ISSN: 1547-5441 (Print) 1547-3341 (Online) Journal homepage: www.tandfonline.com/journals/hlld20

Asymmetries in the Acquisition of Numbers and Quantifiers

Felicia Hurewitz, Anna Papafragou, Lila Gleitman & Rochel Gelman

To cite this article: Felicia Hurewitz, Anna Papafragou, Lila Gleitman & Rochel Gelman (2006) Asymmetries in the Acquisition of Numbers and Quantifiers, Language Learning and Development, 2:2, 77-96, DOI: 10.1207/s15473341lld0202_1

To link to this article: https://doi.org/10.1207/s15473341lld0202_1

LANGUAGE LEARNING AND DEVELOPMENT, *2*(2), 77–96

Asymmetries in the Acquisition of Numbers and Quantifiers

Felicia Hurewitz
Department of Linguistics
University of Delaware

Anna Papafragou
Department of Psychology
University of Delaware

Lila Gleitman
Rutgers University Center for Cognitive Science
Rutgers University–New Brunswick
Institute for Research in Cognitive Science
University of Pennsylvania

Rochel Gelman
Rutgers University Center for Cognitive Science
Rutgers University–New Brunswick

Number terms and quantifiers share a range of linguistic (syntactic, semantic, and pragmatic) properties. On the basis of these similarities, one might expect these 2 classes of linguistic expression to pose similar problems to children acquiring language. We report here the results of an experiment that explicitly compared the acquisition of numerical expressions (*two, four*) and quantificational (*some, all*) expressions in younger and older 3-year-olds. Each group showed adult-like preferences for "exact" interpretations when evaluating number terms; however they did not use the corresponding upper bounded interpretation when evaluating the quantifier *some*. Apparently, children follow different procedures for learning and evaluating numerals and quantifiers. These findings have implications for theories of number representation in child and adult grammars.

Correspondence should be addressed to Felicia Hurewitz, Department of Linguistics, University of Delaware, Newark, DE 19716. E-mail: fel@udel.edu

This article is concerned with the acquisition of the meaning of number words such as *two* or *three*. Number terms are notoriously vexing entities. The word *two* in the phrase *two horses* does not describe any individual in the environment, and it does not refer to a property that any individual could possess; rather it refers to the set of horses. Thus, acquiring the semantics of number terms might pose special difficulties for young language learners. After all, word learning is facilitated by observation for items such as *horse* because the chances of seeing a horse as this word is uttered are quite good; the adult may even relevantly point at that visible horse, saying "See? That's a horse." More generally, cross-situational observation is a good clue to the meanings of object terms. In contrast, there is obviously no coherence of observed object across usages of the word *two*—sometimes it is used as if it is noun and sometimes as if it is an adjective (Frege, 1974), sometimes it describes heterogeneous sets of objects and sometime homogeneous. The child's task is to recover the less than salient fact that in standard uses of *two* in context, there is indeed a recurrent property of the environment, namely that the set-size of relevant entities (whatever their essence) is two.

Thought about this way, it seems astonishing that the number terms are acquired with anything like their veridical semantics in the preschool years of life. Yet it is a fact that 2- and 3-year-olds use number words in systematic ways, apparently never mistaking them for object-reference terms (Shatz & Backscheider, 1991), and indeed their use of number expressions reveals subtle aspects of numerical reasoning (Gelman & Gallistel, 1978).

Still, the acquisition of counting procedures that are consistent with the counting principles of one-to-one, stable ordering and cardinality (Briars & Siegler, 1984; Fuson, 1988; Gelman & Gallistel, 1978; Wynn, 1990, 1992), is very protracted. Variability across tasks is the rule for children younger than 3 years (Gelman & Cordes, in press; Grinstead et al., 1997). Yet, by the time they are 3½, children display a variety of number skills, including the ability to count small sets accurately and to estimate numerosity without counting (Gelman & Gallistel, 1978). This has prompted intense experimentation and theorizing about child semantics for quantification and its expression (see Carey, 2004; Cordes & Gelman, in press; Gelman, 1993; Gelman & Butterworth, 2004; Mitchie, 1984; Mix, Huttenlocher, & Levine 2002; Sarnecka & Gelman, 2003; Wynn, 1990, 1992, for varying perspectives).

We can distinguish broadly between two rather different current proposals about what children think number terms mean and how this knowledge is acquired. According to one proposal, children from the start represent the numbers as pertaining to sets and to the generative and productive concept of exact numerosity (Grinstead, MacSwan, Curtiss, & Gelman, 1997). Proponents of this view are impressed, for example, by preschoolers' implicit appreciation of the cardinality principle that relates counting to numerosity (i.e., the last number counted represents the cardinality of the set, Gelman & Gallistel, 1978; Zur & Gelman, 2004). Children as young as 3 appear to realize that one can "guess" the solution to an arithmetic problem, but only counting assures a precise answer (Zur & Gelman, 2004). Indeed, 2-year-olds will spontaneously resort to counting when a known set size is "magically," or inexplicably, altered (Gelman, 1972). Furthermore, children implicitly recognize that there must always be a logical "next" number in the count list, even if they are not themselves privy to their native language's word for this numeral (Hartnett & Gelman, 1991). This type of evidence has been taken to indicate that children's use of natural language number terms maps onto a rich domain specific nonverbal numerical system (Gelman, 1998).

An alternative to the hypothesis that children tap into domain specific knowledge of numerosity when learning to count may be termed a quantifier-bootstrapping hypothesis. According to this second view, children begin by treating number words as quantificational expressions without true grasp of their connection to set cardinality (Carey, 2004). Therefore, children may appreciate that numbers are quantity expressions without linking quantity to a truly generative and productive notion of numerosity. Instead, they initially associate number words with nonspecific approximate quantity concepts akin to those encoded by natural language quantifiers such as *some* or *many.* This view is supported by experimental findings that suggest a lack of understanding of cardinality in 2- to 3-year-olds. For example, a request such as "Give me *three horses*" often results in a toddler giving three, four, or even five or six toy horses to the requester (Wynn, 1990, 1992). According to one theory, it is only later in development that the integer words branch off from their quantificational cousins, the inexact natural-language quantifiers. Not only does this approach suggest a period in which the representations of these kinds of concepts are of the same categorial type, even more strongly, the position suggests the possibility that children might use their knowledge of the semantics of quantifiers to bootstrap their way into representations of number semantics (Carey, 2004).

In this article, we compare the predictions of the quantifier bootstrapping hypothesis to the view that number words initially map to a preexisting magnitude system, as applied to a small sampling of early-acquired number words and quantifiers. In the following section, we offer a brief overview of those areas where quantificational and number terms overlap in their semantics and pragmatics, and those areas where they diverge. Thereafter, we present an experiment that investigates if 3-year-old children apply identical pragmatic and semantic principles to the number words *two* and *four* and the quantifiers *some* and *all*. We reason that if children go through an initial stage where quantifiers and numbers are functionally and semantically equivalent, they should assign similar interpretations to the two types of quantity expressions found in our sample.

LINGUISTIC SIMILARITIES BETWEEN NUMBERS AND QUANTIFIERS

Number terms and quantifiers in natural language share a range of syntactic, semantic, and pragmatic properties. Syntactically, both numbers and quantifiers in English can co-occur with partitives (*two/some of the apples*) and precede adjectival modifiers (*two/some big apples*); furthermore, all numbers and many quantifiers have count syntax (*two/some men*).[1] Semantically, both numerals and quantifiers are predicates over sets of individuals (Barwise & Cooper, 1981). Furthermore, both sets of expressions are internally ordered, that is, they form a *scale* from weaker to stronger elements (Horn, 1972):

Number scale: < *… three, two, one* >
Quantifier scale: < *all, most, many, some* >

Scalars have several characteristic semantic properties. First, statements containing a stronger numerical or quantificational scalar term entail statements containing a weaker scalar term. (*Three balls are red* entails *Two balls are red*; *All balls are red* entails *Some balls are red*).

Second, members of both the number and the quantifier scales have traditionally been claimed to have lower bounded lexical meanings (Horn, 1972; Gazdar, 1979; Levinson, 2000). For example, *Two balls are red* is logically interpreted as meaning that at least two, and possibly more, balls are red. Similarly, *Some balls are red* encodes the information that at least some and possibly all of the balls are red. Pragmatically, the use of a lower number or quantifier usually excludes a higher ranked member of the same scale: *Some of the balls are red* is typically used to convey that no more than some (i.e., not all of the balls) are red. Such inferences, known as *scalar implicatures*, arise from the conversational presumption that speakers should be relevant and informative (i.e., offer the strongest statement warranted by conversational demands; Grice, 1989). The use of a lower ranked number or quantifier (e.g., *two, some*) often warrants the inference that, as far as the speaker knows, a higher ranked number or quantifier (e.g., *three, all*) does not apply. Thus, for example, if told "Noam Chomsky has one leg," all competent users of English—perhaps barring a few linguists and logicians—will conclude, falsely, that he does not have two legs. Listeners who hear "Some Democrats voted for Bush" will conclude, truly, that not all of them did.

One hallmark of conversational implicatures (and by extension scalar implicatures) is that they may be canceled by the speaker. Furthermore, in certain situations the context renders scalar implicatures nonapplicable. Specifically,

[1]Quantifiers may appear with mass nouns only (*much sand*), count nouns only (*a few men*), or both count and mass nouns (*some sand/men*).

when the contrast between the number word or quantifier and broader set is not deemed relevant within the context of the utterance, a scalar implicature is not triggered (Kratzer, 2003).

A: Did you meet some of Jane's relatives at the party?
B: Yes, I did. (When in fact B met all of them.)

A: I need two quarters for the parking meter.
B: I have two quarters. (When in fact B has 4 quarters.)

In the above examples, traditional analyses would assert that the upper bounding implicature has been waived (based on the contextual irrelevance of the entire set), and therefore a lower bounded ("at-least") interpretation applies. In conclusion, numbers and quantifiers in many circumstances appear to evoke identical rules of interpretation.

How could these broad similarities between numbers and quantifiers impact the learner's task of figuring out the semantics of number terms? Bloom and Wynn (1997) have proposed that children use syntactic and semantic information such as co-occurrence with partitives or position with respect to adjectives to narrow down conjectures about the meaning of number words.[2] The authors conclude that distributional information might help children break into the class of number words even before knowing exactly which number word is paired with which number denotation. By extension, one theory proposes that there is a stage during which children treat numbers and quantifiers as members of the same class followed by a stage during which the former "grow out of" the latter (Carey, 2004). In sum, according to this line of reasoning, numbers and quantifiers should pose similar problems to language learners and should be acquired via the same initial mechanisms.

However, from other perspectives a shared initial route to the acquisition of these two word types seems less plausible. Along with the similarities in semantics and syntax of these scalar terms, there are also important differences that could lead the learner down the wrong acquisition path. The first of these is the context sensitivity (or vagueness) of quantifiers. Individual quantifier terms do not map to specific quantities, or even coalesce around a central, prototypical number. Use of the identi-

---

[2]This claim can be unpacked into either of two diametrically opposed learning conjectures for the lexicon, depending on the implied direction of causal flow. If the logic of quantification projects these linguistic structures and the linguistic novice antecedently understands this logic and expects simplicity in meaning-to-form linguistic mappings, then the observed forms can serve as partial cues to the meanings of terms; this would be an instance of the syntactic bootstrapping hypothesis (e.g., Gleitman, 1990; Landau & Gleitman, 1985). In contrast, if the semantics of quantification are taken to be unknown by the learner in advance, rather, the semantics is constructed "from" the observed linguistic behavior of the terms, then this claim is that linguistic properties play a causal role in conceptual development and achievement, a version of the Sapir-Whorf hypothesis (e.g., Goldberg, in press).

cal noun phrase, such as "many people" does not protect one from the erratic variability in how the quantifier references a magnitude (Moxey & Sanford, 1993). Estimated magnitudes for some quantifiers will depend on the sentential and discourse context in which they are used, as can be seen in the contrast of implied magnitudes between *Many people watched the Oscars* and *Many people came to my tea party.* Other quantifiers such as *all* and *most* reference a proportion of a set. However, people's estimates of what proportion constitutes *most* will range depending on context conditions (Papafragou & Schwartz, in press). A child who is using quantifiers to bootstrap the meaning of numbers would have to learn that numbers do not allow for context dependent or set size manipulation of the represented magnitude.

Within linguistic theory, the traditional view that numbers and quantifier scales are semantically similar has come under criticism. There are linguistic reasons to assume that numbers differ from regular, lower bounded scalars such as quantifiers (Breheny, 2004; Carston, 1995, 1998; Horn, 1992; Koenig, 1991). Most importantly, various linguistic tests indicate that the "at least" and the "exact" interpretations of numbers intuitively belong to the truth-conditional content of numerally modified statements. For instance, in the examples below (from Horn, 2004), the cardinal *two* allows a disconfirmatory response in the case where the "at least two, possibly more" reading would apply. Contrast this with *many*, where the responder may cancel the scalar implicature using an affirmative response.

A:	Does she have two children?	A:	Are many of your friends linguists?
B1:	No, she has three.	B1:	?No, all of them.
B2:	?Yes, (in fact) three.	B2:	Yes, (in fact) all of them.

This type of evidence has led to a general consensus that number words in natural language differ from quantifiers (but see Levinson, 2000). According to some authors numbers should not be considered semantically lower bounded and pragmatically upper bounded (by scalar implicature). Instead, the meaning of *two* is plainly TWO and, depending on a contextual parameter, can yield "exact" or "at least" (and sometimes "at most") interpretations (on the role of context in number interpretation, see Breheny, 2004).[3] Under these accounts natural language quantifiers diverge from numbers. Quantifiers can still best be accounted for under a neo-Gricean (Levinson, 2000) system, where the upper boundary is obtained via scalar implicature.

Summing up, this more recent linguistic approach to number has implications for what children need to learn when they acquire number words; and most rele-

[3]Whether a number word refers to an exact or approximate numerosity will depend on various characteristics of the number. Frequently used numbers that fall on decades and multiples of five tend to be more likely candidates for the approximate sense. Hence, uttering "I have 77 pennies" indicates exactly 77 pennies, whereas "I have 50 pennies" may be taken as either an exact or approximate indication of numerosity (see Pollmann & Jansen, 1996).

vant to the work we next report, it has implications for the process whereby learning is achieved. If numbers have "exact" semantics (with or without a contextual parameter), their acquisition should show signs of early sensitivity to exact numerosities (for related discussion, cf. Musolino, 2004; Papafragou & Musolino, 2003; Papafragou & Schwarz, in press). Importantly, the linguistic theories discussed above maintain that numbers and quantifiers are different in terms of their semantics, although they often appear to have surface similarities once pragmatic implicatures come into play. In the course of acquisition, children and adults may be differentially sensitive to the logical meaning and the pragmatics of utterances (Hurewitz et al., 2000; Noveck, 2001; Papafragou & Musolino, 2003). Given the surface similarities between numerical and quantified expressions, it is of interest to determine when children are able to detect these semantic differences. A developmental stage where numbers and quantifiers are interchangeable (as posited by those who suggest number expression are bootstrapped from quantifiers) may actually hinder the acquisition of quantity terms because the underlying semantics of the various quantity expressions would have to be revised later in development.

A direct comparison between the acquisition of number words and quantifiers in young learners is required to test the two competing hypotheses sketched above. So far, experimental evidence on children's acquisition of the semantics of individual quantifiers is limited, and has not addressed the parallel with numbers systematically (for different perspectives, see Badzinski, Cantor, & Hoffner, 1989; Brooks & Braine, 1996; Crain & Thornton, 1998; Drozd & van Loosbrook, 1998; Gualmini, 2003; Inhelder & Piaget, 1964; Meroni, Crain, & Gualmini, 2000; Neimark & Chapman, 1975; Papafragou & Schwarz, in press; Philip, 1996). Here we report results from an experiment that explicitly compared comprehension of several numerical (*two, four*) and quantificational (*some, all*) expressions in 3-year-olds to determine children's early conjectures about their meaning. We had dual ambitions in this new experimental work. First, we wanted to generate data on how very young children (and adult controls) interpret numerals and quantifiers under conditions where scalar interpretations would be plausible. Second, we hoped that these data, closely analyzed, would help resolve some of the issues about the early nature of number knowledge.

EXPERIMENT

Method

Participants

Twenty-six children ages 3;0 to 4;0 were recruited and tested at day cares in the Philadelphia and New Brunswick metropolitan areas. Data from 2 children were excluded from analysis; 1 left the daycare before completing the second session, and the other was unable to answer any of the pretest items correctly. Children were divided into two groups of equal size based on age: Group 1 consisted of 12 children between 35 and 42 months ($M = 37.8$, $Mdn = 37.4$), and Group 2 included 12 children between 43 and 48 months ($M = 45.2$, $Mdn = 45.1$). Ten University of Pennsylvania undergraduates served as controls; they received extra credit for a psychology course for their participation.

Procedure

The experiment was delivered in two experimental sessions with the interval between sessions ranging from 2 to 10 days. Adults completed the experiment in a single session.

Pretest. Before children responded to questions pertaining to quantities, we tested whether they could understand the rules of a matching task involving stickers. The experimenter described this "sticker game" as follows: "You are going to put a sticker on the picture that is the best match with the sentence that I am going to say." To encourage children to choose the best answer rather than any related picture, pretest items contained both perfect matches and loose but nonoptimal matches: for example, Spoken sentence: *The mouse is eating the cheese*; Pictures: (a) a mouse standing next to cheese, (b) a mouse eating cheese, (c) a mouse with a bicycle, (d) a boy with a bicycle. Those who chose the nonoptimal answer were then asked to choose the "best picture" and given a second opportunity (with a new sentence). If necessary, children were given a third opportunity to try training items before going on to quantity questions. At the second experimental session children were given one or two pretest questions of this kind to review the task.

Experimental sticker task. Children heard sentences with either a quantifier (*some* or *all*) or number (*two* or *four*)—for example, *The alligator took some/all/two/four of the cookies* (see Figure 1 for corresponding pictures). Responses were coded by the experimenter at the end of the session on the basis of the location of the sticker. There were no cases where sticker location was ambiguous.

Stimuli. Twelve groups of pictures were designed, with each related group of pictures printed on a single laminated page of a "sticker book" (each scene was in a different quadrant of the page, as in Figure 1). Each page of pictures contained one picture that could be characterized as indicating the complete set, a response that indicated a partial set, one where none of the objects were affected, and one where the picture contained only the subject of the sentence. Each picture page was presented with each of the possible quantity expressions in a between subjects design, such that all items were run in all conditions (i.e., one child would hear *The alligator took four of the cookies*, whereas another would hear *The alligator took some of the cookies*, etc.). See the Appendix for a complete listing of stimulus sentences.

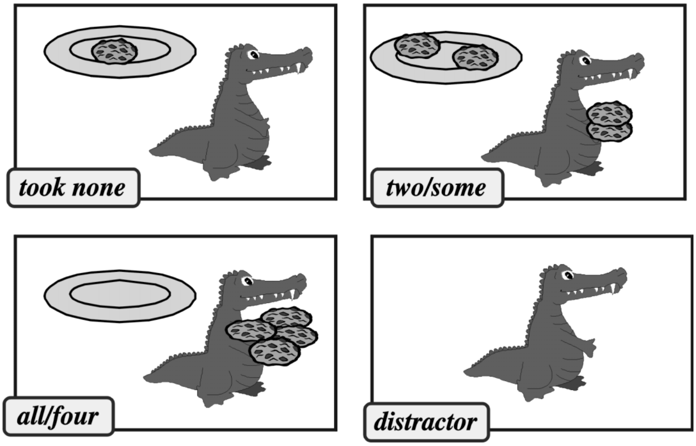

FIGURE 1 The alligator took some/all/two/four of the cookies.

All of the sentences used the partitive construction ("[number/quantifier] *of the Xs*"). Four blocked lists were created, such that each list contained all four quantity words, but in different orders (e.g., *some, some, some, four, four, four, two, two, two, all, all, all*). In each of the two experimental sessions, children were run on two of the four item blocks (one number block and one quantifier block).

Counting task. After the second session of the sticker task was administered, children were asked how many objects were placed in a bucket. They were presented with groups of 1 to 5 objects in a random order. The group of objects remained visible throughout each trial. Responses were recorded after each trial.

Analysis

Unsurprisingly, the adult controls were perfect in these tasks. For each item, they interpreted the sentences using Gricean rules (Grice, 1975); for example, for *all* and *four* they chose the set that included the complete array of all four items, and for *two* and *some* they chose the partial set containing two items (the upper bounded reading).

Child responses were coded as correct if they matched adult response patterns (i.e., upper bound choices were counted as correct). Stimulus list order was found to have no effect on the results and was dropped from further analyses. Child data were submitted to an ANOVA analysis in which the factors were Quantity Type (number or quantifier) × Set Type (partial set or complete set) × Age Group (younger or older threes). Results indicated no main effect of age for the two child age groups, $F(1, 22) = 1.45$, *ns*. Overall, children gave adult-like responses on 82% of the number items, and 65% of the quantifier items, yielding a main effect of Quantity type, $F(1, 22) = 11.97$, $p < .005$. This indicates that children were significantly more accurate in the number condition than in the quantifier condition. For set type, responses were adult-like on 70% of the partial set responses and 77% in the complete set responses, $F(1, 22) = 2.75$, *ns*. Figure 2 shows the proportion of upper bound responses for each condition. As illustrated by this graph, the interaction between set type and quantity type was significant owing to an increase in complete set responses in the *some* condition, $F(1, 22) = 8.74$, $p < .01$. That is, children tended towards an "at least some, possibly all" reading for *some* but a "no more than two" reading for *two*.

Individual results were examined by setting a criterion for competence for each quantity word—children needed to respond correctly to two out of three items for a given quantity word to meet the criterion (see Table 1). Findings indicated that 20 out of 24 children met the criterion for *four* and 22 out of 24 met the criterion for *two*. Although 19 out of 24 met the criterion for *all*, only 9 out of 24 did so for *some*. Notably the vast majority of the errors on *some* (25 out of 32 trials with errors) were selections of the *all* picture. Errors on the *all* condition appeared to be less systematic: Out of the 17 errors, on 9 trials the partial set response was chosen, on 2 trials the distractor item (i.e., no cookies in the picture) was chosen, and in 5 trials the "none" response was chosen (The alligator takes none of the cookies).

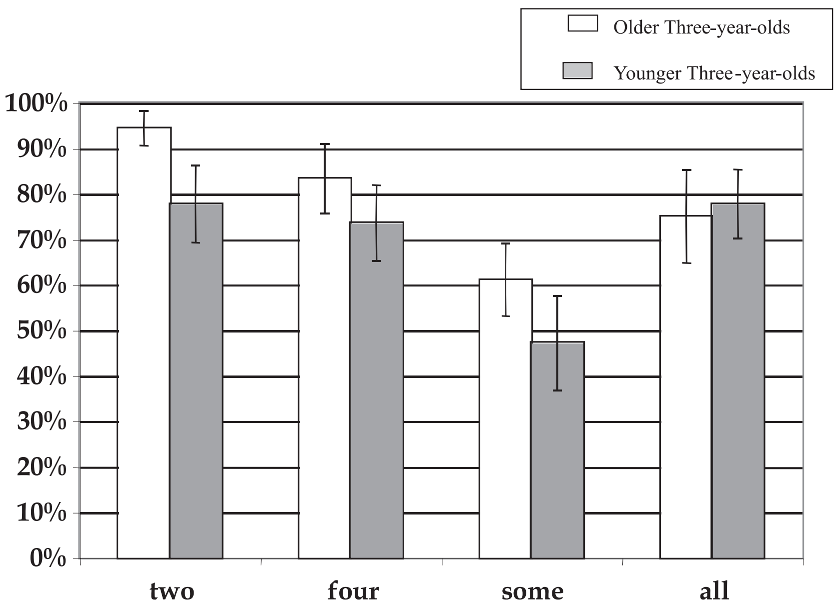

FIGURE 2 Proportion of upper bounded responses by condition.

For *two* trials, 7 out of the 11 errors involved the child choosing the complete set of four, possibly indicating counting confusion. For *four* trials, 13 of the 15 errors referred to the partial set (the set of two). In summary, the child groups were at chance in choosing an upper bounded interpretation for the word *some* but well above chance in choosing an upper bounded (or possibly exact) interpretation of the word *two*. Furthermore, the children demonstrated competence in comprehending *all* and *four*. The "how many" task showed results consistent with the general literature on number acquisition, namely shakiness in counting with even quite small numbers. In general, mistakes were the result of miscounts (i.e., counting one item twice or leaving an item out). A few of the children who had faulty performance failed to count aloud and appeared to "guess" the correct number. For the purpose of analysis, the highest number that the child could correctly respond to was coded. As Table 1 shows, 10 of the 24 children did not accurately count sets of four or five. Performance on this counting task was then compared to performance on the number sticker task (see Table 2). The one to three counters were compared to the four to five counters for each of the four conditions of the sticker task, using the subject criterion described above (children were determined to be *some*-knowers if they got two out of three some trials correct, etc.). None of these comparisons revealed a significant difference across the two groups of counters (using Fishers Exact Test, $p > .2$ in all four comparisons). Of the 10 children who did not accurately count sets of four or five, only 1 failed to reach criterion for the number four in the quantity sticker task. Hence, performance on the "how many" task (in essence, a traditional test for numerical competence) did not predict performance on the numeral items in the sticker task.

TABLE 1
Number of Child Participants Reaching Criterion (2 out of 3 Items Correct)

	Two	*Four*	*Some*	*All*
36 to 41 months	9	10	4	10
42 to 48 months	12	10	7	9

Note. $N = 12$ children in each age group.

TABLE 2
Maximum Number That Received Accurate Count in "How Many" Task

	One	*Two*	*Three*	*Four*	*Five*	*Total*
36 to 41 months	0	2	3	1	6	12
42 to 48 months	1	0	4	2	5	12
Total	1	2	7	3	11	24

Discussion

The experimental tasks just reported showed no dramatic changes over the (admittedly narrow) age range studied: Age-group results do not indicate that children shift strategies in how they interpret number words or quantifiers as they develop from 3-year-olds to young 4-year-olds. Children throughout this age range interpreted number words as exact magnitudes and were able to accurately match numbers to the corresponding pictures. Even those children in our sample who were challenged by the "how many" task (demonstrating an inability to accurately count a group of four objects) had no difficulty with the higher number word (*four*) in the sticker task. This aspect of our results suggests that counting tasks are sufficiently challenging as to mask some protocompetence with numbers possessed by young preschoolers. This is consistent with results from several other laboratories; see, e.g., Sarnecka & Gelman, 2004).

The children demonstrated a less consistent response pattern with quantifiers in the sticker task. In particular, *some* tended to receive varied responses within individual children and across children. These findings suggest that these children interpreted *some* to mean "at least some, possibly all." Only 11 of our 24 three-year-olds reached criterion (i.e., two out of three correct) in responding to s*ome* in its scalar interpretation (as "not all"), whereas our adult controls applied this interpretation invariably.

Although the observed differences between numbers such as *two* and quantifiers such as *some* could be explained as a result of children's failure to obey pragmatic principles (i.e., derive the scalar implicature) from the use of *some* but not *two*, such an account would not be parsimonious. Given that the experimental materials were exactly the same, and making the assumption that the logic of the numerical and quantificational scales is identical, it is hard to see why children would fail with the one but not with the other. This explanatory failure suggests that the assumption of sameness in the logic of these scales must be abandoned.

An alternative possibility is that children have a different pattern of responses from adults because they took the relevant set referenced by the definite description (e.g., *the cookies*) to extend beyond the set of four found in the pictures. When hearing "The alligator took some of the cookies," one may interpret "the cookies" to mean the set of all cookies in the house (including some extra ones hidden in the cupboard), or the even broader set of all cookies currently in existence. A child who chooses such an expansive set as a referent for the definite description would be correct in selecting either the *two–some* picture or the *four–all* picture in response to *some* trials.

Even though initially plausible, this alternative explanation of our findings is not satisfactory. To begin with, our materials were explicitly designed with the aim of excluding this more expansive analysis. The use of the partitive construction made it less likely that an expansive set reading was intended (in "X of the cookies," the noun phrase is presumed to refer to a uniquely identifiable group of cookies, not all of the cookies in the world; Gundel, Hedberg, & Zacharski, 1993). Furthermore, our stimulus pictures, where possible, delineated a well-defined set. For example, in the *four–all* picture for "The girl picked all of the flowers," the girl is shown holding four flowers, and four broken stems can be seen in a flowerpot nearby. Finally, if children were consistently interpreting the definite descriptions to be the set of all such objects in the world, it is unclear why they had reasonable performance on the *all* trials. For these trials the target picture showed the main character taking, eating, or carrying four objects, not more. Therefore we feel it is unlikely that children were consistently responding on the basis of an expanded referential set, although ambiguities in the extension of the sets could of course be responsible for some of the variance in child performance.

Our general findings are consistent with previous experimental results with older children indicating that children have difficulties with the computation of scalar implicatures for some types of quantity expressions (Noveck, 2001; Papafragou, 2003; Papafragou & Musolino, 2003). We take our results to reflect a *difference* in the underlying semantics of the two classes of quantity expressions—specifically, to demonstrate that numbers, unlike quantifiers such as *some*, lack "at least" semantics, and instead map onto a system of exact numerosities. We return to this issue in more detail next.

GENERAL DISCUSSION

The present findings suggest that for children as young as 3 years of age, numerals—unlike quantifiers such as *some*—resist "at least" interpretations and map instead onto precise numerosities. These results coincide with converging evidence that number terms and quantifiers are handled differently by the language acquisition system. Recall that the quantifier boostrapping hypothesis predicts that children will go through a stage when numerals (including, at various points, two, three, and four) mean something akin to "a lot" or "some" (Carey, 2004). Were this to be the case, we would expect the younger group of children in our experiment, or the less adept counters, to be more likely to allow lower bounded interpretations of number terms. This type of error rarely occurred. Even the less competent counters in the "how many" task did not assign vague meanings to numerals in the sticker task. Instead, children appeared uniform in their tendency to keep numbers as exact, while quantifiers were allowed the more vague, "at least" interpretation (resulting in more variance in performance in the "some" condition). The current study fits in well with current theoretical frameworks that suggest that number words are mapped into a dedicated, domain specific counting-to-magnitude system (Gallistel & Gelman, 1992; Gelman & Cordes, in press). In general, we found no evidence of a stage where numbers and quantifiers are collapsed into a single linguistic category.

An early differentiation between numbers and quantifiers makes good theoretical sense: Even though both numbers and quantifiers share an underlying scalar structure, the internal principles governing the numerical and the quantificational scales differ in their combinatorics, and it is these combinatorics that matter for these systems. For instance, although one can add 3 to 7 to produce 10, one cannot add *some* to *many* to produce *all.*[4] The same distinction applies when we note that quantifiers lack a distinctive "next": One can coherently introduce *a lot* between *many* and *most,* but such insertions are not acceptable within the number scale. More generally, unlike the numerical scale, the quantificational scale lacks an ordering rule that strictly and completely determines the internal structure of the scale and the positioning of its members. Relatedly, it lacks the successor function which yields the next member of the scale through the rule "$n + 1$." From a learning standpoint, given that numbers and quantifiers follow different principles, it is difficult to see how the set of number terms could begin life without true numerical denotations and later "grow" such denotations on the basis of input evidence.

The present findings are consistent with other recent results that suggest that very young children acquire numbers and quantifiers using different mechanisms. Sarnecka and Gelman (2004) explicitly set out to test if 2- and 3-year-olds believe that number terms apply to unique specific quantities. In one task participants were asked to assess the quantity of pennies in two bowls, using either a quantifier or a number term. In the number condition, the experimenter initially stated that she was placing *six* pennies in each bowl. Later, the experimenter poured additional pennies into one of the bowls. She then asked the children which bowl contained *six pennies*. These very young children showed a clear tendency to correctly choose the untransformed bowl. However, when the experimenter used the words *a lot of pennies* in both the initial and final phase of the experiment, children preferred the transformed bowl (the one with extra pennies). So *six* plus some other numerical quantity is not itself *six*, whereas *a lot* plus some other numerical quantity is still a lot—in fact it is an even better exemplar of *a lot*.

Crucially, many of the children tested by Sarnecka and Gelman (2004) failed a more traditional cardinality task. Data from the "give-a-number" task (as described by Wynn, 1990, 1992) are often cited to support the view that 2-year-olds initially do not assign cardinal values to numbers beyond two. This task involves having children give a puppet X items from a pile of objects. A standard finding is that younger children can correctly complete the task for the cardinal value *one*, but they randomly grab objects when *two* or more is requested (at a slightly later stage, they are correct for both one and two, but fail at three or more). Sarnecka and Gelman gave the children who participated in the

[4]As pointed out by one anonymous reviewer, *some* plus *many* does not even add up to *most*.

pennies task a rendition of the give-a-number task using the count words *one*, *three*, and *six*. The children were separated into two groups on the basis on their performance on give-a-number results. When separately analyzed, the tendency to assign a specific value for *six* but not for *a lot* held up for both groups of children, even those who failed to indicate specific knowledge of any number over *one* on a give-a-number task pretest.

Summarizing, children who demonstrated poor performance on a give-a-number task still applied different rules to quantifiers and numbers. They were willing to accept that *a lot* may refer to either six pennies, or to more than six pennies. Yet they interpreted the word *six* to mean a constant amount (namely, six) even when the context changes.

Another recent study made a similar point: Huang, Snedeker, and Spelke (2004) presented 2- to 3-year-old children with open boxes of one and three fish, as well as a closed box with an unknown number of fish. Children who were categorized as *two*-knowers via a pretest would not accept that *two* applied to the open three-fish box. In choosing the box that matched the sentence, children preferred the closed box for the number *two* (= "exact" interpretation). Evidently, children primitive enough to be shaky with numerical tasks using numbers as small as two and three nevertheless do not suppose that all quantity terms are of the same kind.

In this context, it is useful to consider the actions of those children who failed to count to four accurately in the "how many" task of our experiment. One might expect under the quantifier-bootstrapping hypothesis that these children, inexpert in the application of the term *four*, would take that term to mean a vague, context dependent magnitude such as *some*. Yet, if this were the case, the results should have indicated chance performance on the *four* condition of the sticker task. This did not occur. The one to three counters were overwhelmingly accurate in the four condition, whereas at chance in the *some* condition. We conclude that the one to three counters that we tested do not equate *four* with *some*, although it is possible that they equate *four* with an unknown, specific magnitude as conjectured by Sarnecka and Gelman (2003).

Finally, we should mention other experiments that have explored scalar properties of numbers and quantifiers in older preschoolers. Papafragou and Musolino (2003) presented 5-year-old children with a scenario in which a group of three horses jumped over a fence. Children almost always rejected the statement *Two of the horses jumped over the fence* as a potential description of the story; however, under the same conditions, children were much less likely to reject the statement *Some of the horses jumped over the fence*. That is, these children overwhelmingly accepted lower bounded interpretations for quantifiers such as *some* but not for number words such as *two*—precisely the asymmetry which surfaced for much younger children in the present experiments. Other recent findings suggest that there is a broader asymmetry between exact and inexact scalars: *half* patterns together with the numbers in prompting "exact" interpretations, whereas *most* patterns together with *some* (Papafragou, 2002; Papafragou & Schwarz, in press).[5]

One crucial question raised by the current and aforementioned studies is why children fail to derive pragmatic implicatures in many cases where adults would compute them. That is, even if we adopt the view that child performance in the present study reflects the basic semantics of the quantifier terms, we still need an explanation for the child's nonapplication of pragmatic principles. Recent studies indicate that children's difficulties with scalar inference extend to a variety of scalar terms including aspectual verbs (e.g., *I started doing my essay* → I didn't finish my essay; Papafragou, 2003) and modals (*The marble may be in the blue box* → It is not certain that the marble is in the blue box; Noveck, 2001). However, the same studies indicate that child-insensitivity to scalar implicatures can be overcome in situations in which prior discourse sets up clear informativeness expectations (Papafragou, 2003; Papafragou & Tantalou, 2004). For instance, children know that a character who was given a few oranges and is later asked whether he ate them has not, in fact, eaten all of the oranges if he responds that he ate some. This evidence is consistent with results from developmental sentence-processing studies (Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Trueswell, Sekerina, Hill, & Logrip, 1999). These studies show that children fail to derive context-based inferences online, even when a rich visual scene supports such an interpretation: For instance, in referential tasks, 4- to 5-year-old children fail to credit speakers with the pragmatic necessity for uniqueness in definite reference based exclusively on visual scene information. However, these children will make adult-like uniqueness inferences when discourse contrast, as triggered by a relevant question, is used (Hurewitz, Trueswell, & Gleitman, 2003; cf. also Musolino & Lidz, 2004). Taken together, this evidence supports the view that although children may derive pragmatic inferences, they require strong discourse support (not just visual scene or situational support) in order to do so.

Given this pattern of results, we may surmise that the performance of the 3-year-olds in our sticker task most likely reflects a semantic or logical value for the quantifier term used (also see Noveck, 2001). Our findings reveal that this value is in fact different for quantifiers and numbers. Converging evidence suggests that despite the distributional and semantic similarities between numbers and quantifiers, children recognize the disparate characteristics of these quantity expressions during early development, prior to the age where they can be said to be completely competent with the use rules of the first five integers.

[5]This is not to say that children lack lower bounded interpretations for exact scalars; in fact, they can access "at least" interpretations for both numbers and *half* (see Musolino, 2004; Papafragou, 2003, respectively). The point is that these interpretations are more dispreferred than "at least" interpretations for the scalar quantifier *some.*

Taken together with other recent research, the present findings for 3-year-old children have implications for children's early conjectures about the meanings of number words. Specifically, they are hard to reconcile with the hypothesis that number terms and quantifiers start out as a single categorial type. Indeed our findings would be more consistent with the idea that number words are mapped onto a dedicated magnitude system with exact semantics (a system which represents exact and unique numerosities; Gelman & Cordes, 2001; Gelman & Gallistel, 1978). On this picture, those processes that underlie the acquisition of number words are distinct from the mechanisms responsible for learning and evaluating (non-numerical) quantified expressions.

Finally, the present results have implications for the semantic representation of number terms and quantifiers in the adult grammar. Recall that, according to standard semantic-pragmatic accounts, scalar expressions have a lower bounded ("at least *n*") semantics that is upper bounded by a scalar implicature ("no more than *n*"); more recent accounts, however, acknowledge that the scalar class is internally split. On these more recent theoretical accounts, number words in natural language are not semantically lower bounded but have an underspecified meaning that typically yields "exact" interpretations in context. Our own data strengthen the case for revising the traditional semantics of numbers in the adult grammar in line with these more recent proposals. Three-year-olds' interpretation of numbers thus turns out to be informative for the formulation of semantic principles underlying number knowledge in the adult grammar.

ACKNOWLEDGMENTS

This research was supported by a training grant to the Rutgers Center for Cognitive Science, NIMH 32 MH019975, postdoctoral funds to Rochel Gelman, and a Spencer Fellowship to Felicia Hurewitz. We would like to acknowledge Julien Musolino for supplying the inspiration and three anonymous reviewers for substantial improvements to the content and form of this article. This research would not have been possible without the generous cooperation of the children, parents, and staff of the Magic Years Preschool, PennWynne Presbyterian Church Preschool, the JCC Kaiserman Branch Preschool, and the Beacon Center for Children.

REFERENCES

Badzinski, D. M., Cantor, J., & Hoffner, C. (1989). Children's understanding of quantifiers. *Child Study Journal, 19,* 241–258.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy, 4,* 159–219.

Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language, 24,* 511–533.

Breheny R. (2004, November). Some scalar implicatures really aren't quantity implicatures—But some are. In *Proceedings of the 9th annual meeting of the Gesellschaft für Semantik (Sinn und Bedeutung VIIII)*. Nijmegen, The Netherlands: University of Nijmegen.

Briars, D., & Siegler, R. S. (1984) A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20,* 607–618.

Brooks, P., & Braine, M. D. S. (1996). What do children know about the universal quantifiers *all* and *each*? *Cognition, 60,* 235–268.

Carey, S. (2004). Bootstrapping and the origins of concepts. *Daedalus, 133*(1), 59–68.

Carston, R. (1995). Quantity maxims and generalised implicature. *Lingua, 96,* 213–244. (Reprinted from *UCL Working Papers in Linguistics, 2,* 1–31, 1990)

Carston, R. (1998). Informativeness, relevance, and scalar implicature. In R. Carston & S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 179–236). Amsterdam: Benjamins.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar.* Cambridge, MA: MIT Press.

Drozd, K., & van Loosbrook, E. (1998). Weak quantification, plausible dissent, and the development of children's pragmatic competence. In *Proceedings of the 23rd Boston University Conference on Language Development* (pp. 184–195). Somerville, MA: Cascadilla Press.

Frege, G. (1974). The foundations of arithmetic: A logic-mathematical enquiry into the concept of number (2nd ed. rev., J. L. Austin, Trans.). Oxford, England: Blackwell. (Original work published 1884)

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

Gallistel, C. R., & Gelman, R. G. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74.

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

Gelman, R. G. (1972). Logical capacity of very young children: Number invariance rules. *Child Development, 43,* 75–90.

Gelman, R. G. (1993). A rational-constructivist account of early learning about numbers and objects. In D. Medin (Ed.), *Learning and motivation* (Vol. 30, pp. 61–96). New York: Academic Press.

Gelman, R. G. (1998). Domain specificity in cognitive development: Universals and nonuniversals. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in psychological science: Vol. 2. Biological and cognitive aspects* (pp. 50–63). Hove, England: Psychology Press.

Gelman, R. G., & Butterworth, B. (2005). Number and language: How are they related? *Trends in Cognitive Science, 9,* 6–10.

Gelman, R., & Cordes, S. A. (2001). Counting in animals and humans. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 279–301). Cambridge, MA: MIT Press.

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

Gleitman, L. R. (1990) The structural sources of word learning. *Language Acquisition, 1,* 3–55.

Goldberg, A. (2004, in press). But do we need universal grammar? Comment on Lidz, Gleitman, and Gleitman, 2003. *Cognition*.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic.

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

Grinstead, J., MacSwan, J., Curtiss, S., & Gelman, R. (1997). The independence of language and number. *Proceedings of the 22nd Boston University Conference on Language Development* (Vol. 1, pp. 303–313). Somerville, MA: Cascadilla.

Gualmini, A. (2003). *Some* knowledge children don't lack. In *Proceedings of the 27th Boston University Conference on Language Development* (pp. 276–287). Somerville, MA: Cascadilla.

Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69,* 274–307.

Horn, L. (1972). *On the semantic properties of the logical operators in English*. Unpublished doctoral dissertation, University of California, Los Angeles.

Horn, L. (1992). The said and the unsaid. *SALT II,* pp. 163–192.

Horn, L. (2004). Implicature. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 3–28). Oxford, England: Blackwell.

Huang, Y. T. , Snedeker, J., & Spelke, E. (2004, August). *What exactly do numbers mean?* Paper presented at the 26th annual meeting of the Cognitive Science Society, Chicago.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research, 29,* 597–626.

Hurewitz, F., Trueswell, J., & Gleitman, L. (2003, April). *The development of sentence parsing ability: A probabilistic multiple-constraint account*. Paper presented at the conference for the Society for Research in Child Development, Tampa Bay, Florida.

Inhelder, B., & Piaget, J. (1964). *Early growth of logic in the child: Classification and seriation.* London: Routledge.

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

Landau, B., & Gleitman, R. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.

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

Meroni, L., Crain, S., & Gualmini, A. (2000). A conservative approach to quantification in child language. In *Proceedings of the 24th Penn Linguistics Colloquium, University of Pennsylvania Working Papers in Linguistics, 7.1* (pp. 171–182). Philadelphia: University of Pennsylvania.

Mitchie, S. (1984). Why preschoolers are reluctant to count spontaneously. *British Journal of Experimental Psychology, 2,* 347–358.

Mix, K., Huttenlocher, J., & Levine, S. (2002). *Quantitative development in infancy and early childhood.* New York: Oxford University Press.

Moxey, L. M., & Sanford, A. J. (1993). *Communicating quantities: A psychological perspective*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Neimark, E. D., & Chapman, R. H. (1975). Development of the comprehension of logical quantifiers. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 135–151). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Papafragou, A. (2003). Aspectuality and scalar structure. *Proceedings of the 27th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

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

Papafragou, A., & Schwarz, N. (in press). *Most* wanted. *Language Acquisition*.

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

Philip, W. (1996). The event quantificational account of symmetrical interpretation and a denial of implausible infelicity. In *Proceedings of the 20th Boston University Conference on language Development* (pp. 564–575). Somerville, MA: Cascadilla Press.

Pollmann, T., & Jansen, C. (1996). The language user as an arithmetician. *Cognition, 59,* 219–237.

Sarnecka, B., & Gelman, S. (2003). *Six* does not just mean *a lot*: preschoolers see number words as specific. *Cognition, 92,* 329–352.

Shatz, M., & Backscheider, A. (1991, November). *Acquiring the normative concepts of color and number.* Paper presented at the meeting of Psychonomic Society, Seattle, WA.

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.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155–193.

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

Zur, O., & Gelman, R. (2004). Young Children can add and subtract by predicting and checking. *Early Childhood Quarterly Review, 19,* 121–137.

APPENDIX
Stimulus Sentences

The horse wore some/all/two/four of the boots.
The alligator took some/all/two/four of the cookies.
The man dropped some/all/two/four of the ice cream cones.
The lion blew out some/all/two/four of the candles.
The child caught some/all/two/four of the butterflies.
The boy colored in some/all/two/four of the stars.
The turtle carried some/all/two/four of the ladybugs.
They washed some/all/two/four of the pigs.
The girl picked some/all/two/four of the flowers.
The bird popped some/all/two/four of the balloons.
The crab knocked over some/all/two/four of the sand castles.
The elephant ate some/all/two/four of the apples.