Did you make the fruit salad?
Child: I peeled the fruit.

Examples such as (1)–(4) have been studied extensively in philosophical and linguistic work on verbal communication and several influential proposals about scalar inferences have been developed beyond Grice's original theory (Horn 1972, 2005; Harnish 1976; Gazdar 1979; Geurts 1998, 2010; Chierchia 2004; Sauerland 2004; Fox 2007; Spector 2007; Chierchia et al. 2009). Three proposals are worth mentioning in some detail. On certain neo-Gricean accounts, SIs based on logical scales such as those in (1b)–(3b) become stored in the lexicon and available immediately every time the weak scalar term is accessed, only to be canceled later in case they are not supported by context (Levinson 2000). Other recent accounts take the generation of SIs to be a grammatical process contributing to truth conditional content (Chierchia 2004; Chierchia et al. 2009): this grammatical account shares with Levinson's (2000) neo-Gricean model the assumption that logical scales are part of the lexicon and are activated every time a weak scalar item is encountered. Relevance theory proposes that all pragmatic inferences, including implicatures, are guided by a general tradeoff between the projected cognitive gains from computing an inference and the amount of cognitive effort necessary to derive it (Sperber and Wilson 1985/1996; see also Carston 1995; Noveck and Sperber 2007). A crucial difference between these three accounts is that Relevance theory sides with traditional Gricean accounts in treating SIs as products of context-driven inference (cf. also Horn 1989), whereas Levinson's (2000) neo-Gricean account and the grammatical view hold that SIs from logical expressions are generated by default, independently from contextual input.

From a developmental perspective, SIs raise several questions about both the nature of early conversational inferences and the mechanisms whereby contextual inferences are computed: Do children generate pragmatic inferences? If not, how do they recover and become pragmatically savvy adults? Are the mechanisms underlying linguistic pragmatics fundamentally the same across children and adults? Moreover, if one assumes continuity across the two populations, which aspects of pragmatic ability actually develop? Finally, how do different scalars contribute to the derivation of SIs? Within semantic theory, there has been considerable discussion of the similarities and differences in the semantic contribution of different expressions which all seem to give rise to scalar effects (Sadock 1984; Horn 1992; Koenig 1993; Carston 1995; and section 4 below). Do different scalar expressions follow distinct acquisition paths? Even though these issues are far from theoretically or empirically settled, they have been the topic of intensive recent experimentation and theorizing.

## 3. SI processing in children

### 3.1 Developmental evidence

Are SIs psychologically real for children (and adults)? Several findings from early studies designed to investigate children's knowledge of quantification and propositional connectives bear on this question. For example, Smith (1980b) observed that preschool children, who had mastered many of the syntactic aspects of quantifiers like *some* and *all*, were likely to respond affirmatively to questions such as "Do some birds have wings?" (i.e. they were likely to interpret *some* as "at least some and possibly all"). In another study, Braine and Rumain (1981) found that adults tended to weakly favor an exclusive interpretation of the disjunction operator *or* (i.e. they tended to interpret *p or q* as "either p or q but not both"); however, 7- and 9-year-old children favored a logical/inclusive interpretation of disjunction on which *p or q* is interpreted as "p or q and possibly both" (see also Paris 1973). However, these early studies mostly focused on the logical meaning of quantifiers/disjunction and the significance of these findings for children's pragmatic development was largely overlooked.

The first study to point out the relevance of this prior work and systematically investigate children's understanding of SIs was Noveck (2001). That study examined comprehension of the modal term *might* in 5-, 7-, and 9-year olds, as well as adults, using a scenario that involved reasoning about the contents of a covered box. In a crucial trial, participants had to say whether they agreed or not with the statement “There might be a parrot in the box” when the evidence made it clear that there *had to* be a parrot in the box. Notice that the target statement was logically true (since *might* is compatible with *have to*) but pragmatically infelicitous (since the use of *might* could be grounds for excluding *have to*). Noveck found that adults tended to disagree with the statement, while children of all three age groups overwhelmingly agreed with it. In another experiment, Noveck tested French-speaking children’s and adults’ comprehension of the existential quantifier *certains* ‘some’. He asked participants whether they agreed or not with statements of the form “Some elephants have trunks.” He found that 8- and 10-year-old children typically treated *certains* logically, as compatible with *tous* ‘all,’ whereas adults were equivocal between the logical and the pragmatic interpretations. As Noveck (2001) pointed out, these data cohere with results from earlier studies from the 1970s and 1980s (Smith 1980b; Braine and Rumain 1981, among others). Taken together, these results strongly indicate that children appear to be more logical than adults in reasoning tasks that involve the use of quantity terms. Specifically, it seems that otherwise linguistically competent children are oblivious to pragmatic inferences from the use of scalar terms such as modals and quantifiers.

What is the nature of children’s failure with the calculation of scalar implicatures? One possibility is that this failure reflects a genuine inability to engage in the computations required to derive such implicatures. Another possibility is that the failure is due to the demands imposed by the experimental task on an otherwise pragmatically savvy child. As Noveck (2001) said, “[t]he tasks described here, which are typical of those found in the developmental literature, demand no small amount of work as they require children to compare an utterance to real world knowledge. This might well mask an ability to perform pragmatic inferencing at younger ages” (2001: 184). Under the second, but not the first, of these alternative hypotheses, children’s ability to derive scalar implicatures could improve under certain experimental circumstances.

Evidence in support of the second hypothesis comes from a set of studies by Chierchia et al. (2001) and Gualmini et al. (2001). These studies investigated preschoolers’ interpretation of the disjunction operator *or*. Extending and confirming previous studies, the authors found that adults were sensitive to the implicature of exclusivity from the use of disjunction (e.g. they interpreted an utterance such as "Every boy chose a skateboard or a bike" as excluding the possibility that *both* a skateboard *and* a bike were chosen); children, by contrast, often showed virtually no sensitivity to the exclusive reading of disjunction. Crucially, children were able to distinguish between a stronger statement containing the conjunction operator *and* and a weaker scalar statement containing the disjunction operator *or* (Chierchia et al. 2001). Specifically, when children were presented with two statements (produced by two puppets) and were asked to reward the puppet who "said it better," they overwhelmingly chose to reward the puppet who produced a stronger/more informative statement with *and* ("Every farmer cleaned a horse and a rabbit") over a puppet who offered a weaker/less informative statement with *or* ("Every farmer cleaned a horse or a rabbit") in a context that made the stronger statement true. More recently, Ozturk and Papafragou (2014) obtained similar results with modals. In one of their experiments, children (and adults) overwhelmingly preferred logical interpretations of the modal *may* in a reasoning task that involved locating a hidden animal; for instance, both groups accepted the statement "The cow may be in the orange box" in a context in which, according to the available evidence, the cow *had to* be in the orange box. A second experiment found that, if given a choice between two statements in the very same context, both adults and 5-year-olds preferred statements with *have to* over statements with *may*.[2] Taken together, these studies show that children have knowledge of the relative information strength of sentences with *or* versus *and*, or *may* versus *have to*, and they use information strength as the basis of their preference for stronger sentences over weaker ones in circumstances which make the stronger sentences true. Thus it seems that a critical prerequisite for calculating SIs (namely, recognizing strength differences between otherwise identical conversational contributions) is in place in young children. What appears to be problematic for children is the step of generating and comparing stronger alternatives to

[2] Other evidence shows that, even in the absence of background evidence, children can differentiate between stronger and weaker scalars. For instance, if presented with two conflicting statements about the possible location of an object ("The x may be under the cup" versus "The x must be under the box"), children typically choose the location associated with the stronger expression (Hirst and Weil 1982; Byrnes and Duff 1989; Moore et al. 1989; Noveck et al. 1996).

a weak statement on-line in cases where the alternatives are not explicitly presented to them.

More direct evidence for the conclusion that children are not entirely insensitive to pragmatic inferences but may be able to compute SIs under certain circumstances is provided by Papafragou and Musolino (2003). In one study, the authors tested Greek-speaking 5-year-olds and adults on three types of scales: the quantificational scale <*all*, *some*>, the numerical scale < *three*, *two*> and the aspectual scale < *finish*, *start*>. Their method was a variation of the Truth Value Judgment (TVJ) task (Crain and McKee 1985) called the Acceptability Judgment task: participants watched a series of acted-out stories along with a puppet. At the end of the story, the puppet was asked to say "what happened" and participants had to say whether the puppet "answered well." On critical trials, the puppet produced a true but underinfomative statement (e.g. "Some/Two of the horses jumped over the fence" in a story in which every horse in a group of three horses jumped over a fence). It was found that 5-year-olds were much more likely than adults to accept the logically true but pragmatically infelicitous statements. Papafragou and Musolino (2003) hypothesized that this adult–child difference might be due to the difficulty of reconstructing the experimenter's goal in this task: in asking whether the puppet "answered well," children (unlike adults) may have been more likely to base their judgments on truth than pragmatic infelicity. To test this hypothesis, in a second study Papafragou and Musolino modified their procedure in several ways. First, to enhance awareness of the goals of the experiment, they initially trained children to detect pragmatically anomalous statements produced by a "silly puppet" (e.g. children were encouraged to say that the statement "This is a small animal with four legs" is "silly" and the puppet should simply say "This is a dog"). Second, to ensure there was a salient informativeness threshold, the experimental scenarios and test question were modified to focus on a character's performance in a task (e.g. in one of the stories, Mr Tough brought back three horses that had run away; when asked how Mr Tough did, the puppet gave the response "He caught some/two of the horses"). Under these conditions, 5-year-olds were more likely to compute scalar implicatures, even though still not at adultlike levels. These data show that, given contextual support, children show some ability to spontaneously generate SIs; furthermore, these results leave open the possibility that different experimental manipulations might reveal higher success with pragmatic inferences.[3]

Building on Papafragou and Musolino (2003), other studies have confirmed the role of training and context in older children's ability to calculate SIs. Guasti et al. (2005) showed that Italian-speaking 7-year-olds accepted underinformative but true statements of the type "Some giraffes have long necks" more often than adults when the statements were presented out of context (thereby replicating Noveck, 2001). However, when the same statements were preceded by training in rejecting infelicitous statements, children behaved exactly like adults (even though the effects of training did not persist when the children were re-tested a week later). Similarly, when underinformative statements were embedded within a story rather than presented out of context, children typically generated SIs at adultlike levels. Taken together, these studies make it increasingly clear that young children have the ability to make pragmatic inferences; however, they are limited in doing so by the cognitive resources they bring to bear on the process of utterance interpretation (see section 3.2 for details).

A particularly striking conclusion from this work is that the type of task used to elicit pragmatic responses affects children's success with scalar inferences. Notice that most or all of the early studies documenting children's limited awareness of SIs primarily relied on judgments about the acceptability of weak scalar expressions in contexts in which a stronger term is warranted. These tasks, however, are different from the actual circumstances in which SIs are computed during naturalistic conversations in several respects (see Papafragou and Tantalou, 2004, for related discussion). First, experimental conditions did not make it clear whether (or why) SIs should be considered as part of what the speaker actually *intended* to communicate. In ordinary cases of intentional communication, the speaker intends the addressee to compute the implicature (and further intends the addressee to recover this intention; cf. Grice 1989). But in the experimental designs discussed so far, the computation of SIs was not similarly constrained by the speaker's intention. In some experiments (Papafragou and Musolino 2003), a "silly" puppet uttered an underinformative statement (probably

[3] Papafragou and Musolino (2003) also discovered that 5-year-olds treated numbers differently from quantifiers in the generation of SIs. We postpone discussion of this finding and its theoretical significance until section 4.

because of incompetence) and might not have noticed that the statement carried the potential for conveying a SI; in others (Noveck 2001), underinformative statements appeared out of context and therefore invited participants to reconstruct a possible situation in which they could have been uttered by an actual communicator. In short, previous tasks measured children's sensitivity to *potential* implicatures in an effort to approximate their performance with *actual* (communicated) implicatures.

Second, previous tasks typically involved situations in which an utterance containing a scalar term (e.g. "Some of the Xs ved") semantically conveyed a true proposition (e.g. "Some and possibly all of the Xs ved") but carried a (potential) implicature that was false ("Not all of the Xs ved"). To perform correctly in these tasks (i.e. to reject the statement), hearers had to take the implicature (rather than simply the proposition expressed) as the basis for their assent/dissent with the target statement. In other words, participants had to estimate the experimenter's goal in setting up the task. This step is non-trivial, especially since, in the absence of cues about whether a logical or a pragmatic response is required, either type of response is acceptable. In fact, adults, who are otherwise able to compute SIs, when presented with underinformative statements (e.g. "Some airplanes have wings") without supporting linguistic or extralinguistic context, agree with the statements about half of the time (Braine and Rumain 1981; Noveck 2001; Guasti et al. 2005; cf. Papafragou and Schwarz 2006). When examined closely, this response pattern turns out to be due to the fact that adults consistently select either the logically true meaning of the utterance or the pragmatically enriched interpretation (Guasti et al. 2005).[4] Given adults' mixed pattern of responses, children's tendency to accept underinformative statements in the same environments cannot offer solid evidence of indifference to pragmatic meaning.

Relatedly, when children are asked to evaluate logically true but pragmatically underinformative statements in a binary (*Yes–No*) task, they may notice pragmatic infelicity but not penalize it as heavily as logical falsehood. Evidence supporting this hypothesis comes from judgment tasks that elicited more fine-grained responses to scalar statements. When 6- to 7-year-olds were asked to offer a 'small', 'big', or 'huge'

[4] When judgment tasks either explicitly (Noveck 2001, Exp.2) or implicitly (Ozturk and Papafragou, 2014, Exp.1) encourage logical responses, adults overwhelmingly accept weak scalar statements (i.e. they appear to disregard pragmatic unacceptability).

strawberry as a reward to a speaker depending on how good the speaker's responses were, children rewarded fully informative responses by giving the speaker 85 (out of 100 trials) 'huge' strawberries, underinformative ones by giving 89 'big' strawberries, and false responses by giving 95 'small' strawberries; crucially, children of this age massively accepted underinformative statements in a standard *Yes–No* judgment task (Katsos and Bishop 2011). Thus what appears to be insensitivity on the part of children to violations of informativeness might be better explained as tolerance towards underinformative statements in a binary task. These data reinforce the suspicion that the observed adult–child differences in pragmatic judgment tasks might be at least partly due to different choices in the way task requirements and goals are understood by young and more mature communicators (for instance, whether the task is interpreted as targeting logical versus pragmatic responses, or whether logical falsehood should be treated on a par with pragmatic infelicity). Interpreting task demands seems less relevant to processes underlying pragmatic processing in ordinary conversation but appears tied to metalinguistic awareness, an ability that is known to develop gradually over the school years (Ackerman 1981).

These observations, taken together, suggest that the family of judgment tasks, even though useful as an initial tool in exploring awareness of SIs, may underestimate young children's ability to compute implicatures "in the wild." More recently, several studies have explored different methods of evaluating children's early pragmatic abilities. In one such study (Papafragou and Tantalou 2004), Greek-speaking children were shown acted-out stories in which animals were asked to perform different tasks. After spending some time "off-stage," each animal was asked whether it performed the task and would sometimes answer with underinformative statements (e.g. Experimenter: "Did you color the stars?", Animal: "I colored SOME"). The reasoning was that, if children interpreted the animals' answer pragmatically, they would arrive at the conclusion that the animals failed the task and would not give them a reward; but if children interpreted the statement logically, they would conclude that the animals might well have performed the task and should be rewarded. Children overwhelmingly interpreted underinformative statements pragmatically (i.e. they denied the animals the reward). Furthermore, children's own reports as to why they had withheld the reward correctly made reference to the stronger scalar alternative (e.g. "The animal did not color ALL the stars”). Interestingly, children succeeded in computing SIs from non-logical scales alongside the more standard quantificational scales (e.g. Experimenter: “Did you eat the sandwich?”, Animal: “I ate THE CHEESE.”) Other work has confirmed that this more naturalistic method leads to higher rates of pragmatic responding than judgment tasks (Papafragou 2006; cf. also Verbuk, 2006b, for evidence that SIs are selectively generated in such question–answer pairs only when relevant).

In another study, Pouscoulous et al. (2007) tested French-speaking children’s ability to generate SIs spontaneously, that is, without the use of training that had been used in earlier work (Papafragou and Musolino 2003; Guasti et al. 2005). A first experiment used a TVJ task: 9-year-olds and adults were shown a series of boxes with animals in or next to them and were asked to evaluate statements such as “Some elephants are in the boxes.” Some of the statements were both true and felicitous, some were true but infelicitous, and some statements were false. It was found that semantically competent 9-year-olds were overwhelmingly logical in their responses to infelicitous statements, while adults were mixed. A second experiment used a similar context but an action-based task. Participants were shown a series of boxes containing two, five, or zero tokens and heard a statement (e.g. “I want some/all/no boxes to contain a token”.). Participants were expected to either add/remove tokens from the boxes, or leave the boxes unchanged so as to satisfy the experimenter’s wish. In critical trials, a *some*-statement was uttered when all of the boxes contained a token. Participants’ actions (i.e. whether they would leave the boxes intact or remove a token from at least one box) would reveal whether they had generated a logical or a pragmatic interpretation of the statement. Under these circumstances, even 5- and 7-year-olds generated SIs with regularity, and adults became more pragmatically oriented compared to Pouscoulous et al.’s (2007) first experiment.

Not all paradigms that lack a judgment component reveal children’s success with SIs. Using a visual world task in which sets of objects were divided among different characters in the display, Huang and Snedeker (2009a) monitored participants’ eye movements to test on-line processing of scalars. In one of their trials, 5-year-olds and adults were asked to “point to the girl that has some of the soccer balls” while inspecting a scene that contained a girl with two soccer balls, a boy with two soccer balls, a girl with three socks, and a boy with nothing. If participants calculated the SI, they should be able to resolve the temporary ambiguity in the noun (*sock . . .s* versus *socc . . . er balls*) and infer that the sentence refers to the girl with the subset of soccer balls shortly after the onset of the quantifier *some*. Although adults generated SIs and used them to identify the correct girl shortly after hearing *some*, children failed to do so and had to wait until the disambiguating phoneme in the noun was heard to identify the referent. Further experimentation showed that children did not seem to be slowed down by contexts in which the SI was subsequently violated—a fact supporting the conclusion that children did not calculate the SI in the first place. It remains an open question how to properly account for children's failures in these studies. One possibility is that children failed to reconstruct the set–subset relationship between soccer balls in the display, despite the experimenters' efforts to highlight these relationships at the beginning of each trial when objects were distributed among characters. Another possibility is that children (and perhaps to an extent adults) initially resisted applying *some* to scenes with only two referents, which might have been more felicitously described by a numeral. The infelicity of *some* may have been heightened by the fact that other trials in Huang and Snedeker's design involved numerals ("Point to the boy that has two of the soccer balls"); on these trials, both children and adults succeeded in quickly disambiguating the referent (see also section 4 below). Small modifications of Huang and Snedeker's procedure that involved, among other things, removing numerally quantified statements from the trials have been shown to greatly increase the speed with which adults calculate the inference from *some* (Grodner et al. 2010;[5] cf. Huang and Snedeker 2009b). It remains to be seen whether similar manipulations might affect children's performance in the same task.

## 3.2. Pragmatic inferences: What develops?

As already discussed, children are not entirely insensitive to pragmatic inferences. Nevertheless, the data surveyed in the previous section present a complex pattern of early successes and failures with SIs. How can this pattern be explained? Based on the results reported , we propose that part of the answer lies with children's difficulty in

---

[5] Grodner et al. (2010) used the term *summa* 'some of' instead of some, with *alla* and *nonna* being the other contextually available alternatives.

accessing and integrating different premises during the computation of speaker meaning—more specifically, the difficulty in inferring expectations of informativeness/relevance and evaluating a linguistic stimulus with respect to other possible alternatives that the speaker could have selected (cf. Papafragou and Tantalou 2004; Papafragou 2006).

Recall that, according to the standard Gricean model, the calculation of an SI requires the hearer to go through the following steps (cf. Section 2):

(i) The speaker has uttered a sentence with a weak scalar item.

(ii) The speaker thus violated the Quantity maxim since he/she chose a relatively weaker term from among a range of items logically ordered in terms of informational strength.

(iii) Assuming that the speaker is trying to be cooperative and will say as much as he/she truthfully can that is relevant to the exchange, the fact that he/she chose the weaker term is reason to think that he/she is not in a position to offer an informationally stronger statement.

(iv) Thus, as far as the speaker knows, the stronger statement does not hold.

To perform these calculations, the hearer needs to access a set of ordered alternatives to the weak scalar expression used by the speaker (Step ii); consider whether any of the stronger members of the ordered set are relevant and true (Step iii); if so, negate the stronger alternative (Step iii) and add the negated proposition to what the speaker intended to communicate (Step iv). Children's failure to compute SIs in previous tasks seems tied to the process of generating stronger relevant alternatives to a weak statement (Steps ii and iii). We know that, if alternative scalar statements are explicitly supplied, children consistently prefer the stronger alternative to the weaker one (Chierchia et al., 2001, on *or* versus *and*; Ozturk and Papafragou, 2014, on *may* versus *have to*). Thus children seem to have access to information about the logical properties of different scalar items, more specifically, the relative ordering of these items in terms of logical entailment (e.g. they recognize that *and* is stronger than *or*). The ability to appreciate the informational ordering of scalar items is a crucial prerequisite to Step (ii) above; nevertheless, this ability does not guarantee that children can spontaneously compute relevant scalar alternatives in the course of conversation. In fact, as shown by the studies reviewed in the previous section, the ability to consider stronger relevant alternatives is fragile: children failed to compute SIs when asked to judge weak scalar statements without specific cues about whether judgments should target pragmatic felicity (for which the target statement needs to be compared to potential alternatives that the speaker could have uttered) or logical truth (for which the presence of alternatives is irrelevant; Gualmini et al. 2001; Noveck 2001; Papafragou and Musolino 2003; Guasti et al. 2005). In experiments that clearly targeted pragmatic felicity and/or boosted the relevance of the stronger alternative, children's performance with SIs improved (Papafragou and Musolino 2003; Papafragou and Tantalou 2004; Guasti et al. 2005; Papafragou 2006; Pouscoulous et al. 2007).

If this line of reasoning is correct, children's difficulty in identifying relevant alternatives to what the speaker said should extend beyond SIs to other environments involving alternative-generation. This prediction is borne out (Barner et al. 2011): when 4-year-olds were presented with three sleeping animals and asked whether "some/only some of the animals are sleeping," they responded affirmatively about 66 percent of the time regardless of the form of the question. Children's failure to respond with *No* to the question with bare *some* is not surprising given that SIs typically do not arise in questions. But children's failure to respond with *No* to the question with *only* reveals that children have difficulty generating scalar alternatives even when the generation of alternatives is triggered by the grammar (since the focus element *only* grammatically requires the generation—and negation—of relevant alternatives). Crucially, when members of the set of animals were explicitly identified within the same displays (thereby making the set of relevant alternatives more salient), children's performance with *only* improved dramatically: when asked whether "only the cat and the dog are sleeping," children correctly gave *No*-responses 86 percent of the time. When members of the animal set were explicitly identified but the grammatical need to generate alternatives was eliminated (i.e. when simply asked whether "the cat and the dog are sleeping"), children accurately responded with an affirmative answer 93 percent of the time.

At present, the extent to which children can monitor linguistic alternatives in different contexts in accordance with informativeness/relevance expectations remains unknown to a large extent (although see Skordos and Papafragou, to appear). Some evidence, however, suggests that the ability to consider contrastive alternatives is not only within young learners' reach but is, in fact, very active in language acquisition (see Papafragou and Tantalou 2004). For instance, 2-year-olds can use the fact that an adult used a novel word (e.g. *dax*) rather than a known word (e.g. *car*) in a context that contains a car and a novel, unlabeled object to infer that the novel word must refer to the novel object (Carey 1978b; Halberda 2006). This inference stems from the assumption that word (especially, noun) meanings are *mutually exclusive* or *contrastive* (Clark 1987, 1988; Markman 1989). The assumption of mutual exclusivity/contrast is suspended for words that belong to different levels of description or to different languages (Au and Glusman 1990; Diesendruck 2005), presumably because in these cases the known label is not considered a relevant alternative to the novel label (Barner et al., 2011. Inferences driven by mutual exclusivity/contrast are distinct from implicatures in several ways; nevertheless, both involve calculating speaker meaning (including speaker intention) and both require holding in mind and negating lexical alternatives to an expression used by the speaker. For these reasons, some researchers consider these processes in early word learning to be essentially Gricean in nature (Gathercole 1989; Clark 1990; Diesendruck and Markson 2001). Even though the precise affinities between the mechanisms underlying mutual exclusivity/lexical contrast and SIs remain open, the fact that young children successfully consult known lexical alternatives in conjecturing meanings for novel words bolsters the conclusion that children's difficulties with SIs are not due to a complete inability to reason about conversational pragmatics.

How do children come to organize lexical alternatives in the form of scales and use such alternatives to compute SIs? Obviously, the first step for children is to acquire the lexical semantics of individual scalar expressions. This step might take place quite early in development: for instance, very young children seem to know that *some* and *all* refer to distinct set relations (Barner et al. 2009a). Beyond this semantic step the ability to treat terms such as *some* and *all* as scalemates requires further learning. The linguistic literature suggests that knowledge underlying which items form scales and which do not is quite complex: scalemates need to be expressions of equal length and complexity that are syntactically replaceable (Horn 1972; Levinson 2000; Katzir 2007), and scales obey several other constraints (Hirschberg 1985; Horn 1989; Matsumoto 1995).[6] To develop sensitivity to what count as scalar alternatives, children need to be able to acquire these fine-grained restrictions on scales. One cue for grouping scale members together may come from explicit contrasts in adult speech ("You can eat some of the cookies, but not all of them."). Another cue may be offered by the syntactic properties of semantically related lexical items (Barner et al., in press). The ability to access scalar alternatives seems to be undergoing development until at least the age of 7 or 9, since some studies find that children are still not quite adultlike at this age, especially if presented with scalar statements in isolation (e.g. Noveck 2001).

Beyond these broad points, there are several possibilities regarding the precise theoretical and psychological status of scales. Recall that *contextualist* accounts such as Relevance theory (Sperber and Wilson 1985/1996; Noveck and Sperber 2007) and traditional Gricean/neo-Gricean accounts (Horn 1989) treat all SIs as products of context-driven inference. In contrast, *defaultist* accounts such as Levinson's (2000) neo-Gricean account and the grammatical view of Chierchia et al. (2009) hold that SIs from logical expressions are generated by default, independently from contextual input, and later canceled if necessary. These two types of account make distinct commitments about the nature of scales. According to contextualist theories, there is no difference between logical (entailment-based) scales such as < *all*, *some* > and encyclopedic/ad hoc contextual scales (see examples (1)–(3) versus (4) in section 2): in both cases, stronger scalemates are accessed only when contextually relevant. According to defaultist accounts, logical scales differ in several respects from contextual scales: for instance, the scale < *all. . . some* > forms part of the lexical entry of individual quantifiers and stronger scalemates are accessed automatically every time a weak quantifier is encountered. One might hypothesize that such a scale might be easier to acquire compared to contextual scales because of the stable logical ordering of its members; furthermore, once established, it might lead to more reliable calculation of SIs compared to more context-dependent scales. So far no advantage has been detected for logical as opposed to non-logical scales: some studies find no difference between logical

[6] According to this literature, synonyms are expected to give rise to similar SIs since SIs are considered *non-detachable* (Horn 1972). Children seem to behave in accordance with this constraint (see Papafragou 2006, on inferences from *start* versus *begin* in Greek).

and contextual scales in children's calculation of SIs (Papafragou and Tantalou 2004), whereas others find that contextual scales are easier for children (Katsos and Smith 2010; Barner et al., 2011). Even within the logical class, there are differences in terms of how often children succeed in deriving SIs from expressions belonging to different scales (Noveck 2001; Papafragou and Tantalou 2004; Papafragou 2006; Verbuk 2006b; Pouscoulous et al. 2007; and section 4). Taken together, this evidence tentatively supports two conclusions: first, the ability to appreciate what counts as a relevant alternative develops in scale-specific ways; second, the fact that logical scales—unlike contextual ones—involve stable semantic orderings of lexical alternatives does not guarantee that children find it easier to build such scales or draw from them when they need to generate lexical alternatives.

The two theoretical positions sketched above also lead to different predictions about the online computation of SIs. According to contextualist theories, since SIs are always inferred as part of the speaker's intended meaning with more or less cognitive effort depending on the literal meaning of the sentence and the available context, the generation of SIs should be cognitively costly: as a result, more time should be required to process a sentence where an SI needs to be generated compared to a sentence without an SI. According to defaultist accounts, since (entailment-based) SIs are generated by default and then canceled when necessary, other things being equal, it should take more time to process a sentence requiring the cancellation of an SI than of a sentence where the SI is simply generated but not canceled. Available evidence from adults appears to fall in line with contextualist accounts. In a reaction time study (Bott and Noveck 2004), adults who judged underinformative sentences such as "Some elephants are mammals" to be false took longer than those who judged them to be true. This shows that adults take longer to compute implicatures than to arrive at an utterance's literal meaning and suggests that additional processes are involved (cf. Rips 1975; Noveck and Posada 2003; but see Feeney et al. 2004). In another study, adults reliably took more time to read sentences in which the generation of the SI was warranted by the context compared to sentences where no SI was warranted (Breheny et al. 2006). Further support for contextualist accounts comes from evidence showing that the amount of cognitive load from a concurrent task affects SI generation: when asked to judge underinformative sentences of the form "Some oaks are trees," adults who were simultaneously engaged in a relatively easy secondary task were more likely to accept the sentences as true compared to other adults who were engaged in a relatively more demanding secondary task (de Neys and Schaeken 2007). Recently, studies of the time-course of quantifier processing using a visual world paradigm have also concluded that the generation of SI-strengthened interpretations of quantifiers involves additional time compared to literal interpretations in both adults and children (Huang and Snedeker 2009a, 2009b; but cf. Grodner et al. 2010 and the end of Section 3.1 above).

## 4. The interface between semantics and implicature

As mentioned in the Introduction, a wide variety of expressions in natural language seem to form entailment-based scales. According to traditional accounts of SIs (Horn 1972; Grice 1975), all such scalar expressions should follow the same logic and should give rise to SIs in a similar way. A striking finding from past developmental studies is that not all scalar items are born equal: specifically, numerals seem to behave differently from quantifiers. This finding has important implications for semantic and pragmatic theories of scalars, so it is worth examining in some detail.

The main observation comes from a study by Papafragou and Musolino (2003) already discussed in Section 3.1. In that study, 5-year-olds were asked to say whether they agreed or not with a puppet's description of the outcome of a story. The study found that, in a story in which a set of three horses jumped over a fence, children were much more likely to reject the statement "Two of the horses jumped over the fence" than the statement "Some of the horses jumped over the fence". Further studies showed that the tendency to interpret *some* as lower-bounded ("at least some") but a number such as *two* as exact emerges early during language acquisition: in a sentence-to-picture matching task, 3-year-olds were more likely to select a picture in which an alligator had taken all four in a set of cookies after hearing the sentence "The alligator took some of the cookies" than the sentence "The alligator took two of the cookies" (Hurewitz et al. 2006; see also Barner et al. 2009a; Huang et al., 2013). Moreover, as mentioned earlier, online studies of how children and adults process numbers versus quantifiers reveal that SI calculation begins to influence referential processing after a delay in the case of quantifiers but impacts the comprehension process more rapidly for numbers (Huang and Snedeker 2009a, 2009b). Despite this preference for exact number meanings, both children and adults display flexibility in their interpretation of number terms adopting, for instance, “at least” interpretations for numerals if the context warrants them (Musolino 2004). Finally, numerals are not the only scalar expressions that behave differently from weak scalars such as *some*: the proportional quantifier *half* also seems to preferentially give rise to exact interpretations. For instance, in a story in which a puppet finished building a tower, Greek-speaking children were much more likely to reject the statement “The puppet built half of the tower” than the statement “The puppet began to build the tower” (Papafragou 2006; cf. also Papafragou and Schwarz 2004, for a similar comparison between *half* and *most*).

How are these differences between numerals and quantifiers to be explained? One interpretation of these findings is that numbers (and *half*) have exact semantics, unlike other weak scalar terms such as the quantifier *some* that have a lower-bounded, “at least” semantics and are pragmatically upper-bounded by an SI (Papafragou and Musolino 2003; Musolino 2004; Hurewitz et al. 2006; Papafragou 2006). According to this view, the meaning of *two* is plainly TWO and, depending on a contextual parameter, can yield “exact” or “at least” (and sometimes “at most”) interpretations (see Breheny, 2008, for a full semantic proposal). The “exact” view of numbers is compatible with evidence that numbers behave differently from regular, lower-bounded scalars such as quantifiers on a number of linguistic tests (Sadock 1984; Koenig 1991; Horn 1992; Carston 1995, 1998; Breheny 2008). For example, the “at least” and the “exact” interpretations of numbers intuitively belong to the truth-conditional content of numerally modified statements. In the dialogs below (based on Horn, 2004), the cardinal *two* allows a disconfirmatory response in case the “at least two, possibly more” reading applies. Contrast this with *some*, where the responder may cancel the SI using an affirmative response:

(5) A: Does she have two children?
B1: No, she has three.
B2: ?Yes, (in fact) three.

(6) A: Are some of your friends vegetarians?

B1: ?No, all of them.

B2: Yes, (in fact) all of them.

The semantic reanalysis of numerals is consistent with proposals within developmental psychology according to which number words are mapped onto a dedicated magnitude system with exact semantics (a system which represents exact and unique numerosities; Gelman and Cordes 2001; Gelman and Gallistel 1978). On this picture, the processes underlying the acquisition of number words are distinct from the mechanisms responsible for learning and evaluating (non-numerical) quantified expressions. Numbers and quantifiers obey different principles: the numerical scale < . . . , *four*, *three*, *two*, *one* >, unlike the quantificational scale < *all*, *most*, *many*, . . . *some* >, has an ordering rule that strictly and completely determines the internal structure of the scale and the positioning of its members. Specifically, the numerical scale is based on the successor function which yields the next member of the scale through the rule "$n + 1$." Children seem to recognize the distinct logic underlying numbers and quantifiers: 2- and 3-year-olds know that numbers such as *six* but not quantifiers such as *a lot* apply to unique specific quantities, even if they do not know any number word beyond *one* (Sarnecka and Gelman 2004). Furthermore, children begin to explicitly memorize the number list before they acquire number meanings, and even 2-year-olds are able to recite at least part of the list (e.g. "one, two, three"; Fuson 1988). The count list is essential for mapping number words onto number meanings. No comparable list exists for quantifiers or other scalars (Hurewitz et al. 2006).

A different interpretation of the quantifier–number asymmetry was recently put forth by Barner and Bachrach (2010). On this proposal, numbers receive a unified semantic analysis with quantifiers and other scalars, that is, they have lower-bounded ("at least") semantics and are upper-bounded by an SI (see also Levinson 2000). The reason young children successfully compute upper-bounded quantity interpretations from the use of numbers such as *two* but not from the use of quantifiers such as *some* is simply that lexical alternatives for numbers are easier to retrieve compared to quantifiers: as already pointed out, numbers belong to an explicitly memorized list of scalar alternatives—the count line—which is available to children even before individual number meanings are learned. More precisely, the count line provides easily accessible scalemates which serve to constrain the interpretation of novel numerals (i.e. numerals for which children have not acquired adultlike semantics). On this account, 2-year-olds initially assign a weak, "at least" interpretation to *one*. Once children acquire a weak ("at least") interpretation for *two*, they use the contrast to *two* to set an upper boundary to their meaning of *one* (i.e. "no more than one"). By this account, assigning *one* an exact interpretation depends on first acquiring *two*, and deriving "exactly one" via pragmatic inference (similarly for assigning *two* an exact interpretation). Thus it is argued that children as young as 2 are able to compute certain types of SIs, since they rely on pragmatic inference to derive exactness in number meanings (but not upper-boundedness in quantifier meanings). One piece of evidence consistent with this account is that very young children use known numerals to constrain the interpretation of unknown numerals: when 2-year-olds who know the meaning of the word *one* (but no higher number) are shown two sets—for example, one balloon and five balloons—they infer that the word *five* refers to the set of five objects, despite not knowing its meaning (Wynn 1992; Condry and Spelke 2008). Crucially, 2-year-olds do not use known numerals to restrict known quantifiers such as *some* (Condry and Spelke 2008) or novel quantifiers such as *toma* (Wynn 1992).[7]

At this point, the number data await a final synthesis. We simply note that a semantic account that treats numbers as lower-bounded expressions on a par with weak quantifiers leaves open several questions. First, this account crucially relies on the assumption that pragmatic inferences used to fix lexical meaning for unknown numerals are truly comparable to SIs—an assumption which requires further empirical scrutiny (see discussion of mutual exclusivity/contrast in section 3.2). Second, it is unclear how a unified semantic account for numbers and quantifiers can explain the fact that the two types of expression seem to make distinct truth-conditional contributions (see (5)–(6)

---

[7] In Wynn's (1992) account, children consider known (exact) number labels in trying to figure out the meaning of novel number words. This proposal is consistent with the linguistic literature on SIs that generally take alternatives invoked for implicature-calculation to be other lexical items the speaker could have used. But in Barner and Bachrach's account (see also Sauerland 2004; Spector 2007), scalar alternatives for numbers are not known number words but *pragmatically enriched interpretations* of known number words. That is, children use implicature-enriched "exact" interpretations of known numbers such as *one* contrastively in trying to constrain the meaning of unknown number words such as *two*. In this sense, implicature-calculation for numbers remains different from implicature-calculation for quantifiers and other scalars.

above). Finally, an explanation of children's early success with exact meanings for numbers needs to extend to proportional quantifiers such as *half*, which nevertheless do not form part of the canonical number line or the standardly rehearsed counting routine. It is plausible that children associate *half* with its stronger scalemate *all* by hearing adults explicitly contrast the two terms ("Eat all your vegetables, not half of them"). Whether such contrasts are more frequent and potent compared, for example, to contrasts between *some* and *all*, remains to be determined.

## 5. Conclusions

Children's ability to compute scalar implicatures is currently a very active area of research, where developmental data and semantic-pragmatic theorizing mutually inform and constrain each other. The emerging consensus from this literature is that children are capable of deriving pragmatic inferences and that the mechanism for deriving such inferences is broadly similar to the mechanism underlying utterance interpretation in adults. Nevertheless, children's computation of SIs is limited in crucial respects, such as the ability to construct relevant alternatives from the use of a weak scalar expression, which seems to undergo development until well into the school years. An important direction for future work is to integrate the detailed findings on SIs with other findings in developmental pragmatics, including relevance implicatures (Verbuk 2006b; de Villiers et al. 2009; Shulze et al. 2010), pragmatic enrichment (Noveck et al. 2009), reference assignment (Nadig and Sedivy 2002), and various forms of non-literal speech (Vosniadou 1987; Bernicot et al. 2007), on which children's performance is similarly mixed. Future work also needs to relate results on SI-calculation in older children to findings from early word learning that indicate—somewhat paradoxically—precocious abilities to interpret communicative intentions in toddlers (Tomasello 1992; Baldwin 1993b; Bloom 2000; cf. also section 3.2). A promising hypothesis in integrating this entire body of work is that non-linguistic cues such as eye gaze can be used early and accurately to make basic inferences about the speaker's referential intent in uttering a novel word; nevertheless, the ability to integrate linguistic and non-linguistic information to go beyond the words uttered and infer additional meanings the speaker had in mind (as is the case for SIs and other complex cases of inference) takes time and protracted linguistic experience to develop.

The present data on children's implicature calculation place important constraints on theories of semantic development. Most approaches to word learning to date have not addressed the fact that the interpretation of words is context-dependent or that words in use give rise to conversational inferences. Furthermore, the literature on semantic development sometimes implicitly assumes that pragmatic ("contextually enriched") interpretations are acquired at a later stage than "pure" semantic meaning—hence it is possible to study the acquisition of semantics independently from pragmatic development. The results on SI reviewed in this chapter show that this approach is unproductive. First, as shown here, pragmatic interpretations are well within the grasp of young children. Second, semantic content is not transparent but is bundled with pragmatic inference in the input to children: to infer the correct ("at least") meaning for scalar expressions, learners need to process adult uses of scalar items that may carry either lower-bounded or upper-bounded interpretations. Third, and most importantly, a pragmatically-informed approach captures important generalizations about how natural-language meaning is structured and learned. In the case of scalars, a single, powerful mechanism is responsible for delivering SIs from a wide variety of expressions (e.g. *some*, *or*, *may*) that all submit to a unified semantic-pragmatic treatment (lower-bounded semantics with upper-bounding implicatures). Furthermore, this mechanism applies universally: as the previous sections showed, implicatures from weak scalars arise cross-linguistically in much the same way (e.g. English *some*, Greek *meriki*, and French *quelques* all give rise to "not all" inferences). Thus the proper division of labor between semantics and SIs has wide explanatory potential, since it provides a unified account of what it is learners need to know about scalar expressions in any language. Together these observations suggest that theories of the acquisition of word meaning should take into serious consideration the pragmatic inferences which words give rise to in context. In certain cases (such as number words), the lexical semantics and pragmatics of child language can even provide evidence for the contents of the adult lexicon.

## Acknowledgments

Preparation of this chapter was partly supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NIH) grant 5R01HD55498-3 to A.P. We thank Ira Noveck for useful comments.


## References


Ackerman, B. (1981). When is a question not answered? The understanding of young children of utterances violating or conforming to the rules of conversational sequencing. *Journal of Experimental Child Psychology,* 31, 487-507.

Au, T., K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development,* 61, 1474–1490.

Baldwin, D., A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20, 395–418.

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

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition,* 188, 87-96.

Barner, D., Chow, K., & Yang, S., J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, 58, 195–219.

Bernicot, J., Laval, V. & Chaminaud, S. (2007). Nonliteral language forms in children: In what order are they acquired in Pragmatics and Metapragmatics? *Journal of Pragmatics*, 39, 2115–2132.

Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

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

Braine, M., & Rumain, B. (1981). Children's 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 numerally quantified Noun Phrases. *Journal of Semantics,* 25, 93-139.

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, 434–463.

Byrnes, J. P., & Duff, M.A. (1989). Young children's comprehension of modal expressions. *Cognitive Development,* 4, 369-387.

Carey, S. (1978). The child as word learner. In J. Bresnan, G. Miller and M. Halle (eds.*), Linguistic Theory and Psychological Reality*, 264-293. Cambridge, MA: MIT Press.

Carston, R. (1995). Quantity maxims and generalized implicature. *Lingua,* 96,213-244.

Carston, R. (1998). Informativeness, relevance, and scalar implicatures. In R. Carston and S. Uchida (eds.), *Relevance theory: Applications and implications*. Amsterdam: Benjamins.

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

Chierchia, G., Fox, D., & Spector, B. (2009). Hurford's Constraint and the Theory of Scalar Implicatures. In: P. Egré & G. Magri. *Presuppositions and Implicatures. Proceedings of the MIT-Paris Workshop*. Cambridge, MA: *MIT Working Papers in Linguistics*, 47–62.

Chierchia, G., Crain, S., Guasti, M., T., Gualmini, A., & Meroni L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. *Proceedings from the Annual Boston University Conference on Language Development,* 25, 157-168.

Clark, E., V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*, 1–33. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, E., V. (1988). On the logic of contrast. *Journal of Child Language,* 15, 317–335.

Clark, E., V. (1990). On the pragmatics of contrast. *Journal of Child Language,* 17, 417–431.

Condry, K., & Spelke, E., S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, 137, 22–38.

Crain, S., & McKee, C. (1985). The acquisition of structural restrictions on anaphora. *Proceedings of NELS,* 15, 94-110.

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments in the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press.

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

DeVilliers, P., A., DeVilliers, J., G., Coles-White, D., & Carpenter, L., A. (2009). Acquisition of Relevance Implicatures in Typically-Developing Children and Children with Autism. In J. Chandlee, M. Franchini, S. Lord & G. Rheiner (eds.), *Proceedings of the 33rd Annual Boston University Conference on Language Development,* 1, 121-132. Boston, MA: Cascadilla Press.

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psycholog*y, 41, 451–463.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology,* 37, 630–641.

Feeney, A., Scrafton, S., Duckworth, A., & Handley, S., J. (2004). The story of "some": everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology,* 58, 121–132.

Fox, D. (2007). Free Choice Disjunction and the Theory of Scalar Implicatures. In: U. Sauerland and P. Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*. New York: Palgrave Macmillan.

Fuson, K., C. (1988). *Children's counting and concepts of number*. New York: Springer.

Gathercole, V. (1989). Contrast: A semantic constraint? *Journal of Child Language,* 16, 685–702.

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

Gelman, R., & Cordes, S. (2001). Counting in animals and humans. In E. Dupoux (ed.), *Language, Brain, and Cognitive Development:* Essays in Honor of Jacques Mehler, 279-303. Cambridge, MA: MIT Press.

Gelman, R., & Gallistel, C., R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Geurts, B. (1998). Scalars. In P. Ludewig and B. Geurts (eds.), *Lexicalische Semantik aus kognitiver Sicht*, 95-117. Tübingen: Narr.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

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

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M., T. (2001). At the semantics/pragmatics interface in child language. *Proceedings of Semantics and Linguistic theory XI*. Ithaca, NY: CLC Publications, Department of Linguistics, Cornell University.

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

Grodner, D., Klein, N., Carbary, K., & Tanenhaus, M. (2010). "Some", and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment, *Cognition*, 116, 42-55.

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

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, 53, 310-344.

Harnish, R. (1976). Logical form and implicature. In T., G. Bever, J., J. Katz & T. Langdon (eds.), *An Integrated Theory of Linguistic Ability*, 313–92. New York: Thomas Y. Crowell. Reprinted in S. Davis, ed. (1991), *Pragmatics: A Reader*, 316–364. Oxford: Oxford University Press.

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

Hirst, W., & Weil, J. (1982). Acquisition of the epistemic and deontic meaning of modals. *Journal of Child Language,* 9, 659-666.

Horn, L., R. (1972). *On the semantic properties of the logical operators in English*. Doctoral diss., UCLA.

Horn, L., R. (1984). Toward a New Taxonomy for Pragmatic Inference: Q-based and R-based Implicature. In D. Schiffrin (Ed.), *Georgetown University Round Table on Languages and Linguistics 1984*, 11–42. Washington, D.C.: Georgetown University Press.

Horn, L., R. (1989). *A natural history of negation*. Chicago: University of Chicago Press. Repr. in 2001. Stanford: CSLI Publications.

Horn, L., R. (1992). The said and the unsaid. *SALT* II, 163-192, Dept. of Linguistics, Ohio State University.

Horn, L., R. (2004). Implicature. In L., R. Horn & G. Ward (eds.), *The Handbook of Pragmatics*, 3–28. Oxford: Blackwell.

Horn, L., R. (2005). The border wars: A neo-Gricean perspective. In K. Turner & K. von Heusinger (eds.), *Where semantics meets pragmatics*, 21–48. Oxford: Elsevier.

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

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

Yi Ting Huang , Elizabeth Spelke & Jesse Snedeker (2013): What Exactly do Numbers Mean?, *Language Learning and Development, 9:2*, 105-129Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development,* 2, 77–96.

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition, 120,* 67-81.

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

Koenig, J., P. (1991). Scalar predicates and negation: Punctual semantics and interval interpretations. *CLS,* 27, 140-155.

Koenig, J., P. (1993). Scalar Predicates and Negation: Punctual Semantics and Interval Interpretations. *Chicago Linguistic Society* 27, part 2: The Parasession on Negation, 140–155.

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

Markman, E., M. (1989). Categorization and naming in children. Cambridge, MA: MIT Press.

Matsumoto, Y. (1995). The Conversational Condition on Horn scales. *Linguistics and Philosophy,* 18, 21-60.

Moore, C., Bryant, D. & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development,* 60, 167-171.

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

Nadig, A., & Sedivy, J. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, 13, 329-336.

Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition,* 78, 165-188.

Noveck, I., Chevallier, C., Chevaux, F., Musolino, J., & Bott, L. (2009). Children's enrichments of conjunctive sentences in context. In P. De Brabanter, & M. Kissine (eds.), *Current Research in Semantics/Pragmatics* (Vol. 20), 211-234. Bingley, UK: Emerald Group.

Noveck, I. A., Ho, S. & Sera, M. (1996). Children's understanding of epistemic modals. *Journal of Child Language,* 23, 621-643.

Noveck, I., A., & Posada, A., (2003). Characterizing the time course of an implicature. *Brain and Language,* 85, 203–210.

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

Ozturk, O., & Papafragou, A. (2014). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development* online: 1-24.Papafragou, A. (2006). From scalar semantics to implicature: Children's interpretation of aspectuals. *Journal of Child Language*, 33, 721-757.

Papafragou, A., & Schwarz, N. (2006). *Most* wanted. *Language Acquisition* 13 (Special Issue: On the Acquisition of Quantification), 207-251.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition,* 86, 253-282.

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition,* 12, 71-82.

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

Pouscoulous, N., Noveck, I., Politzer, G., Bastide, A. (2007). Processing costs and implicature development. *Language Acquisition,* 14, 347-375.

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

Sadock, J. (1984). Whither radical pragmatics? In D. Shiffrin (ed.), *Georgetown University Roundtable on Language and Linguistics,* 139-149. Washington, DC.: Georgetown University Press.

Sarnecka, B., W., & Gelman, S., A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition,* 92, 329–352.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27, 367–91.

Shatz, M. (1983). Communication. In P. H. Mussen (Series Ed.) & J. Flavell & E. Markman (Vol. eds.), *Handbook of Child Psychology: Vol. 3. Cognitive development* (4th ed.), 841-889. New York: Wiley.

Shulze, C., Grassman, S., & Tomasello, M. (2010). Relevance inferences in 3-year-olds. *Proceedings from the 35th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Skordos, D., & Papafragou, A. (to appear) Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*

Smith, C., L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, 30, 191–205.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland & P. Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 71–120. New York: Palgrave Macmillan.

Sperber, D., & Wilson, D., (1986/1995). *Relevance: Communication and Cognition*, 2nd ed. 1995. Cambridge, MA: Harvard University Press.

Tomasello, M. (1992). *First verbs: A case study in early grammatical development*. Cambridge, MA: Cambridge University Press.

Trueswell, J., C., Sekerina, I., Hill, N., M., & Logrip, M.,L. (1999). The kindergarten path effect: studying on-line sentence processing in young children. *Cognition,* 73, 89-134.

Verbuk, A. (2006). Acquisition of scalar implicatures: When some of the crayons will do the job. *Online Proceedings from the 31st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Vosniadou, S. (1987). Contextual and linguistic factors in children's comprehension of nonliteral language. *Metaphor and Symbolic Activity,* 2, 1-11.

Wynn, K. (1992). Children's acquisition of number words and the counting system. *Cognitive Psychology,* 24, 220–251.