

How Infants Use the Words They Know to Learn New Words

George Hollich¹, Peter W. Jusczyk¹, & Michael R. Brent²

¹Johns Hopkins University and ²Washington University

At twelve months of age, infants begin producing their first words, and by all accounts, at this age, word learning is a laborious process (Hollich, Hirsh-Pasek, & Golinkoff, 2000). At the beginning, infants struggle in their quest to attach meaning to sound. However, by 24 months of age, less than one year later, these same infants learn up to nine new words a day and usually have over 50 words in their productive vocabulary (Bloom, 1973). The challenge for developmental psychology is to explain how it is that these infants come to learn so many words, so quickly.

One factor that has often been assumed to play a role in this rapid acquisition is the idea that infants might use the words they know in learning new words. Thus, each word acquired helps infants to better learn the next word, and so on, and so forth, in an exponential growth function (see Ganger & Brent, this volume). However, even if this idea is correct, there are at least two ways, or levels, in which infants could use familiar words in acquiring new ones.

First, at the phonetic level, infants could use words they know to segment the speech stream itself: using known sound patterns to parse out new acoustic combinations in the kind of incremental distributional regularity optimization proposed by Brent (1999; Brent & Cartwright, 1996; Dahan & Brent, 1999). Of course, by now, there is considerable evidence that infants begin segmenting words from fluent speech during their first year (Jusczyk & Aslin, 1995). Moreover, infants have been shown to use a variety of different sources of information in word segmentation, such as prosodic stress (Jusczyk, Houston, & Newsome, 1999), allophonic cues (Jusczyk, Hohne, & Bauman, 1999), phonotactic cues (Mattys & Jusczyk, in press), and statistical cues (Saffran, Aslin, & Newport, 1996). Consequently, it would not be surprising if infants were also making use of their knowledge of existing words in segmenting new words from the speech stream.

Second, at the conceptual level, infants could use the words they know to infer appropriate word-to-world mappings: assuming, for example, that novel names map onto novel categories as has been proposed by Golinkoff, Mervis, and Hirsh-Pasek (1994). This paper explores both possibilities in greater detail: examining infant abilities both to segment the speech stream and to discover new word-to-word mappings on the basis of prior word knowledge.

1. Experiment 1

Brent's (1999) INCRemental Distributional Regularization OPTimization (INCDROP) model was designed as a computer simulation of how learners segment the speech input and build a lexicon for their native language. A key assumption of the model is that familiarity with words allows learners to better segment the speech-stream. Thus, when confronted with the phrase "pretty *kiosk*" infants should be able to extract the familiar word, "pretty," from the utterance and infer that "*kiosk*" must also be a word. In fact, the computer simulation of this model achieved a reasonable level of success in segmenting words from an input corpus of child-directed speech. The success of this model leads to an empirical hypothesis about the performance of infant language learners. Children should learn the meaning of words surrounded by a familiar context faster than those surrounded by an unfamiliar one.

Twenty-four, 24-month-old infants were tested, using the split-screen preferential looking procedure (Hollich, Jusczyk, and Luce, 2000). Similar sorts of tasks have been used successfully in studying on-line word recognition with infants between 15 and 24 months of age (Schafer & Plunkett, 1998; Swingley, Pinto, & Fernald, 1999) on their word learning following familiarization with words that were placed in a familiar or an unfamiliar context. Specifically, infants were first familiarized with a pair of phrases: either "pretty *kiosk*" and "ritzy *pylon*" or vice versa. Next, they were tested with a pair of items already known to most infants (*flower* and *apple*). These warm-up test trials served to inform infants about the nature of the task and provided a baseline measure of their performance. Next, the infants were trained on the meaning of the words *kiosk* and *pylon*. Finally, they were tested on their recall of the meaning of these words (see Table 1 for the design).

Table 1. Design of Experiment 1.

Visual Display	Sound Track
Simulated Lights	"Ritzy Kiosk" (x 3)
Simulated Lights	"Pretty Pylon" (x 3)
Flower & Apple	"Flower" (x 3)
Flower & Apple	"Apple" (x 3)
Pylon	"Pylon" (x 3)
Kiosk	"Kiosk" (x 3)
Pylon & Kiosk	"Pylon" (x 3)
Pylon & Kiosk	"Kiosk" (x 3)

** Note. Each set of trials was counterbalanced.

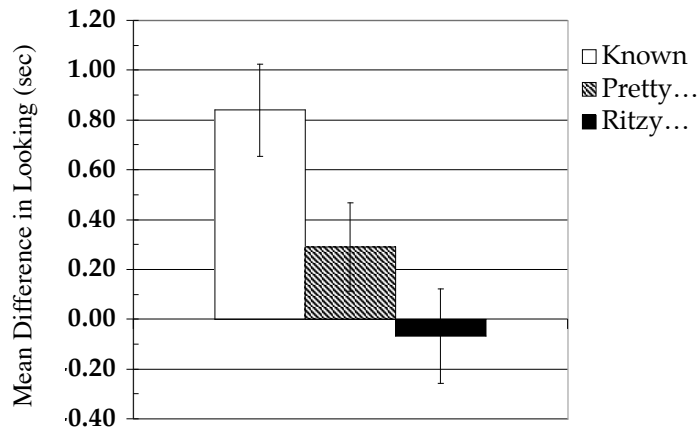


Figure 1. Mean difference in looking to target by condition.

As in all preferential looking procedures (see Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987), the logic was that infants would look longer at the object that matches the auditory stimulus. The dependent variable was the mean difference in looking time between the target and the non-target. An additional dependent measure was the time it took infants to shift, after the onset of the target word, from looking at the non-target to the target object.

It should be noted that this task of learning two words in close temporal proximity is a difficult one. Other studies using this task have found that infants tend to learn one word or the other, but not both (see Hollich, Jusczyk, & Luce, 2000). Specifically, Hollich et al. found that infants learned a word more easily when they had been previously familiarized with just a few similar sounding words, as opposed to many similar sounding words. Hence, familiarization was critical in determining which of the words was learned. Given that only one word tends to be learned, in such circumstances, then factors that favor the easier extraction of one word over the other from fluent speech might well dictate that that word is the one that is learned. One such factor is the context that the target words appear in. Consequently, a task of this sort may be an effective vehicle for investigating lexical effects of the sort predicted by the INCDROP theory. In effect, the task forces infants to choose one word over the other. Thus, in this case, any consistent preference in favor of a word presented in the familiar context would provide evidence of the kind of effect predicted by the INCDROP model.

The results, presented in Figure 1, indicate that infants looked significantly longer at the targeted objects in the case of already known words, such as flower

and apple, $t(23) = 4.55, p < .001$. However, for the two new words, first presented in the familiarization period, the mean difference in looking to the target did not differ significantly from zero, $t(23) = 1.64$ and $0.35, p = .11$ and $.75$ for the familiar and unfamiliar contexts, respectively. Nor was there any significant effect of context when these were directly contrasted, $t(23) = 1.17, p = .25$. Furthermore, the reaction times, for the nineteen subjects who provided usable data, tell a similar story. Thus, while the time it took infants to shift their gaze to the target object after the onset of the target word was slightly faster for the familiar context, 883 ms compared to 929 ms ($SE = 41$ and 56 ms, respectively), this difference was not statistically significant, $t(18) = 0.28, p = .78$.

Although infants indicated that they understand the known words, *apple* and *flower*, they did not show any significant effect of surrounding context in learning the new words, *kiosk* and *pylon*. Although individual infants did learn at least one word, they did not consistently learn the word in the familiar context, “pretty,” as predicted by the INCDROP model. Interestingly, however, the results were in the correct direction for the effect predicted by INCDROP. That is, infants seemed to perform better in the familiar context, although this effect was not marginally significant.

These findings raise a methodological question. Perhaps, with this method and this task, the benefit, such as it was, from a familiar context was mitigated by the presence of other cues to segmentation in the input. That is, metrical stress, allophonic cues, and even the fact that silence occurred after the critical words, all could have signaled to these infants that “ritzy” and “pylon” were two separate words. In this manner, although we found little evidence of a INCDROP effect, it may be that infant’s segmentation abilities were good enough to parse both words easily, thus obscuring any INCDROP effect. Perhaps evidence for INCDROP could be found in a study that minimized other segmentation cues.

2. Experiment 2

In Experiment 2, we attempted to minimize the contribution of the other cues to segmentation in order to encourage the use of the INCDROP strategy in segmenting these words. The target words were changed to *opal* and *onyx*, because prior work had indicated that vowel initial words are more difficult to segment (Mattys & Jusczyk, in press; Nazzi, Jusczyk, & Bhagirath, 1999). In addition, the surrounding contexts during familiarization were increased to the full sentences, “The ___ is pretty,” and “Faux ___ fools misers.” This latter change had the effect of surrounding the target words on both sides and minimizing the prosodic cues to segmentation. In effect, these stimuli were designed to increase the difficulty of extracting the target words from the contexts in order to make it more likely that infants would rely on the INCDROP strategy.

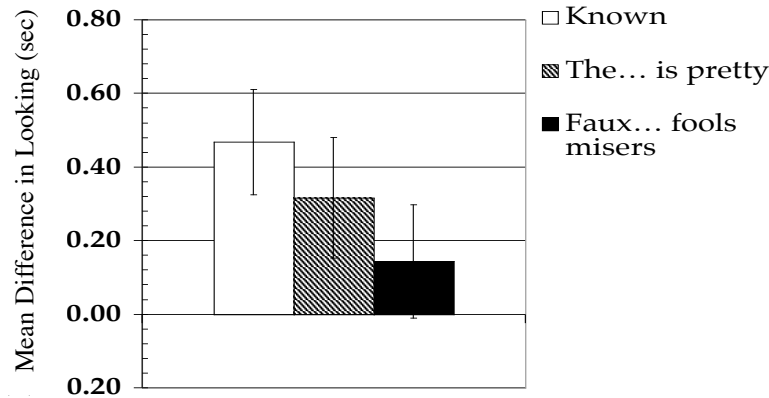


Figure 2. Mean difference in looking to target by condition.

The results for twenty-four 24-month-old subjects are presented in Figure 2. Once again, infants looked significantly longer at the targeted objects in the case of the already known words, *flower* and *apple*, $t(23) = 2.15$, $p = .02$. However, the mean difference in looking to the target for the words taught during familiarization was not significantly different from zero, regardless of context, $t(23) = 1.36$ and 0.66 , $p = .19$ and $.51$ for the familiar and unfamiliar contexts, respectively. Nor was there any significant effect of context when performance on these was directly contrasted, $t(23) = 0.47$, $p = .64$. Further, the reaction times, for the twenty subjects who provided usable data, are nearly identical. Thus, the time it took infants to shift their gaze to the target after the onset of the target word was 938 ms for the familiar context and 929 ms for the unfamiliar context ($SE = 52$ and 57 ms, respectively). This difference was not statistically significant, $t(19) = 0.17$, $p = .86$, and it was actually smaller than the difference observed in Experiment 1.

Once again, these results suggest that, although infants responded appropriately the known words, *apple* and *flower*, they did not show any significant effect of responding faster or more accurately to the taught word that had been presented in the familiar context than to the one presented in the unfamiliar context.

3. Experiment 3

Although the behavior of infants in these first two experiments does not provide support for a segmentation strategy along the lines of the INCDROP model, perhaps there were additional methodological factors concealing any effects that familiar context may have on segmenting new words. Might the task we used have been ineffective in teaching new words to infants? This seems extremely unlikely given previous success in using this very task with 17-

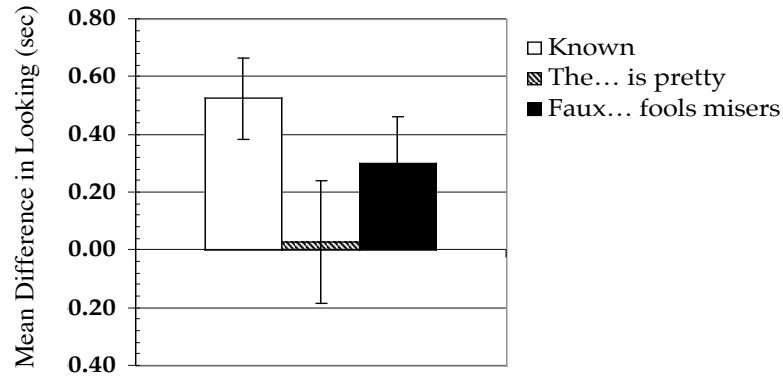


Figure 3. Mean difference in looking to target by condition.

month-olds (Hollich et al., 2000). More plausibly, a critical assumption of the previous two experiments is that segmentation during familiarization leads to facilitation during training. However, perhaps, because infants heard both words, individually, during training, the carry-over of any benefit from correct segmentation during familiarization was minimized. Thus, because each word is repeated in the clear during the training phase, infants might not have needed any advantage provided by the familiar contexts during the initial familiarization period.

If this account is correct, then the lack of any advantage for the familiar contexts we used stems from the fact that the isolated words used in the training phase were sufficient to ensure that infants mapped these words onto the appropriate objects. In retrospect, these two studies did not directly test the idea that a learner can make use of familiar words to learn a new one. Perhaps clearer evidence for the use of an INCDROP strategy could be found if infants were *trained* with words that were embedded in the familiar and unfamiliar contexts.

In Experiment 3, the design was revised to have the familiarization stimuli played during training. Thus, instead of hearing onyx or opal during labeling, infants heard: “faux onyx fools misers,” and “the opal is pretty.” This modification, it was hoped, would further force infants to rely on the INCDROP strategy in order to segment the correct word from these contexts to be attached to the object in the visual display. Only a correct segmentation of the training passage would ensure that infants would attach the correct word to the appropriate referent.

Results for twenty-four 24-month-olds are presented in Figure 3. Again, infants looked significantly longer at the targeted objects in the case of the already known words, flower and apple, $t(23) = 3.75, p = .001$. However, once again, the mean difference in looking to the target for the newly taught words

did not differ significantly from zero, regardless of context, $t(23) = 0.14$ and 1.88 , $p = .89$ and $.07$ for the familiar and unfamiliar contexts, respectively. Nor was there any significant effect of context when these were directly contrasted, $t(23) = 1.00$, $p = .32$. If anything, the results are actually in the opposite direction from that predicted by the INCDROP model, although this effect was not even marginally significant. However, the reaction times, for the nineteen subjects who provided usable data, tell a different story, one that is in the direction predicted by INCDROP. Unfortunately, while the time it took infants to shift their gaze to the target after the onset of the target word was numerically faster for the familiar context, 884 ms compared to 972 ms ($SE = 40$ and 60 ms, respectively), this difference was not close to being statistically significant, $t(18) = 0.41$, $p = .68$.

Given these results, it may be that this method is simply not sensitive enough to detect subtle INCDROP effects, or perhaps INCDROP only comes on-line later in development. Certainly, many models of adults' word recognition abilities assume that known words play an important role in on-line word segmentation (Elman & McClelland, 1986; Luce & Pisoni, 1998; Norris, 1994). In any case, these experiments did not support the notion that infants use previously learned words in learning new ones, at least at the phonetic level.

4. Another way that familiar words may affect word learning

If infants do not use the words they know to segment the speech stream, perhaps they use the words they know to find the correct word-to-world mapping. How might this occur? Golinkoff, Mervis, and Hirsh-Pasek, (1994) proposed a principle of novel-name-nameless category (N^3C), which states that infants will select an unnamed object as the referent for an unfamiliar label rather than an object for which these children already have a word. Thus, for example, when confronted with a rose and another unknown object, infants are more likely to select this unknown object as the referent for the word "dawnoo" than the rose. This idea is similar to Markman's (1989) principle of mutual exclusivity and Clark's (1983) principle of contrast. It implies that infants should be able to infer the referent of a new word if it is paired with a familiar word. Even if the "familiar" word has been recently learned, as in the case of the "pylon" and the "kiosk," learned in Experiment 1.

To explore this possibility, we decided to test the infants from Experiment 1 in an additional experiment, one week after their initial visit to the lab. The design of this experiment is shown in Table 2, and it examines three hypotheses across three sets of trials. The first set of trials asks whether infants even remembered the words across a one-week delay. For these trials, the taught words are paired with each other, as in the test phases of the previous experiments. The second set of trials asks whether infants could infer the meanings of new words when shown a novel object and another object corresponding to one of the newly learned words or one of the already known words, *apple* and *flower*. Notice that successful performance on these trials

would indicate that infants at least encoded some knowledge about the objects and labels of the week prior. Finally, the last set of trials asked whether infants actually showed any evidence of inferring, and remembering, the meaning of the novel objects during the course of the prior trials and despite the absence of any feedback about their responses to these words.

The results (shown in Figure 4) for twenty-four 24-month-olds indicate that although infants did, as usual, correctly identify targeted words, *apple* and *flower*, which were already in their vocabulary, $t(23) = 2.75, p = .005$, these infants did not seem to remember the words from the previous week, $t(23) = 1.46, p = .08$. This latter finding is not surprising given that they showed little evidence of learning these words during the first visit in the previous week. Nevertheless, in accordance with the principle of N³C, these infants were able to use these words to correctly infer the meaning of novel names. That is, looking toward the target object was significantly greater than zero on trials

Table 2. Design of Experiment 1B.

Visual	Sound Track
Flower & Apple	"Flower" (x 3)
Flower & Apple	"Apple" (x 3)
Pylon & Kiosk	"Pylon" (x 3)
Pylon & Kiosk	"Kiosk" (x 3)
Pylon & Modi	"Pylon" (x 3)
Dawnoo & Kiosk	"Kiosk" (x 3)
Pylon & Modi	"Modi" (x 3)
Dawnoo & Kiosk	"Dawnoo" (x 3)
Flower & Modi	"Flower" (x 3)
Dawnoo & Apple	"Apple" (x 3)
Flower & Modi	"Modi" (x 3)
Dawnoo & Apple	"Dawnoo" (x 3)
Modi & Dawnoo	"Modi" (x 3)
Modi & Dawnoo	"Dawnoo" (x 3)

** Note. The middle two sets of trials were counterbalanced across each other. The first and last set of trials were counterbalanced within each other.

where either a known or a taught object was paired with a new object, $t(23) = 4.43$ and 4.84 for the Known/Novel and Taught/Novel trials, respectively, $p < .0001$. Further, they correctly responded to each of these new words when the novel object were paired against each other on test trials, $t(23) = 2.45$, $p = .01$.

These results are interesting for several reasons. First, infants at this age demonstrate that they retain some memory for words taught a week earlier in a laboratory task (see also Markson & Spelke, 2000) this experiment provides evidence that infants at 24 months of age can infer meaning based on their knowledge of prior words. This suggests at least one level for which prior word knowledge affects infants' word learning. However, what is even more striking in these results is that infants' memory for newly learned and even previously known words is apparently strengthened by placing them in contrast against novel objects. That is, the Known/Novel trials and Taught/Novel trials provided the largest mean differences in looking of any of the trials from any of the experiments: even larger, on average than the known trials in the same experiment. Moreover, this increased responding is not simply the result of infant tendencies to look at novel objects. If that were the case, then looking to target would have been substantially blunted when the known and taught labels were requested. They were not. Instead, the mean difference in infant looking to target was universally high in both sets of trials.

This effect was replicated across each of the experiments previously reported, and it implies that infant memory is highly context sensitive. With respect to the words taught during their initial visit, infants were confused as to which label went with which object. However, when the same items were tested against a novel object, the infants were no longer confused. In a way, it is as if,

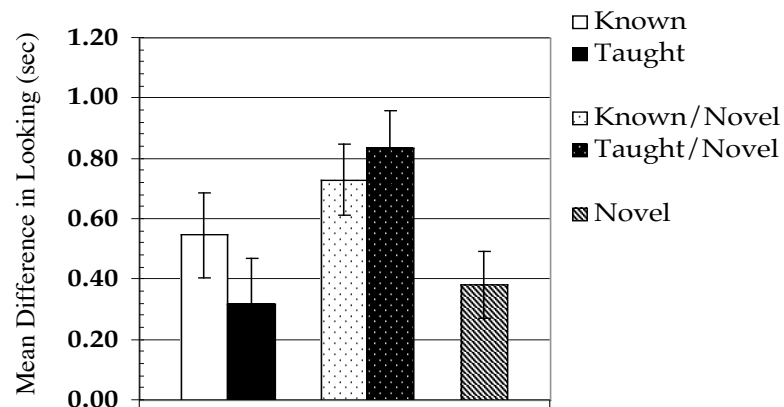


Figure 4. Mean difference in looking to target by condition.

they were saying to themselves, “I know I saw that object last week, and I heard that label last week, so those two definitely go together.

5. Conclusions

These studies examined the levels at which infants might use previously known words to learn new ones. Across three different studies, we failed to find consistent, statistically significant evidence that a familiar context aided infants in segmenting new words from fluent speech, as would be predicted by Brent’s (1999) INCDROP model. Still, there were some tendencies in the predicted direction for four of the six cases, involving accuracy or speed of responding.

At the conceptual level, however, infants appeared more than willing to assign linguistic meaning for novel words to a novel object in accord with Golinkoff, Mervis, and Hirsh-Pasek’s (1994) principle of N³C. Furthermore, infants correctly remembered these newly learned words in subsequent test trials, implying that these learned meanings have some conceptual stability and relevance.

The present findings are consistent with a context dependent model of memory, suggesting that words learned in close temporal proximity to each other might be inherently confusable. Infants appear to be more than capable of tracking novelty in their environment and using it to their advantage in learning a language. Thus, it would appear that previous familiarity with words does play an important role in learning new words, albeit at a slightly higher level of implementation than one suggested by INCDROP.

Endnotes

* The research reported here was supported by a Research Grant from NICHD (#15795) and a Senior Scientist Award from NIMH (#01490) to Peter W. Jusczyk. We thank Ann Marie Jusczyk and Natasha Scheitlin for their assistance in recruiting and testing the children. We also thank Lou Ann Gerken for her suggestion regarding Experiment 3.

References

- Bloom, Lois. (1973). *One word at a time: The use of single-word utterances before syntax*. The Hague: Mouton.
- Brent, Michael. (1999). An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning*, 34, 71-105.
- Brent, Michael R., & Cartwright, Timothy A. (1996). Distributional regularity and phonotactics are useful for segmentation. *Cognition*, 61, 93-125.

- Clark, Eve. (1983). Meanings and concepts. In James Flavell & Ellen Markman (Eds.) *Handbook of Child Psychology: Vol III. Cognitive Development*. New York: John Wiley.
- Dahan, Delphine & Brent, Michael. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General*, 128, 165-185.
- Golinkoff, Roberta M., Hirsh-Pasek, Kathryn, Cauley, Kathleen M., & Gordon, Laura. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23-45.
- Golinkoff, Roberta Michnick, Mervis, Carolyn B., & Hirsh-Pasek, Kathryn. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21, 125-155.
- Hollich, George, Hirsh-Pasek, Kathryn, & Golinkoff, Roberta Michnick. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65 (3, Serial No. 262).
- Hollich, George, Jusczyk, Peter W., & Luce, Paul. (2000, December). *Infant sensitivity to lexical neighborhoods during word learning*. Paper presented at the Meeting of the Acoustical Society of America, Newport Beach, CA.
- Jusczyk, Peter W., & Aslin, Richard N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1-23.
- Jusczyk, Peter W., Hohne, Elizabeth A., & Bauman, Angela. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, 61, 1465-1476.
- Jusczyk, Peter W., Houston, Derek, & Newsome, Mary (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159-207
- Luce, Paul A., & Pisoni, David B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19, 1-36
- Markman, Ellen. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
- Markson, Lori, & Spelke, Elizabeth S. (2000, July). Rapid learning of names and arbitrary properties of objects. Paper presented at the Meeting of the International Conference on Infant Studies, Brighton, England.
- Mattys, Sven L., & Jusczyk, Peter W. (in press). □Do infants segment words of recurring contiguous patterns? □ *Journal of Experimental Psychology: Human Perception & Performance*.
- McClelland, James L., & Elman, Jeffrey L. (1986) The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- Nazzi, Thierry, Jusczyk, Peter W., Bhagirath, Keya M. (1999, April). *Infants' segmentation of verbs from fluent speech*. Paper presented at the meeting of the Society for Research in Child Development, Albuquerque, NM.
- Norris, Dennis. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.

- Saffran, Jennifer R., Aslin, Richard N., & Newport, Elissa L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.
- Schafer, Graham, & Plunkett, Kim. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*, 309-320.
- Swingle, Dan, Pinto, John P., & Fernald, Anne. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*, 73-108.