Cognition 118 (2011) 87–96

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference

David Barner [a,*], Neon Brooks [b], Alan Bale [c]

[a] Department of Psychology, University of California, San Diego, CA, USA
[b] Department of Psychology, University of Chicago, Chicago, IL, USA
[c] Department of Linguistics, Concordia University, Montreal, QC, Canada

ARTICLE INFO

Article history:
Received 8 February 2010
Revised 14 October 2010
Accepted 15 October 2010

Keywords:
Pragmatic development
Inference
Scalar implicature
Word learning
Counting
Exactness
Language acquisition
Semantic development

ABSTRACT

When faced with a sentence like, "*Some of the toys are on the table*", adults, but not preschoolers, compute a scalar implicature, taking the sentence to imply that not all the toys are on the table. This paper explores the hypothesis that children fail to compute scalar implicatures because they lack knowledge of relevant scalar alternatives to words like "some". Four-year-olds were shown pictures in which three out of three objects fit a description (e.g., three animals reading), and were asked to evaluate statements that relied on context-independent alternatives (e.g., knowing that *all* is an alternative to *some* for the utterance "*Some of the animals are reading*") or contextual alternatives (e.g., knowing that the set of all three visible animals is an alternative to a set of two for the utterance "*Only the cat and the dog are reading*"). Children failed to reject the false statements containing context-independent scales even when the word *only* was used (e.g., *only some*), but correctly rejected equivalent statements containing contextual alternatives (e.g., *only the cat and dog*). These results support the hypothesis that children's difficulties with scalar implicature are due to a failure to generate relevant alternatives for specific scales. Consequences for number word learning are also discussed.

1. Introduction

As children acquire language, their task is complicated by the fact that speakers' intended meanings go beyond the literal meanings of their utterances. Word learning is not simply a process of mapping words onto speaker intentions. Instead, children must infer the core lexical meanings of words by distinguishing what is logically entailed from that which is merely implied. For example, in a dialogue like (1), John is likely to infer that Mary did not eat all of his cake.

(1) John: Did you eat my cake?

Mary: I ate some of it.

Although Mary's statement would be literally true if she had in fact eaten the whole cake (eating *all* entails eating *some*), her utterance nonetheless implies that she did not. This inference relies on the assumption that, if Mary had eaten the whole cake and was communicating cooperatively, she would have uttered a more informative statement like "I ate *all* of it" (Grice, 1978, 1989). Thus, although Mary's utterance does not logically rule out the possibility that she ate all of John's cake, this is the intended meaning ascribed to her nonetheless.

The language acquisition literature is filled with examples of children learning words by making inferences about speaker intentions. A classic demonstration of this comes from experiments investigating mutual exclusivity. When

* Corresponding author. Address: Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, USA. Tel.: +1 858 450 0302; fax: +1 858 734 7190.
E-mail address: barner@ucsd.edu (D. Barner).

0010-0277/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2010.10.010

a child is shown two objects, one of which has a known label (e.g., a car), they infer without difficulty that a novel label (e.g., *dax*) refers to the previously unlabeled object. Such an inference follows from the assumption that the speaker would not use two words to denote the same kind of object (i.e., words exhibit *mutual exclusivity*, or *contrast*; Clark 1987, 1988; Markman, 1989). Children apply such strategies not only when learning nouns, but also when interpreting other classes of words, such as numerals. For example, when 2-year-olds who know the meaning of the word *one* (but no higher number) are shown two sets – e.g., one balloon and five balloons – they infer that the word *five* refers to the set of five objects, despite not knowing its meaning (Condry & Spelke, 2008; Wynn, 1992).

Amidst such evidence, and further studies which find that children are sensitive to subtle intentional cues like eyegaze, speaker desires, etc. (Baldwin, 1993; Bloom, 2000; Diesendruck & Markson, 2001 Rapacholi & Gopnik, 1997; Tomasello, 1992, etc.) children also exhibit striking failures in computing some simple inferences, including the inference in (1), which is a type of *scalar implicature*. Following Horn (1972, 1989), it is typically assumed that the quantifier *some* belongs to a larger class of terms called "scalar items". Scales are used to generate sets of alternative meanings, which are ordered according to their informativeness and are implicitly contrasted during interpretation. In the case of *some*, the relevant scale includes other quantifiers – e.g., *a, many, most, all*. Examples of such scales are shown in (2):

(2a) ⟨a, some, many, most, all, etc.⟩
(2b) ⟨warm, hot, boiling, etc.⟩
(2c) ⟨one, two, three, etc.⟩

By most accounts, deriving a scalar implicature involves at least four steps, summarized in I–IV. First, the listener computes the basic, *literal*, meaning of the expression (Step I). Second, they generate the set of alternative sentences that might have been uttered (by substitution of scalar alternatives; Step II). Third, they restrict these alternatives by removing those that are less informative (Step III). Finally, they "strengthen" the interpretation of the sentence by negating the remaining alternatives – e.g., "I ate some (but not all) of the cake" (Step IV).

I. Compute basic meaning of a sentence S containing L, a scalar item.
II. Generate a set of alternatives ($a_1, a_2, \ldots, a_n$) to S, called S_{alt}. These are all the sentences that can be generated by replacing L with its scalar alternatives.
III. Restrict the alternatives in S_{alt} by removing any alternative that is entailed by the original utterance S. Call this restricted set S^*.
IV. Strengthen the basic meaning of S (containing L) with the negation of all of the members of S^*.

A large number of studies starting in the early 1980s have found that children, unlike adults, often fail to derive scalar implicatures. This has been shown for many scalar contrasts, including *might* vs. *must* (Noveck, 2001), *a* vs. *some* (Barner, Chow, & Yang, 2009), *some* vs. *all* (Huang & Snedeker, 2009a, 2009b; Hurewitz, Papafragou, Gleitman, & Gelman, 2006; Musolino, 2004; Noveck, 2001; Papafragou & Musolino, 2003; Smith, 1980), and *or* vs. *and* (Chierchia, Crain, Guasti, Gualmini, & Meroni, 2001; Gualmini, Crain, Meroni, Chierchia, & Guasti, 2001). For example, in a study by Papafragou and Musolino (2003), 5-year-old children were shown a scene with three horses, in which all three animals successfully jumped over a log. When children were asked whether the sentence, "Some of the horses jumped over the log" was a good description of the event, most said yes. Adults, in contrast, denied that this was a good description, since all of the horses jumped over the log. Adults, unlike the 5-year-old children, computed a scalar implicature. Interestingly, children do not always lack a so-called *strengthened* meaning. Papafragou and Musolino found that children provided adult-like responses when tested with numerals. Children denied that "Two of the horses jumped over the log" when three horses did. Thus, although children failed to have adult-like response with *some* and *all*, they interpret numerals with an *exact*-meaning just like the adult controls.

What accounts for children's frequent failure to compute implicatures, and why do they fail with the quantifiers *some* and *all*, but not with numerals? Previous studies have suggested various factors that might affect children's derivation of implicatures, including limitations on working memory, limited understanding of context and metalinguistic tasks, and the salience or availability of relevant scalar alternatives (see Chierchia et al., 2001; Musolino, 2006; Musolino & Lidz, 2006; Papafragou, 2006; Papafragou & Musolino, 2003; Papafragou & Tantalou, 2004; Pouscoulous, Noveck, Politzer, & Bastide, 2007). According to Papafragou and Musolino, since each of these factors might limit children's computation of implicatures, and since children readily assign exact interpretations to numerals, children must not be using implicatures to derive exact meanings of numerals. Instead, by their view, the difference between quantifiers and numerals is due to the fact that numerals, unlike quantifiers, have lexically strengthened, exact, meanings (see also Huang, Snedeker, & Spelke, submitted for publication; Hurewitz et al., 2006; see Breheny, 2008, for an example of an exact semantics).

It is uncontroversial that context affects when and whether children (or adults, for that matter) will compute implicatures (e.g., Guasti et al., 2007; Musolino, 2006; Papafragou & Musolino, 2003). It is also known that working memory capacity grows over the course of development (e.g., Gathercole & Baddeley, 1990). Finally, it is known that adults take longer to compute implicatures than to arrive at an utterance's literal meaning, suggesting that additional processes are involved (Rips, 1975; Noveck & Posada, 2003; Huang & Snedeker, 2009a, 2009b). Nevertheless, the role of these factors in children's pragmatic difficulties has not been empirically established. First, although previous studies find that implicatures are more likely in some contexts than others, the fact that strong contextual cues can push children towards one interpretation over another does not show that their difficulties are due to contextual misunderstanding. Instead, strong contextual cues may compensate or mask difficulties that originate elsewhere in the process of deriving implicatures.[1]

Second, there is currently no direct evidence that working memory or processing constraints are responsible for children's difficulty computing implicatures. Studies that document additional processing time for implicatures do not suggest that such additional processing is taxing, nor that it is sufficient to cause difficulty for children. Instead, the claim is only that implicatures are not immediate (and therefore, by these accounts, not part of the literal semantics). Studies that attribute children's problems to processing limits (e.g., Chierchia et al., 2001; Pouscoulous et al., 2007; Reinhart, 2004) do not actually assess working memory or processing capacity, and thus do not test whether individual differences in capacity predict differences in pragmatic abilities.[2] Also, these studies do not provide a model of how processing limits could generate the effects found – i.e., which specific processes are affected and how.

For example, in their study of implicature and disjunction, Chierchia et al. (2001) tested 3- to 6-year-old children's interpretation of *or*, and whether they could distinguish its weak, inclusive, interpretation from its strong, exclusive one. Unlike adults, when children of this age were told "Every boy chose a skateboard or a bike," they accepted situations in which each boy chose both objects. Thus, they accepted the weak inclusive interpretation of *or*, when adults did not. However, when children were explicitly presented a sentence containing *and* as an alternative, they strongly preferred it over a sentence containing *or*. According to Chierchia et al. (2001), children's failure to compute implicatures under normal circumstances is "due to the processing cost involved in constructing alternative representations of the sentence they were presented with" (p. 167).

This study clearly shows that when children are presented with two alternative descriptions of a scene they prefer stronger, more informative descriptions. Thus, it suggests that if children were able to spontaneously access alternatives and compare them, then they would be able to correctly judge that *and* is stronger than *or*. However, the study does not provide evidence that processing limits are the source of children's difficulty computing implicatures. First, as discussed further below, judging which of two statements is stronger is not equivalent to computing a scalar implicature. Not only are different processes involved, but it is not clear that children actually represented the alternatives provided as scale mates. The problem is that, even for adults, many statements that differ in strength do not involve members of common scales. For example, although adults might judge that *two cakes* is more informative than *some cakes* in a particular context, we would not want to conclude that they were therefore computing an implicature; adults would never infer that *eating some cakes* implies *not eating two cakes*. Second, and related to this first point, processing limits are not the only way to explain why access to alternatives might be limited. As we will argue, the Chierchia et al. result, like others in the literature, is consistent with the hypothesis that children lack knowledge of which lexical items belong on common scales, and thus which words are relevant alternatives in a given context (see Papafragou & Tantalou, 2004). By this view, children may know the meanings of scalar terms and their relative strength as descriptions, but lack knowledge of which items constitute scales, and therefore which words to access as alternatives when processing sentences.

According to this hypothesis, knowledge of scales goes beyond knowledge of lexical meanings and a capacity to identify which of two descriptions is more informative. Children must learn additional restrictions to scales, such that not all possible alternatives are considered when computing implicatures, and such that alternatives can be rapidly and automatically accessed. This hypothesis is motivated not only be developmental data, but also by previous observations that restrictions to scales are necessary (Hirschberg, 1985; Horn, 1972, 1989; Katzir, 2007; Matsumoto, 1985). First, without restrictions beyond informativeness, language users might consider an unbounded number of stronger alternatives to a given utterance, resulting in a problem of computational complexity (i.e., the inference may not be computable because too many alternatives must be considered). Second, without restrictions many spurious and unattested inferences would result. For example, as noted above, although *two apples* is a stronger and more informative description than *some apples* in a context containing two apples, the words *some* and *two* are clearly *not* scale mates. If they were, then we would expect *two* to restrict and strengthen the interpretation of *some* – e.g., for the expression *John ate some apples* to rule out *John ate two apples* (and for that matter, any stronger alternative that contains a numeral, resulting in a situation where *some* can only mean *exactly one*). However, even young children avoid this error, and assign numerals to a unique scale before they know their specific meanings. Although 2-year-olds use known numeral meanings to restrict their interpretation of unknown numerals, they do not use them to restrict known quantifiers like *some* (Condry & Spelke, 2008) or novel quantifiers like *toma* (Wynn, 1992). Thus, at least in the case of numerals, children identify scale members even before meanings have been fully acquired.

A failure to represent lexical items as members of psychological scales could explain numerous results in the literature, and also the apparent discrepancy between children's difficulty with implicatures and their relatively sophisticated use of pragmatic cues elsewhere in language

[1] As shown by Musolino, children readily assign numerals a lower-bounded "at least" interpretation in the context of familiar games (in which the winner must get some minimum number of points). These studies ask children to judge whether a Troll is the winner. However, it is unknown whether these same children would agree that a Troll who scored 5 baskets also scored 3. Children may know how to determine a winner in such contexts without actually thinking that *three* can be assigned a weak interpretation of "at least three".

[2] The evidence from Pouscoulous et al. (2007) for a processing cost comes from the observation that some tasks (i.e., creating a set) are more likely to elicit adult behavior than others (i.e., truth value judgment), and that some lexical items are used in more complex syntactic constructions (like the partitive). However, the studies fail to test whether the tasks or words used actually differ in their demands on processing, or how such a processing limit would explain the findings (i.e., which processes were limited, and how).

acquisition. For example, this account, as noted by Barner and Bachrach (2010), could explain children's ability to assign exact interpretations to numerals, which belong to an explicitly memorized list of scalar alternatives – the count list. As already noted, researchers who claim that children cannot compute implicatures are forced to assign numerals a qualitatively different analysis from other quantifiers, since children readily access exact meanings of numerals but cannot strengthen quantifiers used in equivalent contexts. However, by the view that scales are constructed by associating scale mates to one another, numerals and quantifiers can be assigned a uniform analysis, unlike any previous account.

In their study, Barner and Bachrach (2010) argued that very young children – beginning at 2 years of age – routinely make inferences that are similar in structure to scalar implicatures when interpreting unknown numerals (i.e., numerals for which they have not yet acquired an adult-like meaning). As noted earlier, when a child who knows the meaning of *one* is shown two sets – e.g., one containing one balloon, and the other containing five – they systematically point to the larger set when asked to find *five balloons*. However, they do not do so when asked to find *blicket balloons*. According to Wynn (1992), "Since all the children knew that the word 'one' refers to a single item, then if they knew that, for example, the word 'five' refers to a numerosity, they should infer that it does not refer to a single item since they already have a word for the numerosity one" (p. 229).

This inference – that *five* refers to the larger set by virtue of *not* referring to *one* – requires all of the processing resources that an ordinary implicature would require, as well as several of the same steps. The child must generate a weak meaning for *five* (Step I), generate *one* as an alternative (Step II), and strengthen the interpretation of *five* by negating *one* (Step IV). The only missing component of implicature is that, typically, weaker items are strengthened by appeal to stronger ones (something that is impossible when interpreting unknown numerals, since stronger words have not yet been acquired). Still, once children acquire a meaning for *two*, they should be in a position to compute an implicature for *one*, meaning that even 2-year-olds could compute implicatures to derive exact meanings for numerals (for evidence that this is, indeed, what children do, see Barner & Bachrach, 2010). As far as we can see, there is no difference in the processing demands required by these well-documented inferences and the inferences required by scalar implicature. The only salient difference is the particular scale being processed – i.e., numerals rather than quantifiers.

Critically, studies of early number word acquisition find that children begin to explicitly memorize a count list before they learn any numeral meanings, and recite it like the alphabet (Fuson, 1988). In contrast, no child is taught to recite quantifiers in a song. Thus, when interpreting utterances containing *some*, children may be unaware that their knowledge of *all* is relevant, since *some* and *all* have not been associated as scalar alternatives, their semantic properties notwithstanding. For the same reason that children (or adults) do not treat *two* as an alternative to *some*, they also do not treat *all* as relevant. For them, these two expressions may differ in informativeness, but are simply not relevant alternatives when interpreting sentences.

Currently, there is little evidence outside the cases of numerals and quantifiers to assess the role of scalar knowledge in the development of children's pragmatic abilities. The present study widens the scope of this debate by contrasting children's interpretation of *some*, whose scale members are specified in a context-independent way (i.e., *all* is always a scale member with *some*), with their interpretation of words that have contextually specified alternatives (for discussion, see Hirschberg, 1985). Previous studies find that young children are able to strengthen utterances that rely on accessing contextual alternatives. For example, Goro, Minai, and Crain (2005, 2006) found that English-speaking 3- to 6-year-olds correctly rejected sentences like "Only Bunny Rabbit will eat a carrot or a pepper" in contexts where another character chose a pepper. Crain and colleagues concluded that such children must be computing alternative contrast sets in order to strengthen these sentences. Further, they concluded that "children have adult-like knowledge about the semantics of *only*, and are able to compute its complex semantic interaction with the interpretation of *or*." (Crain, Goro, & Minai, 2005) This is because, without the word *only*, such a sentence would be felicitous whether or not another character was eating a pepper, so long as the Bunny was.[3]

As Crain and others have noted, the semantics of *only* mirrors the algorithm for calculating scalar implicatures. This fact allows us to isolate the role that knowledge of relevant alternatives plays in implicature. As with implicatures, *only* triggers the negation of alternative sentences. For example, consider the sentence in (3).

(3) I ate only some of the cake.

This sentence indicates that the speaker did not eat all of the cake, much like Mary's statement in (1). The difference between the sentences in (1) and (3) is that in (3) the denial of the alternative "*I ate all of the cake*" is logically entailed by the sentence's core, literal meaning (it is not merely implied). Still, in order for this entailment relation to be realized, the listener must generate *all* as a relevant alternative in order to negate it (as in the case of implicature). Thus, although *only* makes the negation of alternatives obligatory and semantic, rather than optional and pragmatic, it still requires that the relevant alternatives be generated before they are negated. Therefore, evidence that children comprehend *only* but fail to strengthen sentences containing *only some* would suggest that their difficulty may be caused by a failure to access scalar alternatives, rather than a general inability to reason pragmatically.

In our study, we contrasted children's interpretation of sentences containing *some* to their interpretation of sen-

[3] Crain and colleagues also note the difficulties children have with *only* when multiple dimensions of contrast are possible – e.g., only the cat is holding a flag (when a duck is holding a flag and a balloon). However, as they show in the Bunny example, such errors do not arise when only one dimension of contrast is salient.

tences containing contextually specified alternatives.[4] We also examined the effect of grammatical strengthening by manipulating the use of *only*. For example, we asked questions like those in (4), for a situation where three animals (a dog, cat, and cow) are sleeping:

(4a) Are some of the animals sleeping?
(4b) Are only some of the animals sleeping?
(4c) Are the dog and the cat sleeping?
(4d) Are only the dog and the cat sleeping?

If children's difficulty computing implicatures for context-independent scales like ⟨*some, all*⟩ is due to a failure to generate relevant alternatives, then they should accept statements like (4a) and (4b) regardless of whether *only* is present. They should fail to construct the alternative sentence containing *all*, and therefore be unable to strengthen either sentence. In contrast, children should have no difficulty strengthening a sentence like (4d), since the alternative contrast set is contextually specified.[5] Evidence of this kind would suggest that knowledge of alternatives critically limits children's ability to interpret scalar items like *some*.

2. Methods

2.1. Participants

We recruited 60 4-year-old native speakers of English (27 girls; *M* = 4;6, range = 4;1–5;0) by phone, through daycares, or in children's museums in the greater San Diego area. Two additional children were excluded due to failure to complete the task.

2.2. Stimuli

Stimuli were 12 pictures, each depicting three items. Four cards were used during familiarization, and eight in the test phase. Familiarization cards depicted sets of animals with distinct characteristics, such as color or clothing. The test cards depicted four scenes (in 1–4).

a. Cookie Monster holding fruit (an orange, an apple, and a banana).
b. Animals sleeping (a dog, a cat, and a cow).
c. Animals reading (a dog, a cat, and a rabbit).
d. Toys on a table (a ball, a drum, and a train).

Two versions of each scene were created: one in which all three items shared a common property (e.g., Cookie Monster is holding all three fruit in his hands; all three animals are sleeping), and one in which two of the three items shared a common property (e.g., Cookie Monster is holding two fruits, and one is on the floor; two animals are sleeping and one is reading). An example is provided in Fig. 1.

2.3. Procedure

Children were seated with an experimenter at a child-sized table and were told that they would play a game with pictures. First, they were shown the four familiarization cards one at a time and asked to identify each animal on the card ("What's this? That's right, it's a cow!"). If the child labeled an animal incorrectly, they were told the correct response and encouraged to repeat it after the experimenter ("That's a cow, can you say 'cow?"). Children were then asked a yes/no question, for which there was only one right answer (e.g., "Is the cow wearing a hat?" when the *fish* is wearing a hat). This familiarization exercise was designed to accustom children to answering both 'yes' and 'no' to questions. If a child answered any of these questions incorrectly, the experimenter moved onto the next familiarization card, but returned to the problematic card after completing the remaining familiarization trials. If a child failed twice on any single familiarization trial, the experimenter ended the testing session.

At test, children were given nine trials using the test cards, presented in one of two counterbalanced orders. As in the familiarization phase, children were first asked to identify all of the items in the picture and then to evaluate the truth-value of a statement.

Children were randomly assigned to one of four test conditions in a 2×2 between-subjects design. The first factor, Alternative Type, determined whether children heard context-independent alternatives (i.e., *some* and *all*), or contextual alternatives (object labels, such as *dog* and *cat*) in the test questions. The second factor, grammatical strengthening, determined whether or not test questions contained the word *only*.

There were nine questions in each condition. On "2-Item False" questions, children were shown a picture in which only two of the three items fit a description, and were asked whether the description was true for all the items (e.g., "Are all of the animals sleeping?" or "Are the dog, the cat and the cow sleeping?"). These were used as control trials, assuring that 'no' would be the correct answer on at least 1/3 of trials, and were presented without only in all conditions. On "2-Item True" questions, children saw pictures where two of the three items fit a description, and were asked whether the description was true for that subset of items (e.g. "Are (only) some of the animals sleeping?" or "Are (only) the cat and the cow sleeping?"). Lastly, on "3-Item Test" questions, children were shown pictures in which all three items fit the description, and asked whether the description was true for a subset of the items (e.g. "Are (only) some of the animals sleeping?" or "Are (only) the cat and the cow sleeping?"). A complete list of questions each condition are listed in Appendix A.

[4] Note that the distinction between context-dependent and context-independent scales is not identical to the distinction between generalized vs. particularized implicatures. As shown here, both types of scale can be deployed in a situation that does not involve implicature at all, as is the case with sentences involving *only*, which entail rather than imply strengthening. As a result, we believe that the distinction is important whether or not a single pragmatic system is responsible for both generalized and particularized implicatures.

[5] Typically sentences like 6c are only strengthened if a preceding context is provided in which a presupposition exists that all three might be reading. For example, if asked, "Are the cat, the dog, and the cow reading?", a response of, "The cat and the dog are reading" would imply that the cow is not. This confirms that contextual alternatives are created only when required contextually, either by preceding sentences or the use of a focus element like *only*.

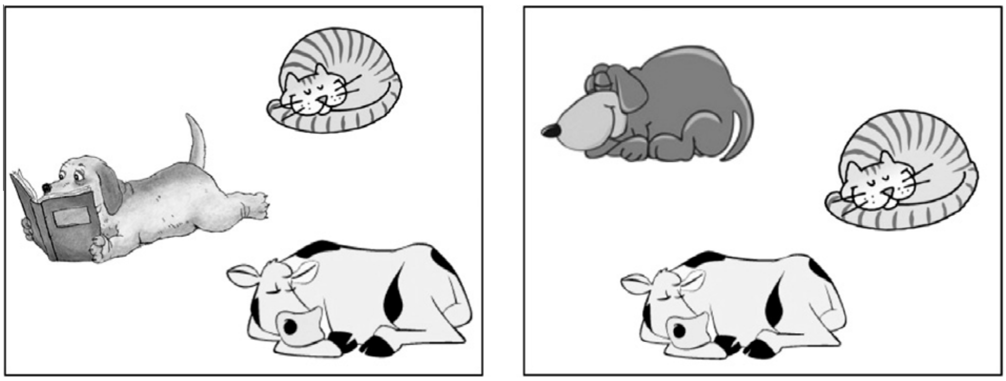

Fig. 1. Examples of test stimuli in which two out of three, and three out of three, animals are sleeping.

Neither the word *only* nor the quantifier was emphasized by the experimenter's prosody.

3. Results

The use of the word *only* had a significant effect on how children interpreted sentences involving contextual alternatives, but had no effect on their interpretation of sentences involving context-independent alternatives (*some* and *all*).

First, consider the data for context-independent alternatives (see Fig. 2). In contexts where two of three items fit a description (e.g., two out of three animals are sleeping), children correctly agreed to sentences like, "Are some of the animals sleeping?" (2-Item True trials) on 80.0% of trials, and correctly denied that all of the animals were sleeping (2-Item False trials) on 87.2% of trials. On 2-Item True trials, use of the word *only* had no significant effect on children's judgments ($t(28) = 0.00$, $p > 0.05$). *Only* was never used in 2-Item False trials, so it is not surprising that there was no difference between children's responses in the *only* and *no-only* conditions ($t(28) = 0.96$, $p > 0.05$). Critically, on 3-Item Test trials, children in the context-independent alternatives conditions behaved like children in previous studies of scalar implicature: when all three animals were sleeping and children were asked, "Are some of the animals sleeping?" they did not strengthen the utterance. Children in this condition accepted 3-Item Test trials on 66.6% of trials. Their responses did not differ significantly between 2-Item True trials and 3-Item Test trials ($t(14) = 1.0$, $p > .3$), suggesting that they were equally likely to agree that *some* animals were sleeping when all three of them were as when only two animals were sleeping. The insertion of *only* did not

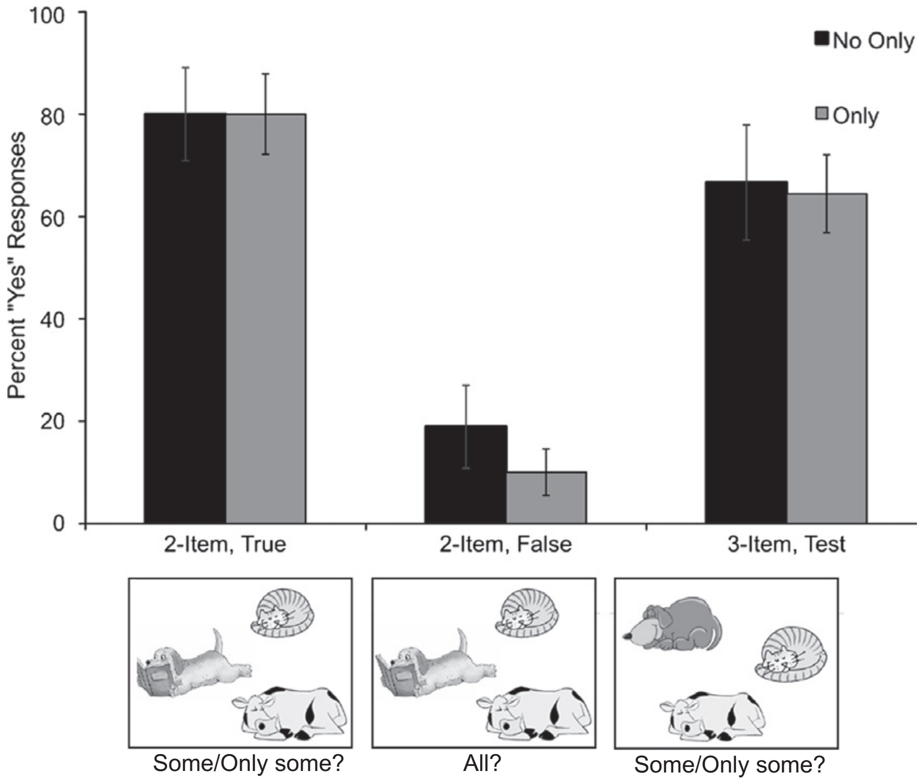

Fig. 2. Percentage of children who said "yes" to questions in the context-independent alternatives conditions. Error bars represent standard error.

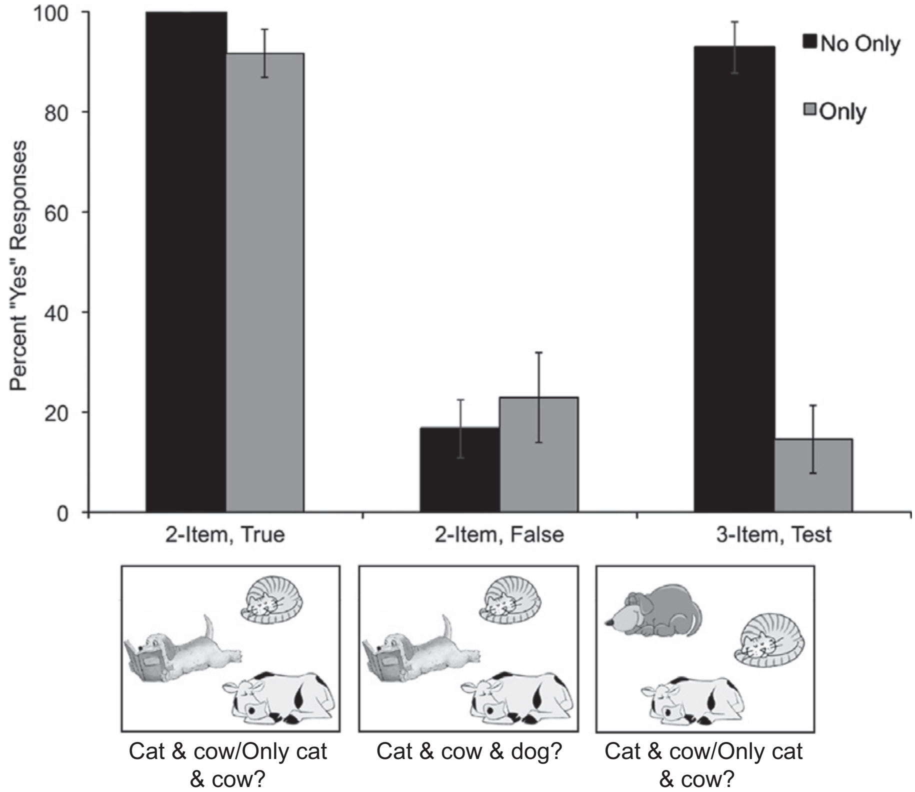

Fig. 3. Percentage of children who said "yes" to questions in the contextual alternatives conditions. Error bars represent standard error.

improve matters: there was no significant difference between strengthened and un-strengthened conditions on 3-Item Test trials ($t(28) = .16$, $p > .8$). For example, when three animals were sleeping, children were equally likely to say "yes" when asked, "Are only some of the animals sleeping" as when asked "Are some of the animals sleeping".

Next consider the data for contextual alternatives (Fig. 3). As in the context-independent condition, children answered control questions correctly: they agreed to 2-Item True trials 95.8% of the time, and correctly rejected 2-Item False trials 80.5% of the time. Use of the word *only* had no significant effect on children's judgments on 2-Item True trials ($t(28) = 1.62$, $p > .05$) nor on 2-Item False trials ($t(28) = 0.57$, $p > 0.05$). In contrast, on the critical 3-Item Test trials, children tested with contextual alternatives were highly sensitive to the presence of *only*. Children almost always said "yes" when asked, "Are the cat and the cow sleeping?" (92.9% of trials) but rarely said "yes" when *only* was added: "Are only the cat and the cow sleeping?" (14% of trials; $t(28) = 8.98$, $p < .001$). When *only* was added, there was also a significant difference between the number of "yes" responses in the 2-Item True and 3-Item Test conditions ($t(30) = 9.27$, $p < 0.001$).

To compare results across conditions, data were entered into a linear logistic regression with Trial Type ("2-Item False" vs. "2-Item True" vs. "3-Item Test") as a within-subjects variable and Scale Type (context-independent vs. contextual) and grammatical strengthening (*only* vs. *no-only*) as between-subjects variables. The model found significant main effects for all factors and all two-way interactions (all $ps < 0.001$). Critically, there was also a highly significant three-way interaction between Trial Type, Scale Type, and grammatical strengthening ($z = 3.29$ $p < .001$). Children were significantly more likely to accept a critical 3-Item test sentence if the sentence was grammatically strengthened and had contextual alternatives than if it was strengthened and was presented with context-independent alternatives. Whereas *only* had a very large impact on children's interpretation of utterances including contextual alternatives, it had no effect at all when children interpreted utterances containing the word *some*.[6]

4. General discussion

The results from this study are consistent with the hypothesis that children's knowledge of scalar alternatives places a significant constraint on their ability to compute scalar implicatures. In the experiment, children were able to assign strengthened interpretations to utterances when they included the focus element *only*, as long as alternatives were provided contextually. For context-independent scales – e.g., *some/all* – children failed to compute strengthened interpretations most of the time, even when *only* was added. Since *only* forces strengthening grammatically (and clearly did so for contextual alternatives), no pragmatic inference was required. Thus, children's failure to derive

[6] The difference in children's sensitivity to *only* across conditions rules out the possibility that they had systematic difficulties interpreting it, a problem that children *do* encounter when multiple dimensions of contrast are present in a context (see studies by Crain et al., in the Introduction).

strengthened interpretations for *some* is consistent with a hypothesized failure to generate relevant scalar alternatives – in this case the quantifier *all*.

Besides showing that children have difficulty generating relevant alternatives for scales like ⟨*some, all*⟩, these data also suggest that, contrary to speculation in the literature (e.g., Chierchia et al., 2001; Pouscoulous et al., 2007; Reinhart, 2004), children's difficulties are not straightforwardly attributable to limitations of working memory, task demand, or other types of processing constraints. In our study, we found that children were perfectly capable of deriving strengthened interpretations for utterances that involved contextual alternatives, but failed for identical sentences that involved *some*. This is despite the fact that similar processes were involved across conditions. In both cases, children were required to represent the basic meaning of the utterance (Step I), generate relevant alternatives (Step II), restrict alternatives to those that are stronger/more informative (Step III), and negate the stronger alternatives to strengthen the utterance (Step IV). Our suggestion is that whereas a clear difference between conditions is found for Step II, there is no clear way that differences between conditions at other steps in the process could predict children's behaviors. This is because when children are freed from the need to generate alternatives – i.e., they are specified contextually – they succeed without difficulty.

Could a processing limit explain a failure to access alternatives? As noted in the introduction, Chierchia et al. (2001) speculate that children's failure to compute implicatures for *or* vs. *and* is due to a failure to access alternatives, and that this failure is related to processing limits. Although Chierchia et al. do not describe a specific mechanism in their account, there are several possibilities. One solution, consistent with our account, is that if scalar items are not associated with one another as scale mates, then children may need to expend additional resources to construct scales in an *ad hoc* fashion, and fail in their effort to do so on most occasions. This might explain why children become more likely to compute implicatures when they are first trained to reject descriptions that are underinformative (e.g., the use of *animal* to describe a dog). By encouraging children to actively search alternatives – which they might otherwise avoid due to the additional effort required – such training might compensate for weak knowledge of scales, allowing children to succeed during the course of the experiment. Consistent with this, training effects are transient, and do not persist after a delay of 1 week (Guasti et al., 2007). By this account, if quantifiers were associated as members of scales (not a processing cost in itself) children would access alternatives without problem, as they do for numerals. A second possibility, which we consider less plausible, is that children have difficulty holding scalar alternatives in working memory as they compute inferences, thereby explaining why quantifiers would be harder than contextual alternatives (which remain visible to children as they process the sentence). The problem here is that such an account cannot explain children's relative success with other scales. When scale members are strongly associated via memorization, as with numerals, children have no difficulty holding alternatives in working memory and computing inferences.

From a broader perspective, the idea that children's difficulties are scale-specific, rather than due to pragmatic immaturity, is also consistent with reports of pragmatic sophistication in other domains, such as noun learning (see Baldwin, 1993; Brooks, Audet, & Barner, in preparation; Clark, 1987, 1988; Markman, 1989; Tomasello, 1992). For example, as noted in the Introduction, children assume that nouns have contrasting meanings from very early in language learning (Carey & Bartlett, 1978; Clark, 1987, 1988; Markman, 1989). Children do not assume that two words contrast if they believe that the words are at different levels of description, or if they are told that one word is from another language (Au & Glusman, 1990; Diesendruck, 2005). In these cases, the known label is not considered a relevant alternative to the novel label. These simple inferences, though distinct from implicatures in many ways, nonetheless require both pragmatic understanding (including ascription of speaker intent), and the processing abilities needed to simultaneously entertain and negate potential alternatives. Indeed, by some accounts, these inferences can only be explained as Gricean inference (for discussion see Clark, 1990; Diesendruck & Markson, 2001; Gathercole, 1989). Such abilities would be difficult to explain if children's difficulties with scalar implicature were due to general processing limits or an insensitivity to conversational pragmatics.

What must children learn about scales to use them for implicature? As already noted, children must of course begin by learning the core meanings and syntactic properties of words, since both are prerequisites to being treated as scale mates for the purposes of pragmatic inference. Past studies show that, in the case of quantifiers, such knowledge is acquired early in childhood, and that children distinguish *some* from *all* shortly after their second birthday (Barner, Chow, et al., 2009; for similar results from Japanese, see Barner, Libenson, et al., 2009). When asked to give *some* of a set, children are significantly less likely to give all items than when asked for *all*. Also, when asked if *some* of a set is in a container, children are less likely to say 'yes' when all of the items are in the container than when only some are. Thus, despite their pragmatic limits, 2-year-old children know that *some* and *all* denote different set relations. Our suggestion is that, beyond this semantic prerequisite, children must perform additional learning in order to rapidly and automatically access lexical items as scalar alternatives. One possibility is that, in absence of brute memorization as with numerals, children group semantically related lexical items as scale mates by a gradual association of syntactically replaceable alternatives (see Katzir, 2007). Scale members may also be associated as children hear them explicitly contrasted in conversation – e.g., "Give me some of the cake, but not all of it". In each case, some form of gradual learning would be required, on top of acquiring the semantics of individual lexical items. Currently, no empirical studies address how scales might be constructed, and even in adults the psychological status of scales is poorly understood – i.e., which words form scales and which do not. Future studies should explore the nature of scales both in adults and children, to better understand how scales are constructed in acquisition to support pragmatic inference.

Acknowledgements

We would like to thank Asaf Bachrach, Charles Reiss, Jennifer Audet, and two anonymous reviewers for their help and useful comments on previous drafts of this paper.

Appendix A

Condition 1 (context independent, not strengthened):

3-Item Test	Is Cookie Monster holding some of the food? (holding 3 of 3 fruits) Are some of the animals reading? (3 of 3 are reading) Are some of the animals sleeping? (3 of 3 are sleeping)
2-Item False	Is Cookie Monster holding all of the food? (holding 2 of 3 fruits) Are all of the animals sleeping? (2 of 3 are sleeping) Are all of the toys on the table? (2 of 3 toys are on the table)
2-Item True	Is Cookie Monster holding some of the food? (holding 2 of 3 fruits) Are some of the animals reading? (2 of 3 are reading) Are some of the toys on the table? (2 of 3 toys are on the table)

Condition 2 (context independent, strengthened):

3-Item Test	Is Cooking Monster holding only some of the food? (holding 3 of 3 fruits) Are only some of the animals sleeping (3 of 3 are sleeping) Are only some of the toys on the table? (3 of 3 toys are on the table)
2-Item False	Is Cookie Monster holding all of the food? (holding 2 of 3 fruits) Are all of the animals sleeping? (2 of 3 are sleeping) Are all of the toys on the table? (2 of 3 toys are on the table)
2-Item True	Are only some of the animals reading? (2 of 3 are reading) Are only some of the animals sleeping? (2 of 3 are sleeping) Are only some of the toys on the table? (2 of 3 toys are on the table)

Condition 3 (context-dependent, not strengthened):

3-Item Test	Is Cookie Monster holding the banana and the apple? (holding 3 of 3 fruits) Are the cat and the cow sleeping? (3 of 3 are sleeping) Are the rabbit and the dog reading? (3 of 3 are reading)
2-Item False	Are the train, the ball and the drum on the table? (2 of 3 toys are on the table) Is Cookie Monster holding the apple, the banana and the orange? (holding 2 of 3 fruits) Are the dog, the cat and the cow sleeping? (2 of 3 are sleeping)
2-Item True	Is Cookie Monster holding the banana and the apple? (holding 2 of 3 fruits) Are the rabbit and the dog reading? (2 of 3 are reading) Are the drum and the ball on the table? (2 of 3 toys are on the table)

Condition 4 (context-dependent, strengthened):

3-Item Test	Are only the cat and the cow sleeping? (3 of 3 are sleeping) Are only the drum and the ball on the table? (3 of 3 toys are on the table) Is Cookie Monster holding only the banana and the apple? (holding 3 of 3 fruits)
2-Item False	Are the dog, the cat and the cow sleeping? (2 of 3 are sleeping) Are the train, the ball and the drum on the table? (2 of 3 toys are on the table) Is Cookie Monster holding the banana, the apple, and the orange? (holding 3 of 3 fruits)
2-Item True	Are only the dog and the rabbit reading? (2 of 3 are reading) Are only the cat and the cow sleeping? (2 of 3 are sleeping) Are only the drum and the ball on the table? (2 of 3 toys are on the table)

References

Au, T. K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development, 61*(5), 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., Chow, K., & Yang, S. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology, 58*, 195–219.

Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology, 103*, 421–440.

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

Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics, 25*(2), 93–140.

Brooks, N., Audet, J., & Barner, D. (in preparation). Pragmatic inference, not semantic competence, guides children's interpretation of unknown number words.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford child language conference, 15*, 17–29.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H.-J. Do, L. Domingues, & A. Johansen (Eds.), *Proceedings of the 25th Boston University conference on language development* (pp. 157–168). Somerville, MA: Cascadilla Press.

Clark, E. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33).

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

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

Condry, K. F., & 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., Goro, T., & Minai, U. (2005). Hidden units in child language. In A. C. Schalley & D. Khlentzos (Eds.), *Mental states. Evolution, function, nature* (Vol. 1, pp. 275–294). Amsterdam: John Benjamins.

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

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

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

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

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*(3), 336–360.

Goro, T., Minai, U., & Crain, S. B. (2005). Two disjunctions for the price of only one. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th Boston University conference on language development* (pp. 228–239). Somerville, MA: Cascadilla Press.

Goro, T., Minai, U., & Crain, S. B. (2006). Bringing out the logic in child language. In L. Bateman & C. Ussery (Eds.), *Proceedings of the 35th annual meeting of the north east linguistic society* (pp. 245–256). Amherst, MA: GLSA Publications.

Grice, H. P. (1978). Further notes on logic and conversation. In P. Cole (Ed.), *Syntax and semantics: Pragmatics* (Vol. 9, pp. 113–128). New York: Academic Press.

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

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

Guasti, M., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., Meroni, L. (2007). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes* 20, 667–696.

Hirschberg, J. (1985). A theory of scalar implicature. Doctoral dissertation, University of Pennsylvania, Philadelphia, PA.

Horn, L. (1972). On the semantic properties of the logical operators in English. Doctoral dissertation, UCLA, Los Angeles, CA.

Horn, L. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Huang, Y., Snedeker, J. & Spelke, E. (submitted for publication). What exactly do numbers mean?.

Huang, Y., & Snedeker, J. (2009a). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology, 58*(3), 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*(6), 1723–1739.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development, 2*, 77–96.

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

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

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

Musolino, J. (2006). On the semantics of the subset principle. *Language Learning and Development, 2*(3), 195–218.

Musolino, J., & Lidz, J. (2006). Why Children aren't universally successful with quantification. *Linguistics, 44*(4), 817–852.

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

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

Papafragou, A. (2006). From scalar semantics to implicature: Children's interpretation of aspectuals. *Journal of Child Language, 33*, 721–757.

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.

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

Rapacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 34*, 1017–1025.

Reinhart, T. (2004). The processing cost of reference-set computation: Acquisition of stress shift and focus. *Language Acquisition, 12*(2), 109–155.

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

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

Tomasello, M. (1992). The social bases of language development. *Social Development, 1*, 67–87.

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