Quantification and learnability

Early mastery of the weak–strong distinction

Irene Krämer
Radboud University Nijmegen

1. Introduction

This paper discusses the question of whether we should assume that children, like adults, distinguish two classes of determiner quantifiers that are commonly referred to as "strong" and "weak". This issue has been raised as a consequence of recent proposals that children initially have only one class of quantifiers, or, minimally, that the border between the classes is fuzzy (Drozd 2001, Musolino et. al. 2000, Geurts 2003).

The present paper points out disadvantages to a one-class approach and presents an experiment which compares children's interpretations of strongly quantified Noun Phrases (NPs) like *all eggs* and *most flowers*, to the interpretation of weakly quantified *many eggs*. The results do not support the one-class approach, but rather indicate that children at the relevant ages distinguish between weak and strong quantification.

2. Strong and weak NP quantification

Weak determiners like *one, two, some, many* are logically distinct from strong determiners like *all*, *every*, or *most*. The former, but not the latter, are existential, symmetrical and intersective (Barwise and Cooper 1981, Keenan and Stavi 1986). The difference can be illustrated by the role that intersectivity can play in a verification procedure for the sentence *Two cats are black*. To establish whether this sentence is true, one needs to consider only the intersection of the sets denoted by both of the quantifier's arguments. If one thus looked at

Linguistics in the Netherlands 2005, **111–123**.
ISSN 0929–7332 / E-ISSN 1569–9919 © Algemene Vereniging voor Taalwetenschap

the intersection set of cats and black individuals, and if one were to find two cats in this set, one would rightly consider the sentence true. The procedure of only considering the set of black cats cannot be applied to a statement like *All cats are black*. For a strong quantifier, we need to relate the intersection set of cats and black individuals to the entire set denoted by the first argument: cats. If all individuals in the latter set are also in the intersection set, the statement is true.

The distinction between strong and weak quantifiers is reflected in the form and interpretation of language in several ways. For example, weakly quantified NPs may appear in the existential construction, whereas strongly quantified NPs are not, or only rarely, felicitous in this construction (*There were some/two/three/*all/*most women in the street*). In addition, weakly quantified NPs can be extracted from syntactic islands, whereas strongly quantified NPs cannot. This accounts for the different interpretive possibilities of sentences (1a) and (1b):

(1) a. Every English teacher praises a pupil who reads a classic.
    b. Every English teacher praises a pupil who reads each classic.

For sentence (1a), we can obtain the interpretation that there is an English classic, say, Jane Eyre, and every teacher praises a pupil who reads Jane Eyre. A similar interpretation, in which the embedded object NP *each classic* is extracted from the relative clause island, is impossible for (1b): *For each classic (Jane Eyre, A Pilgrim's Progress, Hamlet, …) every teacher praises a pupil who reads it.

We thus see that strong and weak quantifiers are logically distinct classes that show coherent behaviour on a linguistic level. However, certain phenomena in children's interpretation of quantified NPs have led to the hypothesis that this is not the case, or at least not entirely so, for child language.

3. The acquisition of quantification

3.1 Weakly quantified NPs

Several studies report on the interpretation of weakly quantified NPs by children aged 4 and older. They find that these children have a preference for a narrow scope interpretation of weakly quantified object NPs such as *two pizzas* and *some horses*: the children generally reject a wide scope interpretation in a context in which such readings are accepted by adults. For example, the following sentence from Musolino et. al. (2000) is accepted by adults in a situation in which Cookie Monster, though he has eaten two pizzas, has left two pizzas untouched. As a result, there are two pizzas that he did not eat.

(2) Cookie Monster didn't eat two pizzas.

Four- and 5-year old children do not accept the sentence, and in their explanations clearly provide evidence of narrow scope interpretations. Similar results were found for Dutch (Klein 1996, Krämer 2000) and the Dravidian language Kannada (Lidz and Musolino 2002). The explanation of Musolino and colleagues is that children initially assume that all quantifiers belong to the class of strong quantifiers, such that they only take scope in situ.

3.2 Universal quantification

When children aged 4 to, roughly, 7 are presented with one of the following pictures, and asked whether "all children are riding an elephant", their responses may differ from those of adults.

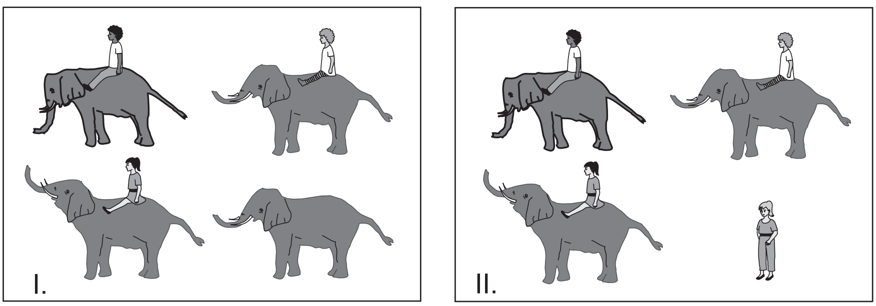

Figure 1. Two picture types that lead to children's errors in universal quantification.

Whereas adults will say "yes, true" in the case of picture I, many children (between 35% and 80%, depending on the study) will say "No", and, as an explanation, point to the elephant that no child is riding. This error has been called "overexhaustive search". Picture II sometimes leads to "underexhaustive search", with children agreeing that the statement is true, whereas it is false for adults (this error is less well researched, but the error rate is usually considerably lower than that of overexhaustive search). The two analyses I will discuss attempt to explain both errors from the same underlying difference between children and adults.

Drozd (2001) proposes that the errors are due to some difficulty connected to the presupposition carried by the strong quantifier. As a result of this difficulty, children interpret the strong quantifier as if it were weak: they only consider the intersection set of children and elephant-riders. Any children who are not part of this intersection set are not taken into consideration. This procedure would lead to underexhaustive search. In order to explain overexhaustive search, Drozd makes the additional assumption that *all* is interpreted in a similar way to the weak quantifier *many*. To evaluate the statement *Many children are riding an elephant*, one must employ a context-dependent expected value. If the intersection set of children and elephant-riders contains more elements than one would have expected, the statement is true (Westerståhl 1985). According to Drozd, children judge the universal statements according to an expected value based on the number of elephants.

Geurts (2003) proposes that children experience difficulty parsing universally quantified NPs. As a result, they apply the interpretation procedure appropriate to weak quantification to universal quantifiers, because this allows a less complicated mapping from syntax to semantics. The core details of Geurts' proposal are different and elaborated in more detail than the proposal by Drozd, but both converge on the point that children treat strong quantifiers as if they were weak.

Drozd's and Geurts' proposals are orthogonal to Musolino's. While the former propose that children treat strong quantifiers as weak, the latter proposes that children treat weak quantifiers as strong. However, the similarity between the approaches is clear: for children, the borders between the classes of strong and weak quantifiers are minimally vague, and possibly non-existent. Such a one-class approach may explain the specific errors in children's interpretation discussed above, but it has numerous disadvantages. These will be discussed in the following section.

4. Disadvantages of a one-class approach to the acquisition of quantification

The assumption that children's errors in quantification are due to a failure to distinguish weak and strong determiner quantifiers has three kinds of disadvantages. These concern the theory of grammar, learnability, and the position of the acquisition of (non-lexical) semantics within the whole of language acquisition.

4.1 An economical theory

As discussed in Section 2, weak and strong quantifiers can be distinguished both by logical properties and by their behaviour with respect to syntax and interpretation. The assumption that quantifiers of one class behave as if they belonged to the other class thus breaks up two independently motivated, coherent classes, diminishing the elegance and economy of our model of the grammar. Reinhart (1995) pointed this out when discussing theories of quantification in adult language, but the point can equally be made for child grammar.

4.2 Learnability

The proposals by Drozd and Geurts face a classical learnability problem. Both analyses assume that the children's weak interpretations are motivated by the relative ease of interpreting weak quantifiers as compared to strong ones. On this reasoning, weak quantification should be easier for adults as well, and it probably is (Just 1974). Nevertheless, adults do not display the interpretations that children do. If children indeed interpreted strong, universal quantifiers in a weak manner, how would they be able to learn on the basis of positive evidence that this interpretation procedure is not allowed in the adult language?

Musolino's analysis of children's preference for narrow scope interpretations does not face such a classical no-negative-evidence problem. However, it faces a problem of decreased learnability as a result of a blurring of the distinction between two classes. For example, a child who discovers that the quantified NP *two pizzas* can take wide as well as narrow scope in example (2), can generalize this finding to other weakly quantified NPs, such as *a couple of pizzas*, or *infinitely many pizzas*, thus obviating the need to encounter similar instances of each of these quantifiers paired to wide scope interpretations. If, however, there are no distinct and coherent classes of quantifiers to the child, the behaviour of individual elements of these classes cannot be predicted from the behaviour of their class members. This increases the learning burden.

4.3 The acquisition of semantic knowledge

A third point of concern regarding a one-class approach to the acquisition of quantification is the position of the acquisition of semantics relative to other grammatical knowledge. It is generally assumed that all of adult syntax has been acquired by the age of five. We further know that children distinguish between grammatical categories such as Subject and Object, and between nouns, verbs and adjectives as early as ages two and three. The one-class approach implies rather a large lag of semantics behind such other kinds of grammatical knowledge, as errors in universal quantification have been reported to exist even at age 7 (Philip and Coopmans 1995, Kang 2001), and the absence of wide scope readings for indefinites is found as late as age 11 (Unsworth 2004).

It is certainly possible that much of semantics is acquired relatively late, just as a child seems to acquire much of morphosyntax after she has acquired most of the phonological system of her language. However, the sheer size of the lag that follows from the one-class approach may raise questions that are larger than the questions the approach attempts to answer.

5. The empirical basis for a one-class approach

As the one-class assumption has rather serious implications concerning our theory of child grammar, it would be wise to see if it finds any empirical support in contexts other than the ones presented in Section 3.

The proposal by Musolino and colleagues, that children's weakly quantified object NPs take scope in situ because they are treated as strong quantifiers, is contradicted by facts from the acquisition of Dutch. Dutch children show a similar preference for narrow scope, even though, in Dutch, this is not scope in situ (Krämer 2000, Unsworth 2004). However, the Dutch facts do not pertain to the specific issue of whether there are indeed two classes of quantifiers for children.

Smith (1980) investigated the interpretation of weak *some* and strong *all* by children aged 4 to 7. She found that children who had first heard a block of questions containing *some* were more likely to respond incorrectly to sentences containing *all* than children presented with the opposite order. This finding may be an indication that children may use a "weak" verification procedure for strong *all*, but it may also be the result of a carry-over effect that only occurs in this specific experimental set-up. Another point rendering Smith's results inconclusive with respect to the current issue is that the sentences used in her experiments were all generic, hence not very well comparable with the concrete situations in which the errors occur that were discussed above. So far, no experiments have been reported in which children were offered the opportunity to apply strong or weak verification procedures to quantified NPs in a concrete, non-generic context. The following section presents such an experiment.

6. Experiment

6.1 Method

In a Truth Value Judgment Task, Dutch children aged 4;0 to 8;4 were asked to judge sentences like (3) to (5). Each of these sentences was paired to the picture shown in Figure 2.

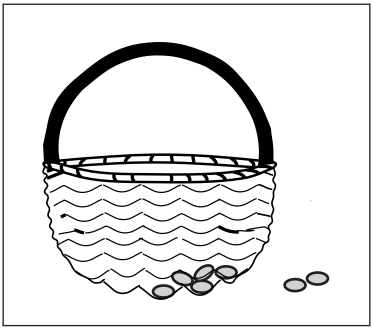

Figure 2. Picture for test-sentences (3) to (5).

(3) Alle eieren zitten in de mand.
all eggs sit in the basket
'All eggs are in the basket'

(4) De meeste eieren liggen naast de mand.
the most eggs lie next to the basket
'Most eggs are next to the basket'

(5) Veel eieren zitten in de mand.
many eggs sit in the basket
'A lot of eggs are in the basket'

To an adult, sentence (3) is false. To arrive at this judgment, one relates the intersection set of eggs in the basket to the total set of eggs in the picture. Should subjects treat *alle eieren* "all eggs" like a weak quantifier, this could lead to their checking only the intersection set. In such a case, the statement would be judged true, because the remaining eggs are ignored.

Sentence (4) is also false. The set of eggs next to the basket is related to the total of eggs, and found to contain less than half of this total. It is hard to imagine what response exactly a non-relational interpretation of *most* would lead to.

To sentence (5), there is no one correct response. The sentence can be assigned a truth value by comparing the number of eggs in the basket to some contextually determined expected value. This might be, for instance, "as much as I can eat". In that case, the sentence would be true for most speakers, as five eggs make more than one meal. The expected value may also be a cardinal value: someone may consider any number of eggs smaller than 10 to be few. This would lead to a "false" judgment. The expected value may also be related to the container of the experimental context: for *Many eggs are in the basket* to be true, the basket has to be at least somewhat full. In this case, the statement is again false of the picture shown.

All of the interpretations of *many* mentioned so far are weak, as any elements that are not in the intersection set of both of the quantifier's arguments are irrelevant to the truth value of the statement. Interestingly, *many* also allows a reading on which the size of the set of eggs in the basket is compared to the total of eggs. On this "strong" or "proportional" reading (Partee 2003), (5) is true, since most of the eggs are inside the basket. As the quantified NP is sentence-initial, this is the preferred reading, at least for adults (cf. Milsark 1977, De Hoop 1996, Krämer, in press). In sum, on a weak reading of *many*, sentence (5) may be either true or false, while on a strong reading, sentence (5) is always true.

6.2 Design, subjects and procedure

The interpretation of *all* and *many* was contrasted in a between-subjects design. Eighteen children aged between 4;0 and 5;1, and 8 children between 6;4 and 8;1 (a total of 26 children) were presented with sentences containing *alle* "all", 16 children aged between 4;0 and 5;1, and 7 children between 6;4 and 8;1 (a total of 23 children) were presented with sentences containing *veel* "many".

Subjects were tested one-by-one. The experimenter played a Sesame Street hand puppet that looked at the pictures and told the child what it saw. The child's task was to tell the puppet whether it was right or wrong. Whenever the child had responded, the puppet would ask how he or she could tell. The sessions were videotaped and transcribed.

In both conditions, there were 6 test items, which all involved pictures of either a basket or a flower pot containing 5 small items, with 2 similar items placed next to the container. They were preceded by 4 training items and interspersed with 4 filler items, two of which should lead to a true judgment, and two to a false judgment. At the end of the session, subjects in the *many*-

condition were presented with 3 control items testing for a rudimentary understanding of *many* as opposed to *few*.

Most and *many* were contrasted in a within-subjects design. Twelve children participated. Their ages were exactly in between the two age groups from the *all/many* comparison: 5;4 to 6;3. Three items with *veel* "many" preceded 3 items with *de meeste* "most". The remainder of the procedure was the same as described above. After the first three items, the three *many*-control items were presented. After the final three items, these control items were presented once more, with sentences containing *most*.

6.3 Results

6.3.1. *All versus Many*

For the *all*-condition, the data of 23 children were analysed, after 3 children from the youngest age group had been excluded due to incorrect responses on at least 2 of the 4 filler items. Twenty-two of these 23 children correctly rejected the sentences containing *all* on at least 5 out of 6 occasions. One child from the 4-year old group incorrectly accepted all test sentences. Twenty children offered explanations for their rejection responses that one might also expect adults to give, such as "No, because there are also flowers outside (of the pot)". The responses by the two remaining children were idiosyncratic, such as "No, because they are four". Thus, for 20 out of 23 child subjects (87%), we have a clear indication of an adult-like and thus presumably strong interpretation of the universally quantified NP.

For *many*, the results are different. All children showed a basic knowledge of the meaning of *veel* "many". The data of 21 children were analysed after excluding 2 children from the youngest age group on the basis of incorrect responses to the filler items. In the youngest age group, out of 14 children, 6 consistently rejected the sentences, 5 children consistently accepted the sentences, and 3 children had mixed responses. In the oldest age group, 5 children consistently rejected, and 2 children consistently accepted the sentences.

As there is no clearly correct response, we must look to the explanations that subjects provided to establish whether the interpretations are strong (proportional) or weak. Two of the children had a pattern of strong interpretations, as evidenced by responses like "Yes (they are a lot), because there are five inside, and only two outside". Sixteen children had weak interpretations, employing mainly cardinal expected values, such as "Yes (they are a lot) because they are five" or "No (they are not a lot), because they are five". Occasionally the "container" type of weak interpretation appeared, as in "No, (it is not a lot, because) there are few flowers here, and there should be ones over here because otherwise it is not full enough". The responses of 3 children could not be classified as they had idiosyncratic explanations. Thus, for *many*, 2 children (10%) have strong interpretations, and 16 children (76%) have weak interpretations.

6.3.2. *Most versus Many*

In the within-subjects comparison of *most* and *many*, the results of the *many*-condition resembled the previous comparison. Six children rejected all three test sentences, 4 accepted all three test sentences, and 2 had mixed responses. Only 1 child (8%) had strong interpretations, 7 children (58%) provided evidence of weak interpretations. Four subjects remained unclassified.

Eleven of the 12 children correctly rejected all *most*-sentences, one child accepted them. This child had rejected the *many*-sentences. She failed one of the final control items for *most*, which indicates that her incorrect responses may have been due to a lack of understanding of the lexical item itself. Thus, while on the *many*-items only 50% of the children consistently rejected the test-sentences, the *most*-items were correctly rejected by 92% of the subjects.

The children's explanations of their judgments of the *most*-items are not enlightening as to the use of either a strong or a weak verification procedure. Mostly children referred to there being (only) two elements next to the container and sometimes they referred to the elements inside the container — both of which are sensible responses from an adult point of view. Nevertheless, as rejection is the correct, adult-like response, we may assume that the children's rejection responses result from the appropriate strong interpretation procedure.

The outcomes of the *all/many* comparison and the *most/many* comparison are strikingly similar. Both for *all* and for *most*, over 90% of the children consistently rejected the test sentences, whereas for *many*, for both comparisons we find a considerable proportion of children consistently rejecting the *many*-sentences, a considerable proportion of children accepting them and some children who had mixed response patterns. Both for *all* and for *most*, around 90% of the children had strong interpretations. For *many*, only around 10% of the children had strong interpretations, and a clear majority had weak interpretations.

7. Discussion and conclusion

In the experiment presented above, children interpreted strong and weak quantifiers in the appropriate manner. The *all/many* comparison does not give us any reason to believe that the procedure leading to the adult-like responses to statements containing *alle* "all" was any different than the adult procedure. Had the children applied an intersective procedure, leaving the objects outside the container out of consideration, this would have led to accepting the test sentences. Instead, the children rejected them. Had the children interpreted *all* in a similar way to *many*, as proposed by Drozd (2001), we would have found similar response rates and explanations for *all* and *many* — but we find that they are very different.

Similarly, for the *most/many* comparison, there is no reason to assume that the adult-like responses to *most* involved any non-adultlike interpretation procedures. In addition, the results for *many* were different than those for *most*. Although *many* allows a strong reading which is very similar to *most*, the subjects preferred weak, cardinal readings. It seems that children prefer a weak reading when this is allowed, and obey the rules of adult grammar when it is not allowed.

In sum, if children aged 4 and above had the option of treating weakly and strongly quantified NPs alike, we would have expected to find some evidence of this in the experiment presented above — but we did not. Considering this finding, and considering the disadvantages of a one-class approach to the acquisition of quantification, we should conclude that these children have adult-like knowledge of the existence of two classes of determiner quantifiers. Thus, there is no need to assume that this aspect of the acquisition of grammar lags many years behind other types of grammatical knowledge.

This conclusion again raises the question of the origin of children's errors in quantification. If the source of the errors does not lie in the absence of the distinction between two classes of determiners, where, then, should we look to find an explanation? Recently, some studies have proposed that the children's errors reflect pragmatic development. De Hoop and Krämer (to appear) present an analysis of child and adult interpretations of indefinite NPs, on which the wide scope readings of object NPs involve a departure from the preferred interpretation. A hearer will make such a departure on the basis of contextual cues and insight into the speaker's intentions. De Hoop and Krämer claim that children partly lack the latter. Gualmini et. al. (2003) suggest that the origin of children's errors in universal quantification lies in a failure to determine which elements in the extension of the first argument ("children" in our example) are to be considered in a given context. Krämer (2005) presents a detailed proposal of the same kind, proposing that the errors are related to children's difficulty in handling shared knowledge.

All three analyses place children's problems with quantification squarely in the realm of pragmatics. Such an approach avoids the problems beleaguering the one-class approach. Further research will have to show whether the pragmatic approach to the acquisition of quantification will prove fruitful. For now, it is a promising direction to take, as it will allow us to relate these late errors not just to the absence or presence of grammatical principles, but also to further cognitive development, such as the development of Theory of Mind or perspective shifting.

Note

1. The results from this many-condition were also used in a comparison of the effects of visual context and syntactic structure on the interpretation of *many* (Krämer, in press). This comparison involved additional age groups and conditions, which are not relevant to the present issue.

References

Barwise, J. and R. Cooper (1981) 'Generalized Quantifiers and natural language.' *Linguistics and Philosophy* 4, 149–219.

Drozd, K.F. (2001) 'Children's weak interpretations of universally quantified questions.' In: Bowerman, M. and S.C. Levinson, eds., *Language Acquisition and Conceptual Development*. Cambridge University Press, Cambridge, 340–376.

Geurts, B. (2003) 'Quantifying Kids.' *Language Acquisition* 11, 197–218.

Gualmini, A., L. Meroni and S. Crain (2003) 'Children's Asymmetrical Responses.' In: *Proceedings of the Third Tokyo Conference on Psycholinguistics*. Tokyo, Hituzi Publishing Company.

Just, M.A. (1974) 'Comprehending quantified sentences: the relation between sentence-picture and semantic memory verification.' *Cognitive Psychology* 6, 216–236.

de Hoop, H. (1996) *Case Configuration and Noun Phrase Interpretation*. New York, Garland Publishing.

de Hoop, H. and Krämer, I. (to appear) 'Children's optimal interpretations of indefinite subjects and objects.' *Language Acquisition*.

Kang, H.K. (2000) 'Age differences in the acquisition of quantifiers: Evidence from English and Korean.' *UCL working papers in linguistics* 12.

Keenan, E.L. and J. Stavi (1986) 'A semantic characterizaton of natural language determiners.' *Linguistics and Philosophy* 9, 253–326.
Klein, Wenda (1996) Specificity in child Dutch: An experimental study. MA thesis, Utrecht University.
Krämer, I. (2005) 'Quantification takes two minds: Children's errors in Universal quantification.' Ms., Radboud University Nijmegen.
Krämer, I. (in press) 'When Does Many Mean a Lot? Discourse Pragmatics of the Weak–Strong Distinction.' *Proceedings of the 29th Boston University Conference on Language Development*.
Krämer, I. (2000) Interpreting Indefinites. An experimental study of children's language comprehension. PhD thesis, Utrecht University. Nijmegen: Max Planck Institute Series in Psycholinguistics.
Lidz, J. and J. Musolino (2002) 'Children's command of quantification.' *Cognition* 84, 113–154.
Milsark, G. (1977) 'Toward an Explanation of Certain Peculiarities of the Existential Construction in English.' *Linguistic Analysis* 3, 1–29.
Musolino, J. (1998) Universal grammar and the acquisition of semantic knowledge: An experimental investigation into the acquisition of quantifier negation interaction in English. PhD thesis, University of Maryland.
Musolino, J., Crain, S. and Thornton, R. (2000) 'Navigating negative quantificational space.' *Linguistics* 38, 1–32
Partee, B. (2003) 'Many quantifiers.' In: *Compositionality in semantics*, 241–258. Malden, USA: Blackwell publishers.
Philip, W. and P. Coopmans (1995) 'Symmetrical Interpretation and scope ambiguity in the acquisition of universal quantification in Dutch and English.' Ms., Utrecht University.
Reinhart, T. (1995) Interface strategies. Utrecht University: *OTS working papers*.
Smith, C. (1980) 'Quantifiers and question answering in young children.' *Journal of Experimental Child Psychology* 30, 191–205.
Unsworth, S. (2004) 'On the late acquisition of scrambled indefinites in L1 Dutch: How late is late?' Ms., Utrecht Institute of Linguistics.
Westerståhl, D. (1985) 'Logical constants in quantifier languages.' *Linguistics and Philosophy* 8, 387–413.