



# Words in a sea of sounds: the output of infant statistical learning

Jenny R. Saffran\*

*Department of Psychology and Waisman Center, University of Wisconsin – Madison,  
Madison, WI 53706, USA*

Received 9 August 1999; received in revised form 14 February 2001; accepted 14 March 2001

---

## Abstract

One of the first problems confronting infant language learners is word segmentation: discovering the boundaries between words. Prior research suggests that 8-month-old infants can detect the statistical patterns that serve as a cue to word boundaries. However, the representational structure of the output of this learning process is unknown. This research assessed the extent to which statistical learning generates novel word-like units, rather than probabilistically-related strings of sounds. Eight-month-old infants were familiarized with a continuous stream of nonsense words with no acoustic cues to word boundaries. A post-familiarization test compared the infants' responses to words versus part-words (sequences spanning a word boundary) embedded either in simple English contexts familiar to the infants (e.g. "I like my *tibudo*"), or in matched nonsense frames (e.g. "zy fike ny *tibudo*"). Listening preferences were affected by the context (English versus nonsense) in which the items from the familiarization phase were embedded during testing. A second experiment confirmed that infants can discriminate the simple English contexts and the matched nonsense frames used in Experiment 1. The third experiment replicated the results of Experiment 1 by contrasting the English test frames with non-linguistic frames generated from tone sequences. The results support the hypothesis that statistical learning mechanisms generate word-like units with some status relative to the native language. © 2001 Elsevier Science B.V. All rights reserved.

*Keywords:* Words in a sea of sounds; Output; Infant statistical learning

---

\* Tel.: +1-608-262-9942; fax: +1-608-262-4029.

*E-mail address:* [jsaffran@facstaff.wisc.edu](mailto:jsaffran@facstaff.wisc.edu) (J.R. Saffran).

## 1. Introduction

By the end of their first year, infants have learned a great deal about their native language. For months, they have been able to distinguish their native language from others based on its prosodic patterns (e.g. Mehler et al., 1988). They have acquired phonological (e.g. Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Houston, & Newsome, 1999) and phonotactic (e.g. Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993) structure, and can even distinguish frequently exemplified phonotactic structures from relatively unfamiliar but equally legal patterns (e.g. Jusczyk, Luce, & Charles-Luce, 1994). Native language experience has shaped their speech perception abilities (e.g. Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984). Of most significance to their parents, they are beginning to understand words (e.g. Hallé and de Boysson-Bardies, 1994; Tincoff & Jusczyk, 1999, 2000).

How do infants learn so much, so quickly? Recent research has endeavored to uncover the learning mechanisms underlying the beginnings of language learning (e.g. Gómez & Gerken, 1999; Goodsitt, Morgan, & Kuhl, 1993; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Saffran, Aslin, & Newport, 1996). One type of mechanism that has been implicated in early language acquisition is statistical learning: the process of detecting linguistic units by tracking patterns of sounds in the input. Statistical learning mechanisms may be particularly useful in solving one early learning task facing the infant – word segmentation. Unlike the white spaces available in written text, speech does not contain consistent physical cues marking word boundaries (e.g. Cole & Jakimik, 1980). Although word segmentation is an extremely complex learning problem, infants as young as 7.5 months of age can extract words from continuous speech (Jusczyk & Aslin, 1995).

In previous research, we have suggested that infants are able to use statistical properties of the input language to discover word boundaries (Saffran et al., 1996). For example, the two word sequence *pretty baby* contains three transitional probabilities computed between syllable pairs. Across a corpus of English, the probability that *ty* will be followed by *ba*, as in this example, is lower than either of the two word-internal transitional probabilities (*pre* followed by *ty* or *ba* followed by *by*). In order to directly ask whether infants can use statistical cues in the service of word segmentation, we briefly exposed 8-month-old infants to a small nonsense language, generated by a speech synthesizer such that the only available cues to word boundaries were statistical. A sample of the speech stream was analogous to the following orthographic string: *pabikugolatudaropitibudo...*, etc. The infants successfully distinguished the familiar words from “part-words” (strings consisting of sequences spanning a word boundary, like *ty#ba* in the example above), preferring to listen to the relatively novel part-words (Saffran et al., 1996; Experiment 2). These results strongly suggest that infants are able to detect and use the sequential probabilities which characterize auditory sequences (see also Aslin, Saffran, & Newport, 1998; Saffran, Johnson, Aslin, & Newport, 1999).

While we and a number of other investigators (e.g. Brent & Cartwright, 1996; Christiansen, Allen, & Seidenberg, 1998; Elman, 1990; Goodsitt et al., 1993; Perru-

chet & Vintner, 1998; Redington & Chater, 1998) have suggested that various statistical cues (along with other types of information such as prosodic regularities and words in isolation) may subserve word segmentation, it is not yet known whether infants actually treat the output of the statistical learning process as candidate words. Consider the sequence *golabu*. Once segmented from a continuous stream of speech by virtue of its statistical properties, infants might continue to treat *golabu* as a relatively coherent sound sequence with high internal probabilities, but with no particular status with respect to their native language. Alternatively, infants might now treat *golabu* as a potential word,<sup>1</sup> available for integration into their native language.

We attempted to tease apart these two hypotheses experimentally by manipulating the predicted direction of preferential listening following a word segmentation statistical learning task. There are two classes of variables that may be separately manipulated to affect infant preferential listening performance: variables internal to the learning task, and variables concerning the relationship between the task and the infant's prior knowledge. We consider each in turn.

A number of different *task factors* influence the direction of infant preferences for novel versus familiar stimuli. These include the infant's age, the length of the delay between exposure and test, the closeness of the match between exposure and test items, and the complexity of the task (e.g. Hunter & Ames, 1988; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982; Spence, 1996; Wetherford & Cohen, 1973). The number of exposures to target stimuli during familiarization also affects the direction of preference (e.g. Hunter & Ames, 1988; Hunter, Ames, & Koopman, 1983; Rose et al., 1982). This variable is particularly critical in preferential listening tasks because unlike habituation techniques, in which the infant controls the extent of familiarization, all infants in a preferential listening task receive the same number of exposures. The number of times an item is repeated or processed during familiarization can lead to either familiarity preferences after relatively little exposure, or novelty preferences after a greater number of exposures (e.g. Aslin, 1999, 2000). Prior to any learning during familiarization, neither familiar nor novel stimuli should be inherently more interesting. Once learning has begun to occur, the currently encoded stimulus draws the infants' attention; only after the stimulus has reached a high level of encoding does its interest level begin to wane (Aslin, 2000; Hunter & Ames, 1988). This analysis is consistent with the novelty effects reported in our previous research using both linguistic and non-linguistic materials (Aslin et al., 1998; Saffran et al., 1996, 1999; Saffran & Griepentrog, 2001), as well as the novelty effects found in studies using segmentation tasks and familiarization procedures closest to our own (e.g. Echols, Crowhurst, & Childers, 1997). Studies which have revealed familiarity preferences typically include far fewer exposures to the target words (e.g. Jusczyk & Aslin, 1995; Jusczyk et al., 1999). For example, infants in the Jusczyk and Aslin (1995) study heard an average of 12 instances of each of the two target words during familiarization, while infants in our studies heard 45

---

<sup>1</sup> We use the term 'word' loosely to refer to unit-like representations that have been segmented and are available to the infant for subsequent language learning and processing.

instances of each word. It thus appears that extensive familiarization leads to novelty preferences, even when no habituation criterion is employed.

A second class of factors known to influence preferential listening pertains to the infant's *prior knowledge*. Infants consistently prefer to listen to the sound patterns of their native language (for a recent overview, see Jusczyk, 1997). Infants also prefer to listen to prosodic structures consistent with native language structure: when artificial pauses are placed within test items, infants prefer to listen to materials where the pauses occur at structurally relevant junctures, such as clause, phrase, and word boundaries (e.g. Gerken, Jusczyk, & Mandel, 1994; Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Myers et al., 1996). Materials which are interrupted at unnatural breaking points are dispreferred by infants in this paradigm. Musical stimuli segmented at structurally-relevant boundaries (e.g. Jusczyk & Krumhansl, 1993; Krumhansl & Jusczyk, 1990), or played in the musical contexts in which they were originally learned (Saffran, Loman, & Robertson, 2000), are also preferred by infant listeners. This literature suggests that infants typically prefer to listen to materials which best match their prior knowledge, as determined by the effects of segmenting test stimuli at natural versus unnatural boundaries.

The current studies were designed to use these two classes of effects on infant preferential listening to probe infants' representations following statistical learning. In particular, can we counteract the novelty effects due to extensive familiarization by segmenting test materials at natural versus unnatural boundaries with respect to the native language? As noted above, infants prefer to listen to relatively novel part-words (sequences spanning word boundaries) rather than words when these test items are played in isolation following extensive exposure (Aslin et al., 1998; Saffran et al., 1996). What happens when the tested words and part-words are embedded in English sentences during testing? For example, after exposure to a segmentation task in which *golabu* is a word and *pigola* is a part-word, a sentence such as "I like my *golabu*" would be more natural and coherent than "I like my *pigola*", which contains word fragments – *if* infants are treating *golabu* as a possible word in their native language. Based on the studies described above that suggest a preference for materials consistent with the infant's prior knowledge, we would expect infants to prefer English sentences containing newly segmented words over English sentences containing part-words. However, if infants treat *golabu* as a highly probable sequence of sounds, but not as a possible English word, then infants' prior knowledge of English should not affect their performance on this task, and we should see the same pattern of results regardless of whether the newly segmented words are embedded in English sentences or non-English sentences.

To determine whether infants treat the output of the statistical learning process as candidate words, we combined a word segmentation familiarization phase with a subsequent test in which we embedded word and part-word targets in either English or nonsense sentence frames. If the statistical learning process renders highly probable sound sequences, but not possible English words, then the English and nonsense test frames should elicit equivalent listening preferences. If, however, infants are treating the highly probable sound sequences as possible English words, then we

would expect to see different patterns of performance in the English and nonsense test frames. In particular, we predicted that infants would prefer words over part-words when these targets are embedded in English frames. According to this hypothesis, the newly segmented words should be more readily integrated into the infant's knowledge of English than the part-words. The sentences containing the new words should thus be segmented more naturally and preferred over the sentences containing the part-words, which will contain incorrectly segmented words. However, performance in the nonsense frame should *not* generate a familiarity preference for words over part-words, because there is no prior knowledge with which these test items must be integrated. Instead, the primary influences governing listening patterns in the nonsense condition are task factors. Given that the number of exposures is identical to the stimuli used by Saffran et al. (1996), we might expect a novelty preference. Alternatively, the presence of the nonsense frames might dilute the dishabituation response relative to the Saffran et al. results, muting the novelty preference. Regardless, the principle question is whether infants show different patterns of performance in the English and nonsense frame test conditions. If so, the results would suggest that the output of the statistical learning process consists of word-like representations, rather than probabilistically-related strings of sound with no status with respect to English.

## 2. Experiment 1

### 2.1. Method

#### 2.1.1. Subjects

Fifty-six full-term 8-month-old monolingual infants with no history of recurrent ear infections were tested (mean age 8 months, 3 days; range 7:3 to 8:2). Half of the infants were randomly assigned to the English Frame condition and half were assigned to the Nonsense Frame condition. Within each condition, half of the infants were assigned to each of two counterbalanced exposure languages (Conditions A and B). Thirty-one additional infants were tested but not included in the analysis for the following reasons: fussiness (23), looking times averaging less than 3 s to one or both sides (5), parental interference (1), not looking at the side lights (1), and listening to all 12 trials for the maximum amount of time (1). All infants were solicited from local birth announcements and hospital records, and parental consent was obtained prior to testing in accordance with the guidelines of the local human subjects review committee and the principles of ethical treatment established by the American Psychological Association.

#### 2.1.2. Stimuli

For the familiarization phase, we used the two exposure audio files from Saffran et al. (1996; Experiment 2). Each language consisted of four nonsense words: Condition A: *pabiku*, *tibudo*, *golatu*, *daropi*; Condition B: *tudaro*, *pigola*, *bikuti*, *budopa*. The speech stream was generated by a speech synthesizer (MacinTalk on an Apple

Quadra 650 computer) in a monotone female voice at a rate of 270 syllables/min. Each word was repeated 45 times in random order, with the constraint that the same word could not occur twice in a row, for a total of 180 words. The speech was digitized and stored on disk at a 22 kHz sampling rate for playback during the experimental session. Each speech stream lasted 2 min. There were no acoustic or prosodic cues to word boundaries. The only cues to word boundaries were the transitional probabilities between syllable-pairs, which were higher within words (1.0 in all cases) than between words (0.33 in all cases).<sup>2</sup>

As in Saffran et al. (1996; Experiment 2), the test items included the following four sound sequences: *pabiku*, *tibudo*, *tudaro*, and *pigola*. For infants in Condition A, *pabiku* and *tibudo* were words, and *tudaro* and *pigola* were part-words, which were created by joining the last syllable of one word to the first two syllables of another word. For example, the part-word *pigola* spanned the word boundary between *daropi* and *golatu*. Thus, the part-word sequences were heard during familiarization, but contained different statistical properties than the words. For infants in Condition B, *pabiku* and *tibudo* were part-words, and *tudaro* and *pigola* were words. This between-subjects counterbalanced design ensured that any observed preference for words or part-words across the two languages was due to learning.

Unlike the test used by Saffran et al. (1996; Experiment 2), in which words and part-words were repeated in isolation, targets in the present test were placed at the ends of sentence frames (see Table 1). Sentence-final position was used to maximally highlight the targets (Aslin, 1999). Also unlike the Saffran et al. study, the test items were produced by a human speaker. In the English Frame condition, each of the four targets (two words and two part-words) were embedded in English sentence frames likely to contain words familiar to the infants. In the Nonsense Frame condition, the same four targets were embedded in nonsense sentential frames. The nonsense words were chosen to rhyme with the English words to make the phonological and prosodic structure of the two different types of test frames as similar as possible.

The test sentences were recorded by a trained female speaker. All recording was done digitally, using a NetSet PC microphone and CoolEdit software running on a Dell PC. We used a human speaker rather than the speech synthesizer to create test stimuli to ensure that the test sentences sounded natural. Each sentence was spoken with infant-directed prosody; the Nonsense Frame sentences were prosodically matched to their English counterparts. In order to avoid any acoustic differences between words/part-words in the two types of frames, each of the four targets were recorded in isolation. The speaker then recorded each of the four English and Nonsense Frame sentences, including the target in final position. These targets were then digitally excised and replaced by the targets recorded in isolation. This editing allowed us to retain the appropriate coarticulatory cues to the final word/part-word while ensuring that the targets were acoustically identical in the English and Nonsense Frame conditions. Overall, the two types of test sentences were compar-

---

<sup>2</sup> Frequency of co-occurrence cues to word boundaries was also available in these stimuli; however, prior results suggest that infants use probabilities rather than frequencies in this task (Aslin et al., 1998).

Table 1  
English and Nonsense Frame test items<sup>a</sup>

	Test items	Condition A target	Condition B target
English Frame	I like my <i>pabiku</i>	Word	Part-word
	We saw the <i>tibudo</i>	Word	Part-word
	You play with <i>tudaro</i>	Part-word	Word
	What a nice <i>pigola</i>	Part-word	Word
Nonsense Frame	Zy fike ny <i>pabiku</i>	Word	Part-word
	Gee baw fuh <i>tibudo</i>	Word	Part-word
	Foo dray miff <i>tudaro</i>	Part-word	Word
	Vut luh kife <i>pigola</i>	Part-word	Word

<sup>a</sup> Items containing target words for infants in Condition A contain part-words for infants in Condition B, and vice versa.

able in amplitude (English Frame RMS energy: 23.15 dB; Nonsense Frame RMS energy: 23.15 dB) and duration (English Frame mean: 1.46 s; Nonsense Frame mean: 1.53 s). Approximately, 0.5 s of silence was added to the end of each sentence so that all test items were a total of 2 s in duration.

### 2.1.3. Procedure

Infants were tested individually while seated in a parent's lap in a sound-attenuated booth. An observer outside the booth monitored the infant's looking behavior on a closed-circuit TV system and coded the infant's behavior using a button-box connected to a PC. This button-box was used to initiate trials and to enter the direction of the infant's head turns, which controlled the duration of each test trial. Both the parent and the observer listened to masking music over headphones to eliminate bias.

At the beginning of the 2 min familiarization phase, the infant's gaze was first directed to a blinking light on the front wall in the testing booth. Then the sound sequence for one of the two languages was presented without interruption from two loudspeakers (one located on each of the two side walls in the booth). During this familiarization period, to keep the infants' interest, a blinking light above one of the two loudspeakers (randomly selected) was lit and extinguished dependent on the infant's looking behavior. When this blinking side light was extinguished, the central blinking light was illuminated until the infant's gaze returned to center, and another blinking side light was presented to elicit the infant's gaze. During the entire familiarization phase there was no contingency between lights and sound, which played continuously.

Immediately after familiarization, 12 test trials were presented (three trials for each of the four targets, presented in random order). Six of these trials were words embedded in frames and six were part-words embedded in frames. Each test trial began with the blinking light on the front wall. When the observer signaled the computer that the infant was fixating this central light, one of the lights on the two

side walls began to blink and the central light was extinguished. When the observer judged that the infant had made a head turn of at least 30° in the direction of the blinking side light, a button press signaled to the computer that one of the test sentences should be presented from the loudspeaker adjacent to the blinking light. This test item was repeated until the observer coded the infant's head turn as deviating away from the blinking light for 2 s consecutively. When this look-away criterion was met, the computer extinguished the blinking side light, turned off the test stimulus, and turned on the central blinking light to begin another test trial. The computer randomized the order of test trials (three for each of the four test sentences) and accumulated total looking time to each of the two test word and part-word sentence frames. Trials automatically ended after a maximum of eight test sentence repetitions.

## 2.2. Results and discussion

Prior to analyzing the data, we excluded trials with listening times less than 1.5 s in duration, because the head turn on these trials was initiated prior to the onset of the word/part-word target. This criterion led to the exclusion of ten of the 672 test trials (1.5%). The first analysis contrasted infants assigned to the two exposure languages, Conditions A and B. As there were no significant differences as a function of exposure language in either the English Frame condition ( $t(26) = 0.54$ , NS) or the Nonsense Frame condition ( $t(26) = 0.59$ , NS), the data from Conditions A and B are pooled in the subsequent analyses.

The principal hypothesis concerned the differential effects of the test frame (English versus Nonsense) on listening preferences for familiar (words) versus novel (part-words) items. An ANOVA revealed that the main effects of Frame ( $F(1, 54) = 1.04$ , NS) and Familiarity ( $F(1, 54) < 1$ , NS) were not significant. However, the Frame  $\times$  Familiarity interaction was significant ( $F(1, 54) = 5.83$ ,  $P < 0.05$ ) (see Fig. 1). This analysis suggests that preferences for words versus part-words were affected by the frame in which these items are embedded during testing. Overall, 21 of the 28 infants in the English Frame condition, but only 12 of the 28 infants in the Nonsense Frame condition, showed a preference for the test sentences containing words over part-words. Paired  $t$ -tests revealed that for infants tested on English Frame test items, there was a significant preference for words over part-words ( $t(27) = 2.16$ ,  $P < 0.05$ ) whereas infants tested on Nonsense Frame items showed a non-significant trend to prefer part-words over words ( $t(27) = -1.26$ , NS). The lack of significant effects in the latter condition may reflect the dilution of the dishabituation response to the part-words due to the presence of the nonsense frame, which makes even the familiar test items somewhat novel with respect to the exposure corpus.

An additional analysis tested the hypothesis that the results from the English Frame condition, but not the Nonsense Frame condition, would differ from the pattern of data found by Saffran et al. (1996; Experiment 2). In that study, infants tested on words and part-words presented without sentence frames showed a preference for the relatively novel part-words, presumably as a function of dishabituation



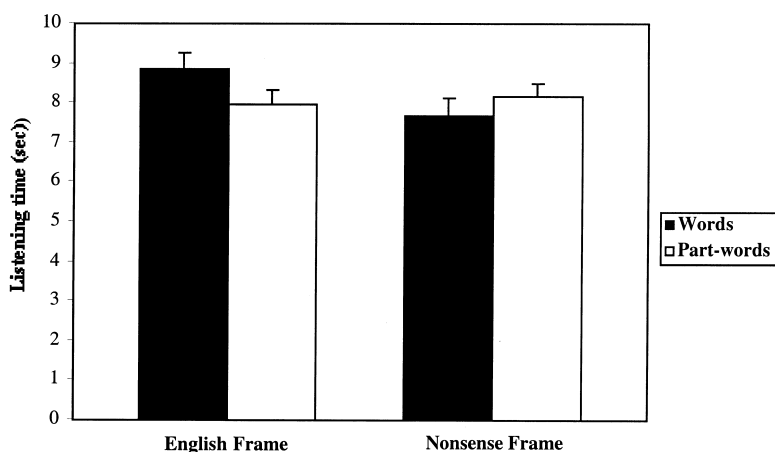


Fig. 1. Experiment 1: listening times for words and part-words for infants in the English Frame and Nonsense Frame test conditions.

following extensive exposure to the words. We performed a second ANOVA including the data from Saffran et al. (1996; Experiment 2), referred to below as the No Frame condition, and contrasted the difference scores (words minus part-words) for the three different types of frames. The ANOVA revealed that the difference scores were significantly different as a function of test frame ( $F(2, 77) = 5.4, P < 0.01$ ). Bonferroni comparisons of the two conditions from the current experiment with the results from Saffran et al. (1996; Experiment 2) revealed significant differences between the English Frame and the No Frame conditions ( $P < 0.01$ ), but no difference between the Nonsense Frame and No Frame conditions ( $P > 0.5, NS$ ). This analysis supports the hypothesis that infants treat the words and part-words from the exposure session differently when they are embedded in English sentences than when they are presented in nonsense frames or in isolation.

The results of this study suggest that the output of the statistical learning process underlying word segmentation consists of units with some status relevant to English. Unlike infants tested on word and part-word targets embedded in nonsense frames, or played in isolation, infants tested on targets placed in English sentential frames preferred to listen to the sentences containing words over the sentences containing part-words. If infants were instead treating the segmented sequences as statistical patterns containing high probabilities, but without any lexical status or relationship to English, we should have found the same pattern of listening preferences in the English and Nonsense Frame conditions. This pattern of results suggests that the representations emerging from statistical learning may serve as candidate lexical items for infants, available for integration into the native language.

An additional finding which emerged from Experiment 1 is that infants treated prosodically and phonotactically legal nonsense utterances differently from English utterances. The significant Frame  $\times$  Familiarity interaction provides indirect evidence that infants can distinguish short common English phrases like “I like

my” from corresponding nonsense phrases like “Zie fike ny”. Prior studies have demonstrated that 6-month-old infants know the referents of common English words, such as “Mommy” and “feet” (Tincoff & Jusczyk, 1999, 2000). The present findings indirectly suggest that 8-month-olds know enough about common phrases to distinguish them from matched nonsense phrases. Because the interpretation of Experiment 1 hinges on this indirect finding, we directly tested the claim that 8-month-olds can discriminate common English phrases from nonsense phrases in the next experiment.<sup>3</sup>

### **3. Experiment 2**

The significant interaction between test frame (English versus Nonsense) and familiarity (word versus part-word) in Experiment 1 supports the hypothesis that infants treat words and part-words differently when they are incorporated into native language materials than when they are not. For this to be the case, infants must be able to discriminate common English phrases from nonsense phrases that resemble English. To address this hypothesis directly, we tested 8-month-old infants on the English and Nonsense frames used in Experiment 1. Unlike Experiment 1, the current experiment did not include a familiarization period, and all infants heard both English and Nonsense test items. The question of interest is whether infants discriminate English and Nonsense frames. If infants cannot perform this discrimination, then we would not expect to see any differences in listening times for English versus Nonsense sentences. However, if infants do discriminate English and Nonsense frames, we would have direct evidence that infants process English and nonsense differently, as suggested by the results of Experiment 1.

Prior studies investigating infant discrimination of native and non-native language materials have typically elicited familiarity preferences: infants prefer materials maximally similar to their native language (e.g. Jusczyk et al., 1993, 1994; Jusczyk, Friederici et al., 1993). In general, this literature would lead to the prediction that if infants can discriminate English from nonsense, they should prefer the English materials. However, the present study is methodologically quite different from other preferential listening studies investigating infant native language knowledge, and these differences might lead to preferences for the nonsense materials, as described below.

Studies designed to tap native language knowledge in infants between 6 and 12 months of age have typically used lists of items as the test materials, minimizing repetition in order to assess general knowledge of the native language rather than item knowledge.<sup>4</sup> For example, in the Jusczyk, Friederici et al. (1993) investigation of preferences for native versus non-native phonetic and phonological patterns, infants heard six 15-word English lists and six 15-word Dutch lists, with no

---

<sup>3</sup> Thanks to an anonymous reviewer for this suggestion.

<sup>4</sup> Exceptions are the studies by Tincoff and Jusczyk (1999, 2000). However, their method required cross-modal matching, which is likely to elicit familiarity preferences due to task difficulty (e.g. Hunter & Ames, 1988; Wagner & Sakovits, 1986).

words repeated over the course of the experiment. In the current experiment, however, we wanted to assess infant's knowledge of the specific English phrases used in Experiment 1. Thus, instead of stimulus materials including 90 different English items and 90 different non-English items, as in Jusczyk, Friederici et al. (1993), infants in the current study heard only four different English items and four different Nonsense items, presented on separate trials (e.g. a trial consisted of a single repeated phrase: "I like my... I like my... I like my...", etc.). Because each trial consisted of repetitions of a single item, listening times reflected interest or disinterest in particular items. One factor known to influence preferential listening is the ease with which infants process incoming stimuli (e.g. Hunter & Ames, 1988). Due to their prior familiarity with English words and phrases, infants may process and encode the English phrases more rapidly than the nonsense phrases. If this is the case, repeated English phrases may become boring more quickly than repeated Nonsense phrases, eliciting a preference for the nonsense materials.

### *3.1. Method*

#### *3.1.1. Subjects*

Twenty full-term monolingual 8-month-old infants with no history of recurrent ear infections were tested (mean age 8 months, 2 days; range 7:3 to 8:2). Two additional infants were tested but not included in the analysis for the following reasons: fussiness (1), and exceeding a maximum time of 6 min allowed for the test (1).

#### *3.1.2. Stimuli*

The materials consisted of the four English frames and four Nonsense frames from Experiment 1, without the final words and part-words manipulated in Experiment 1. The original recordings from Experiment 1 did not contain appropriate intonational contours when the final target word/part-word was removed, as the words just prior to the target were function words (or nonsense function words) and were highly reduced in the Experiment 1 stimuli. We thus re-recorded all of the frames from Experiment 1, spoken with utterance-appropriate intonational contours and utterance-final lengthening. Overall, the two types of test frames were comparable in amplitude (English Frame RMS energy: 24.96 dB; Nonsense Frame RMS energy: 24.97 dB) and duration (English Frame mean: 0.84 s; Nonsense Frame mean: 0.84 s). Approximately 0.5 s of silence was added to the end of each frame so that all test items were a total of 1.5 s in duration. Two additional English and Nonsense frames were recorded for use during practice trials prior to the test trials.

#### *3.1.3. Procedure*

The procedure was identical to Experiment 1, except that there was no familiarization period prior to testing. To familiarize infants with the pairing of lights and sound used during testing, all infants received four practice trials (two additional English frames and Nonsense frames) prior to the test trials. Following the practice

trials, each infant heard the four test English frames and the four test Nonsense frames on two different trials, for a total of 16 test trials.

### 3.2. Results and discussion

The average listening time was 5.86 s (SE 0.32) for the English frames and 6.53 s (SE 0.33) for the Nonsense frames. A two-tailed paired *t*-test indicated that this difference was significant ( $t(19) = 2.28$ ,  $P < 0.05$ ). Fourteen of the 20 infants listened longer to the Nonsense frames. These results corroborate the results from Experiment 1 suggesting that infants discriminate English from English-like but novel nonsense materials.

One explanation for the direction of preference observed in this experiment – a preference for the nonsense frames over the English frames – is a processing advantage for English materials relative to nonsense materials. On this view, repeated English fragments are rendered less interesting than repeated nonsense fragments because they are processed and acquired more rapidly. An alternative account concerns the nature of the English frames used in this experiment. It is ungrammatical in English to end sentences with grammatical function words such as ‘my’ and ‘the’.<sup>5</sup> Infants are sensitive to the prosodic differences between function and content words from birth (Shi, Werker, & Morgan, 1999), and by 10.5 months of age, infants detect ungrammatical orderings of function words (Shady, Gerken, & Jusczyk, 1995). It is possible that infants in Experiment 2 noticed the ungrammaticality of the English frames, two of which ended in function words. The nonsense frames were not “ungrammatical”, as they contained neither content nor function words. On this view, infants preferred the nonsense frames because they detected the ungrammaticality of the English frames, a discrimination which requires knowledge of English function words (relative to the phonologically-matched nonsense words in the nonsense frames) and their legal locations. Both of these explanations for the novelty preference in Experiment 2 require infants to detect differences between the English and nonsense frames, consistent with the results from Experiment 1. Moreover, the results of Experiment 2 lend support to the claim that infants are knowledgeable about the words of their native language well before their first birthday (e.g. Hallé and de Boysson-Bardies, 1994; Tincoff & Jusczyk, 1999, 2000).

While the results of Experiment 2 buttress the claim from Experiment 1 that the representations emerging from the statistical learning process contain lexical status with respect to the infant’s native language, alternative explanations for the Experiment 1 results remain. One possibility concerns the degree to which the English and nonsense frames are dissimilar from the familiarization stimuli used in Experiment 1.<sup>6</sup> During the segmentation task, the infants were presented with sequences of nonsense sounds. When tested in the Nonsense Frame condition, the test frames also consisted of nonsense; however, the English Frame test items did not. When test

---

<sup>5</sup> Thanks to an anonymous reviewer for this suggestion.

<sup>6</sup> Thanks to an anonymous reviewer for suggesting this alternative hypothesis, and the experiment designed to test it.

items differ in kind from exposure items, it is possible that infants may show a preference for familiar stimuli; on this view, the overall novelty of the stimulus domain during testing might lead infants to preferentially attend to any familiar material presented within the novel frames. If the English test items are more different from the exposure material than the nonsense test items, infants might be listening longer to words in English frames due to an orienting response to familiarity not evoked by the nonsense frames. To the extent that the pattern of results is due to the global familiarity of the test frames relative to the exposure material, we cannot assess whether newly segmented words serve as candidate English lexical items.

We thus designed Experiment 3 to ask whether the familiarity of the test frames relative to the exposure stimuli is the primary factor influencing the results obtained in our first experiment. Following a segmentation task, infants were tested with words and part-words embedded either in an English frame, as in Experiment 1, or a non-linguistic tone sequence frame. If infant performance is affected by the degree of similarity between the exposure and test stimuli, then the tone sequence frame should elicit the same familiarity preference as the English frame, since both English and tone sequences are dissimilar from the nonsense syllables heard during familiarization. Alternatively, if native language status is the key variable influencing the direction of preferential listening, then we would expect to find preferential listening responses like those from Experiment 1: different patterns of listening for English versus non-English materials.

#### **4. Experiment 3**

This study was designed to replicate and extend the results from Experiment 1. After listening to the segmentation stimuli used in Experiment 1, infants were tested with word and part-word targets in one of two types of test frames. Half of the infants heard targets in English frames; the rest of the infants heard targets in tone frames, consisting of tone sequences generated to match the English frames in amplitude, duration, and average pitch. Because the pitch contours of the four English frames from Experiment 1 were similar, the four resulting tone frames would have been difficult to distinguish from one another, which would have led to a difference in the variability of the frames across conditions, possibly affecting direction and degree of preference. We thus decided to use one English test frame and one matched tone test frame for all of the word and part-word targets. One consequence of using the same test frame for all test trials in each condition was increased repetition over the test trials; pilot studies suggested that this repetition diminished both overall listening times and differences over the course of 12 trials. Thus, while infants received all 12 test trials, only the listening times for the first eight trials were analyzed. In addition, the target words and part-words were generated synthetically. Other than these changes, the two experiments were identical.

#### 4.1. Method

##### 4.1.1. Subjects

Forty full-term 8-month-old infants with no history of recurrent ear infections were tested (mean age 7 months, 28 days; range 7:3 to 8:1). Half of the infants were randomly assigned to the English Frame condition and half were assigned to the Tone Frame condition. Within each condition, half of the infants were assigned to each of two counterbalanced exposure languages (Conditions A and B). Thirteen additional infants were tested but not included in the analysis for the following reasons: fussiness (8), exceeding a maximum time of 6 min allowed for the test (2), looking times averaging less than 3 s to one or both sides (1), not looking at the side lights (1), and falling asleep (1).

##### 4.1.2. Stimuli

Materials consisted of the familiarization sequences from Experiment 1. The two target words and two part-words used on the test were generated in citation form by a speech synthesizer (MacinTalk running on a Power PC), using the same parameters as the familiarization sequences. For the English Frame test, the target words and part-words were placed at the end of the carrier sentence “I like my \_\_\_”, which was one of the carrier sentences used in Experiment 1. The same acoustic token of this sentence was used for all trials on the English Frame test. The carrier phrase for the Tone Frame was generated using the sine-wave tone generator included in the CoolEdit PC software. The tone frame consisted of three 0.2 s tones synthesized to approximate the mean *F0* of the three spoken words in the English frame. The tone frame was equivalent to the English frame in amplitude (English Frame RMS energy: 23.44 dB; Tone Frame RMS energy: 23.40 dB), duration (English Frame: 1.44 s; Tone Frame: 1.42 s), and pitch (English Frame average *F0* for each word: 187, 250, 210 Hz; Tone Frame pitches for each tone: 187, 250, 210 Hz). Approximately 0.5 s of silence was added to the end of each sentence so that all test items were a total of 2 s in duration.

##### 4.1.3. Procedure

The procedure was identical to Experiment 1.

#### 4.2. Results and discussion

Pilot studies revealed that when a single test frame was used for all test trials, the increased repetition led to diminished listening times and listening time differences over the course of the experiment. Thus, unlike Experiment 1, we included only the first eight trials in our analyses. As in Experiment 1, we excluded trials with listening times less than 1.5 s in duration prior to analyzing the data, because the head turn on these trials was initiated prior to the onset of the word/part-word target. This criterion led to the exclusion of six of the 320 test trials (1.8%).

The principal hypothesis concerned the differential effects of the test frame (English versus Tones) on listening preferences for familiar (words) versus novel

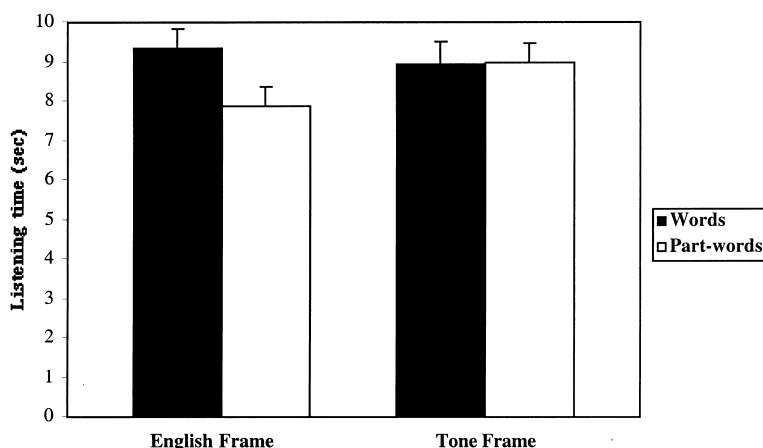


Fig. 2. Experiment 3: listening times for words and part-words for infants in the English Frame and Tone Frame test conditions.

(part-words) items. An ANOVA revealed that the main effects of Frame ( $F(1, 38) = 0.37$ , NS) and Familiarity ( $F(1, 38) = 3.4$ , NS) were not significant. However, the Frame  $\times$  Familiarity interaction was significant ( $F(1, 38) = 4.1$ ,  $P < 0.05$ ) (see Fig. 2). As in Experiment 1, preferences for words versus part-words were affected by the frame in which these items are embedded during testing. Overall, 16 of the 20 infants in the English Frame condition, but only seven of the 20 infants in the Tone Frame condition, showed a preference for the words over the part-words. Paired  $t$ -tests revealed that infants tested on English Frame test items showed a significant preference for words over part-words ( $t(19) = 5.91$ ,  $P < 0.05$ ), while infants tested on Tone Frame items showed no significant preferences ( $t(19) = -0.02$ , NS). Like the Nonsense Frame results from Experiment 1, the lack of significant effects in the Tone Frame condition may reflect the dilution of the dishabituation response to the part-words relative to testing targets in isolation, as in Saffran et al. (1996). The presence of the novel materials in the test frames presumably enhances infants' interest in test items containing both familiar and novel targets, diminishing the impact of prior habituation.

As in Experiment 1, we performed an additional analysis to test the hypothesis that the results from the English Frame condition, but not the Tone Frame condition, would differ from the pattern of data found by Saffran et al. (1996; Experiment 2) using test words in isolation (No Frame condition). To do so, we contrasted the difference scores (words minus part-words) for the three different types of frames. The ANOVA revealed that the difference scores were significantly different as a function of test frame ( $F(2, 61) = 6.3$ ,  $P < 0.01$ ). Bonferroni comparisons of the two conditions from the current experiment with the results from Saffran et al. (1996; Experiment 2) revealed significant differences between the English Frame and the No Frame conditions ( $P < 0.01$ ), but no difference between the Tone Frame and No Frame conditions ( $P > 0.2$ , NS). This analysis supports the hypothesis that

infants treat the words and part-words from the exposure session differently when they are embedded in English sentences than when they are presented in tone frames or in isolation.

A final set of analyses compared the results from Experiments 1 and 3. If infants are segmenting potential English lexical items from the input, the ‘Englishness’ of the test frames should affect performance: infants tested on words and part-words in English frames should show preferences for words, whereas infants tested with non-English frames (nonsense in Experiment 1, tones in Experiment 3) should not. Alternatively, if infant preferences during testing are influenced by the similarity between the exposure corpus (nonsense) and the test frames, then infants tested with either English sentence frames (Experiment 1) or non-English tone frames (Experiment 3) should show familiarity preferences for words, but infants tested on nonsense frames should not. The former hypothesis predicts an interaction between test frame (English versus non-English) and test item familiarity, but no differences between Experiments 1 and 3. The latter hypothesis does not predict any overall English versus non-English effects, but instead predicts an interaction between experiments (English versus nonsense as opposed to English versus tones) and test item familiarity. A three-way ANOVA including Familiarity (tested words versus part-words), Frame (English versus non-English), and Experiment (1 versus 3) tested these two competing hypotheses. The only significant effect was a significant Familiarity  $\times$  Frame interaction (which also emerged in the individual analyses for Experiments 1 and 3) ( $F(1, 92) = 9.84, P < 0.01$ ). This finding is consistent with the hypothesis that infants are segmenting word-like units from the input. The lack of differences for the Experiment factor suggests that the Nonsense and Tone frames functioned equivalently across these experiments, eliciting a different pattern of performance than the English test frame. These findings, taken together, support the hypothesis that infants are engaged in learning English-relevant items when performing a statistical language learning segmentation task. While future research is needed to clarify the nature of these representations, the current findings suggest that the input and output of statistical learning consist of representations at different grains; the units of input may be syllables, but the units of output supercede syllables.

## 5. General discussion

Our experiments were designed to probe the output of infant statistical learning mechanisms engaged in tracking syllable sequences in a word segmentation task. In particular, do the representations emerging from this learning process consist solely of statistical information in the form of syllables with strong sequential predictive relationships, or do these representations contain lexical status with respect to the infant’s native language? To address this question, we took advantage of factors known to influence infant listening preferences. In particular, we reasoned that if the syllable sequences acquired during familiarization are just that – syllable sequences – then they should not enjoy any sort of privileged relationship with respect to the



infant's prior knowledge of English. Under these circumstances, we would expect that the 'Englishness' of the test frames would not affect performance: the same patterns of listening preferences would emerge for English test frames as for nonsense or tone sequence test frames. Alternatively, if infants treat the newly segmented sound sequences as potential lexical units, available for integration into their native language, then we would expect to see different patterns of performance when words and part-words were embedded in English test frames than when they were embedded in non-English (nonsense or tone sequence) test frames. In particular, we expected infants to prefer words over part-words when embedded in English test sentences, due to the prior observation that infants prefer to listen to materials that are segmented according to the structure of their native language. The results of Experiments 1 and 3 support the latter hypothesis; infants preferred words over part-words when they were embedded in English test frames, but not when they were embedded in test frames consisting of nonsense words or tone sequences. Experiment 2 buttressed these claims by demonstrating that infants can discriminate English sentence fragments from matched nonsense word sequences, providing further evidence that infants may process and represent native language materials differently than non-native language materials, even when matched for their prosodic and phonotactic legality.

If our account is correct, infants are not just detecting the statistical properties of sound sequences. Instead, they are using the statistical properties of sound sequences to acquire linguistic knowledge: possible words in their native language. This conclusion raises the question of whether infants exposed to monotone continuous speech believe that they are hearing, and learning, English, or whether they are treating the new materials as a second language, which they then integrate into English. Because the synthesized familiarization materials incorporated English phonology, infants may have treated them as English despite their lack of prosody. Future experiments manipulating the prosodic and segmental content of the familiarization stimuli will help to resolve this question, with potential interesting insights into infant bilingual learning.

A possible implication of these results is that statistical learning underlies language learning by generating linguistic representations. However, it is not the case that statistical learning of sound sequences *necessarily* renders linguistic knowledge. Both infants and adults readily segment non-linguistic tone sequences by virtue of their statistical properties (Saffran & Griepentrog, 2001; Saffran et al., 1999); similar findings have been reported in the visuospatial and visuomotor domains (Fiser & Aslin, in press; Hunt & Aslin, 1998). Is there a conflict between the current findings of linguistically-relevant output from statistical learning, and related findings suggesting that sequential statistical learning is not limited to language?

We suggest that far from conflicting, the growing body of evidence concerning statistical learning abilities and the domains of knowledge which such mechanisms serve illuminates the relationship between learning and eventual knowledge (see also Saffran et al., 1999). Given our word segmentation task, infants apparently attempt to integrate the output of statistical learning with knowledge of English

acquired prior to the experimental task. Similarly, when infants are exposed to novel tone sequences in statistical learning tasks, their responses during testing are influenced by their prior knowledge of the tonal structures prevalent in Western music (Saffran, 2001). These findings suggest that, at least for familiar domains, the nature of the input determines the fate of the output of statistical learning. That is, linguistic input is transformed into linguistic output, musical input is transformed into musical output, etc. The output representations of this learning process thus must obey the existing constraints imposed by the domains into which this new information is integrated. Crucially, however, this transformation is not mediated by separate learning mechanisms. Instead, domain-general implicit learning abilities are hypothesized to render domain-specific knowledge. The learning mechanisms themselves are subject to internal constraints: selective perceptual and computational abilities that may play a role in determining the structural properties of domains whose acquisition they subserve. Learners exposed to syllable sequences are constrained to acquire some regularities but not others (Newport & Aslin, 2000). Similarly, given a tone sequence segmentation task, infants and adults rely on different aspects of pitch in their statistical computations (Saffran & Griepentrog, 2001). Statistical learning is thus both a general learning device and a quite specific one, depending on the master to whom these mechanisms are enslaved.

In summary, the current results support the hypothesis that infant statistical learning mechanisms generate word-like representations available for integration into the native language. These findings raise the possibility, to be addressed in future studies, that infants may more readily map referents to statistically-defined words, and that these emerging units are now available for the discovery of grammatical patterns. While these findings do not address the actual structure of these newly emerging ‘words’, they do strengthen the claim that statistical learning mechanisms are actually used by young language learners (see also Gómez & Gerken, 2000). In part because the artificial languages used in studies of early learning are relatively simple, it is possible that the learning processes revealed in laboratory learning experiments are not actually used by infants learning their native language. However, the current findings suggest that the statistical learning process does contribute to real-world language acquisition: sound sequences defined only by their statistical properties were treated differently in English contexts than in non-English contexts. In this fashion, the statistical structure of the input heard during a brief laboratory learning experience interacted with infants’ current knowledge of English. These learning mechanisms are apparently able to shape what is initially purely statistical information into the beginnings of lexical representations, thereby laying the groundwork for the infant’s subsequent linguistic achievements.

### **Acknowledgements**

This research was supported by grants from the University of Wisconsin – Madison Graduate School, NICHD (HD03352 and HD37466), and NSF (BCS-9983630). We thank Maria Boardman, Susan Bunton, Emily Dannemiller, Phillis George, Greg

Griepentrog, Michelle Loman, AnnMarie Rian, Melissa Schneider, Maya Thet, Erik Thiessen, Diana Wilson, and especially Rachel Robertson for all of their assistance in the conduct of this research, the parents of the infants for their time and generosity, Dick Aslin and three anonymous reviewers for helpful suggestions, and Henry Gleitman for asking the question that inspired this study.

## References

- Aslin, R. N. (1999). *Utterance-final bias in word recognition by 8-month-olds*. Poster presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM.
- Aslin, R. N. (2000). *Interpretation of infant listening times using the headturn preference technique*. Talk presented at the International Conference on Infancy Studies, Brighton, UK.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.
- Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, *61*, 93–125.
- Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: a connectionist model. *Language and Cognitive Processes*, *13*, 221–268.
- Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.
- Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*, *36*, 202–225.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.
- Fiser, J., & Aslin, R. N. (in press). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*.
- Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception and Psychophysics*, *54*, 287–295.
- Gerken, L. A., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: nine-month-olds' sensitivity to phonological vs. syntactic structure. *Cognition*, *51*, 237–265.
- Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109–135.
- Gómez, R. L., & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186.
- Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, *20*, 229–252.
- Hallé, P. A., & de Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: infants' recognition of words. *Infant Behavior and Development*, *17*, 1–23.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.
- Hunt, R. H., & Aslin, R. N. (1998). Statistical learning of visuomotor sequences: implicit acquisition of sub-patterns. In M. A. Gernsbacher, & S. J. Derry, (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.
- Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*, 338–352.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.
- Jusczyk, P. W., Cutler, A., & Redanz, L. (1993). Infants' sensitivity to predominant stress patterns in English. *Child Development*, *64*, 675–687.

- Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402–420.
- Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252–293.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159–207.
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 627–640.
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630–645.
- Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants perception of phrase structure in music. *Psychological Science*, 1, 70–73.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606–608.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Mehler, J., Jusczyk, P. W., Lambert, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 144–178.
- Myers, J., Jusczyk, P. W., Kemler Nelson, D. G., Charles-Luce, J., Woodward, A. L., & Hirsh-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language*, 23, 1–30.
- Newport, E. L., & Aslin, R. N. (2000). Innately constrained learning: blending old and new approaches to language acquisition. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Perruchet, P., & Vintner, A. (1998). A model of word segmentation. *Journal of Memory and Language*, 39, 246–263.
- Redington, M., & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: a distributional perspective. *Language and Cognitive Processes*, 13, 129–191.
- Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: implications for information processing. *Developmental Psychology*, 18, 704–713.
- Saffran, J. R. (2001). Absolute pitch in infancy and adulthood: the role of tonal structure, submitted for publication.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: evidence for developmental reorganization. *Developmental Psychology*, 37, 74–85.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Saffran, J. R., Loman, M. L., & Robertson, R. W. (2000). Infant memory for musical experiences. *Cognition*, 77, 15–23.
- Shady, M., Gerken, L. A., & Jusczyk, P. W. (1995). Prosody serves as a linguistic marker to local co-occurrence patterns in 10-month-olds. In D. MacLaughlin, & S. McEwen (Eds.), *19th Annual Boston University Conference on Language Development*, 2. Somerville, MA: Cascadilla Press.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72, 11–21.
- Spence, M. J. (1996). Young infants long-term auditory memory: evidence for changes in preference as a function of delay. *Developmental Psychobiology*, 29, 685–695.
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in six-month-olds. *Psychological Science*, 10, 172–177.
- Tincoff, R., & Jusczyk, P. W. (2000). *Do six-month-olds link sound patterns of common nouns to new*

*category exemplars?* Poster presented at the biennial meeting of the International Society for Infant Studies, Brighton, UK.

- Wagner, S. H., & Sakovits, L. J. (1986). A process analysis of infant visual and cross-modal recognition memory: implications for an amodal code. In L. P. Lipsitt, & C. Rovee-Collier (Eds.), *Advances in infancy research* (Vol. 4, pp. 195–217). Norwood, NJ: Ablex.
- Werker, J., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.
- Wetherford, M. J., & Cohen, L. B. (1973). Developmental changes in infant visual preferences for novelty and familiarity. *Child Development*, 44, 416–424.