Published in final edited form as:

Lang Learn. 2014 September 1; 64(Suppl 2): 106–120. doi:10.1111/lang.12057.

Sounds and meanings working together: Word learning as a collaborative effort

Jenny Saffran

Department of Psychology, University of Wisconsin – Madison

Abstract

Over the past several decades, researchers have discovered a great deal of information about the processes underlying language acquisition. From as early as they can be studied, infants are sensitive to the nuances of native-language sound structure. Similarly, infants are attuned to the visual and conceptual structure of their environments starting in the early postnatal period. Months later, they become adept at putting these two arenas of experience together, mapping sounds to meanings. How might learning sounds influence learning meanings, and vice versa? In this paper, I will describe several recent lines of research suggesting that knowledge concerning the sound structure of language facilitates subsequent mapping of sounds to meanings. I will also discuss recent findings suggesting that from its beginnings, the lexicon incorporates relationships amongst the sounds and meanings of newly learned words.

Introduction

Over the course of the first months of infants' post-natal lives, they are bathed in a sea of sounds (or signs, if they are acquiring a signed language). Indeed, even before birth, fetuses have the opportunity to begin to learn about their linguistic environment (e.g., DeCasper & Spence, 1986). By the time infants begin to map sounds and meanings onto one another, they have already amassed a tremendous amount of experience in both domains. Nevertheless, the vast majority of studies on word learning involve pairing a novel label with a novel object, with little attention to learners' prior experience with those particular sounds. We have clearly learned a great deal about the process of word learning from such studies. But could it be the case that infants' prior experiences – before entering the lab – guide how they approach word-learning tasks? That is, word learning likely does not occur in a vacuum; prior experience with the world seems very likely to impact what is most readily acquired in the lab.

A case in point comes from research on the shape bias in word learning – namely, infants' expectation that novel words tend to be extended to objects that are similar in shape (e.g., Landau, Smith, & Jones, 1988). The body of evidence that has emerged surrounding the shape bias suggests that while there may be some innate predispositions to attend to shape (domain-specific and/or domain-general), word learning itself provides “on the job” training for later word learning (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). That is, the shape bias in word learning emerges out of the nascent lexicon, because the preponderance of early nouns are themselves extended to other category members on the basis of shape.

Taking this type of view and applying it to word learning more generally, it becomes easier to imagine circumstances in which prior learning might facilitate later learning in the domain of lexical acquisition. In particular, we will consider the converse of the shape bias (a constraint on the types of referents likely to be mapped onto labels). Here, we will consider how prior learning might shape the types of labels that are most likely to be mapped onto referents. We will then ask how these two types of processes – representations of sounds and representations of meanings – might work together as the infant lexicon begins to take shape.

Knowledge of sound sequences facilitates the process of mapping sounds to meanings

Even before birth, fetuses become attuned to certain sound properties of their native language, predominantly rhythmic patterns (e.g., DeCasper & Spence, 1986). During the first postnatal year, their native language knowledge is increasingly fine-tuned, including characteristics of native phoneme categories, phonotactic regularities, and lexical stress patterns (for review, see Kuhl, 2004; Saffran, Werker, & Werner, 2006). Importantly, this information plays a functional role for infant learners. Sound sequences that are most similar to the native language are most easily segmented from fluent speech. For example, during the second half of the first year of postnatal life, English-learning infants are better at detecting trochaic (strong-weak) bisyllabic patterns in fluent speech than iambic (weak-strong) patterns, reflecting the fact that bisyllables in English are overwhelmingly trochaic (e.g., Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003). Similar results obtain for sequences that match the phonotactic regularities of the native language; sequences that follow English-like patterns (including locations of segments relative to word boundaries) are more readily segmented than sequences that violate those patterns (e.g., Mattys, Jusczyk, Luce, & Morgan, 1999).

This research has also revealed potential mechanisms that support infants' auditory linguistic discoveries. One such mechanism is known as statistical learning: the ability to track regularities of various sorts in the input. These regularities might take the form of frequencies of occurrence, frequencies of co-occurrence, pair-wise probabilities (forward and backward), and other statistical patterns that live in languages (for recent reviews, see Romberg & Saffran, 2010; Thiessen & Erickson, 2013).

How might information about sound patterns provide clues to word learners? Saffran & Graf Estes (2006) lay out a number of possibilities. Perhaps sound patterns that are more familiar are easier to represent, and thus easier to map onto referents. Perhaps sound patterns that more closely resemble known words are more likely to be treated by infants as candidate words, available for mapping to meanings. Perhaps sound patterns that are already segmented from fluent speech are simply easier to map to meanings than sound patterns that have not yet been segmented (though see Shukla, White, & Aslin, 2011, for evidence that the processes of segmentation and mapping to meaning can occur simultaneously as well). In the next section, I will review several recent studies from our lab that speak to these issues.

From word segmentation to word learning

Unlike written language, where the boundaries between words are demarcated by spaces, spoken language does not contain pauses as reliable cues to word boundaries. This is easy to observe when listening to a foreign language: speech appears to fly by very quickly, with no breaks other than breaths. Even when speaking to infants, however, we do not place consistent pauses between words. This raises a fascinating problem in the arena of infant language development: how do infants figure out where words begin and end, given that they are not marked by clear acoustic boundaries? If learners can't figure out where words begin and end, it would be very difficult to learn much else about the language (what words mean, how words combine into sentences, etc.).

We have focused on a specific hypothesis about how infants might discover boundaries between words: **statistical learning**. The idea is that infants can use the distributional properties of speech to segment speech into words (e.g., Goodsit, Morgan, & Kuhl, 1993; Hayes & Clark, 1970). Consider a sequence like “*pretty baby*”. As native speakers, we know there is a boundary between *ty* and *ba*. But to a novice learner, those four syllables might each be a word. Or perhaps the four syllables are a single four-syllable word. How do learners figure out that the correct boundary lies between *ty* and *ba*?

Intuitively, syllables that are part of the same word co-occur together more frequently than syllables spanning a word boundary. Returning to the “*pretty baby*” example, syllables from the same word, like *pre-tty*, are more likely to co-occur than syllables spanning a word boundary, like *ty-ba*. Notably, infants are sensitive to these types of regularities. In a 1996 paper, we tested the hypothesis that infants are sensitive to syllable co-occurrence statistics (Saffran, Aslin, & Newport, 1996; see also Aslin, Saffran, & Newport, 1998). To do so, we exposed 8-month-old infants to a novel language. This “artificial language” was created for the purpose of the experiment to ensure that there were no cues to word boundaries other than the statistical properties of the speech: syllables that followed one another reliably were part of the same word, while syllables that did not predict one another spanned word boundaries. After 2 minutes of exposure to the artificial language, we tested infants’ listening preferences for words from the language versus sequences that spanned word boundaries (part-words, like *tyba* in the *prettybaby* example). The results revealed a reliable pattern of preferences, indicating that the infants had picked up the statistical properties of the speech we had played for them.

Of course, statistics aren't the only possible cue to word boundaries. Infants are also attuned to other properties of speech, including rhythm and pauses, which might provide information about word boundaries (e.g., Johnson & Jusczyk 2001). In one follow-up study, we exposed infants to a new artificial language that had the pitch patterns characteristic of infant-directed speech (which tends to be far more sing-song than speech to adults, and which even newborns prefer over adult-directed speech). We found that infants were better at tracking the statistical properties of speech when it was presented with infant-directed pitch contours (Thiessen, Hill, & Saffran, 2005). In more recent studies, we have found that infants similarly track the statistics of speech when presented with real languages with which they are unfamiliar, such as Italian (Pelucchi, Hay, & Saffran, 2009a, 2009b).

These findings are typically described as evidence for word segmentation: discovering word-like units in fluent speech. However, they do not actually present evidence for word segmentation per se. These studies – and many others – typically show that after exposure to fluent speech, infants can discriminate between sequences that were words in the fluent speech and sequences that were not words (e.g., part-words spanning word boundaries, like *tyba* in the *prettybaby* example above). This type of finding in itself is not evidence for word segmentation. It is definitely evidence that infants are sensitive to the properties that distinguish these sequences. But studies using discrimination tests cannot tell us very much about the representations that emerge from infants' learning mechanisms.

To address this issue, Graf Estes designed an ingenious set of experiments involving multiple stages of language input (Graf Estes, Alibali, Evans, & Saffran, 2007). In these studies, she asked whether prior exposure to the statistics of the speech stream facilitates subsequent mapping of novel words to novel objects. Seventeen-month-old infants were first exposed to a fluent stream of speech from an artificial language. The only cue to word boundaries was the statistical structure of the stream: the transitional probability within each bisyllabic word was 1.0, with a dip to 0.5 at word boundaries. After this exposure, they were not tested on their ability to discriminate between words and part-words, as in previous experiments. Instead, they entered a novel-word training phase. Each infant was exposed to two novel labels paired with two novel referents, with trials continuing until the infant had habituated to the pairings. Crucially, however, there was a between-subject manipulation concerning the labels. For half of the infants, the labels were words from the fluent speech heard at the beginning of the study. For the other half of the infants, the labels were part-words from the fluent speech. The words and part-words were frequency-matched, having occurred equally often in the fluent speech presented at the start of the experiment. The only feature of the stimuli that distinguished the words from the part-words were their internal transitional probabilities (1.0 for the words, 0.5 for the part-words).

Following habituation, the infants were tested using the Same-Switch procedure (e.g., Stager & Werker, 1997). Half of the test trials were identical to the training trials (Same trials), while the other half of the test trials switched the labels and objects to create unfamiliar mappings (Switch trials). The logic of the task is that if learners have acquired the mappings between labels and objects, then the Switch trials should be novel, and thus relatively interesting. However, if infants haven't yet acquired the mappings, then the Switch trials should not be novel to them, given that the labels and objects are all familiar. Learning is thus measured by increased looking (dishabituation) on Switch trials relative to Same trials.

The results of this experiment revealed that only those infants for whom the labels were words from the speech stream showed a significant Switch preference on the test. Infants for whom the labels were part-words showed no evidence of learning. This pattern of results is particularly striking because the words and part-words were frequency matched; that is, both types of items occurred equally often during the fluent speech presented at the beginning of the experiment. The only difference between the words and part-words was their internal statistical structure. The words had high internal transitional probabilities, whereas the part-words spanned word boundaries. This information appears to have been relevant to the subsequent label-object mapping task.

The results of Graf Estes et al. (2007) thus support the contention that statistical learning processes provide, as output representations, candidate words, available for mapping to meanings. A notable feature of this study is that it demonstrates a cascade of effects, such that sequences with higher transitional probabilities in the first phase of learning served as better labels for objects in the second phase. This type of multi-level learning, while likely highly pertinent to natural language acquisition, is not usually tested in lab tasks, which tend to isolate a single aspect of language in any given study.

One potential concern about the generalizability of the study by Graf Estes et al. (2007), however, is that the materials were highly artificial. As in the original studies by Saffran et al. (1996) and Aslin et al. (1998), the language contained only four words, and was presented as synthesized speech in a monotone, isochronous, pause-free stream. It is thus possible that we only saw effects of statistical structure on word learning because the materials were so impoverished. We addressed this issue in a subsequent study (Hay, Pelucchi, Graf Estes, & Saffran, 2011). The materials in the Hay et al. (2011) study were drawn from natural Italian corpora previously developed for use in studies of natural-language word segmentation by Pelucchi et al. (2009a, 2009b). The fluent speech was presented in an infant-directed register, with a far larger phonemic and syllabic inventory and markedly fewer repetitions of each target word than in the prior artificial language studies. Hay et al. (2011) performed a conceptual replication of Graf Estes et al. (2007), replacing the artificial materials with the Italian materials (which were unfamiliar to our American infants). The results were highly consistent with the prior study, suggesting that even given rich natural materials, learners continued to treat sequences with high internal statistics as better possible labels for objects than sequences with lower internal statistics.

These studies, taken together, suggest that the array of knowledge infants acquire about sound patterns is not just useful for learning about speech. Instead, this knowledge base facilitates the process of mapping those sound sequences onto meanings. These processes may occur sequentially, as in the studies by Graf Estes et al. (2007) and Hay et al. (2011). This ordering – learning about sounds, followed by mapping to meaning – is consistent with what we previously believed about word learning: namely, that infants didn't really engage in much word learning prior to 1 year of age. Recent studies, however, suggest that even 6-month-olds, who don't yet have a full grasp on their native language's sound system, already understand certain highly frequent words (Bergelson & Swingley, 2012). This finding suggests that learning about sounds and mapping sounds to meaning likely also occurs simultaneously. Supporting this contention, Shukla et al. (2011) demonstrated that 6-month-olds can simultaneously segment speech from an artificial language and map sounds to novel referents, at least under certain circumstances. Indeed, analyses of child-directed speech suggest that it is a particularly good stimulus to afford such simultaneous processes (Yurovsky, Yu, & Smith, 2012).

If infants are sensitive to the properties of the speech stream, and if this information influences their ability to map units from that speech stream to meanings, what are the implications for natural language learning? We would expect that based on what infants have learned about the sound patterns of their native language, they should find some native language sequences easier to learn as labels than others. This general pattern has been

observed in studies with pre-school children, who are more facile at learning novel words that conform to the phonotactic probability structure of English (e.g., Storkel, 2001). Would novice word learners similarly show effects of English phonotactics on novel word learning? We tested this hypothesis by manipulating the phonotactics of novel labels in a word-learning study with typically developing 18-month-olds (Graf Estes, Edwards, & Saffran, 2011). The infants each received training on a pair of novel words, mapped to novel objects. Half of the infants heard labels that were phonotactically legal in English (*dref*, *sloob*). The other half of the infants heard labels that were phonotactically illegal in English (*dlef*, *sroob*). Infants were then tested on their word learning using a looking-while-listening task. We hypothesized that because the legal items are consistent with English sound statistics, infants should successfully acquire mappings between those labels and their referents. The illegal items, on the other hand, violate English sound statistics, and should thus be more difficult for infants to map to their referents.

In fact, the results of this study were somewhat more nuanced. We found an interaction between the participants' level of English attainment and their pattern of successful word learning. Only those infants with larger native-language vocabularies showed successful word learning with the legal labels; infants with smaller English vocabularies (as measured by parental report on the Bates-MacArthur Child Language Development Inventory) failed to learn the legal labels. However, the infants with smaller English language vocabularies successfully acquired the illegal labels, while the infants with larger vocabularies did not. This pattern of results is intriguing. It's not that the infants with smaller English language vocabularies were "worse" word learners in this task; they actually learned some items (the illegal labels for objects) that the larger-vocabulary infants failed to learn. Instead, the results suggest that for infants with larger vocabularies, there is a greater impact of native-language statistics: labels that are consistent with English sound structure are easier to learn. The infants with smaller vocabularies do not show facilitory effects of English knowledge. Instead, we hypothesized that they found the illegal labels easier to learn because they were somewhat more salient than the legal labels.

It thus appears that native sound structure impacts word learning in the lab. Indeed, the statistics of the real world appear to impact word learning in multiple ways. For example, when the manipulation involves phonotactic probability (high versus low – as opposed to legal versus illegal as in Graf Estes et al., 2011), infants showed better label-object mapping when the labels consisted of more probable sequences relative to less probable sequences (Graf Estes & Bowen, 2012). Similarly, English-learning infants find novel labels consistent with English lexical stress (bisyllabic trochees) easier to map to objects than iambic bisyllables, which are inconsistent with the typical English stress pattern (Graf Estes & Bowen, 2012).

We have also observed other types of relationships between native-language regularities and word learning. In a recent study, we tested English-learning infants in a word learning task where the labels contrasted only in lexical tone (Hay, Graf Estes, Wang, & Saffran, 2014). While this dimension of sound is crucial to meaning contrasts in tonal languages like Mandarin or Thai, it does not carry lexical meaning in English. Thus, for English learners, the label /kʊ/ produced with a rising pitch contour and the label /kʊ/ produced with a falling

pitch contour should have the same meaning. However, for infants learning a tonal language, these two versions of /kʊ/ could have entirely different meanings. In our study, we asked whether there was an effect of native language exposure such that younger infants might be open-minded about whether or not tone contour signaled different meanings, with older infants following the English pattern (tone contour is irrelevant to meaning). Infants at 14 and 19 months of age were tested in a Same-Switch word learning paradigm. There were two labels (rising /kʊ/ and falling /kʊ/), mapped to two distinct objects.

The results of this study were extremely interesting. The younger infants, but not the older infants, successfully mapped these two labels onto their referents. Thus, unlike most studies, we found greater word learning success in younger infants than older infants. This pattern of results suggests that the younger infants, with less knowledge about English, were indeed more open-minded about the dimensions of sound that are relevant to meaning contrasts. The older infants, however, were more bound by English-relevant patterns, as has been found in other studies with older toddlers and adults using pitch contrasts (Quam & Swingley, 2010). These findings are broadly consistent with prior literature suggesting that perceptual learning processes shape what infants consider to be possible labels (e.g., Galle, Apfelbaum, & McMurray, 2014; Rost & McMurray, 2010; Yeung, Chen, & Werker, 2013).

In sum, it appears that experience with sound structure does more than help infants learn sound structure. This knowledge base provides constraints on subsequent learning processes. Studies suggest that these constraints arise at multiple levels. The most specific level is the exact phoneme/syllable sequences that are likely to be words, as shown in the studies contrasting words with part-words as labels for objects (e.g., Graf Estes et al., 2007; Hay et al., 2011). The next level entails generalizations about the types of phoneme/syllable sequences that are likely to be words, as shown in the studies contrasting legal/illegal phonotactic patterns, frequent/infrequent phonotactic patterns, or trochaic/iambic patterns as potential labels (e.g., Graf Estes et al., 2011, Graf Estes & Bowen, 2012). At the most general level, the distribution of sounds in the native language provides information about which dimensions of sound are meaning-relevant and which are meaning-irrelevant, as shown in the tone contrast study by Hay et al. (2014). All of this accumulated knowledge likely facilitates word learning in important ways, constraining potential labels and simplifying the combinatorial explosion that young word learners must face. Indeed, this cascade continues at higher levels of language structure. For example, the distributions of words, and their sound structure, influences toddlers' abilities to discover lexical category information (e.g., Lany & Saffran, 2010; 2011). It thus appears that what infants have already learned about sound has a potentially profound effect on their ability to glean information about meaningful patterns in the input.

The ontogeny of lexical networks

As learners begin to glean knowledge about new words, how is that knowledge organized? That is, are words learned and represented quite separately, with little interaction between nascent lexical items? Or is there something like a lexicon – organization based on form and function – from the beginning?

We know that on the sound side of things, there is a neighborhood structure into which words are organized. Words are related based on their beginnings (cohort similarity) and ends (rhyme similarity). There are also effects of phonotactic probability and neighborhood density on word learning – these are effects of the sounds of known words on newly learned words, as discussed in the previous section (e.g., Graf Estes et al., 2011; Hollich, Jusczyk, & Luce, 2002). Once learned, the sounds of words do not appear to exist in a vacuum. Instead, they exert reciprocal effects on one another.

We see similar effects on the meaning side. Words are organized into semantic networks, which are characterized by rich interconnections amongst meanings (e.g., Rogers & McClelland, 2004). These relationships can be investigated in adults and older children using semantic priming tasks, which can be used to probe specific types of connections amongst words (e.g., semantic category membership, thematic relations, feature overlap, etc.). Importantly, toddlers also show broadly effects of semantic relatedness when tested using tasks designed to assess age-appropriate variants of priming (e.g., Arias-Trejo & Plunkett, 2009; Willits, Wojcik, Seidenberg, & Saffran, 2013). These studies suggest that at least for highly familiar words, toddlers represent semantic relationships in a lexical network.

How and when do these relationships emerge? As toddlers acquire new words, are they also acquiring the relationships amongst those words? To address the issue of the ontogeny of lexical networks, we developed a new word-learning task in which we manipulated the similarities amongst the referents of those words (Wojcik & Saffran, 2013). For this first set of studies, we operationalized similarity in the simplest possible way: degree of visual overlap amongst the referents. Two-year-old toddlers were trained on a miniature artificial lexicon containing four novel label-object pairings. The lexicon was structured such that each object was similar to one of the other objects, and dissimilar from the other two. In this fashion, we created the opportunity for learners to represent the relationships amongst the meanings of the novel words during the word learning process.

Following exposure to the four label-object pairs, the toddlers were tested using a variant of the Headturn Preference Procedure. We had previously developed this method to investigate semantic priming using a purely auditory task with familiar words, avoiding contamination from visual similarity (Willits et al., 2013). In the current study, participants heard repetitions of a pair of novel words on each trial (e.g., *tursey, coro, tursey, coro...*) with no visual referents present. On half of the trials, the word pairs referred to similar-looking objects. On the other half of the trials, the word pairs referred to dissimilar-looking objects. The question of interest was whether toddlers would show differential looking to the two types of trials, which differed only in the similarity of the referents (which were not pictured during testing). In our previous study with familiar English words (Willits et al., 2013), we found that toddlers discriminated between pairs with similar meanings (*dog, kitty, dog, kitty...*) versus pairs with more distant meanings (*dog, shoe, dog, shoe...*).

The results of the novel word learning study showed that indeed, the toddlers did distinguish word pairs based on the visual similarity of their referents (Wojcik & Saffran, 2013). Toddlers discriminated between novel word pairs with similar referents versus novel word

pairs with dissimilar referents. This finding is exciting for several reasons. The first is methodological. It appears that our task is sensitive to learners' internal representations of the meanings of words, without needing to depict the meanings visually during testing. This suggests that our task is suitable to address numerous previously-unstudied questions concerning the dimensions of similarity pertinent to young learners. Indeed, we are currently investigating the manner in which sentential positions of novel words render them similar/dissimilar to one another using this method (Wojcik & Saffran, 2014).

The second implication of this line of research is that as toddlers acquire words, they also represent the relationships amongst those words. Words with similar meanings are treated differently than words with dissimilar meanings. And simply hearing the sound portion of these words activates meanings in a way that is relevant to cross-word relationships, even for words that have just been acquired. These results suggest continuity in language processing across development; semantic networks are being grown from the outset of word learning.

Conclusions

For a variety of reasons, developmental researchers have tended to investigate “levels” of language independently of one another. Some researchers focus on the sounds of language, some on mapping to meaning, others on syntax, yet others on pragmatics, etc. In part, this division of labor has evolved for methodological reasons: it's not been obvious how to design interpretable studies that integrate multiple levels of linguistic structure. There are also historical roots: different labs and research traditions have focused on different aspects of the language acquisition puzzle (e.g., scholars focused on the logical problem of language acquisition, inductive reasoning, etc., have been largely distinct from those who focus on speech sound learning, who have been largely distinct from those who focus on social influences on language development).

These distinctions, however, have begun to break down in very productive ways. For example, some of the most influential work on word learning has come from a research group that was previously focused on speech sound acquisition (e.g., Stager & Werker, 1997; Werker & Yeung, 2005). Similarly, researchers who previously focused primarily on speech sound acquisition have drawn important connections to the role of social interaction (e.g., Kuhl, Tsao, & Liu, 2003). As these interconnectivities increase, we will find the field of language development in the position that is currently enjoyed by adult psycholinguistics, which has now acknowledged the role of mutual interactive constraints for two decades (e.g., MacDonald, Pearmutter, & Seidenberg, 1994).

The research program described in this manuscript has taken steps in the direction of acknowledging – and exploiting – the multiple levels of structure that live in language to better understand the processes underlying language development. The results suggest that far from occurring in a vacuum, word learning is facilitated by the knowledge that infants have gathered about the sound patterns of their native language. By tracking sound patterns in speech, infants gain valuable information about specific sound sequences that might be words, about general patterns of sounds that typify words in the native language, and about

the aspects of sound that are pertinent to words in the native language. These processes constrain the possible words to be mapped to meanings.

Meanings also help to constrain the relationships amongst words as they are learned. The lexicon has its roots in the early connections and relationships amongst words. Toddlers are sensitive to these connections in both well-learned words and just-learned words. Overlapping meanings influence how toddlers process words, even in the absence of referents. Sound and meaning are thus integral to one another from the earliest phases of word learning.

By considering multiple dimensions of language simultaneously, it becomes possible to investigate the connections amongst them. The dynamic network that is emerging from research in our field is giving us new perspectives on previously distinct language learning problems. Like language learners, who must consider multiple dimensions of language simultaneously, researchers will benefit from the mutual constraints that emerge when the richness of language input is taken seriously.

Acknowledgments

I am grateful to the current and former members of the Infant Learning Lab at the University of Wisconsin – Madison for their many contributions to the development of these ideas. I am especially grateful to Katie Graf Estes, Jessica Hay, and Jill Lany, whose work formed the basis for the first sections of this manuscript, and to Erica Wojcik and Jon Willits, whose work formed the basis for the final sections of the manuscript. Preparation of this manuscript was funded by grants from the National Institute of Child Health and Human Development (to J. R. Saffran; R37HD037466), the Waisman Center (P30HD03352), and the James F. McDonnell Foundation (to J. R. Saffran).

References

- Arias-Trejo N, Plunkett K. Lexical semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2009; 364:3633–3647.
- Aslin RN, Saffran JR, Newport EL. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*. 1998; 9:321–324.
- Bergelson E, Swingley D. At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*. 2012; 109(9):3253–3258.
- DeCasper AJ, Spence MJ. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*. 1986; 9:133–150.
- Galle M, Apfelbaum K, McMurray B. Within-speaker variability benefits phonological word learning. *Language Learning and Development*. in press.
- Goodsit JV, Morgan JL, Kuhl PK. Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*. 1993; 20:229–229. [PubMed: 8376468]
- Graf Estes K, Bowen S. Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of Experimental Child Psychology*. 2012
- Graf Estes K, Edwards J, Saffran JR. Phonotactic constraints on infant word learning. *Infancy*. 2011; 16(2):180–197. [PubMed: 21297877]
- Graf Estes K, Evans JL, Alibali MW, Saffran JR. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*. 2007; 18(3):254–260. [PubMed: 17444923]
- Hay J, Graf Estes K, Wang T, Saffran JR. From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development*. 2014
- Hay JF, Pelucchi B, Estes KG, Saffran JR. Linking sounds to meanings: infant statistical learning in a natural language. *Cognitive Psychology*. 2011; 63(2):93–106. [PubMed: 21762650]

- Hayes, JR.; Clark, HH. Experiments in the segmentation of an artificial speech analog. In: Hayes, JR., editor. *Cognition and the development of language*. New York: Wiley; 1970. p. 221-234.
- Hollich, G.; Jusczyk, PW.; Luce, PA. Lexical neighborhood effects in 17-month-old word learning. *Proceedings of the 26th Annual Boston University Conference on Language Development*; Boston, MA. Cascadilla Press; 2002. p. 314-323.
- Johnson EK, Jusczyk PW. Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*. 2001; 44(4):548–567.
- Kuhl PK. Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*. 2004; 5:831–43.
- Kuhl PK, Tsao F-M, Liu H-M. Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*. 2003; 100:9096–9101.
- Landau B, Smith LB, Jones SS. The importance of shape in early lexical learning. *Cognitive Development*. 1988; 3(3):299–321.
- Lany J, Saffran JR. From Statistics to Meaning: Infants' Acquisition of Lexical Categories. *Psychological Science*. 2010; 21(2):284–291. [PubMed: 20424058]
- Lany J, Saffran JR. Interactions between statistical and semantic information in infant language development. *Developmental Science*. 2011; 14(5):1207–1219. [PubMed: 21884336]
- MacDonald MC, Pearlmutter NJ, Seidenberg MS. The lexical nature of syntactic ambiguity resolution. *Psychological Review*. 1994; 101(4):676. [PubMed: 7984711]
- Mattys SL, Jusczyk PW, Luce PA, Morgan JL. Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*. 1999; 38(4):465–494. [PubMed: 10334878]
- Pelucchi B, Hay JF, Saffran JR. Statistical learning in a natural language by 8-month-old infants. *Child Development*. 2009a; 80:674–685. [PubMed: 19489896]
- Pelucchi B, Hay JF, Saffran JR. Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*. 2009b; 113(2):244–247. [PubMed: 19717144]
- Quam C, Swingle D. Phonological knowledge guides two-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language*. 2010; 62:135–150. [PubMed: 20161601]
- Romberg A, Saffran JR. Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*. 2010; 1:906–914. [PubMed: 21666883]
- Rogers, TT.; McClelland, JL. *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press; 2004.
- Rost GC, McMurray B. Finding the signal by adding noise: The role of non contrastive phonetic variability in early word learning. *Infancy*. 2010; 15(6):608.
- Saffran JR, Aslin RN, Newport EL. *Science*. 1996; 274:1926–1928. [PubMed: 8943209]
- Storkel HL. Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*. 2001; 44:1321–1337.
- Thiessen ED, Erickson LC. Beyond Word Segmentation A Two-Process Account of Statistical Learning. *Current Directions in Psychological Science*. 2013; 22(3):239–243.
- Thiessen ED, Hill E, Saffran JR. Infant-directed speech facilitates word segmentation. *Infancy*. 2005; 7:53–71.
- Saffran JR, Graf Estes K. Mapping sound to meaning: Connections between learning about sounds and learning about words. *Advances in Child Development and Behavior*. 2006; 34:1–38. [PubMed: 17120801]
- Saffran, JR.; Werker, J.; Werner, L. The infant's auditory world: Hearing, speech, and the beginnings of language. In: Siegler, R.; Kuhn, D., editors. *Handbook of Child Development*. New York: Wiley; 2006. p. 58-108.
- Shukla M, White KS, Aslin RN. Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*. 2011; 108(15): 6038–6043.
- Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, Samuelson L. Object name learning provides on-the-job training for attention. *Psychological Science*. 2002; 13(1):13–19. [PubMed: 11892773]

- Stager CL, Werker JF. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*. 1997; 388(6640):381–382. [PubMed: 9237755]
- Thiessen ED, Hill E, Saffran JR. Infant-directed speech facilitates word segmentation. *Infancy*. 2005; 7:53–71.
- Thiessen ED, Saffran JR. When cues collide: Statistical and stress cues in infant word segmentation. *Developmental Psychology*. 2003; 39:706–716. [PubMed: 12859124]
- Werker JF, Yeung HH. Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*. 2005; 9(11):519–527. [PubMed: 16202639]
- Willits JA, Wojcik EH, Seidenberg MS, Saffran JR. Toddlers Activate Lexical Semantic Knowledge in the Absence of Visual Referents: Evidence from Auditory Priming. *Infancy*. 2013
- Wojcik EH, Saffran JR. The Ontogeny of Lexical Networks Toddlers Encode the Relationships Among Referents When Learning Novel Words. *Psychological Science*. 2013; 24(10):1898–1905. [PubMed: 23938274]
- Wojcik EH, Saffran JR. Toddlers encode the lexical structure of novel words from meaningful sentences. Manuscript under editorial review. 2014
- Yeung HH, Chen KH, Werker JF. When does native language input reorganize phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*. 2013; 68(2):123–139.
- Yurovsky D, Yu C, Smith LB. Statistical speech segmentation and word learning in parallel: scaffolding from child-directed speech. *Frontiers in Psychology*. 2012; 310.3389/fpsyg.2012.00374