
**Please Note:** This manuscript is posted for noncommercial use. It is the pre-publication version and, as such, may contain errors that are not present in the final publication.

Number Words, Quantifiers, and Principles of Word Learning

Jessica Sullivan and David Barner

University of California, San Diego
Abstract

How do children learn the meanings of words like *many* and *five*? Although much is known about the mechanisms that underlie children’s acquisition of nouns and verbs, considerably less is understood of how children begin to learn the meanings of words that refer to sets (e.g., number words and quantifiers). Here we argue that children’s acquisition of quantity expressions relies on learning mechanisms typically associated with learning content words. In particular, we argue that the Whole Object Assumption and the Principle of Contrast are special cases of more general principles that guide not just the acquisition of nouns and verbs, but also the acquisition and interpretation of number words and quantifiers.
Introduction

When children learn a language, many of the words they acquire refer to things that can be seen and explored with the senses. For example, nouns, verbs, and adjectives often denote things with perceptible properties—e.g., *dog, jump,* and *red.* However, many words do not correspond to concrete things in the world, but instead convey abstract logical meanings. Number words like *twelve,* for example, can apply to a set of things without applying to any of its members in isolation (e.g., in a set of twelve cats, none of the cats has the property ‘twelve’). Similarly, quantifiers (e.g., *all, many*) and number morphology like the plural (e.g., *a cat/some cats*) denote the properties of sets and not types of things in the world.

Perhaps because the meanings of quantity expressions differ so much from those of nouns and verbs, relatively few parallels have been drawn between these case studies in the word learning literature. While it may seem intuitive that children recruit common learning mechanisms for acquiring many different classes of words, the literatures on the acquisition of nouns, verbs, and quantifiers have largely diverged, positing unique learning constraints for acquiring different kinds of words. In some cases, which we describe, radically different conclusions about children’s early knowledge have been reached in different literatures. Here, we explore how quantity expressions are acquired and suggest that quantifiers and number words, while different in critical ways from other words, nonetheless draw on at least some common word learning mechanisms. Starting with two case studies—the Whole Object Assumption and the Principle of Contrast—we argue that common principles govern children’s learning and interpretation of diverse linguistic forms, from nouns to numbers.

Default Units Of Reference And Quantification

At first pass, the problem of word learning seems simple. A parent points to a visible
object in the world (e.g., a rabbit), the child looks at it, and the parent utters a new word like
*Rabbit!* Over time, the correlation between word and object gradually focuses the meaning of the
word, resulting in adult-like knowledge of the word *rabbit*. However, as researchers have long
noted, the simplicity of this example is deceptive, since it hides both the complexity of real
learning situations and also the cognitive work contributed by the child ‘behind the scenes’.
Word learning appears easy not because it is a simple problem, but because children are so good
at it.

For an unconstrained learner a new word could mean almost anything. Take again the
case of ‘rabbit’. As philosopher Willard Quine observed,¹ when a parent points to a rabbit and
says a novel word like, *Gavagai!*, this utterance could mean ‘rabbit’, but it could also refer to the
animal’s tail, it’s fur, it’s color, or the act of running. It could also mean ‘dinner!’ or ‘animal!’ or
even ‘Hey look!’ These are all meanings that each of us might consider, since each is expressed
by words in English. However, there are also many meanings for *Gavagai* that we would never
consider—e.g., a collection of rabbit parts, leftward moving animal, and rabbit after dark.
Despite the large number of possible meanings for *rabbit* or *Gavagai*, children converge on an
adult-like meaning for thousands of nouns within the first few years of life. This suggests that, in
some important way, word learning is constrained.

Although the problem of word learning is complex and remains in many ways unsolved,
previous studies suggest that children bring rich assumptions to the table when learning the
names for things.²³ For example, in cases like our rabbit above, children typically infer that the
novel word refers to the whole object and not its parts, color, texture, or activities, a behavior
that Markman has called the Whole Object Assumption.²⁵ Consistent with this, when shown an
unfamiliar object paired with a novel label like *koba*, children as young as 2 assume that the
word applies to other objects that share the same shape, rather than the same substance or color.\textsuperscript{6–8} Similar behaviors are reported in languages like Japanese and Chinese,\textsuperscript{9–10} suggesting that children learning diverse languages make similar assumptions about the meanings of novel words.

While considerable research has been devoted to how children use the Whole Object Assumption to learn nouns, this learning mechanism may not be specific to noun learning. Instead, we believe that it is a special case of a more general word learning strategy that is used when learning number words, quantifiers, and even singular–plural morphology. The idea behind this proposal is that when children focus on whole objects, they do so because they privilege units that are picked out by their nonlinguistic perceptual and conceptual systems. We will call these units ‘default units’.

Studies of preverbal humans find evidence that, in the absence of language, infants can represent and discriminate quantities of discrete individuals. For example, 5-month-old infants can keep track of small sets of objects when occluded\textsuperscript{11, 12} and use spatiotemporal information like continuity, cohesiveness, and solidity to identify and track objects.\textsuperscript{13, 14} Infants are also sensitive to the approximate numerical magnitude of sets; when habituated to a set of 8 dots, infants subsequently look longer at 16 dots than they do at 8, even when other factors, like total surface area and contour length, are controlled.\textsuperscript{15, 16} Similar prelinguistic representations of individuals and sets have been documented for actions\textsuperscript{13} and sounds (although less work has examined the specific perceptual and conceptual principles that govern individuation in these cases).\textsuperscript{16, 17} Together, these studies indicate a nonlinguistic capacity to represent and quantify discrete individuals.\textsuperscript{18} It is these units, we suggest, that act as defaults in word learning.

Consider, first, the case of counting. In a surprising set of experiments, Shipley and
Shepperson\textsuperscript{19} presented children with arrays that included both whole objects (e.g., forks), and pieces of broken objects (e.g., half forks), and asked the children to count them. They found that, up to the age of 6, most children responded by counting the pieces of broken objects. For example, when shown three whole forks and one fork cut into two pieces, children counted the set as ‘five forks’. According to Shipley and Shepperson, this behavior did not differ significantly from when children were asked to count the ‘things’, suggesting that, although nouns tell children what kinds of things to count, they did not appear to help them decide \textit{how} to count them. Children treat discrete physical objects as default units for quantification, rather than using nouns to identify units.

Interestingly, recent studies have found similar results for other quantity expressions in language, suggesting that children’s reliance on nonlinguistic default units is not specific to counting. Children treat pieces of broken objects as units when labeling plural sets (e.g., they call a broken fork ‘some forks’), when deciding which of two sets is more (a fork broken in three pieces is judged as more than two whole forks), and when interpreting words like \textit{every}, \textit{a}, \textit{both}, etc. (Figure 1).\textsuperscript{20} Further, similar results have been found for children’s counting of events, suggesting that default units are not limited to discrete physical objects, but may also be used in the acquisition of verb meanings.\textsuperscript{21} When children were asked how many times a bunny jumped into a bucket, they counted the number of individual jumps leading to the bucket, rather than the number of times a bucket was reached. Finally, other studies find similar results for collective nouns (nouns that refer to a collection of individuals, such as ‘army’). For example, one study found that 4-year olds often counted individual objects (e.g., counting each soldier) when counting the referents of a collective noun (e.g., \textit{army}), instead of counting each collection as a unit.\textsuperscript{22} Other studies show that 4- and 5-year olds exhibit more adult-like interpretations of
collective nouns when the individuals are either physically connected or spatially overlapping. Together, these studies suggest that for children as old as 6, the referents of quantifiers, number words, and number morphology are, by default, individuals defined by nonlinguistic systems.

Our proposal is that this use of nonlinguistic default units best explains children’s early acquisition not only of number words and quantity expressions but also of a wide range of words, including common nouns. On this view, the Whole Object Assumption is a special case of a more general linguistic assumption that children make about the meanings of words. However, as typically stated, the Whole Object Assumption, with its focus on single whole objects, cannot explain many of the phenomena that the more general appeal to default units

**Figure 1** – Sample stimuli used to assess children’s use of default units. For stimulus (a) children are asked, e.g., ‘Who has more boots?’ (and choose the one ‘broken’ boot); for stimulus (b) children are asked, e.g., ‘Count the forks’ (and count 5)
allows. For example, *three* can never refer to a single whole object. Also, the fact that children count parts of things instead of whole things seems at odds with the idea that children restrict noun labels to whole things. In contrast, a use of default units explains not only children’s acquisition of object nouns but also their treatment of number words, quantifiers, number morphology, and even words that denote actions. These facts suggest that children treat nonlinguistic individuals as units for counting, and only later in acquisition (around age 5 or 6), use more abstract linguistically defined units.

**Contrast And Communicative Inference**

So far, we have argued that the Whole Object Assumption may be a special case of a broader mechanism that children use when acquiring language. Our proposal is that, by focusing on default units, children can quickly map new nouns to their intended referents, thereby speeding the acquisition of both nouns and quantity expressions early in development. At the core of this proposal is the idea that important principles that guide noun learning may also apply language-wide, to parts of language that are often treated as distinct learning problems.

Recently, we have begun to consider how another principle often associated with noun learning—the Principle of Contrast—might be used to learn and interpret quantity expressions. This principle, closely related to what Markman called the Mutual Exclusivity Assumption, proposes that when children (or adults) hear a new word, they assume that its meaning somehow differs from the meanings of other words they have previously acquired. For example, if a child has learned the word *rabbit* and now hears the word *furry* used to describe a passing bunny, they will assume that these two words differ somehow in meaning: whatever this new word means, it does not mean ‘rabbit’. By relying on the Principle of Contrast, children could vastly narrow the range of possible meanings they consider as they learn new words, leading to
quicker learning as their vocabularies grow. Our proposal, based on previous studies in the literature, is that the Principle of Contrast extends well beyond learning nouns, and reflects a more general pragmatic capacity that children use to learn and interpret number words, and that adults use to interpret quantifiers like *some* and *all*. Although perhaps intuitive to researchers who believe that pragmatic inference permeates communication in early development, this proposal is controversial to researchers who study quantity expressions, and who believe that pragmatic understanding emerges late in acquisition.

The starting point for our proposal is an observation made by Karen Wynn in her discussion of how children learn and interpret number words in early acquisition. In her groundbreaking studies, Wynn found that children learn number word meanings in distinct stages, learning the meanings of *one, two, three, and four* in sequence over a span of many months, before finally learning how counting represents numbers (i.e., that the last number in a count represents the cardinality of a set). Critical to the current discussion, Wynn found that before children acquire a number’s specific meaning (e.g., *five*), they draw on already known numbers (e.g., *one*) to constrain how they interpret unknown numbers. First, Wynn noted that although children who have an exact meaning for *one* (‘one-knowers’) systematically give one object when asked for one, they virtually never give one object when asked for two or three. Children know, in other words, that ‘one’ cannot be the meaning of another number word, because it is the meaning of the number word *one*. Second, Wynn showed that when children are presented with two sets—e.g., one containing one balloon, and the other containing five—one-knowers always point to the larger set when asked to ‘Point to five balloons!’ To explain this finding, Wynn suggested that children must be using the Principle of Contrast: ‘Since all the children knew that the word ‘‘one’’ refers to a single item, then if they knew that, for example,
the word ‘five’ refers to a numerosity, they should infer that it does not refer to a single item since they already have a word for the numerosity one’ (p. 229).

However, Wynn also noted that, taken alone, the principle of contrast cannot explain children’s behavior. Critically, children only treat words as contrasting if they belong to the same semantic class, and thus are relevant alternatives to one another. For example, children do not appear to think that a nonsense word (e.g., toma) or a quantifier (e.g., some) should contrast with ‘one’.33, 34 Instead, they restrict their inferences about word meanings by appealing specifically to alternative expressions that contrast along a specific dimension or scale (numbers, quantifiers, etc.). For example, in noun learning experiments, children do not assume that nouns contrast in meaning if they are at different levels of description (e.g., animal vs cat) or if they are told that a novel word is from another language.35

Together, these findings suggest that children interpret new words by contrasting them with alternatives that belong to common classes of words. When children hear an unknown word (e.g., five) they consider the meanings of known alternatives (e.g., one) and negate known meanings as possible interpretations for the unknown word (e.g., ‘five is equal to some quantity, but not one’). This idea, although perhaps at first pass unsurprising, is in fact controversial because it suggests not only a common mechanism for interpreting number words, nouns, and other novel words but it also suggests that children’s pragmatic abilities greatly exceed what is sometimes ascribed to them by psycholinguists. In particular, the pragmatic inferences documented by Wynn closely resemble inferences that children routinely fail to make when interpreting quantifiers like some and all. A large body of research investigating quantifiers has argued that children have relatively poor pragmatic competence early in development, and cannot make inferences similar to those documented by Wynn.30-32, 36
In one such study of quantifier development, Papafragou and Musolino\textsuperscript{30} showed 5-year-old children a scene in which three horses jumped over a log, and then asked them to evaluate descriptions of the scene. When presented with the sentence, ‘Some of the horses jumped over the log’ most children agreed that it was a good description, since it was, after all, true. Adults, however, did not agree, since all of the horses jumped over the log. In this context, the utterance containing some was rejected by adults because a better, more informative alternative was available. Interestingly, when the study was repeated using number words, a different result was obtained. Now children, like adults, rejected the sentence ‘Two of the horses jumped over the log’ when three horses had jumped. Thus, although adults always rejected true statements when more informative descriptions were possible, children only did so for numbers. In another study, Barner et al.\textsuperscript{37} found a similar distinction between quantifiers and number words. When children were shown three items in a container and asked, ‘Is there a banana in the container?’, they responded ‘yes’, even though some bananas would be a better description of the set. However, when they were asked, ‘Is there one banana in the container?’, children responded ‘no’, presumably because they thought that three bananas would be a better description. In both cases a ‘yes’ response would be literally true, but only in the case of number words did children reject descriptions that were literally true, but pragmatically underinformative. Thus, like in the study conducted by Papafragou and Musolino, children were adult-like in their interpretation of number words, but not quantifiers, see also Ref 38. This raises the question: if children can use contrast to interpret nouns and number words, why do they fail to make these simple inferences for quantifiers, which are so similar in structure?

By most accounts, children’s difficulty with quantifiers like some and all results from a failure to make a type of pragmatic inference, typically called ‘scalar implicature’.\textsuperscript{39} This
inference involves interpreting a word relative to other known words that occur on a common semantic scale. A scale is a set of alternatives that are ordered along a measuring dimension according to their relative strength or informativeness.\textsuperscript{40} For quantifiers, words like \textit{some}, and \textit{all} are on a shared scale, and \textit{all} is the strongest and most informative member. When a speaker utters a sentence that contains a relatively weak scalar item (e.g., \textit{some}) the listener assumes that sentences that contain stronger scalar alternatives (e.g., \textit{all}) must be false. This is because they assume that the speaker is speaking cooperatively, and thus should use the strongest, most informative, scalar item available. In the case of numbers, if a set contains exactly three objects, then the speaker should say \textit{three} rather than saying \textit{two} (because although \textit{two} is true, it is not as informative). In the case of quantifiers, if \textit{all} the horses jumped over a fence then the speaker should say \textit{all} rather than saying \textit{some} (even though \textit{some} is literally true). In each case, the listener expects the speaker to be cooperative and informative. Therefore, they assume that the speaker will use the strongest, most informative alternative available, whether the alternatives are quantifiers or number words.

Some researchers have suggested that children’s difficulties with scalar implicature may be the result of processing limits or a poor understanding of context and metalinguistic tasks.\textsuperscript{32, 36, 38} However, such accounts have difficulty explaining children’s inferences for number words, which are almost identical in structure.\textsuperscript{41-43} Our suggestion is that children are in fact quite good at computing pragmatic inferences, and that their difficulty instead lies in their knowledge of scales, particularly in the case of quantifiers. The evidence for this is that scalar implicature involves steps that are analogous to those used by children when interpreting unknown number words, like in the experiments described by Wynn. A key difference, we suggest, is that although children begin their mastery of counting by memorizing the count list as a set of alternatives
(one, two, three, four, etc.), children never memorize sets of quantifiers as a list. As a result, children are able to consider alternatives when interpreting number words (e.g., inferring that five cannot mean ‘one’ because one does), but are unable to do so when interpreting quantifiers (e.g., failing to generate all as a relevant alternative to some).

In support of this idea, when children are tested in a situation where the relevant alternatives are provided, they have no difficulty considering these alternatives and providing adult-like responses. In a recent study, we showed children a scene in which a cat, a dog, and a cow were all sleeping and then asked them either ‘Are only some of the animals sleeping?’ or ‘Are only the dog and the cat sleeping?’ (Figure 2). Adults overwhelmingly responded ‘no’ to each question. In order to correctly respond ‘no’ to the first sentence, the participant needs to know that all is a relevant alternative to some—i.e., ‘No, because all of the animals are sleeping’. For the second sentence, however, the participant can construct an alternative online by examining the scene, rather than using stored knowledge of relevant alternatives—i.e., ‘No, because the cow is also sleeping’. Crucially, these two conditions differ in the availability of relevant alternatives—if children’s failure to make adult-like inferences about quantifiers can be attributed to a lack of knowledge of the quantifier scale,
then they should fail for the first sentence (where the child must generate the relevant alternatives from memory), but succeed for the second sentence (where relevant alternatives were visually provided). In fact, children did just this. They failed to generate adult-like responses for the first type of sentence, but had no problem with the second type. This suggests that their knowledge of scales, rather than their ability to compute pragmatic inference, is the source of their problems with quantifiers early in development.

To summarize, our proposal is that the Principle of Contrast reflects a more general capacity that underlies both children’s reasoning about number words and also aspects of the adult ability to compute scalar implicatures. While computing scalar implicature depends on additional knowledge not required when using the Principle of Contrast for word learning, we argue that children may use the basic pragmatic capacity that underlies these two forms of inference to acquire and interpret words. This proposal unites two previously separate literatures that have reached opposite conclusions about children’s early pragmatic abilities. Also, it is consistent with the idea that children reason pragmatically from an early age not only when interpreting novel nouns but also when interpreting language more generally.

**Conclusion**

Two principles that are often invoked to explain children’s acquisition of nouns are also implicated in their acquisition of quantity words. First, we suggest that the Whole Object Assumption may be a special case of children’s more general reliance on ‘default units’ in language acquisition. Evidence from the word learning and counting literatures demonstrates that children privilege prelinguistic default units (e.g., discrete objects, discrete actions) when describing and quantifying individuals and sets. By relying on default units, we suggest, children may not only narrow their hypothesis space for learning nouns but also may speed their
acquisition of number words, quantifiers, verbs, and number morphology.

Second, we argued that pragmatic inferences, like the Principle of Contrast, guide children’s acquisition and interpretation of quantity expressions, and that children’s pragmatic abilities are not as poor as sometimes claimed. Previous studies show that, when learning the meanings of new words, children rely on the Principle of Contrast in order to constrain their meaning hypotheses. These inferences, we argue, bear a striking resemblance to inferences that children typically fail to make when interpreting quantifiers. In such cases, it appears that knowledge of specific contrast sets—in the case of quantifiers, scales—may restrict children’s abilities to deploy their inferential capacities. However, once children have learned these scales—or they are made available contextually—children can and do make pragmatic inferences like those of adults.

Researchers often posit distinct mechanisms for learning about quantity expressions and other types of words. We suggest that some previously attested mechanisms have a broader application than typically thought. Future research into language-wide learning mechanisms will be vital to understanding language learning not as a series of distinct learning problems, but rather as a process that draws on common principles and inferential abilities for learning across a broad set of semantic domains.

Note

Notice that here, unlike previous studies, the word ‘only’ is used. As a result, the sentences do not require pragmatic inference at all—‘only some’ logically entails ‘not all’ (e.g., you can say, ‘I ate some of the cake; in fact I ate all of it,’ but you cannot say ‘I ate only some of the cake; in fact I ate all of it’). Thus, any failure on the part of the children cannot be attributed to
pragmatics, but must be due to some other source, like knowledge of scales.

Acknowledgements

The authors thank Eleanor Chestnut, Alan Bale, and versions of the manuscript. This work was supported by a Jacobs Graduate Fellowship and through an NSF Graduate Fellowship to JS.
References


