



# Meaning matters in children's plural productions

Jennifer A. Zapf<sup>a,\*</sup>, Linda B. Smith<sup>b</sup>

<sup>a</sup> *Department of Human Development and Psychology,*

*University of Wisconsin-Green Bay, 2420 Nicolet Drive MAC C310, Green Bay, WI 54311, USA*

<sup>b</sup> *Indiana University, Department of Psychology, 1101 East Tenth Street, Bloomington, IN 47405, USA*

Received 30 August 2006; revised 15 February 2008; accepted 21 March 2008

---

## Abstract

The English plural is about the *number of individuals* in a *set of like kinds*. Two-year-old children use the plural but do not do so in all obligatory contexts. The present report asks whether the limitations on their production of the plural are related to aspects of meaning. In two Experiments plural productions were elicited from 2-year-old children for sets of size two and four and for instances of basic-level categories that were either similar or identical. Children were much more likely to produce the plural of these well-known nouns when there were four rather than two and when the instances were identical rather than merely similar. The results provide new evidence on children's acquisition of the English plural, showing that children's early productions are not just limited by knowledge of the noun and its plural form but also is limited by properties of the labeled sets in ways that are relevant to the underlying meaning of the plural.

© 2008 Elsevier B.V. All rights reserved.

*Keywords:* Cognitive development; Morphological development; Language acquisition; English plural

---

## 1. Introduction

Two-year-old children learning English spontaneously produce the plural forms of nouns, using them to label sets containing multiple instances of the same kind.

---

\* Corresponding author. Tel.: +1 920 465 2490; fax: +1 920 465 5044.

E-mail address: [zapfj@uwgb.edu](mailto:zapfj@uwgb.edu) (J.A. Zapf).

However, they do not use the plural in all required contexts (Berko, 1958; Brown, 1973; Cazden, 1968; Mervis & Johnson, 1991). Past explanations of this limited productivity have focused on morphological rules, exceptions to those rules, and phonological constraints (Marchman, Plunkett, & Goodman, 1997; Marcus et al., 1992; Winitz, Sanders, & Kort, 1981). The meaning of the plural, however, has not been studied in relation to its growing productivity. This report presents new evidence that meaning matters.

Formally, the English plural partitions sets into those with precisely one individual versus those with more than one. Although this seems natural to mature English speakers, other meaning distinctions are possible. For example, the Russian plural distinguishes sets of one, sets with few members, and sets with many members. Many Indo-European languages have a separate plural form for sets of exactly size two. Sursurunga, an Austronesian language, has five plural forms that are dependent on the number of instances in the referred to set. Other languages such as Japanese have no plural but quantify sets through the use of quantifiers that depend on the kind of thing in the set. These cross-linguistic differences mean that learners have to figure out the relevant meaning categories for their language. This paper examines two factors that might be relevant to young children's figuring out the meaning of the English plural: the similarity of the instances in the set and the number of things in a set.

### 1.1. Hypothesis 1: Similarity matters

The plural is about the number of things *in a set* and thus requires that children think about the individuals as forming a coherent group, and not just as unrelated distinct things. This should be easier if those things are highly similar (e.g., Markman & Gentner, 2005; Mix, Huttenlocher, & Levine, 2002). Thus, although the plural forms of English count nouns can be used to label sets of highly similar things and sets of very different things (e.g., *poodles*, *dogs*, *animals*, and *things*), for young children the similarity of the instances in the to-be-labeled set may be critical. Accordingly, Experiment 1 presented children with sets that could be labeled by the same basic-level noun (e.g., “dog”) but that varied in the similarity of the presented instances. Are 2-year-old children more likely to use the plural to label four identical dogs, for example, than to label four individually recognizable but distinct dogs?

### 1.2. Hypothesis 2: Number matters

Considerable research suggests that infants and toddlers accurately and categorically distinguish small set sizes such as one versus two versus three (e.g., Schaeffer, Eggleston, & Scott, 1974), and treat sets of one and also sets of two as unique categories (e.g., Wynn, 1990). As Barner and Snedeker (2005) recently suggested, this sophisticated and categorical apprehension of small numerosities could present problems for learning the English plural in that the plural requires children to treat sets of two, four, and one hundred as equivalent. In light of this, two opposing ver-

sions of the “number matters” hypothesis make developmental sense. First, young learners could preferentially attach the plural to sets of two, since these are well-differentiated from sets of one and easily quantifiable. If so, children might use the plural first for pairs of things before they use the plural for larger set sizes. Alternatively, if children first understand the plural as being about many instances, and if they understand many as distinct from both one and two, they may be most likely to use the plural when there are more things than just two in the set. This is also tested in Experiment 1. Are young children more likely to use the plural to label two dogs than four, or perhaps, are they more likely to use the plural to label four dogs than to label just two?

## 2. Experiment 1

The experimental task, borrowed from Johnston, Smith, and Box (1997), elicits productions by asking children to describe things to a blind-folded teddy bear. In the present version, the child was presented on each trial with an array of objects

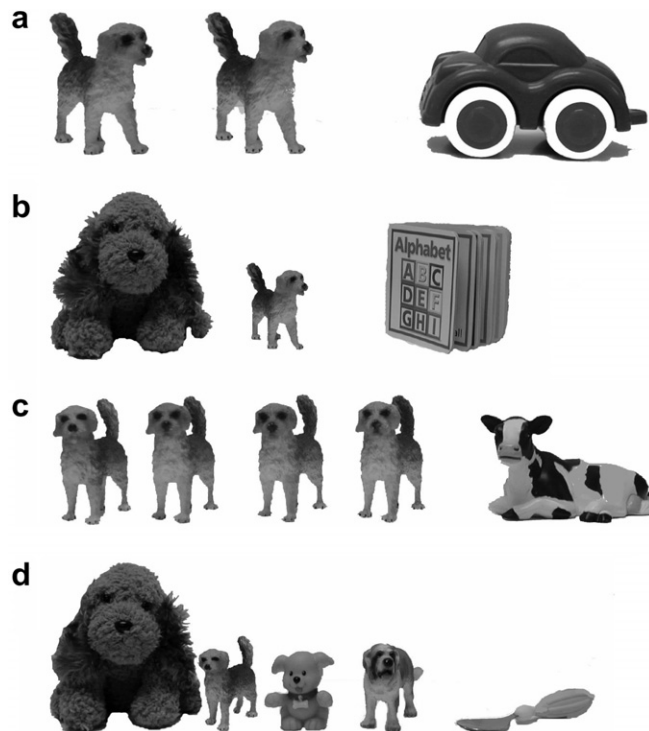


Fig. 1. Sample stimuli sets from Experiment 1: (a) two identical dogs, (b) two similar dogs, (c) four identical dogs, and (d) four similar dogs. In the placement of objects during the experiment, the position of the target was counter-balanced.

as illustrated in Fig. 1. Each array included two sets: a set of one (S set) and a set of more than one (P set). The child's task was to tell the bear "to get" one of the sets. On some trials, this was the S set and so a likely word to indicate that set would be the singular basic-level category term (e.g., "tree"). On the critical test trials the target set was the P set, potentially generating a plural noun (e.g., "pigs"). The size of the P set (two or four) and the similarity of instances within the P set (identical versus merely similar) was varied.

## 2.1. Method

### 2.1.1. Participants

Participants were 34 children between 23 and 30 months of age (mean age = 26 months). They were assigned randomly to the two between-subjects conditions, P set size of two and P set size of four.

### 2.1.2. Stimuli and design

Four unique three-dimensional instances (varying from 9 to 25 cm on the longest dimension) were selected from the following categories: dogs, flowers, monkeys, turtles, babies, birds, spoons, planes, boats, forks, cows, cars, hats, keys, dolphins, horses, trucks, bears, pigs, bunnies, bottles, crayons, giraffes, and books. These were chosen from a preliminary study such that all instances were readily recognizable by 2-year-old children. In that preliminary study, 28 children (mean age = 26 months) who did not participate in the main experiment indicated a named object in a three-choice forced-choice task. The four unique instances selected for each category were recognized by at least 90% of these children.

The four different instances formed the Similar-Instance P set size four. Two unique instances were randomly selected to form the Similar-Instance P set size two. Two or four identical replications of one instance were used to form the Identical-Instance P sets. Different unique instances were selected to form these Identical-Instance P sets for different children so that across children performance on identical sets could not be due to the particular instance used.

### 2.1.3. Procedure

To begin a trial, the experimenter placed an S set and a P set in the table. The child was allowed to play with the objects for 15 s to minimize subsequent choices based on toy preferences. The experimenter then arranged the objects into the two segregated sets (P and S) and placed the blindfold over Teddy's eyes. The child was reminded that Teddy could not see but could hear. The experimenter then told the child, "Tell Teddy to get. . ." at which time the experimenter gestured to the S set or the P set. After the child told Teddy what to "get", the experimenter took the blindfold off and Teddy "got" what the child said. On each trial, Teddy retrieved whatever the child verbally specified. If the child said nothing or something ambiguous (e.g., "that") the question was repeated once and if the child still did not respond or said something that did not unambiguously indicate one set, the experimenter went on to the next trial without Teddy retrieving anything. If the child used

the singular form when a plural was called for, Teddy “got” one instance. No other feedback of any kind was given.

The session began with “warm up” trials in which the experimenter demonstrated the task twice (once indicating an S set and once a P set) and had the parent perform the task and then had the child immediately imitate the parent. Once these two “warm up” trials were complete, the 10 experimental trials began.

On two of the trials the experimenter targeted the S set and on eight of the trials she targeted the P set. On half the P set trials the instances in the set were similar and on half they were identical. The identical and similar trials were blocked such that half the children received the identical trials first and the similar trials second and half received the similar trials first and the identical trials second. Size of the P set was two or four and varied between subjects. No child saw instances of the same basic-level category more than once.

## 2.2. Results and discussion

Children primarily labeled the sets with basic-level category nouns, either in the plural or singular form, doing so on 73% of trials of the P sets and on 76% of the S sets (here only one child on one trial ever used a plural form). The frequency of basic noun responses (plural plus singular) did not differ for the P sets and S sets,  $t(33) = -1.192$ ,  $p = .24$ , indicating that children were equally able to recognize the presented instances of these categories when presented as a single individual or in sets of multiple instances.

On the critical P set trials, children used the plural 51% of the time, using the singular form 22% of the time. They also sometimes said nothing (26%), rarely, they used some alternative expression might be construed as correct (e.g., “these”) less than 1% of the time. The number of plural productions on P trials were submitted to a 2 (set size)  $\times$  2 (order – similar/identical first)  $\times$  2 (similarity) ANOVA with similarity as a within subject factor. The analysis yielded three main effects: (1) number,  $F(1, 30) = 5.36$ ,  $p < .03$ , (2) order (identity trials first or second),  $F(1, 30) = 7.24$ ,  $p < .02$ , and similarity,  $F(1, 30) = 4.58$ ,  $p < .05$ . As is evident in Table 1, overall children produced the plural more often when labeling a set of four than a set of two and more often when labeling identical rather than similar instances. Set size and set

Table 1  
The proportion of P trials (sets with multiple instances) in Experiment 1 on which children produced the plural

Set size	Order	Identical trials	Similar trials
2	Similar trials first	.18	.21
	Identical trials first	.60	.50
	Overall	.39	.36
4	Similar trials first	.60	.40
	Identical trials first	.85	.70
	Overall	.73	.55

composition thus matter to children's early plural productions. For young children, sets with four items are a more potent elicitor of the plural than are sets with two; and sets of identical things are a more potent elicitor of the plural than are sets of similar things. Set size and similarity did not interact reliably. As is apparent in Table 1, the effect of similarity was only weakly evident on set size two due, perhaps, to the infrequent plural productions in both conditions and that the effect of similarity was more pronounced with set size four, when plural productions were more common.

The effect of order was unexpected and dramatic. Children more frequently produced the plural, for sets with identical instances as well as similar instances, when the experiment began with identity trials as opposed to similar trials. This pattern was obtained for both set sizes. It is as if first seeing multiple *identical* instances directed attention to the set and its quantity and this generalized to the Similar-Instance sets. In contrast, first seeing sets composed of basic-level category instances that varied in their individual properties seems to have directed attention away from the set as a whole yielding fewer plural productions even on subsequent sets containing identical things.

These results indicate that children's early productions of the plural form depend on both the similarity and number of items in a basic-level noun category. These new findings potentially challenge current assumptions about the nature of developing knowledge of the plural and relevant processes to the early production of the plural. Accordingly, Experiment 2 provides a replication of these results

### 3. Experiment 2

The main findings of Experiment 1 are that two aspects of sets that – by adult standards – should not matter in plural productions do matter for young children: the number and the similarity of like kinds. These results, however, were embedded in carry-over effects involving whether identical or merely similar instances were judged first. Accordingly, Experiment 2 sought to replicate the main findings in a completely between-subject design. Children were randomly assigned to one of four conditions: plural sets of two identical items, four identical items, two similar items, and four similar items. In addition, given this between-subject design, and the extent of individual differences in the rate of language development in this period, we collected data on children's vocabulary size.

#### 3.1. Method

##### 3.1.1. Participants

Fifty-five children between two and three years of age (31 males, 24 females) were randomly assigned to four groups. Productive vocabulary was measured by the MacArthur Communicative Developmental Inventory, a parental checklist widely used to measure productive vocabulary (Fenson et al., 1993). Table 2 gives the mean age and age ranges, as well as the mean vocabulary and vocabulary ranges, per group. A 2 (number)  $\times$  2 (similarity) ANOVA revealed no differences in age or

Table 2  
The proportion of P trials (sets with multiple instances) in Experiment 2 on which children produced the plural

Size	Identical stimuli	Similar stimuli
2	Mean age = 25.6 Age range: 22–35 Mean vocabulary: 334 Vocabulary range: 49–668	Mean age = 26.0 Age range: 22–36 Mean vocabulary: 296 Vocabulary range: 72–530
4	Mean age = 27.2 Age range: 22–35 Mean vocabulary: 374 Vocabulary range: 121–658	Mean age = 25.6 Age range: 22–33 Mean vocabulary: 429 Vocabulary range: 80–667

vocabulary size across the conditions. In the results section we will consider further possible effects of vocabulary size and plural productions.

### 3.1.2. Stimuli and design

The stimuli and design were the same as in Experiment 1.

### 3.1.3. Procedure

There were two test trials in which the experimenter targeted the S set and four trials in which she targeted the P set. All P set trials adhered to the condition in which the child was randomly placed. All other aspects of the procedure were identical to Experiment 1.

## 3.2. Results and discussion

As in Experiment 1, children primarily labeled the sets with basic-level category nouns, either in the plural or singular form, doing so on 70% of the trials of the P sets and on 78% of the S sets. Although they were reliably more likely to use the basic-level noun for the S sets compared to the P sets, the high level of performance overall indicates that children were able to recognize the presented instances.

The left most portion of Table 3 shows children's plural productions on the critical P trials in the four conditions. Again, when children did not offer the plural, they sometimes said the singular (26%) and sometimes said nothing (7%). The number of plural productions was subjected to a 2 (number)  $\times$  2 (similarity) ANOVA. The analysis revealed a main effect of number,  $F(1, 51) = 5.87, p < .02$ . As in Experiment 1, children produced the plural more for P sets of size four than size two. The analysis also revealed a main effect of similarity,  $F(1, 51) = 5.47, p < .02$ . Again, as in Experiment 1, children produced the plural more for P sets of identical rather than merely similar like kinds. There was no interaction,  $F(1, 51) = 1.10, p = .29$ . These results indicate that the effects of number and similarity on young children's early plural productions is a robust phenomenon.

Because vocabulary size and knowledge of language varies greatly in this period of rapid developmental change, we also partitioned the children into two groups,

Table 3

The proportion of P trials (sets with multiple instances) in Experiment 2 on which children correctly produced the plural

Set size	All participants		High vocabulary		High vocabulary	
	Identical	Similar	Identical	Similar	Identical	Similar
2	.55	.19	.65	.25	.48	.18
4	.70	.55	.83	.63	.61	.33

high and high vocabulary with the boundary defined by the grand mean of vocabulary size for all participants in the experiment (351). As can be seen, although high-vocabulary children produced more plurals than low vocabulary children, the basic pattern was the same for both language groups: more plural productions for sets of four than set of two and more plural productions for identical items than merely similar kinds. These results again suggest that children's early productions of the plural form depend on both the similarity and number of items in a basic-level noun category.

#### 4. General discussion

The present results indicate that children's early use of the plural is limited in part by the properties of the sets being labeled. Two-year-olds were more likely to use the plural when the multiple instances of a category were identical rather than merely similar and when there were four rather than two instances. The English plural marks sets of *multiple* instances of *like* kinds, and thus similarity and number are relevant dimensions of meaning, so children seem to attending to the relevant information for the plural but may not have yet fully worked out when the use of the plural is appropriate. One open question is the locus of the observed limitations on plural productions: are these limitations strictly about the *plural*, or are they about the contexts in which one chooses to talk about and quantify sets of individuals, or perhaps, do they reflect general conceptual growth in the coordination discrete individuals as both individuals and as members of sets?

For example, young children's greater plural productions for identical than similar items could reflect a general tendency of people to attend to (and thus talk about) the differences among otherwise similar things, a bias that has been documented in a variety of domains from perception, to similarity judgment, to learning adjectives (Garner & Sutliff, 1974; Sandhofer & Smith, 2004; Tversky & Gati, 2004; Waxman & Braun, 2005). Thus the pattern observed here, may not be strictly limited to their understanding of the plural. Nonetheless such a general tendency to focus on different individuals as distinct and not as members of the same class could be a limiting factor in children's quantification of sets more generally. If this were so, then this limitation should be seen in other quantification tasks such as counting or success in number conservation tasks, for example. As far as we know, there is no explicit evidence suggesting such a limitation, but clearly the question deserves further study.



Alternatively, the limitation could reflect children's understanding of the communicative intent of the experimenter when gesturing to the items to be labeled. These gestures – although made to be obvious to include all the items in the P set – might nonetheless be ambiguous to children. If this is so, then to explain the present results, children must have been more likely to view the intended referent as the whole set when those items were identical rather than similar and when there were more of them rather than view of them, a result that again suggests the potential importance of similarity and number in children's conceptualization of discrete items as forming a coherent (and thus quantifiable) set. The present task was spontaneous production and thus may be viewed as tapping biases in children's system and thus children may well show the ability to apply the plural to small and diverse sets of like kinds in other more constraining tasks. But, at the very least, these results indicate that these biases exist, that they limit children's plural productions, and thus may also play a role in a protracted discovery of the meaning of the English plural.

In this context, we note that the effect of similarity on children's plural productions fits with other findings on early cognitive development. Specifically, the plural requires that children represent the elements in a set in two ways, as discrete individuals that therefore can be counted and also as members of a single set that is being quantified. Such dual representations are known to be particularly difficult for 2-year-olds (see DeLoache & Sharon, 2005; Mix, 2002). For young children, and perhaps also for adults, there may be a psychological tension between conceptualizing individuals and simultaneously conceptualizing them as members of a unified set. The former may be fostered by any factor that separates the individuals, such as non-overlapping boundaries, space, time, and perceptual differences (see Kellman & Shipley, 1992 and Barner & Snedeker, 2005 for relevant discussions and data). The second may be helped by factors, such as similarity, that highlight membership in a single nameable set. Although these ideas are conjectures at this point, they suggest the value of the experimental study of the factors that would influence individualization of items and their conceptualization as a class in children's emerging use and understanding of the plural and the comparison of that understanding to other quantification tasks.

The systematic comparison of the role of similarity and numerosity in the development of the plural and in other quantification tasks is particularly critical given the finding that of more plural productions for sets of size four than sets of size two. Past research on early number concepts provides ample evidence that infants and young children quantify sets of size two (e.g., Wynn, 1990). Thus deficiencies in this regard seem an unlikely explanation of the observed pattern. Instead, the result may reflect competition between two developing systems of quantification. Indeed, as Barner and Snedeker (2005) proposed, children's considerable knowledge about and categorical apprehension of the numerosity of small set sizes could limit their early understanding of the plural. One possibility is that children initially quantify sets as “one”, “two”, “three”, and “more” (see Gelman & Gallistel, 1978; Ginsburg & Oppen, 1988; Wynn, 1990) and as a consequence have the English plural more strongly linked to larger set sizes than to “two” or “three”. With regard to this issue, a key question for empirical study is whether it is specifically the link between the

plural and “two” (and perhaps three) that is dampened or whether the plural is perhaps graded with respect to number for young children and thus generally stronger for larger than smaller numbers. The latter effect would be consistent with evidence on children’s acquisition of dimensional terms such as “big” and “little” which are often first understood as applied to extreme values (e.g., Sera & Smith, 1987; Smith, Cooney, & McCord, 1986).

As a first demonstration of an effect of number and similarity on children’s early plural productions, we acknowledge there are many unanswered questions about the phenomenon itself, its meaning for children’s understanding of the English plural, and its implications for the development of quantification. But the investigation of these new effects should provide a new and informative window on the development of the plural and the role of meaning in that development. At the very least, then, the results show that children’s early productions of the English plural are not just limited by knowledge of the noun and its plural form (Marcus et al., 1992) but also by the properties of the labeled sets in ways that are relevant to the underlying meaning of the plural. In this way, the results may also point us to the conceptual issues that children need to work out in learning the English plural, including: (1) the construal of instances both as individuals and as members of a set and (2) the formation of a category of “more than one” that includes already distinguished number categories such as “two”.

### Acknowledgements

This research was supported by a National Institute of Mental Health Grant (R01 MH60200). The first author was supported by a National Institute of Mental Health Grant (T32 HD 07475). We thank Jennifer Benson, Elizabeth Dahlberg, and McKenzie Skipper for assisting in data collection.

### References

- Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97(1), 41–66.
- Berko, J. (1958). The child’s learning of English morphology. *Word*, 14, 150–177.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development*, 39, 433–448.
- DeLoache, J. S., & Sharon, R. (2005). Symbols and similarity: You can get too much of a good thing. *Journal of Cognition and Development*, 6(1), 33–49.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., et al. (1993). *The MacArthur communicative development inventories: User’s guide and technical manual*. San Diego: Singular Publishing Group.
- Garner, W. R., & Sutliff, D. (1974). The effect of goodness on encoding time in visual pattern discrimination. *Perception & Psychophysics*, 16(3), 426–430.
- Gelman, R., & Gallistel, C. R. (1978). *The child’s understanding of number*. Cambridge, MA: Harvard University Press.
- Ginsburg, H. P., & Oppen, S. (1988). *Piaget’s theory of intellectual development* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

- Johnston, J. R., Smith, L. B., & Box, P. (1997). Cognition and communication: Referential strategies used by preschoolers with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(5), 964–974.
- Kellman, P. J., & Shipley, T. F. (1992). Perceiving objects across gaps in space and time. *Current Directions in Psychological Science, 1*(6), 193–199.
- Marchman, V. A., Plunkett, K., & Goodman, J. (1997). Overregularization in English plural and past tense inflectional morphology: A response to Marcus (1995). *Journal of Child Language, 24*(3), 767–779.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57* (4, Serial No. 228).
- Markman, A. B., & Gentner, D. (2005). Nonintentional similarity processing. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 107–137). New York, NY: Oxford University Press.
- Mervis, C. B., & Johnson, K. E. (1991). Acquisition of the plural morpheme: A case study. *Developmental Psychology, 27*(2), 222–235.
- Mix, K. S. (2002). The construction of number concepts. *Cognitive Development. Special Issue: Constructivism Today, 17*, 1345–1363.
- Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood*. New York, NY: Oxford University Press.
- Sandhofer, C. M., & Smith, L. B. (2004). Perceptual complexity and form class cues in novel word extension tasks: How 4-year-old children interpret adjectives and count nouns. *Developmental Science, 7*(3), 378–388.
- Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology, 6*, 57–379.
- Sera, M., & Smith, L. B. (1987). Big and little: “Nominal” and relative uses. *Cognitive Development, 2*(2), 89–111.
- Smith, L. B., Cooney, N. J., & McCord, C. (1986). What is “high”? The development of reference points for “high” and “low”. *Child Development, 57*(3), 583–602.
- Tversky, A., & Gati, I. (2004). *Studies of similarity*. Cambridge, MA, USA: MIT Press.
- Waxman, S., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*(3), B59–B68.
- Winitz, H., Sanders, R., & Kort, J. (1981). Comprehension and production of the /schwa z/Plural allomorph. *Journal of Psycholinguistic Research, 10*(3), 259–271.
- Wynn, K. (1990). Children’s understanding of counting. *Cognition, 36*, 155–193.