

Research Article

Rapid Word Learning Under Uncertainty via Cross-Situational Statistics

Chen Yu and Linda B. Smith

Department of Psychological and Brain Sciences and Program in Cognitive Science, Indiana University

ABSTRACT—*There are an infinite number of possible word-to-word pairings in naturalistic learning environments. Previous proposals to solve this mapping problem have focused on linguistic, social, representational, and attentional constraints at a single moment. This article discusses a cross-situational learning strategy based on computing distributional statistics across words, across referents, and, most important, across the co-occurrences of words and referents at multiple moments. We briefly exposed adults to a set of trials that each contained multiple spoken words and multiple pictures of individual objects; no information about word-picture correspondences was given within a trial. Nonetheless, over trials, subjects learned the word-picture mappings through cross-trial statistical relations. Different learning conditions varied the degree of within-trial reference uncertainty, the number of trials, and the length of trials. Overall, the remarkable performance of learners in various learning conditions suggests that they calculate cross-trial statistics with sufficient fidelity and by doing so rapidly learn word-referent pairs even in highly ambiguous learning contexts.*

Quine (1960) famously presented the core problem for learning word meanings from their co-occurrence with perceived events in the world. He imagined an anthropologist who observes a speaker saying “gavagai” while pointing in the general direction of a field. The intended referent (rabbit, grass, the field, or rabbit ears, etc.) is indeterminate from this experience. The solution to this indeterminacy problem requires that the learning system be somehow constrained.

Address correspondence to Chen Yu, Department of Psychological and Brain Sciences, 1101 East 10th St., Indiana University, Bloomington, IN 47405, e-mail: chenyu@indiana.edu.

Research on children’s word learning has concentrated on how this learning might be constrained in a single trial, such that the word is correctly mapped to the referent on that trial. This literature suggests that attentional (Smith, 2000), social (Baldwin, 1993; Tomasello, 2000), linguistic (Gleitman, 1990), and representational (Markman, 1990) constraints enable learners to “fast map” words to referents in a single encounter. However, the indeterminacy problem may also be solved cross-situationally, not in a single encounter with a word and potential referent but across multiple encounters and learning trials. A learner who is unable to unambiguously decide the referent of a word on any single learning trial might nonetheless store possible word-referent pairings across trials, evaluate the statistical evidence, and ultimately map individual words to the right referents through this cross-trial evidence. There has been very little systematic investigation of whether human learners do this kind of learning, and if so, what the underlying learning processes are.

This constitutes a significant gap in current understanding of human learning in general, and word learning in particular. Not all opportunities for word learning outside the laboratory are as uncluttered and as constrained as the experimental settings in which fast mapping has been demonstrated. Instead, in everyday scenarios, there are typically many words, many potential referents, limited cues as to which words go with which referents, and rapid attentional shifts among the many entities in the scene. Such highly ambiguous learning contexts could nonetheless play the dominant role in real-world word learning if learners calculate and use statistical information across multiple encounters with words and referents.

Several formal simulations suggest the plausibility of cross-situational word learning (Siskind, 1996; Vogt & Smith, 2005; Yu & Ballard, in press). In these simulations, learners keep track of many words and many referents over many trials, accruing evidence as to the word-referent pairings. Given the infinite number of potential meanings, cross-situational learning

mechanisms must also be constrained, and there are a variety of potential constraints that work reasonably well in simulations studies. Further, Akhtar and her colleagues (Akhtar, 2002; Akhtar & Montague, 1999) have shown that human learners (children) use information about the labels of two objects within a single learning trial (see also Namy & Gentner, 2002) and that when a single object is unambiguously labeled prior to an ambiguous trial, learners will combine information across trials in order to discover the relevant referent or property (see also Carey & Bartlett, 1978; Markman, 1990). These are both critical components of cross-situational learning. However, there is no evidence as to whether human learners are capable of learning from highly ambiguous contexts involving many words and referents, and whether they are able to compute statistics over many possible word-referent pairs and in so doing close in on the right word-referent mappings. The following experiments provide evidence for such a learning mechanism in adults.

EXPERIMENT 1

Our goal was to capture in a laboratory task some of the complexity and ambiguity of real-world word learning and to examine adult learners' ability to make word-referent mappings under those conditions. To these ends, we asked adult learners to simultaneously learn relatively many word-referent pairs (18 at a time) from individual learning trials that were highly ambiguous. On each trial, the learner was presented with multiple labels and multiple referents, with no information as to which label went with which referent. The experiment included three conditions that differed in their degree of within-trial ambiguity: In the 2×2 condition, each trial presented two words and two possible referents; in the 3×3 condition, each trial presented three words and three possible referents; and in the 4×4 condition, each trial presented four words and four possible referents. The 2×2 condition yielded 4 possible word-referent associations per trial, the 3×3 condition yielded 9 potential word-referent associations per trial, and the 4×4 condition yielded the seemingly overwhelming number of 16 potential word-referent associations per trial.

Although there was no information on any individual trial as to which label went with which word, the underlying word-referent mappings were certain in that an individual label was present in a training trial if and only if the referent was also present. Could learners keep track of the simultaneous co-occurrences of many labels and referents across trials and learn these mappings? Would they accomplish this easily in relatively few learning trials, from relatively few highly ambiguous exposures to each individual word?

The key ingredient to learning from highly ambiguous individual trials would seem to be keeping track of and comparing information across trials. This point is illustrated in Table 1 for the case in which the learner hears two words while viewing two objects (with neither spatial nor temporal cues linking the words

TABLE 1
An Example of a Cross-Situational Learning Situation

Trial	Words	Potential referents in scene
1	A B	b a
2	C D	d c
3	E F	e f
4	G A	g a

Note. There are one-to-one correspondences between the words and referents (e.g., A-a, B-b), but there is no information within a trial that indicates which word goes with which referent. Thus, the only way to find correct word-referent mappings is to compute cross-trial statistics.

to particular referents). On Trial 1, the learner could mistakenly link word A to referent b (and possibly also link word A to referent a). Notice, however, that on Trial 4, this mistake can be corrected; the cognitive system can rule out A-b as a possible word-referent pair if the system registers that word A occurred on Trial 4 without possible referent b. If the cognitive system remembers prior word-referent pairings, if it registers both co-occurrences and non-co-occurrences, and if it calculates the right statistics, it should be able to learn as many as 18 word-referent pairs from relatively few and highly ambiguous individual learning trials.

Method

Subjects

Thirty-eight students at Indiana University received course credit or \$7 for their participation.

Stimuli and Design

The stimuli were slides containing pictures of uncommon objects (e.g., canister, facial sauna, and rasp) paired with auditorily presented pseudowords. These artificial words were generated by a computer program to sample English forms that were broadly phonotactically probable; they were produced by a synthetic female voice in monotone. There were 54 unique objects and 54 unique pseudowords partitioned into three sets of 18 words and referents for use in the three conditions. The training trials were generated by randomly pairing each word with one picture; these were the word-referent pairs to be discovered by the learner. The three learning conditions differed in the number of words and referents presented on each training trial. In the 2×2 condition, each trial presented 2 words and 2 pictures; in the 3×3 condition, each trial presented 3 words and 3 pictures; and in the 4×4 condition, each trial presented 4 words and 4 pictures. There was no indication of which picture went with which word. Each trial began with the simultaneous visual presentation of the referents on a computer monitor. The names were then presented auditorily over the computer's speakers. The temporal order of the spoken names was not

related in any systematic way to the spatial location of the referents.

To form each trial, we randomly selected several (2, 3, or 4, depending on condition) word-referent pairs from the 18 word-referent pairs for that condition; across trials in a condition, each word and referent were presented six times. That is, over training trials, the learner experienced six repetitions of each word-referent pair. However, given that multiple words and referents were presented on each trial, the learner experienced spurious associations that might be expected to make learning from these ambiguous individual trials difficult. Specifically, on average, each word co-occurred with 5.09 incorrect referents in the 2×2 condition, 8.78 incorrect referents in the 3×3 condition, and 12.22 incorrect referents in the 4×4 condition; these numbers are proportional to within-trial ambiguity in the three conditions. During training, the probability of the correct referent given its name, $p(a|A)$, was 1.0 in all conditions, whereas the average probability of irrelevant but co-occurring referents was .205, .231, and .247 in the 2×2 , 3×3 , and 4×4 conditions, respectively. Notice that despite the considerable differences in within-trial uncertainty across conditions, the strength of the spurious correlations varied only moderately across these conditions.

Because the same number of word-referent pairs (18) was taught in each condition, and because we sought, across conditions, to keep the number of exposures to each word-referent pair constant, other presentation factors necessarily varied across conditions. These are summarized in Table 2. Across conditions, the number of repetitions of each unique word and referent and the total time of the training session were kept constant; thus, the total number of trials differed across conditions, as did the duration of each trial. Order of trials within a condition was determined randomly. Order of the three conditions (a within-subjects manipulation) was counterbalanced across subjects.

Procedure

The pictures were presented on a 17-in. computer screen, and the sound was played by the speakers connected to the same computer. Subjects were instructed that their task was to learn the words and referents, but they were not told that there was one referent per word. They were told that multiple words and pictures would co-occur on each trial and that their task was to

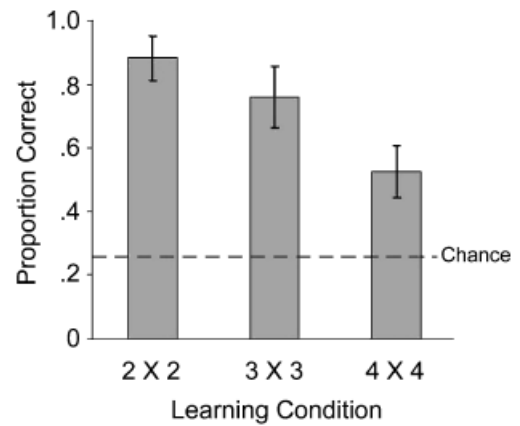


Fig. 1. Test results for the three learning conditions in Experiment 1. Error bars reflect standard errors. In the 2×2 condition, each trial presented two words and two possible referents; in the 3×3 condition, each trial presented three words and three possible referents; and in the 4×4 condition, each trial presented four words and four possible referents.

figure out across trials which word went with which picture. After training in each condition, subjects received a four-alternative forced-choice test of learning. On the test, they were presented with 1 word and 4 pictures and asked to indicate the picture named by that word. The target picture and the 3 foils were all drawn from the set of 18 training pictures.

Results and Discussion

Figure 1 shows that in each condition, subjects learned more word-referent pairs than expected by chance, smallest $t(37) = 8.785, p < .001, p_{rep} > .99, d = 1.425$, one-tailed (4×4 condition). They discovered on average more than 16 of the 18 pairs in the 2×2 condition and more than 13 of the 18 pairs in the 3×3 condition—all this in less than 6 min of training per condition. Even in the 4×4 condition, with 16 potential associations per trial, subjects discovered almost 10 of the 18 word-referent pairs. Indeed, 9 subjects discovered more than 75% of the pairs in this condition. The level of performance in the three conditions is remarkable: In a very short time, over relatively few trials, each highly ambiguous, subjects nonetheless found the underlying word-referent pairs. The degree of within-trial uncertainty clearly mattered, $F(2, 74) = 76.07, p < .001, p_{rep} < .99, \eta^2 = .631$. But just as clearly, subjects calculated

TABLE 2
The Three Learning Conditions in Experiment 1

Condition	Number of words	Number of occurrences per word	Number of trials	Time per trial (seconds)	Total time
2×2	18	6	54	6	324
3×3	18	6	36	9	324
4×4	18	6	27	12	324

cross-trial statistics with sufficient fidelity that they were able to acquire a significant number of word-referent associations and demonstrate this knowledge at test, despite the ambiguity of the individual learning trials.

This experiment was designed as a first demonstration of the general viability of cross-situational learning given highly ambiguous individual learning trials, and does not indicate the precise mechanisms that underlie this learning. The spurious correlations in the training data are relevant to understanding these mechanisms and may suggest what subjects actually learned from their experience. At test, subjects were presented with four alternatives—the correct referent for the tested word and three alternatives. It seems highly likely that subjects simply chose the most strongly associated item among the presented alternatives. If subjects were able to track all word-referent co-occurrences in the training trials, they should have been able to respond perfectly in all conditions, because the association between each word and its referent was 1.0 and much greater than even the strongest spurious correlation in the training sets. However, if learners kept track of all co-occurrences during learning, and if they chose the alternative most associated to the tested word, then any errors would have been related to the spurious correlations that arose given the presentation of multiple words and referents on specific learning trials. Specifically, errors would have been related to what we call *foil probability*, the probability that the foil at test had co-occurred with the word during training. For example, a foil that had occurred with the tested word on three of the six repetitions of that word during training should have been wrongly selected more often than a foil that had occurred with the tested word only once. The probability that the foils were spuriously associated with the tested words was, on average, not high: .056, .115, and .155 in the 2×2 , 3×3 , and 4×4 conditions, respectively.

Because the probability that a tested foil had been associated with a target word was both greatest and most variable across foils in the 4×4 condition, and because subjects in this condition made the most errors, we selected this most ambiguous condition to more closely examine the relation between choices of foils and their strength of association with the tested word. Table 3 shows the probability that subjects chose the tested foils

TABLE 3
Foil Probabilities and Probabilities of Incorrect Answers in the 4×4 Condition of Experiment 1

Foil probability ^a	Probability of incorrect answer
0/6	.162
1/6	.216
2/6	.222
3/6	.077

^aThis column indicates the fraction of training trials on which the foil had co-occurred with the tested word.

as a function of their association to the tested word (accumulated across multiple test trials and subjects). There appears to be little systematic relation. A strong conclusion that foil probability does not matter, however, is not warranted, as the strength of spurious associations of foils to test words was overall quite low. The low level of spurious correlations was the natural result of the large training set (and the random selection of co-occurring pairs during training). These characteristics, however, may well also describe real-world word learning. We pursued the issues of spurious correlations, foil probability, and size of the data set in Experiment 2.

What did subjects learn from these brief experiences? Word-referent pairings were uncertain within a trial but (if subjects tracked all the information) certain across trials. However, given the real-time processing demands of attending to and remembering many words and referents and the relatively brief training regimen, subjects' knowledge of most word-referent pairs may not have been certain. Indeed, many subjects volunteered (quite wrongly) prior to test that they were sure they knew none of the pairings. Thus, subjects may well have not learned, for example, that word A mapped only to referent a. Rather, their knowledge may have been more of the form "word A is associated with referent a and b, but not with anything else." Such partial knowledge could explain the present results. It could also play a powerful role in real-world word learning. Our main point is this: The acquisition of this kind of knowledge (even if imperfect) requires calculations on cross-trial co-occurrences. Experiment 1 shows that adult human learners perform such calculations for relatively many word-referent pairs, despite within-trial uncertainty of the pairings. Cross-situational statistical learning is within the repertoire of human learners.

EXPERIMENT 2

Experiment 2 was designed to replicate the findings of Experiment 1 and to further explore learning in conditions of high within-trial ambiguity as a function of the number of word-referent pairs to be learned. Accordingly, each condition in this experiment was a version of the original 4×4 condition of Experiment 1. We manipulated (a) the total number of word-referent pairs to be learned and (b) the number of repetitions of each word-referent pair. In the 9-words/8-repetitions condition, subjects attempt to discover a total of 9 word-referent pairs each repeated 8 times over the course of training. In the 9-words/12-repetitions condition, subjects attempt to discover 9 word-referent pairs but were given 4 additional repetitions of each word-referent pair. Finally, the third condition was a replication of the 4×4 condition of Experiment 1; the 18 word-referent pairs to be learned were repeated 6 times each.

Intuitively, the 9-words/12-repetitions condition would be expected to improve learning performance because, compared with the 18-words/6-repetitions condition, the number of words to be learned was reduced and the frequency of their occurrence

TABLE 4
The Three Learning Conditions in Experiment 2

Condition	Number of words	Number of occurrences per word	Number of trials	Time per trial (seconds)	Total time
9 words, 8 repetitions	9	8	18	12	216
9 words, 12 repetitions	9	12	27	12	324
18 words, 6 repetitions	18	6	27	12	324

was doubled. However, for statistical learners, smaller data sets may not be as good as large ones because spurious correlations are more likely to occur.

Method

Subjects

Twenty-eight students at Indiana University received course credits for their participation. None had participated in Experiment 1.

Stimuli and Procedure

In all aspects except for the composition of the three training conditions, Experiment 2 was identical to Experiment 1. All three conditions used the 4×4 presentation of 4 words and 4 pictures on each trial, but they differed in the number of word-referent pairs to be learned (9, 9, and 18) and in the number of repetitions of each word-referent pair (8, 12, and 6). The conditions are summarized in Table 4. The random selection of co-occurring word-referent pairs during training and the random selection of foils at test led naturally to differences in foil probabilities, that is, in the associations of the alternatives at test with the tested word. The foil probabilities were higher when 9 word-referent pairs were to be learned (.375 in both 9-pair conditions) and lower (.247) when 18 word-referent pairs were to be learned.

Results and Discussion

The three conditions presented equivalent within-trial uncertainty but differed in the number of word-referent pairs. In terms of the proportion of word-referent pairs discovered, subjects performed comparably in the three conditions, $F(2, 54) = 0.52$, $p > .5$, $p_{\text{rep}} = .42$, $\eta^2 = .03$, discovering more pairs than expected by chance, as shown in Figure 2, $t(27) > 6.4$ in all three conditions, $p < .001$, $p_{\text{rep}} > .99$, $d = 1.249$. Again, adult learners acquired lexical knowledge from highly ambiguous exposure to words and potential referents. Together, the results from Experiments 1 and 2 suggest that within-trial uncertainty is a more critical factor in learning than is the number of pairs in the learning set.

Subjects actually learned more pairs in the 18-pair condition ($M = 9.461$, $SD = 2.907$) than in the two 9-pair conditions

(8 repetitions: $M = 5.111$, $SD = 1.706$; 12 repetitions: $M = 5.481$, $SD = 2.089$). The 18-pair condition presented the same within-trial ambiguity as the other two conditions, with more word-referent pairs to be learned and fewer repetitions of the individual pairs. If number of co-occurrences was all that mattered, this condition should have led to the poorest overall performance. The advantage of this condition lay in its fewer spurious correlations (and thus also its lower foil probabilities at test). Herein lies the power of cross-situational statistical learning: Even when the referent of a word cannot be unambiguously determined on any single learning trial, across multiple trials involving many different words and many different potential referents, the word will co-occur with its referent more systematically than with any other potential referent. The more words and referents that there are to learn and that may co-occur on any learning trial, the more discernible is the systematicity—across trials—of the underlying correct mappings. Could bigger lexicons (more pairs) really be easier to learn than smaller ones? Learning requires multiple processes, some of which (e.g., memory for particular items and attention) will be negatively affected by increasing size of the learning set. However, within these constraints, statistical learning of a *system* of word-referent pairs may well benefit from larger as opposed to smaller data sets.

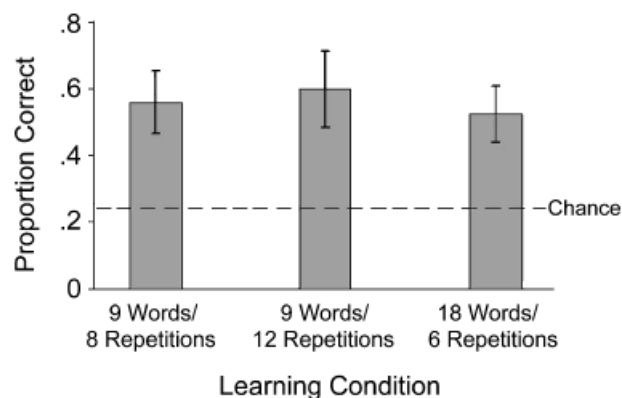


Fig. 2. Test results for the three learning conditions in Experiment 2. Error bars reflect standard errors.

GENERAL DISCUSSION

There is no doubt that human learners (including young children) fast-map names to things by solving the indeterminacy problem in a single trial—linking a novel word correctly to the intended referent through the use of social (Baldwin, 1993; Bloom, 2000; Tomasello, 2000), linguistic (Gleitman, 1990), attentional (Smith, 2000), and conceptual (Gentner, 1982) constraints. The present results suggest the importance of an additional kind of learning that does not require such in-the-moment certainty but instead allows for substantial learning from far more ambiguous learning environments in which the correct mapping of a word to an intended referent cannot be guaranteed.

The robustness of the learning our subjects demonstrated, despite being given brief training, suggests a possible role for cross-situational learning in vocabulary development. Studies indicate that parents, on average, direct 300 to 400 words an hour to their children (Hart & Risley, 1995). Presenting so many words in so little time would seem likely to generate considerable ambiguity about intended referents. Yet this kind of learning environment with much in-the-moment ambiguity may, precisely because of the sheer amount of statistical data provided, yield considerable word learning. The present experiments constitute a first step in understanding the role of cross-situational statistical learning by showing robust learning of relatively many words from the co-occurrence data available in brief exposures (less than 6 min). The findings are reminiscent of recent evidence on adults' and infants' ability to discover segmental units in the sequential probabilities of sounds or visual events (Conway & Christiansen, 2005; Gomez & Gerken, 1999; Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996). Like the present results, the findings on learning sequential probabilities and segmentation suggest that the solution to fundamental learning problems central to language may be found by studying the statistical patterns in the learning environment and the statistical learning mechanisms in the learner (Newport & Aslin, 2004; Saffran, Newport, & Aslin, 1996).

What is the mechanism that gives rise to the effects we observed? One possibility is a simple associative process that counts the number of co-occurrences and on test trials chooses the object most strongly associated with the test word. Alternatively, more complicated associative models that include competition and inhibition of competing associations might be required. Finally, statistical learning that explicitly compares alternative hypotheses and rules out wrong hypotheses might be needed to generate the fast learning of so many word-referent pairs from such minimal training data. Questions regarding underlying mechanisms cannot be answered without formal modeling, the next step in our research agenda.

It is also important to consider the kinds of constraints that must be imposed on the system to account for human learning. In

the present experiments, the subjects were not explicitly told there was one word for each picture. Nonetheless, their comments after the experiments indicated that almost all of them adopted a one-word/one-object strategy, what is known as the mutual-exclusivity assumption in children's word learning (Clark, 1987; Markman, 1990). Does the learning mechanism require this constraint to succeed? Is it an explicit hypothesis-testing strategy? Many subjects indicated that they had been quite sure they had learned nothing from the training and were amazed at their own success. This suggests that cross-situational learning may go forward nonstrategically and automatically, steadily building a reliable lexicon.

A further critical question concerns the availability of these cross-situational learning mechanisms to infants and young children. It seems highly plausible that these mechanisms are in fact available to such young learners, as considerable research suggests strong continuity in general learning mechanisms in infants, children, and adults (Gillette, Gleitman, Gleitman, & Lederer, 1999). At the very least, the present results point to the value of the systematic study of cross-situational learning and its mechanisms.

In conclusion, the human learning environment is data rich. Past analyses questioned the quality of that data for language learning (Quine, 1960) because each datum is highly ambiguous in and of itself. But the data set as a whole—if human learners possess the right learning mechanisms—may readily solve this indeterminacy problem. The present results suggest that human learners may well possess these needed mechanisms.

Acknowledgments—This research was supported by National Science Foundation Grant BCS 0544995. Samantha Brandfon was a great help in running these experiments. We would like to thank Peter Gordon, Nick Chater, and an anonymous reviewer for helpful comments on a previous version of this article.

REFERENCES

- Akhtar, N. (2002). Relevance and early word learning. *Journal of Child Language, 29*, 677–686.
- Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language, 19*, 347–358.
- Baldwin, D. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology, 29*, 832–843.
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.
- Carey, S., & Bartlett, E. (1978, August). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17–29. Retrieved March 2007 from <http://www.wjh.harvard.edu/~lds/index.html?carey.html>
- Clark, E.V. (1987). The Principle of Contrast: A constraint on language acquisition. In B. MacWinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.
- Conway, C., & Christiansen, M.H. (2005). Modality constrained statistical learning of tactile, visual, and auditory sequences.

- Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 24–39.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj, II (Ed.), *Language development* (Vol. 2, pp. 301–334). Hillsdale, NJ: Erlbaum.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73, 135–176.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 1–55.
- Gomez, R.L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109–135.
- Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.
- Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35–B42.
- Markman, E.M. (1990). Constraints children place on word learning. *Cognitive Science*, 14, 57–77.
- Namy, L.L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, 131, 5–15.
- Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48, 127–162.
- Quine, W.V.O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621.
- Siskind, J.M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61, 39–61.
- Smith, L.B. (2000). How to learn words: An associative crane. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Breaking the word learning barrier* (pp. 51–80). Oxford, England: Oxford University Press.
- Tomasello, M. (2000). Perceiving intentions and learning words in the second year of life. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 111–128). Cambridge, England: Cambridge University Press.
- Vogt, P., & Smith, A.D.M. (2005). Learning colour words is slow: A cross-situational learning account. *Behavioral and Brain Sciences*, 28, 509–510.
- Yu, C., & Ballard, D.H. (in press). A unified model of early word learning: Integrating statistical and social cues. *NeuroComputing*.

(RECEIVED 4/27/06; REVISION ACCEPTED 7/21/06;
FINAL MATERIALS RECEIVED 8/1/06)