BUCLD 38 Proceedings
To be published in 2014 by Cascadilla Press

Scalar Inferences in 5-year-olds: The Role of Alternatives

Dimitrios Skordos[*] and Anna Papafragou[*]

1. Introduction

Scalar inferences are pragmatic inferences generated in the presence of lexical items such as quantifiers (e.g. *some*, *all*), where the use by a speaker of one item (*some*) typically leads the hearer to infer that a logically stronger item (*all*) would be false. For instance, (1a) is typically understood to implicate (1b).

(1) a. *Some* of the gnomes have a pie.

b. *Not all* of the gnomes have a pie.

Such inferences are called 'scalar' since linguistic terms like 'some' and 'all' are thought to form sets of alternatives ordered by informational strength (Horn, 1972). The notion of informational strength used here is based on asymmetrical logical entailment, where the sentence with the stronger scalar item (2a) logically entails the sentence with the weaker scalar item (2b), but not vice versa.

(2) a. *All* of the monkeys left the cage.

b. *Some* of the monkeys left the cage.

The philosopher Paul Grice (1975) provided an account of how these inferences are derived in discourse. He described communication as a co-operative effort governed by rational expectations about discourse. These expectations are formalized as four 'Maxims', which specify that communicators should contribute to the discourse by offering statements that are truthful, informative, relevant and appropriate to the common goals of the conversation. These maxims are not inviolable; openly and intentionally violating one or more of them leads the hearer to infer that the speaker intends

---

[*] Dimitrios Skordos, Department of Linguistics & Cognitive Science, University of Delaware, dskordos@udel.edu; Anna Papafragou, Department of Psychology, University of Delaware, APapafragou@psych.udel.edu.

the hearer to recover additional meaning beyond what is semantically encoded in the linguistic message. In other words, flouting a maxim typically carries an 'implicature'. For example in (1a) above the speaker seems to violate the maxim of Quantity:

<u>*Quantity maxim*</u>
i. Make your contribution as informative as is required.
ii. Do not make your contribution more informative than is required.

In (1a), the speaker flouts the Quantity submaxim (i) by using a weaker term from a set of ordered alternatives (< *some, …, all* >). The speaker is expected to say as much as he/she truthfully can, as long as it is relevant to the discourse. The choice of the weaker term leads the hearer to infer that the speaker is not able to commit to an informationally stronger statement ("All of the gnomes have a pie."). Therefore, as far as the speaker is able to say, the stronger statement does not hold. Thus (1b) is an implicature added to the semantic content of (1a) (see also Horn, 1972).

In this account, weak scalar terms such as the quantifier 'some' have lower-bounded semantics ('at least some and possibly all'). Upper-bounded meanings are considered pragmatic enrichments derived from the implicature process described above and can be canceled without logical contradiction ("Some of the gnomes have a pie. In fact all of them do"). This kind of scalar inference, or more specifically scalar implicature (SI), is a subtype of conversational implicature (Grice, 1975).

2. Scalar Implicatures in Children

The computation of SIs has been consistently shown to cause problems for children up to the age of 10. Noveck (2001) showed that children clearly favored a logical interpretation of the modal term 'might' and the quantifier 'some' ('certains' in French), while adults were more ambivalent between the logical and pragmatic interpretations. For instance, children and adults were presented with statements such as "Some giraffes have long necks" and were asked whether they agreed with the statements or not. As Noveck and others noted, the failure observed in children was likely to be due to task demands, since participants had to evaluate the truth value of an out-of-context statement against world knowledge and then decide whether a logical or pragmatic interpretation of the statement was warranted.

Further studies showed that children's ability to calculate scalar implicatures improved under different experimental conditions. In Chierchia, Crain, Guasti, Gualmini and Meroni (2001) and Gualmini, Crain, Meroni, Chierchia and Guasti (2001) preschoolers initially appeared insensitive to the implicature of exclusivity from the use of the disjunctive operator 'or'. Specifically, it appeared that children interpreted weak statements such as "Every boy chose a skateboard or a bike" as compatible with stronger statements such as “Every boy chose a skateboard and a bike”. However, when children were presented with two statements and were asked to choose the best among them, they overwhelmingly chose the stronger/more informative statement with ‘and’ (“Every farmer cleaned a horse and a rabbit”) over the weaker/less informative statement with ‘or’ (“Every farmer cleaned a horse or a rabbit”) under conditions that made the stronger statement true. Similar results were obtained by Ozturk and Papafragou (under review) with modal expressions (‘may’, ‘have to’).

Other studies showed that training, context and more naturalistic tasks improved children’s generation of SIs (Papafragou & Musolino, 2003; Papafragou & Tantalou, 2004; Guasti, Chierchia, Crain, Foppolo, Gualmini, & Meroni, 2005; Pouscoulous et al., 2007; Katsos & Bishop, 2011). These findings support the hypothesis that children have the ability to make pragmatic inferences and suggest that children’s problem lies with generating scalar alternatives spontaneously when faced with a weak scalar term (Gualmini et al., 2001; Papafragou & Skordos, in press).

3. The Role of Accessibility of Alternatives

A recent study by Barner, Brooks and Bale (2011) offered more direct evidence for the role of the accessibility of unspoken lexical alternatives on children’s SI calculation. When 4-year-old children were asked to reason about statements describing a group of animals, they appeared unable to detect the SI typically triggered by the quantifier ‘some’ ($\rightarrow$ ‘not all’). This persisted even when the statement was strengthened by the inclusion of the focus element ‘only’ which is predicted to grammatically trigger the exclusion of the stronger alternative (“Only some of the animals are sleeping” $\rightarrow$ “Not all of the animals are sleeping”). However when set members were individuated (“Only the cat and the dog are sleeping”) children would show evidence of comprehending the statement as excluding a third animal (a cow) from the set of sleeping characters. These findings offered evidence that children’s problem with SIs lies mainly in realizing what terms can come together to form a scale. In other words, when scalemates were explicitly provided, children’s generation of SIs improved.

In previous work (Skordos & Papafragou, 2012), we manipulated the accessibility of the scalar alternatives ‘some’ and ‘all’ within a single task to directly test the idea that children’s difficulty with generating SIs lies with accessing the appropriate relevant alternatives that need to be contrasted in order to compute a SI. If the accessibility of alternatives is a limiting factor in children’s computation of SIs, then children’s SI generation should improve when a contrast between strong (‘all’) and weak (‘some’) scalar alternatives is available. In an Acceptability Judgment Task (AJT), five-year-old children and adult controls were provided with ‘some’ and ‘all’ statements by a ‘silly’ puppet who described scenes where cartoon characters had a number of items in their possession. In these scenes, either some (3 out of 4) or all (4 out of 4) characters would possess an item. The critical trials were a set of infelicitous 'some' statementsthat were used to describe scenes where a statement with 'all' would have been a better description (e.g., statements such as "Some of the Xs have Ys" where 4 out of 4 Xs had a Y).

We manipulated the accessibility of the scalar alternatives by manipulating the order of the 'some' and 'all' statements between subjects within the study. In the 'Mixed' group, participantsencountered infelicitous 'some' statements intermixed with statements featuring the stronger alternative 'all'. In the 'Some-First' group, participants heard the infelicitous 'some' statements before they had a chance to encounter the stronger 'all' statements. We found that 85% of the children in the 'Mixed' group but only 47% of the children in the 'Some-First' group rejected the infelicitous 'some' statements (adults consistently did so across conditions) . We concluded that young children face significant difficulties with accessing the necessary scalar alternatives in order to generate a SI (see also Bagassi, D'Addario, Macchi, and Sala, 2009, for related evidence from older children).

4. The Present Study

While prior studies provide evidence that accessibility of the stronger alternative is an important factor in SI generation for children, they do not shed much light onto the precise nature of children's difficultyor on the way the accessibility of lexical alternatives helps children compute SIs. There are two hypotheses open: The first is that children simply do not know what alternatives enter into a scalar relationship. We will call this *the lean hypothesis*. On this hypothesis, the difference between adult and child performance in SI tasks can be explained by adults' experience with scales and scalemates. It is this experience that develops as children mature into having an adult-like performance. If this is the case, one would expect that, other things being equal, simply providing alternatives should lead children to derive SIs.

The second hypothesis is that children do not always realize when alternatives are relevant in order to generate a SI at a given time. We will call this *the rich hypothesis*. If this is correct, the fundamental difference between children and adult communicators is their ability to home in on the conversational goal at any given time and consequently identify scalar alternatives that are relevant within the identified conversational goal. In this case, simply providing scalar alternatives should not necessarily lead children to derive SIs. What is needed for SI generation is a realization that the alternatives provided (and the SI to be generated) are relevant for furthering the goal(s) of discourse.

Here we present a test of these two hypotheses. We manipulate the degree to which alternatives can be easily recognized as relevant by children. We achieve this by comparing a case where the conversational goal remains stable (in a way that helps establish the relevance of the stronger alternative) with another case where it switches during the experiment (in such a way that it renders the stronger alternative irrelevant). In both cases, we make stronger alternatives relevant to 5-year-olds using a method very similar to that of Skordos and Papafragou (2012). If the lean hypothesis is correct, the presence of a single vs. multiple conversational goals should not affect SI computation (since accessibility of the stronger scalar alternative is guaranteed). If the rich hypothesis is correct, 5-year-olds should benefit from the accessibility of the stronger scalar term only when this term is relevant to the conversational goal.

5. Experiment

5.1 Method

5.1.1 Participants

We tested a group of 50 typically developing 5-year-old children (4;9 – 5;8, M=5;0) and 24 adult controls, all monolingual speakers of English. The children were recruited from daycare centers in Newark, DE. The adults were college students recruited from the University of Delaware and received course credit for their participation. An additional group of 4 children were tested but excluded from the analysis for failure to follow instructions.

5.1.2. Materials and Procedure

The task was an Acceptability Judgment Task very similar to that in Skordos and Papafragou (2012). Children sat in front of a laptop PC computer and were shown the slides depicting the experimental stimuli. A first experimenter introduced a hand-held puppet, Max the silly gorilla, “who says silly things sometimes”, and explained that children and the experimenter would see some pictures on the computer together. Participants were told that the puppet would describe the pictures and that they would have to say whether the puppet “said it well or not”. They would also have to justify their answer in case they rejected the puppet’s statement. A second experimenter animated the puppet and provided the appropriate statements, while the first experimenter marked children’s answers on an answer sheet. Adults were tested in a very similar way with the only differences being that (a) they had to write down their own responses, and (b) they were tested in groups without the presence of a puppet (they were shown a cartoon character, Max the silly gorilla, that supposedly provided the statements that the experimenter read).

Participants first went through 4 pre-test trials. These consisted of slides depicting cartoon animals or objects (e.g., a cow, an ice cream cone). Two of the pre-test trials were erroneously described by the puppet and two of them were correctly described, so that participants would have evidence that the puppet was capable of providing both ‘silly’ and accurate statements. For pre-test trials, participants were also provided with feedback when they failed to reject a false statement. For example, if participants agreed with the puppet when it described the cow as an “elephant”, the experimenter would explain that the puppet “didn’t say it well’, and that in fact the picture depicted a cow.

After the pre-test trials were concluded, participants were introduced to a cartoon character, Ben the Wizard on an introductory slide. Ben was shown to use his magic wand to create the 4 “blickets”, novel animate creatures that would appear on all test slides. Participants were informed that these are the only blickets “in the whole world”. This step was taken to narrow down the universe of discourse to the scene at hand.

Blickets had several everyday items (crayons, flashlights, paintbrushes, etc.). In half of the slides 4 out of 4 blickets would have an item each (*full set* scenes) and in the other half 3 out of 4 blickets would have an item each (*subset* scenes). For each slide Max offered a statement containing a quantifier (‘some’ or ‘all’). Scene type (full set vs. subset) was crossed with quantifier type (some vs. all) to provide 4 types of trials: a) In Full set/All trials, 4 out of 4 blickets had the item and participants heard: “All of the blickets have an X.”. We will refer to these as *True-All* trials. b) In Partial set/All trials, 3 out of 4 blickets had the item and participants heard: “All of the blickets have an X.” These were the *False All* trials. c) In Partial set/Some trials, 3 out of 4 blickets had the item and participants heard: “Some of the blickets have an X.” (*True-and-Felicitous-Some* trials). Finally, in Full set/Some trials, 4 out of 4 blickets had the item and participants heard: “Some of the blickets have an X.” (*True-and-Infelicitous-Some* trials). The first three types of trials tested participants’ semantic judgments about ‘some’ and ‘all’. The last type of trial tested participants’ pragmatic judgment (i.e., their ability to generate SIs). Examples of the visual scenes and statements for each trial type can be found in Figure 1.

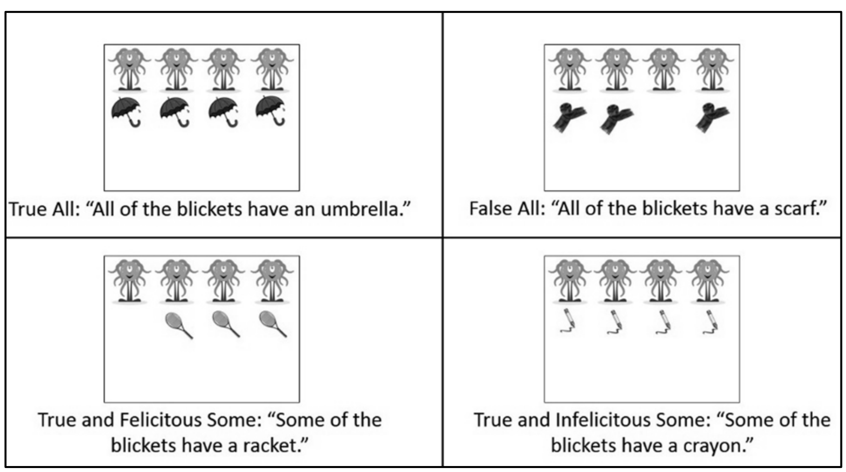

Figure 1. Scene and statement examples.

These 4 types of trials were repeated 4 times each, with different items so that no subject saw the same scene paired with more than one statement, for a total of 16 test trials. Pairings of scenes with statements were rotated to create 4 different batteries so that each scene was paired with a different statement type in each battery.

The ‘all’-statements were always presented in a first block, so that lexical contrast between the stronger (‘all’) and weaker (‘some’) scalar terms could be established. The ‘some’-statements came in a second block. There were two between subjects conditions that differed in a single respect: In the *Quantity* condition, the False-All trials would involve 3 out of 4 blickets having the correct item (for example a scarf) and be accompanied by a statement such as “All of the blickets have a scarf”. Figure 2 gives example scenes and statements for the Quantity Condition.

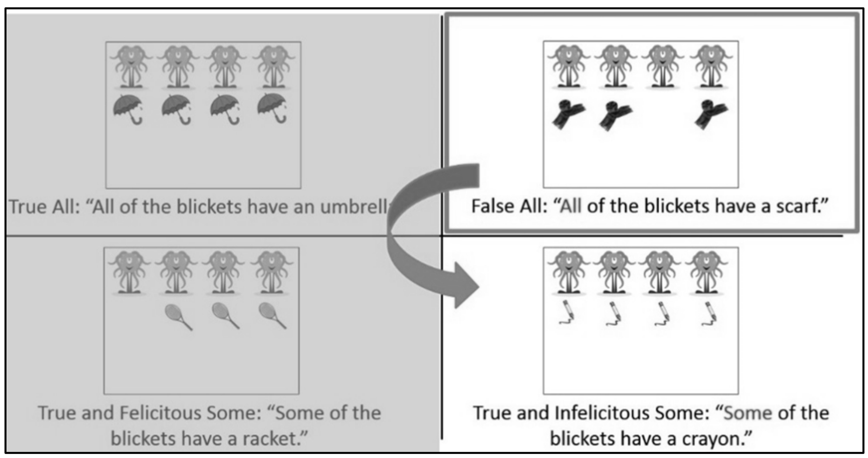

Figure 2. Scene and statement examples: Quantity condition.

In the *Object* condition the same trial would involve 4 out of 4 blickets having an item different from the one mentioned in the statement. For example, each of the 4 blickets would have a shovel and the statement would be: “All of the blickets have a scarf”. The statement was still false in the context of the visual scene, but the reason was different from the previous condition: here, the object was wrong (in the Quantity condition, the quantity of the objects was wrong). Figure 3 shows example scenes and statements for the Object condition.

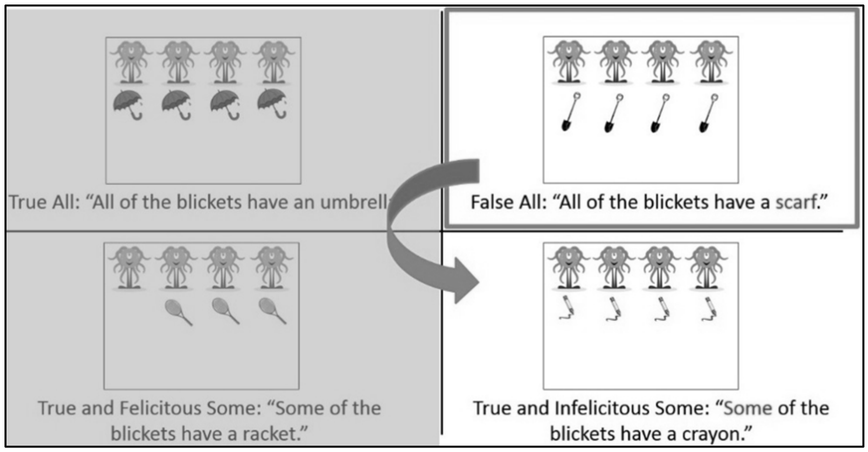

Figure 3. Scene and statement examples: Object condition.

In the Quantity condition, the implicit basis upon which the participants were called to evaluate each statement, was whether the quantity of blickets shown to possess an item in the scene matched the quantity of blickets mentioned in the statement. Thisconversational goal remained stable throughout the experiment: it was established in the first (All) block through the False-All trials and could later be brought to bear on judgments of the True-and-Infelicitous-Some statements. In the Object condition, however, this evaluation basis changed between the first and the second block. The first (All) block, especially the False-All statements, should arguably lead participants to identify object identity as the conversational goal (i.e., whether the blickets possessed the stated object kind or not). In the second (Some) block, however, if participants were to detect the infelicity of the True-and-Infelicitous-Some trials, they would have to recover a different conversational goal (namely, whether the quantity of blickets in possession of a certain object was as stated in the sentence or not).

5.2 Predictions

If the lean hypothesis is correct, then children should successfully reject True-and-Infelicitous-Some statements in both conditions, since 'all' is lexically accessible in both conditions by virtue of being present throughout the first Block. However, if the rich hypothesis is correct, then children should be more successful in the Quantity than the Object condition, since the stronger alternative 'all' is only relevant in the Quantity condition where it is integrated in the same conversational goal with the weak counterpart 'some' that should trigger the SI.

Consider the true/infelicitous "Some of the blickets have a racket" (Fig.1) when all of the blickets have a racket. If children believe that the conversational goal is to evaluate whether the puppet got the quantity of blickets right (Quantity condition) and already have access to the stronger 'all' term, they should easily reject the 'some'-statement (since all of the blickets have a racket). But if children believe the conversational goal is to evaluate whether the puppet got the object owned by the blickets right (Object condition), even if they have access to the stronger 'all' term, they might not reject the statement (since some of the blickets indeed have a racket).

5.4 Coding

Yes answers were coded as correct in the case of true or true-and-felicitous statements. *No* answers were coded as correct in the case of false or true-and-infelicitous statements. A mean of correct answers from 0 to 1 was calculated for each participant for each of the 4 trial types (True-All, False-All, True-and-Felicitous-Some, True-and-Infelicitous-Some). Those scores were used to categorize participants according to their performance on each trial type as either Passers (if they had a score of .75 or greater), or Failers (if they had a score of .50 or less). For example, someone who had at least .75 correct on Felicitous-Some trials was categorized as a Passer for that trial type.

5.5 Results

Adult performance was very high for all conditions and trial types (Table 1 below). Fisher's Exact test analyses on 2x2 contingency tables for each trial type revealed no significant difference in the numbers of Passers vs. Failers across conditions (*True-All*-trials, p = 1; *False-All*-trials, p = 1; *True-and-Felicitous-Some*-trials, p = 1; *True-and-Infelicitous-Some*-trials, p = 1). Adults were overwhelmingly pragmatic in the *True-and-Infelicitous-Some* trials, i.e., they rejected these statements and correctly justified their responses in terms of the stronger available alternative.

		Quantity	**Object**
True All	Passers	12	12
	Failers	0	0
False All	Passers	12	12
	Failers	0	0
True and Felicitous Some	Passers	12	12
	Failers	0	0
True and Infelicitous Some	Passers	11	12
	Failers	1	0

Table 1. Adults' performance.

Children performed well with the 3 semantic trial types (see Table 2). Fisher's Exact Tests on 2x2 contingency tables did not reveal significant differences in the numbers of Passers vs. Failers across the two conditions for either the *True-All*-trials (p = 1) or *False-All*-trials (p = 0.357). We did observe a difference approaching significance in the numbers of Passers vs. Failers in the *True-and-Felicitous-Some-trials* (p = 0.05). Turning to the critical *True-and-Infelicitous-Some*-trials, children were more pragmatic in the Quantity condition (17 Passers and 9 Failers) and more logical (non-pragmatic) in the Object condition with only 8 Passers and 16 Failers (see Table 5). A Fisher's Exact test on a 2x2 contingency table revealed a significant difference (p = 0.046) between the numbers of Passers and Failers for *True-and-Infelicitous-Some* trials between the two conditions.

When asked to justify their rejections of *True-and-Infelicitous-Some* statements, children referenced either the stronger scalar term ("*All* of them have an X"), or used the focus element *only* ("He said that *only some* of them have an X").

		Quantity	**Object**
True All	Passers	26	24
	Failers	0	0
False All	Passers	22	23
	Failers	4	1
True and Felicitous Some	Passers	19	23
	Failers	7	1
True and Infelicitous Some	Passers	17	8
	Failers	9	16

Table 2. Children's performance.

After this initial analysis and to make sure that we are looking at participants who have a solid grasp on the semantics of the quantifiers, we used the semantic trials as controls and conducted a second analysis excluding children who were Failers (had under .75 correct) in either the *True-All*, *False-All*, and *True-and-Felicitous-Some* statements. This resulted in n=9 children being excluded in the Quantity condition, and n=1 child excluded in the Object condition[1] (see Table 3). All of the remaining children can be assumed to have the correct semantics for *some*and *all.* A Fisher's Exact test on the 2x2 contingency table in the True-and-Infelicitous-Some trials of Table 3 revealed a significant difference between the numbers of Passers vs. Failers for the two different conditions (p = 0.000007) with the Quantity condition having

[1] The exclusion criterion for the Object condition was based solely on the *Felicitous Some* trials, since the *True All* and *False All* trials had to be evaluated based on the identity of the items present and not on quantification and thus would not necessarily offer any evidence of a participant's knowledge of quantifiers. That said, even if we included the *True All* and *False All* trials in the exclusion criteria for the *Object* condition for the sake of uniformity, only one additional participant would have been excluded, and our analyses would not have been affected.

significantly more Passers than the Object condition, confirming the results of the first analysis.

		Quantity	Object
True All	Passers	17	23
	Failers	0	0
False All	Passers	17	23
	Failers	0	0
True and Felicitous Some	Passers	17	23
	Failers	0	0
True and Infelicitous Some	Passers	17	7
	Failers	0	16

Table 3. Semantic Passers' performance.

We then compared the performance of the children who have a solid grasp of quantifier semantics with that of adults in the True-and-Infelicitous-Some trials, using Fisher's Exact Test on 2x2 contingency tables. We found no difference between age groups in the Quantity condition (p = 0.414) but a significant difference in the Object condition, with the adult group having significantly more Passers than the child group (p = 0.00006).

6. Discussion

The starting point of the present investigation was the idea that at least part of the problem children face in SI generation lies in failing to access the appropriate stronger alternative when a weak scalar term is used (see Barner et al., 2011, Skordos & Papafragou, 2012, Papafragou & Skordos, in press, among others). Of interest was whether accessibility of the stronger alternative could bear the explanatory burden of children's failure with SIs alone (lean hypothesis) or whether accessibility of alternatives was sensitive to the role of alternatives within the conversational goals of the exchange (rich hypothesis).

We presented children (and adults) with contexts in which they had to judge a true but infelicitous 'some' statement (e.g., 'Some of the blickets have a shovel' in a scene in which all of the blickets had a shovel). We ensured that these infelicitous statements were always preceded by a block of 'all' statements (i.e., the stronger scalar alternative was always made accessible). However, we manipulated the degree to which the scalar alternative was relevant to the conversational goal. In the Object condition, the first block that contained the stronger alternative 'all' hinted at a conversational goal which placed emphasis on object identity ("Do some/all of the blickets have X, or not?"). In the Quantity condition, the conversational goal suggested by the first, 'all' block involved the quantity of blickets ("Do X of the blickets have an item, or not?") and was therefore better aligned with the judgment for the infelicitous 'some' statements in the second block.

Our findings were quite striking: Despite the fact that the stronger alternative ('all') was available in exactly the same way in both conditions, children's performance with infelicitous 'some' statements was much lower in the Object compared to the Quantity condition. In fact, to the extent that they realized that the accessible alternative was relevant, children performed like adults (Quantity condition). We take this as strong evidence for the rich hypothesis that states that scalar alternatives need to be viewed as relevant within a conversational goal in order to lead to SI generation.

It is quite possible that the results of previous studies that have shown improved SI generation can be re-interpreted using this approach. For example, in studies that provided an under-informative and a fully informative alternative and asked children to choose between the two, both alternatives offered were relevant (Chierchia et al., 2001, Gualmini et al., 2001). In other studies, contextual support in the form of background information essentially drew attention to a contextually relevant stronger alternative (Papafragou & Musolino, 2003; Guasti et al., 2005). Further studies will be required to determine the potency of conversational goal in determining the computation of conversational inferences in young children.

Acknowledgements

We would like to thank the Early Learning Center at the University of Delaware and all the other preschools for their cooperation, as well as the members of the Language and Cognition Lab at the University of Delaware for their comments and suggestions. The first author's work was supported in part by a College of Arts and Sciences Dean's Doctoral Student Summer Scholarship and a Dissertation Fellowship by the Office of Graduate and Professional Education at the University of Delaware.

References

Bagassi, M., D'Addario, M., Macchi, L., & Sala, V. (2009). Children's acceptance of underinformative sentences: The case of some as a determiner. Thinking and Reasoning, 15(2), 211-235.

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

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. 12

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.

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.

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

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. Cognition, 120 (1), 67-81.

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

Ozturk, O., & Papafragou, A. (under review). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation.

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

Papafragou, A. & Skordos, D. (in press). Implicature. In J. Lidz, W. Snyder and J. Pater (eds.), The Oxford Handbook of Developmental Linguistics. Oxford: Oxford University Press.

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

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

Skordos, D. & Papafragou, A. (2012). Lexical alternatives improve 5-year-olds' ability to compute scalar implicatures. Online supplement to the proceedings from the 36th Annual Boston University Conference on Language Development, Boston, MA.