*Language and Linguistics Compass* 3/6 (2009): 1379–1389, 10.1111/j.1749-818x.2009.00164.x

# From Tracking Statistics to Learning words: Statistical Learning and Lexical Acquisition

Katharine Graf Estes*
*Department of Psychology and Center for Mind and Brain, University of California, Davis*

## Abstract

By age 1, infants display remarkable sensitivity to the sound structure of their native language. Statistical learning, the process of detecting structure in the environment by tracking patterns in the input, is hypothesized to contribute to infants' early learning about sound. The present paper explores how infants' ability to track distributional information in the speech signal contributes to a fundamental aspect of language development, linking sounds with meanings in word learning. Previous research has demonstrated that infants detect several cues that mark where words begin and end in the fluent stream of speech (e.g., transitional probability, phonotactic regularities). Tracking such patterns may allow infants to isolate individual words, making them available to be associated with referents. Even very early in vocabulary development, statistical learning about which sound sequences are likely or unlikely to occur within words in the native language may also shape word learning. We propose that early experience with sound sequence regularities provides infants with a foundation for lexical acquisition.

Infants' ability to detect sound sequence regularities in linguistic input is remarkable. From Saffran and colleagues' (1996) initial demonstration of infant statistical learning, there has been a great deal of interest in the role the mechanism plays in language acquisition. Statistical learning, the process of detecting structure in the environment by tracking patterns in the input, is not limited to linguistic or auditory information, and is not limited to infants or even to humans. Infants, children, and adults have been shown to track distributional information in non-speech auditory sequences (i.e., pure tones, Creel et al. 2004; Saffran et al. 1999) as well as in visual sequences (i.e., shapes, Fiser and Aslin 2002; Kirkham et al. 2002; Turk-Browne et al. 2008). Non-human animals such as rats and monkeys can also perform statistical learning tasks (Hauser et al. 2001; Toro and Trobalón 2005; Saffran et al. 2008). However, despite the input- and species-generality of statistical learning, the mechanism is particularly well-suited to supporting language acquisition. Natural speech is rich with distributional cues to many levels of linguistic structure, including phoneme distinctions and phoneme combinations, word boundaries, syntactic categories, and permissible orderings of words (e.g., Brent and Cartwright 1996; Cartwright and Brent 1997; Mintz et al. 2002; Vallabha et al. 2007). Thus, the ability to track distributional patterns holds potential to be highly useful to young language learners. Recent investigations of statistical learning have sought to extend understanding of the mechanism by examining how infants take advantage of distributional information to solve real challenges facing language learners. Here, we explore how infants might apply statistical learning to a fundamental task in language acquisition, linking sounds with meanings to learn new words.

To associate a meaning with a word form, infants must first be able to isolate the word from the fluent stream of speech. Even speech addressed to infants contains only a small proportion of words in isolation (Woodward and Aslin 1990; Brent and Siskind 2001).

Segmenting words from continuous speech is not a trivial problem, as it does not contain any fully reliable or obvious acoustic markers of word boundaries. However, there are several probabilistic cues that point to word onsets and offsets (e.g., stress, Curtin et al. 2005; and phonotactics, Jusczyk et al. 1999; Mattys and Jusczyk 2001; McQueen 1998). One such cue is transitional probability information. Over a corpus of speech, the transitional probability from one sound to the next tends to be higher for sounds that occur within the same word, whereas the probability across sounds that span a word boundary tends to be lower. For example, Saffran (2003) pointed out that within the phrase 'pretty baby' the transitional probability across the syllables *pre-* to *–ty* is .8. That is, 80% of the time infants hear *pre-* it is followed by *–ty*. In contrast, the transitional probability from *–ty* to *ba-* is .003; only .03% of the time infants hear *–ty* it precedes *ba*. This difference in probabilities provides a cue that *pretty* is an individual word, but *ty-ba* is not.

Saffran et al. (1996) tested infants' ability to detect transitional probability cues to word boundaries. In this experiment, infants first listened to an artificial language consisting of a set of four three-syllable nonsense words concatenated in a fluent speech stream (e.g., orthographically *golabupabikutibudodapiku*). Artificial languages are commonly used in tests of learning processes because they allow for careful control over the distributional information available in the speech signal. In this case, the language was designed to eliminate pauses, stress, and all other cues to word boundaries except transitional probabilities. Within the words of the language, the transitional probability between syllables was 1.0 (e.g., the syllables in *golabu* always occurred in sequence). Across word boundaries the transitional probability was .33 (e.g., because *golabu* was followed by three different words). After only 2 min of exposure to the artificial language, infants showed that they could detect these sound sequence regularities. In testing using a listening time measure, infants successfully discriminated *words* from the language (sequences with perfect transitional probability) from *nonwords*, novel sequences of syllables from the language (with zero transitional probability). They showed this discrimination by listening longer to the novel sequences (for further discussion of infant novelty and familiarity preferences, see Hunter and Ames 1988; as well as Houston-Price and Nakai 2004; and Thiessen et al. 2005). In addition, infants distinguished between *words* and *part-words*, syllable sequences that the infants had heard in the language but possessed low transitional probability because they spanned word boundaries (see also Aslin et al. 1998). Saffran et al. proposed that the process of tracking statistical regularities in sound sequences may allow infants to extract individual words from fluent speech, providing an early means for young learners to break into the speech signal.

Saffran et al.'s study focused on syllable-level statistical cues to word boundaries; there are also statistical regularities in the phonemes and phoneme sequences that are likely or unlikely to occur at certain positions within the words of a language. For example, English words can begin with /t/ or /ʃ/, but more words begin with /t/. No English words begin with /ŋ/. For consonant clusters, [tr] occurs word-initially, but not word-finally; [ts] occurs word-finally, but not word-initially. The sequence [tr] also begins more words than the sequence [ʃr] (see Vitevitch and Luce 2004; for a phonotactic probability calculator). As discussed in more detail below, by 9 months of age infants are sensitive to their native language phonotactic patterns: the frequencies of and constraints on phoneme sequences in the language (Friederici and Wessels 1993; Jusczyk et al. 1993, 1994). This information is important because these phonotactic patterns can provide word boundary information. Sets of phonemes vary in how likely they are to occur within versus across words. For example, the sequence [mr] does not occur word-initially or word-finally in any English words (McQueen 1998). When a listener encounters the sequence, it must indicate a syllable boundary, and syllable boundaries correlate highly with word boundaries. McQueen (1998) demonstrated that adults can take advantage of this cue to speed word recognition.

Mattys and Jusczyk (2001; see also Mattys et al. 1999); investigated whether infants' precocious knowledge of native-language sound structure (reviewed in Saffran et al. 2006) would allow them to detect phonotactic markers of word boundaries. They played 9-month-olds sets of sentences in which novel target words were embedded in phonotactic contexts that supported segmentation of the target word (i.e., phoneme combinations that typically occur across word boundaries) and sentences in which target words were embedded in contexts that do not support segmentation (i.e., phoneme combinations that often occur within words). For example, some infants heard the novel word *tove* presented after words ending in /v/ (e.g., 'brave *tove*'*)*; the sequence [vt] is unlikely within English words. For other infants, *tove* was preceded by /f/ (e.g., 'gruff *tove*'); the sequence [ft] occurs within English words like *gift* and *lift*. Infants successfully discriminated words from across-word and within-word phonotactic contexts. They listened longer to repetitions of words that had occurred in the contexts supporting segmentation. This finding indicates that infants are attuned to the sound sequence regularities that typically occur within versus across words in their native language, cues that are useful for identifying where words begin and end. Thus, the ability to detect probabilistic word boundary cues is not limited to in-lab experience with artificial languages.

Saffran et al. (1996) and Mattys and Jusczyk's (2001) experiments suggest that infants track statistical word boundary markers. Additional investigations have demonstrated infants' ability to detect other probabilistic cues, such as syllable stress (Curtin et al. 2005) and the allophonic variations that occur at different word positions, like the unaspirated [t] that occurs at the ends of English words versus the aspirated [$t^h$] that occurs at the beginnings of words (Jusczyk et al. 1999). This substantial group of studies has largely been discussed as demonstrating that infants can use such cues to segment words from fluent speech. However, the nature of the representations infants form in segmentation tasks is not apparent; these studies demonstrate that a given cue can motivate listening time differences (e.g., for zero probability *non-words* versus high probability *words*). It is not clear that infants have segmented 'words' per se. One means of testing whether statistical learning allows infants to extract individual words from fluent speech is to investigate whether infants can use the output of statistical learning in a linguistically relevant task—mapping the sounds of words to meanings.

Graf Estes et al. (2007) examined the connection between statistical word segmentation and lexical acquisition. They tested whether infants can take advantage of statistical learning about sound sequences to facilitate word learning. Seventeen-month-olds first listened to several minutes of an artificial language (similar to Saffran et al. 1996, but using 2-syllable words rather than 3-syllable words), designed so that transitional probability information provided the only reliable indicator of word boundaries. After this opportunity to detect word boundary cues in the language, the infants immediately participated in an object label-learning task (Werker et al. 1998). In the task, infants were habituated to two label-object pairings. During the test trials that followed habituation, the original label-object pairings were switched (i.e., object 1 played with label 2). This design is based on the prediction that if infants learned the original label-object pairings they should look longer when those pairings are violated (Werker et al. 1998). In the key manipulation, for some infants the labels were high internal transitional probability *words* from the artificial language, sound sequences that infants had the chance to segment from the speech stream before they ever occurred as object labels. For other infants, the labels were zero-transitional probability *non-words* or low transitional probability *part-word* sequences (that crossed word boundaries in the artificial language). Infants successfully learned the high probability *word* labels, but failed to learn the zero- and low-probability sequence labels. In a control experiment, infants also failed to learn the labels without any artificial language exposure. Thus, finding the words in the speech stream subsequently made it possible for infants to learn them as object labels. These results help to clarify the nature of the representations infants form by tracking statistical regularities of sound sequences. They suggest that statistically segmented words are rapidly available to support lexical acquisition. Furthermore, infants may use this process to detect words in native language input. They may discover individual lexical items that are then stored and ready to link to meaning.

One question to come from Graf Estes et al.'s (2007) experiment is whether the connection from statistical learning about sounds to word learning is consistent across development, or whether only young learners take advantage of statistical word segmentation to acquire new lexical items. In a recent experiment, Mirman et al. (2008) examined the relation between statistical segmentation and learning of object labels in adults. Like in the infant experiment, participants first listened to an artificial language, followed by a label-learning task. In this case, the label-learning task involved an artificial lexicon learning measure (Magnuson et al. 2003) that allows for the comparison of learning rates for different labels. On each trial, participants viewed a pair of novel objects (geometric figures) and were asked to select the appropriate object for an auditorily presented label. They then received feedback regarding whether the selection was correct. There were four label-object pairings in the learning set; each object served as the correct referent and as a distractor object across the trials. Initially, the participants' responses were at chance level—they were never explicitly instructed about the appropriate pairings—but performance improved gradually as they learned the associations. Like Graf Estes et al.'s infant experiments, the crucial manipulation was that some participants heard statistically-defined *words* from the segmentation stream presented as object labels; other participants heard novel syllable sequence *nonword* labels, and the final group heard low transitional probability *part-word* labels. Mirman et al. found that participants presented with *word* labels and *nonword* labels learned at similar rates, whereas participants presented with *part-word* labels learned more slowly.

The findings from Mirman et al. indicate that in adulthood, distributional information about likely and unlikely sound sequences continues to affect lexical acquisition, but the nature of the influence differs from during infancy. The infants and adult experiments are similar in many aspects of the design (statistical segmentation followed by label-learning task; artificial language consisting of 2-syllable words), but by necessity, the studies differ in their label-learning measures. As described above, Graf Estes et al. (2007) used a habituation-based measure in which infants must detect violations of label-object pairings to show learning, whereas Mirman et al. (2008) used a lexicon learning task in which adults must learn labels for shapes based on feedback about their own assignments of labels to objects. Despite the methodological differences, the findings have interesting developmental implications. For the relatively inexperienced word learners in the infant study, prior segmentation experience allows infants to learn object labels that are otherwise difficult to learn. For sophisticated word learners, the additional support of prior segmentation experience is not necessary to buttress learning. However, labels that violate the recently acquired distributional information are difficult to learn. The patterns of performance for infants and adults raise the possibility that across development, learners differ in how they take advantage of statistical learning about sound sequences. Ongoing experiments are testing the developmental course of this shift.

Although studies of statistical learning often use artificial languages to carefully control the distributional information available in the speech stream, it is also important to consider the sound sequences infants might extract from the vast input they receive from the natural linguistic environment. As with controlled experimental language exposure (Graf Estes et al. 2007), natural experience may provide infants with segmented sound sequences, or candidate words that are ready to be mapped to meanings. Swingley (2005) performed a corpus analysis of infant-directed speech to investigate whether distributional information in the speech signal is sufficient to support infants' segmentation of real words and avoid mis-segmentations. In analyses of Dutch and English infant-directed speech, Swingley examined how tracking the probability and frequency of syllable co-occurrences would produce sound sequence clusters. The results of the analysis suggest that use of such a mechanism would largely yield appropriately segmented real words. Swingley proposed that word segmentation based on distributional information may provide infants with a 'proto-lexicon' of sound sequences that have been isolated from continuous speech and are stored, ready to be added to the lexicon once meaning is assigned. Prior segmentation of word forms should facilitate the association of sound and meaning in word learning because part of the problem is already solved; the new lexical item already has an initial stored phonological representation. Thus, early learned words may come from a stockpile of previously segmented word forms.

Another means of investigating how natural language distributional information supports lexical development is to examine how phonotactic patterns shape learning of new words. Although infants' listening performance suggests that they can use phonotactic patterns to detect words in sentences (Mattys and Jusczyk 2001), there is evidence suggesting that phonotactic knowledge facilitates processing in ways that extend beyond its role in segmentation. Numerous studies have shown that in nonword repetition tasks, adults and children repeat novel sound sequences more accurately and faster when they consist of high phonotactic probability sounds and sound sequences (e.g., *ged*) rather than low probability sounds and sequences (e.g., *moid*; Coady and Aslin 2004; Edwards et al. 2004; Gathercole 1995; Munson et al. 2005; Vitevitch and Luce 2005; Zamuner et al. 2004). That is, high phonotactic probability sound patterns confer a processing advantage. One explanation for this effect is that nonword repetition tasks tap phonological working memory, the ability to store novel phonological sequences in memory (Gathercole 2006; see also Coady and Evans 2008). New words that share sound sequences with established lexical items are easier to process because they receive support from stored long term representations of the phonological forms of known words.

Storkel (2001) found a parallel effect in a word learning task with preschoolers. She presented 3- to 6-year-olds with a set of four consonant-vowel-consonant (CVC) novel object labels consisting of high phonotactic probability sounds sequences (i.e., common sequences in English) and four CVC labels consisting of low probability sequences (i.e., rare sequences). The phonotactic probabilities were based on the frequency with which each phoneme occurred at a given word position across a corpus of English words, as well as the frequencies with which pairs of phonemes occurred together within words. Using measures of label comprehension and production, Storkel found more robust and faster learning of the common sequence labels compared with the rare labels. In addition, children with larger receptive vocabularies showed a greater advantage for common over rare sound sequences than children with smaller vocabularies.

The accumulation of language experience provides children with information about the distributional patterns of phonological sequences in native language words. Storkel's (2001) findings indicate that childhood linguistic knowledge affects how new words are acquired; existing phonological representations support lexical acquisition. When children encounter new words that consist of common sound sequences, the words may be relatively easy to process because of their relation to extant phonological knowledge. This could allow children to allocate greater attention to forming and storing the association between the sound of the word and its referent. In contrast, new words that consist of rare sound sequences may require greater attention to processing the phonological form, tying up resources so that they cannot be used for learning about word meaning or the sound-meaning association. The effects of phonological representations on word learning may be particularly relevant for young learners for whom the basic process of linking sounds with meanings is still quite challenging. Furthermore, phonotactic patterns could affect many points in the process of word learning—the initial encoding of the phonological form, the association with meaning, the storage of the phonological form and/or the link to the referent, as well as retrieval of the item. The locus (or loci) of the effect of phonotactic regularities on word learning has yet to be specified.

Werker and Curtin's (2005) PRIMIR (Processing Rich Information from Multidimensional Interactive Representations) model is also consistent with the notion that phonotactic regularities should affect word learning. PRIMIR is a model of early phonological and lexical development that proposes that as learners gain experience with words, phonological representations strengthen. It follows that representations of sounds or sound sequences that occur frequently should develop more rapidly than rare sounds or sound sequences. Robust phonological representations should facilitate the association of word forms with referents and the maintenance of this linkage.

An important question that follows from Storkel's demonstration of phonotactic effects on children's word learning is whether phonotactic knowledge in infancy also affects lexical acquisition. There is ample evidence that by age 1, infants learn about the likely and unlikely sound combinations of their native language. Jusczyk et al. (1993) reported that 9-month-old English-learning infants listened longer to lists of words that contained phoneme sequences that occur in English, but not in Dutch. Dutch infants showed the opposite pattern. At 6 months, infants did not yet display this pattern of preference. Friederici and Wessels (1993) also found that infants listened longer to lists of monosyllabic words that contained consonant clusters at phonotactically legal word positions rather than illegal word positions (but see Zamuner 2006; for limitations on infants' detection of word-final phonotactic constraints). Jusczyk et al. (1994) showed that infants can make an even more fine-grained distinction between words with legal sound sequences that consist of frequent versus infrequent sound patterns. Nine-month olds, but not 6-month olds, preferred to listen to phoneme combinations that are likely in their native language. Friedrich and Friederici (2005) demonstrated further developmental changes in phonotactic knowledge between 12 and 19 months using electrophysiological measures. In this task, event-related brain potential (ERP) responses were measured while infants viewed images of objects likely to be known by 1-year olds (see Friederici 2005 for further discussion of the use of ERP in language acquisition research). While viewing, the infants listened to phonotactically legal or illegal novel words or real words that matched the current image (i.e., correct labels) or did not match it (i.e., incorrect labels). Only 19-month olds seemed to differentiate between legal and illegal sequences as possible object names. That is, 19-month olds, like adults, showed a brain response that is typical of semantic integration (increased activity to words that do not match visual information) when viewing objects presented with incorrect labels or phonotactically legal novel words, but not while hearing phonotactically illegal novel words. Twelve-month-olds did not show the semantic integration response for any words. This finding suggests that by 19 months, infants do not treat phonotactically illegal sequences as possible object labels.

Studies of infants' phonotactic knowledge indicate that by the age at which children are starting to produce their first words, they have gathered a great deal of distributional information about the sound patterns present in the ambient language. There have been few demonstrations of how infants *apply* this knowledge to address real problems facing language learners. One possibility is that phonotactic patterns support word segmentation, as discussed above (Mattys and Jusczyk 2001). However, the processing advantage seen for likely sound sequences in nonword repetition and word learning tasks with older children provokes the following question: Is infants' phonotactic knowledge sufficiently robust to affect how new words are added to the lexicon? By the age at which children are tested in nonword repetition and word learning tasks, they typically produce thousands of words. It is not clear whether such a large base of lexical knowledge is necessary for phonotactic patterns to shape new learning or whether the learning that takes place in infancy exerts similar effects.

Graf Estes (2007) recently tested the effects of phonotactic patterns on infants' acquisition of new lexical items. In this experiment, 19-month-old infants were presented with two novel object labels. For half of the infants, the labels were phonotactically legal sound sequences (e.g., *dref* and *sloob*). For the remaining infants, the labels included phonotactically illegal word-initial consonant clusters (e.g., **dlef* and *sroob*). All of objects were labeled in simple, common labeling frames (e.g., Look at the dref!) and thus were unlikely to introduce a significant word segmentation challenge. In testing, infants presented with the legal labels showed successful recognition of the object labels; infants presented with illegal labels did not. This result demonstrates that a large lexicon on the order of thousands of words is not necessary for phonotactic patterns to shape word learning. By 1.5 years of age, infants are more likely to learn new words that are consistent with native language sound patterns.

Graf Estes's (2007) result suggests that infants' distributional learning about ambient language sound patterns does more than motivate listening preferences. Furthermore, phonotactic knowledge may affect infant's lexical acquisition beyond its effects on word segmentation. Infants may track phonotactic regularities from hearing, segmenting and storing word forms, in addition to information gathered from the words they can produce and understand. This information is then available to support learning of new lexical items that contain likely sound sequences and to inhibit learning of lexical items that contain unlikely or illegal sound sequences.

Another form of distributional information about words that is available in the input is neighborhood density. Neighborhood density is a measure of the number of words that sound similar to a given word. It is often measured as the number of words that differ from a given word by the addition, deletion, or swapping of a single phoneme. Neighborhood density correlates with phonotactic probability. That is, high probability sound sequences also occur in many different words, and low probability sound sequences occur in few words. However, each exerts a different effect on adults' lexical processing. Generally, processing of high probability sound sequences is faster and more accurate than low probability sequences (Vitevitch and Luce 1998, 1999; Frisch et al. 2000; Luce and Large 2001). In contrast, recognition of high density words (those with many neighbors) is slower and less accurate than low density words (Vitevitch and Luce 1998, 1999; Luce and Large 2001). However, high-density items are produced and recalled more effectively (Roodenrys and Hinton 2002; Vitevitch and Sommers 2003; Vitevitch et al. 2004). The relative influences of these two factors and how they interact remains a matter of investigation. Storkel et al. (2006) proposed that in adult word learning, phonotactic probability and neighborhood density influence different aspects of learning. They claimed that phonotactic probability primarily influences triggering of learning, whereas density affects how new words are integrated with known words.

The process of identifying to-be-learned words and integrating new and known words are likely to be quite different for infants and adults, and the effects of phonotactics and density might differ as well. Infants encounter unfamiliar words with greater frequency and have smaller lexicons in which to integrate new items. As described above, research is only starting to uncover how knowledge of phonotactic probability affects word learning in infants. There is also some evidence concerning how infants' acquisition of new words is affected by similarity to known words. Swingley and Aslin (2007) found that for 1.5-year-olds, learning a new object label was hindered when the label was a neighbor of a known word (e.g., *tog* and *dog*). In the same setting, infants successfully learned a label that was dissimilar to known words. Swingley and Aslin proposed that interference could cause difficulty in learning new words that are highly similar to stored words. Hearing the novel word (*tog*) would activate the representation of the known words (*dog*), preventing the formation of a new association between the sound and meaning.

In addition, Hollich et al. (2002) found that prior listening to a set of novel word forms inhibited learning of a new object label that overlapped with the words in the set. In particular, this effect occurred when 17-month-olds listened to the word set several times. After hearing the list once, label learning was facilitated rather than inhibited. This result suggests that some familiarity with sound sequences can promote new lexical acquisition, but when a new word is too close to an existing item, learning is hindered. Thus, knowledge of stored word forms, even those not yet associated with referents, may induce the interference effects described by Swingley and Aslin (2007). Future experiments will be necessary to directly compare the effects of prior word form knowledge with and without associated meaning representations. At this point, the findings regarding neighborhood density and phonotactic probability suggest that there is a complex relation between prior learning and new learning. The accumulation of distributional information seems to shape learning by facilitating the acquisition of sound sequences that are consistent with that information, making inconsistent sound sequences (those not likely to be words in the language) difficult to acquire. However, substantial overlap with existing stored lexical items may also hinder learning.

In conclusion, there is substantial evidence that very young learners possess powerful mechanisms for detecting distributional regularities present in the input. Statistical learning allows infants to gather a great deal of information about the sound structure of their native language. In the present paper, we have reviewed how infants apply this ability to track statistical regularities in the sounds of their language to solve a fundamental task in early language development, lexical acquisition. We propose that early experience with sound sequence regularities provides infants with a foundation for linking sounds with meanings. Infants' knowledge of transitional probabilities, phonotactic patterns, and other probabilistic word boundary cues may allow infants pluck individual words from continuous speech, making isolated sound sequences then available to be linked with referents. There is preliminary evidence that early phonotactic knowledge also shapes how readily new words are acquired. Phonotactic regularity and neighborhood density knowledge may promote learning of likely sound sequences, while inhibiting learning of sequences that are unlikely to be words in the native language or words that are highly similar to known words. Investigators are just beginning to explore how infants take advantage of their precocious speech perception skills and remarkable learning abilities to support lexical acquisition. By understanding how infants use early-emerging knowledge of their language's sound system, we are coming to a better understanding of how development continuously builds on prior development.

## *Short Biography*

Katharine Graf Estes's research investigates the processes underlying early language acquisition and the connection between early phonological and lexical representations. She received her Ph.D. in Developmental Psychology from the University of Wisconsin-Madison in 2007. She is now an Assistant Professor at the University of California, Davis, where she has established the Language Learning Lab at the Center for Mind and Brain. She has received funding from the National Institutes of Health and the National Science Foundation.

## *Note*

* Correspondence address: Katharine Graf Estes, Ph.D., Assistant Professor, Department of Psychology and Center for Mind and Brain, University of California, Davis, 267 Cousteau Place, Davis, CA 95618, USA. E-mail: kgrafestes@ucdavis.edu

## *Works Cited*

Aslin, R. N., J. R. Saffran, and E. L. Newport. 1998. Computation of conditional probability statistics by 8-month-old infants. Psychological Science 9(4).321–4.

Brent, M. R., and T. A. Cartwright. 1996. Distributional regularity and phonotactic constraints are useful for segmentation. Cognition. Special Issue: Compositional language acquisition 61(1–2).93–125.

——, and J. M. Siskind. 2001. The role of exposure to isolated words in early vocabulary development. Cognition 81(2).B33–44.

Cartwright, T. A., and M. R. Brent. 1997. Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. Cognition 63(2).121–70.

Coady, J. A., and R. N. Aslin. 2004. Young children's sensitivity to probabilistic phonotactics in the developing lexicon. Journal of Experimental Child Psychology 89(3).183–213.

——, and J. L. Evans. 2008. Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). International Journal of Language & Communication Disorders 43(1) 1–40.

Creel, S. C., E. L. Newport, and R. N. Aslin. 2004. Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. Journal of Experimental Psychology: Learning, Memory, and Cognition 30(5).1119–30.

Curtin, S., T. H. Mintz, and M. H. Christiansen. 2005. Stress changes the representational landscape: evidence from word segmentation. Cognition 96(3).233–62.

Edwards, J., M. E. Beckman, and B. Munson. 2004. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. Journal of Speech, Language, and Hearing Research 47(2).421–36.

Fiser, J., and R. N. Aslin. 2002. Statistical learning of higher-order temporal structure from visual shape sequences. Journal of Experimental Psychology: Learning, Memory, and Cognition 28(3).458–67.

Friederici, A. D. 2005. Neurophysiological markers of early language acquisition: from syllables to sentences. Trends in Cognitive Sciences 9(10).481–8.

——, and J. M. Wessels. 1993. Phonotactic knowledge of word boundaries and its use in infant speech perception. Perception & Psychophysics 54(3).287–95.

Friedrich, M., and A. D. Friederici. 2005. Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. Journal of Cognitive Neuroscience 17(11).1785–802.

Frisch, S. A., N. R. Large, and D. B. Pisoni. 2000. Perception of wordlikeness: effects of segment probability and length on the processing of nonwords. Journal of Memory and Language 42(4).481–96.

Gathercole, S. E. 1995. Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. Memory & Cognition 23(1).83–94.

—— 2006. Nonword repetition and word learning: the nature of the relationship. Applied Psycholinguistics 27(4).513–43.

Graf Estes, K. 2007. Is there anything wordlike about statistical word segmentation? Paper presented at the Workshop on Current Issues in Language Acquisition: Artificial Languages and Statistical Learning, Calgary, Alberta, Canada.

Graf Estes, K., J. L. Evans, M. W. Alibali, and J. R. Saffran. 2007. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. Psychological Science 18(3).254–60.

Hauser, M. D., E. L. Newport, and R. N. Aslin. 2001. Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. Cognition 78(3).B53–64.

Hollich, G., P. W. Jusczyk, and P. A. Luce. 2002. Lexical neighborhood effects in 17-month-old word learning. Paper presented at the 26th Annual Boston University Conference on Language Development, Boston, MA.

Houston-Price, C., and S. Nakai. 2004. Distinguishing novelty and familiarity effects in infant preference procedures. Infant and Child Development 13(4).341–8.

Hunter, M. A., and E. W. Ames. 1988. A multifactor model of infant preferences for novel and familiar stimuli. Advances in Infancy Research 5. 69–95.

Jusczyk, P. W., A. D. Friederici, J. M. Wessels, V. Y. Svenkerud, and A. M. Jusczyk. 1993. Infants' sensitivity to the sound patterns of native language words. Journal of Memory and Language 32(3).402–20.

——, E. A. Hohne, and A. Bauman. 1999. Infant's sensitivity to allophonic cues for word segmentation. Perception & Psychophysics 61(8).1465–76.

——, P. A. Luce, and J. Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Language 33(5).630–45.

Kirkham, N. Z., J. A. Slemmer, and S. P. Johnson. 2002. Visual statistical learning in infancy: evidence for a domain general learning mechanism. Cognition 83(2).B35–42.

Luce, P. A., and N. R. Large. 2001. Phonotactics, density, and entropy in spoken word recognition. Language and Cognitive Processes 16(5–6).565–81.

Magnuson, J. S., M. K. Tanenhaus, R. N. Aslin, and D. Dahan. 2003. The time course of spoken word learning and recognition: studies with artificial lexicons. Journal of Experimental Psychology: General 132(2).202–27.

Mattys, S. L., and P. W. Jusczyk. 2001. Phonotactic cues for segmentation of fluent speech by infants. Cognition 78(2).91–121.

——, ——, P. A. Luce, and J. L. Morgan. 1999. Phonotactic and prosodic effects on word segmentation in infants. Cognitive Psychology 38(4).465–94.

McQueen, J. M. 1998. Segmentation of continuous speech using phonotactics. Journal of Memory and Language 39.21–46.

Mintz, T. H., E. L. Newport, and T. G. Bever. 2002. The distributional structure of grammatical categories in speech to young children. Cognitive Science: A Multidisciplinary Journal 26(4).393–424.

Mirman, D., J. S. Magnuson, K. Graf Estes, and J. A. Dixon. 2008. The link between statistical segmentation and word learning in adults. Cognition 108(1).271–80.

Munson, B., B. A. Kurtz, and J. Windsor. 2005. The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. Journal of Speech, Language, and Hearing Research 48(5).1033–47.

Roodenrys, S., and M. Hinton. 2002. Sublexical or lexical effects on serial recall of nonwords? Journal of Experimental Psychology: Learning, Memory, and Cognition 28(1).29–33.

Saffran, J. R. 2003. Statistical language learning: mechanisms and constraints. Current Directions in Psychological Science 12(4).110–4.

——, R. N. Aslin, and E. L. Newport. 1996. Statistical learning by 8-month-old infants. Science 274(5294).1926–8.

Saffran, J., M. Hauser, R. Seibel, J. Kapfhamer, F. Tsao, and F. Cushman. 2008. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. Cognition 107(2).479–500.

Saffran, J. R., E. K. Johnson, R. N. Aslin, and E. L. Newport. 1999. Statistical learning of tone sequences by human infants and adults. Cognition 70(1).27–52.

——, J. F. Werker, and L. A. Werner. 2006. The infant's auditory world: hearing, speech, and the beginnings of language. Hoboken, NJ, US: John Wiley & Sons Inc.

Storkel, H. L. 2001. Learning new words: phonotactic probability in language development. Journal of Speech, Language, and Hearing Research 44(6).1321–37.

——, J. Armbrüster, and T. P. Hogan. 2006. Differentiating phonotactic probability and neighborhood density in adult word learning. Journal of Speech, Language, and Hearing Research 49(6).1175–92.

Swingley, D. 2005. Statistical clustering and the contents of the infant vocabulary. Cognitive Psychology 50(1).86–132.

Swingley, D., and R. N. Aslin. 2007. Lexical competition in young children's word learning. Cognitive Psychology 54(2).99–132.

Thiessen, E. D., E. A. Hill, and J. R. Saffran. 2005. Infant-directed speech facilitates word segmentation. Infancy 7(1).53–71.

Toro, J. M., and J. B. Trobalón. 2005. Statistical computations over a speech stream in a rodent. Perception & Psychophysics 67(5).867–75.

Turk-Browne, N. B., P. J. Isola, B. J. Scholl, and T. A. Treat. 2008. Multidimensional visual statistical learning. Journal of Experimental Psychology: Learning, Memory, and Cognition 34(2).399–407.

Vallabha, G. K., J. L. McClelland, F. Pons, J. F. Werker, and S. Amano. 2007. Unsupervised learning of vowel categories from infant-directed speech. PNAS Proceedings of the National Academy of Sciences of the United States of America 104(33).13273–8.

Vitevitch, M. S., J. Armbrüster, and S. Chu. 2004. Sublexical and lexical representations in speech production: effects of phonotactic probability and onset density. Journal of Experimental Psychology: Learning, Memory, and Cognition 30(2).514–29.

——, and P. A. Luce. 1998. When words compete: levels of processing in perception of spoken words. Psychological Science 9(4).325–9.

——, and ——. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory and Language 40(3).374–408.

Vitevitch, M.S., and P.A. Luce. 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, and Computers 36.481–7.

Vitevitch, M. S., and P. A. Luce. 2005. Increases in phonotactic probability facilitate spoken nonword repetition. Journal of Memory and Language 52(2).193–204.

——, and M. S. Sommers. 2003. The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. Memory & Cognition 31(4).491–504.

Werker, J. F., L. B. Cohen, V. L. Lloyd, M. Casasola, and C. L. Stager. 1998. Acquisition of word-object associations by 14-month-old infants. Developmental Psychology 34(6).1289–309.

——, and S. Curtin. 2005. PRIMIR: a developmental framework of infant speech processing. Language Learning and Development 1(2).197–234.

Woodward, A. L., and R. N. Aslin. 1990. Segmentation cues in maternal speech to infants. Paper presented at the Biennial Meeting of the International Conference on Infant Studies, Montreal, Quebec, Canada.

Zamuner, T. S. 2006. Sensitivity to word-final phonotactics in 9- to 16-month-old infants. Infancy 10(1).77–95.

Zamuner, T. S., L. Gerken, and M. Hammond. 2004. Phonotactic probabilities in young children's speech production. Journal of Child Language 31(3).515–36.