



Stress changes the representational landscape: evidence from word segmentation

Suzanne Curtin^{a,*}, Toben H. Mintz^b, Morten H. Christiansen^c

^a*Departments of Linguistics and Psychology, University of Pittsburgh, 2806 Cathedral of Learning,
Pittsburgh, PA, USA*

^b*Departments of Psychology and Linguistics, and Program in Neuroscience, University of Southern California,
Los Angeles, CA, USA*

^c*Department of Psychology, Cornell University, Ithaca, NY, USA*

Received 12 August 2003; revised 27 August 2004; accepted 31 August 2004

Abstract

Over the past couple of decades, research has established that infants are sensitive to the predominant stress pattern of their native language. However, the degree to which the stress pattern shapes infants' language development has yet to be fully determined. Whether stress is merely a cue to help organize the patterns of speech or whether it is an important part of the representation of speech sound sequences has still to be explored. Building on research in the areas of infant speech perception and segmentation, we asked how several months of exposure to the target language shapes infants' speech processing biases with respect to lexical stress. We hypothesize that infants represent stressed and unstressed syllables differently, and employed analyses of child-directed speech to show how this change to the representational landscape results in better distribution-based word segmentation as well as an advantage for stress-initial syllable sequences. A series of experiments then tested 9- and 7-month-old infants on their ability to use lexical stress without any other cues present to parse sequences from an artificial language. We found that infants adopted a stress-initial syllable strategy and that they appear to encode stress information as part of their proto-lexical representations. Together, the results of these studies suggest that stress information in the ambient language not only shapes how statistics are calculated over the speech input, but that it is also encoded in the representations of parsed speech sequences.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Stress; Infant speech segmentation; Corpus analysis; Distributional learning

* Corresponding author. Tel.: +1 412 624 5933; fax: +1 412 624 6130.

E-mail address: scurtin@pitt.edu (S. Curtin).

1. Introduction

Many characteristics of a child's input language shape the child's perceptual development. For example, early on, infants are able to distinguish a variety of native and nonnative contrasts (see Werker & Tees, 1999 for a review of the literature), but around 10 months of age this ability diminishes for contrasts not found in the native language (Werker & Tees, 1984), suggesting that native phonetic categories are formed as a result of experience with the target language. In addition to forming language-specific phonetic categories, infants are attending to the rhythmic properties of the input language. This sensitivity appears quite early in development and continues to be refined with experience. For example, newborn infants can discriminate languages based on their rhythmic properties (Mehler & Christophe, 1994; Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998). Changes in the rhythmic patterns of alternating strong–weak (SW) syllables are detected by infants between one and four months of age (Jusczyk & Thompson, 1978), and at nine-months of age English infants demonstrate a trochaic bias (SW). However, this preference does not exist at six months suggesting that infants begin to orient to the predominant stress pattern of the target language sometime between 6 and 9 months of age (Jusczyk, Cutler, & Redanz, 1993; Echols, Crowhurst, & Childers, 1997). Infants not only demonstrate a preference for stress patterns in the speech input, but also utilize this information. Mattys, Jusczyk, Luce, and Morgan (1999) found that stress was beneficial for learning phonetic contrasts such that minimally distinct sounds were detected in stressed syllables, but not in unstressed ones. These results suggest that the stress pattern of adult language may influence the level of detail encoded in early lexical representations.

Although much is known about infants' sensitivity to lexical stress patterns, questions still remain as to the exact role of lexical stress in the development of language. For instance, what factors influence the preference for the predominant stress pattern in a given language, and does such a bias provide any advantage in language development? Moreover, if infants use stress information to help segment the continuous speech stream, is stress also encoded in the representations of the resulting segmented units? In this paper we take the first step toward addressing these questions using a combination of corpus analyses and behavioral studies with infants. We start by discussing the acoustic properties of English stress, and proposals about how stress is represented. We then review studies examining stress in early speech processing and segmentation. Our first experiment provides statistical analyses of a corpus of child-directed English speech. We show that, if infants represent stressed syllables differently from their unstressed counterparts, then certain cues to word boundaries become more reliable, and moreover, that words with trochaic stress patterns become easier to discover in the speech stream. Experiments 2–4 employ an artificial language learning paradigm to demonstrate that 9- and 7-month-old infants are indeed very sensitive to lexical stress, and exhibit a bias toward trochaic stress patterns even in the absence of other cues to word boundaries. Finally, Experiment 5 confirms that infants appear to encode stress as part of their preliminary lexical representations. Together, our experiments suggest that stress information in the ambient language not only shapes the representational landscape used by infants in segmentation, but that this information is also encoded in the representations of parsed sequences.

1.1. English stress

Stress is a salient part of the speech signal. There are several acoustic cues to English stress: pitch, duration, amplitude and vowel quality (Hayes, 1995). Changes in pitch contour, vowel duration and amplitude occur on the stressed syllable, whereas, changes in vowel quality often, but not always, occur on the unstressed syllable of a word.

An important step in understanding the role of word-level stress in language development requires us to address how stress is represented. Is stress encoded as part of the lexical entry or is it computed? Cutler and colleagues have suggested that stress is part of the representation of familiar words (Cutler, 1979; Cutler & Isard, 1980), however, they leave open the possibility that rules are applied in the case of unfamiliar words (Cutler & Isard, 1980). In contrast, Chomsky and Halle (1968) proposed a set of rules for assigning stress to English words, and these rules are stored separately from the phonological representation of the word (Halle, 1998). Patterns of morphologically related words comprise part of the evidence in favor of the hypothesis that stress is computed rather than stored with the lexical entry stems ([səˈlaɪvə] *salíva* → [ˈsæliˌvet] *sálivàte*, [ˈfɒtəˌgræf] *phótogràph* → [fəˈtagrəfi] *photógraphy*). For example, since the related forms share the same stem, it is reasonable to assume that the representation of the verb has a long vowel in the second syllable and the noun has a full vowel in the initial syllable (/səlaɪvə/, /fɒtəgræf/). Based on these types of alternations, some researchers argue that stress information is not part of the lexical representation of a word (e.g. Halle, 1998).

Some theories of spoken word recognition similarly posit relatively abstract representations of lexical items, in which stress and unstressed syllables are not differentiated (Lahiri & Marslen-Wilson, 1991; McClelland & Elman, 1986; Studdert-Kennedy, 1976). For example, Lahiri and Marslen-Wilson (1991) propose that spoken word recognition is achieved with respect to abstract, underspecified phonological representations in which all predictable and nondistinct information is absent. They further propose that while information such as stress may play a role in language comprehension, it is not part of the lexical entry. The system is only sensitive to information encoded in the abstract underlying representation, which does not contain any prosodic structures such as syllables and feet.

Church and Schacter (1994), however, found that in implicit memory tasks, surface details persist, and memory is affected by changes in intonation and fundamental frequency. Results such as this provide evidence for an exemplar (or episodic) theory of representations (Goldinger, 1992, 1996, 1998), in which details from the input are retained in memory for each token. These details include such aspects of the input as the speaker's voice characteristics, and the context in which the utterance was produced. In other words, the physical forms of auditory and visual stimuli are not filtered out. Rather, these properties are encoded and remain as part of long-term memory representations.

One goal of the present study was to investigate whether, for very young infants, stress is part of the representation formed of newly parsed sequences, and we conclude that it is.

1.2. *Word segmentation and lexical stress*

Distributional information present in the co-occurrence of syllables has been shown to provide infants with cues to segment the speech stream (Saffran, Aslin, & Newport, 1996). However, syllables carry other information as well, such as prominence information, which contributes to the perceived rhythm of the language. The alternation of strong and weak syllables is a particularly salient property to which infants are sensitive at a very young age (Mehler & Christophe, 1994; Mehler et al., 1988; Nazzi et al., 1998). Indeed, when infants first start segmenting words at seven and a half months, their language-specific rhythmic biases guide segmentation (Jusczyk, Houston, & Newsome, 1999; Polka, Sundara, & Blue, 2002).

In English, the stress pattern is overwhelmingly trochaic (SW). According to Cutler and Carter (1987), an examination of the metrical structure of words in the English conversational vocabulary corpus (Svartvik & Quirk, 1980) demonstrates that a listener has approximately a three to one chance of encountering a strong syllable at the onset of a content word. Moreover, they found that based on frequency-of-occurrence statistics, within the set of polysyllabic words listed in Kucera and Francis (1967), words with strong initial syllables occur much more frequently than words with weak initial syllables. These results suggest that English-raised infants are exposed to more trochaic patterns in the language than to other stress patterns. The behavioral data provides support that English infants at this age successfully segment only SW words, demonstrating a trochiac bias (Jusczyk et al., 1999).

Further evidence that infants are sensitive to distributional information comes from research on Canadian French, which is predominately an iambic language (WS). In this case, infants segment only WS words (Polka, et al., 2002), while English infants of 7.5 months mis-parse iambic WS patterns and associate the S syllable with the word initial syllable. Thus in a word like “guiTAR”, they will segment out “TAR” as the initial syllable. If “TAR” is consistently followed by an unstressed word, such as “is”, infants of this age will treat “TAR” and “is” as a single unit, “TARis” (Jusczyk, et al., 1999), violating the word boundary.

A majority of studies examining the role of stress in segmentation use natural speech in natural speech contexts, which often provides other cues word boundaries. For example, Echols et al. (1997) embedded trochaic units into a sentence comprised of three elements with pauses between each one. They found that infants prefer novel trochaic distractors to familiar trochaic ones, but do not prefer novel iambic distractors to familiar iambic ones. While these naturalistic studies demonstrate that stress plays a role in helping the learner determine boundaries between words, they do not tell us how various potential cues to word boundaries, including stress, interact. To ascertain whether properties such as coarticulation and stress could override statistical information based on transitional probabilities, Johnson and Jusczyk (2001) pitted these cues against each other. Their results show that at 8-months, both coarticulatory cues and stress information override transitional probabilities, indicating that those cues influence the parsing strategy. What has not yet been investigated is how useful stress is for segmenting the speech stream in the absence of other cues. Does stress need to co-occur with pause boundaries, phonotactic information, or transitional probabilities in order for it to play a role in segmentation?

This is another question we address in the current research, and we conclude that stress, alone, triggers segmentation of the speech stream.

Finally, the interplay between stress and statistical information appears to change over development. For example, Thiessen and Saffran (2003) have shown that with infants around 6-months of age, when stress and transitional probabilities are pitted against each other, transitional probabilities override stress. One possible explanation for the difference between 6- and 8-month-olds is that the nature of the cue weighting changes. However, an intriguing alternative account is that the computation of transitional probabilities is not separate from the consideration of stress information, but rather segmentally identical stressed and unstressed syllables are treated differently with respect to transitional statistics. On this view, what might change between six and eight months is the degree to which infants consider stress information in their statistical calculations. This clearly is an important possibility to consider if one argues, as we do, that infants encode stress information in their representation of potential word forms. Hence, another question we address in this paper is how word segmentation would proceed if stress were considered when calculating transitional probabilities.

To begin to answer these questions, we carried out a series of computational and behavioral studies. In Experiment 1, we examined how the statistical properties of the speech input change depending on whether or not stressed and unstressed syllables are treated differently. We then, ran a series of experiments to see how infants use stress information in the absence of any other cues to potential word boundaries. In particular, we isolated stress such that it was the only available cue for parsing the speech stream, and found that English-learning infants transferred their predominant, native language rhythmic pattern to our artificial language task. Finally, we found that, in addition to using stress in parsing, infants encode stress in the representations of the parsed sequences. This result is important because it tells us that infants do not simply store sequences of segments, but rather they encode the salient information about the parsed sequence. Thus, stress is not only a cue for discovering word boundaries, but it is a relevant part of the representation of word forms. This, in turn, further supports the possibility that stressed syllables are uniquely represented from their unstressed counterparts.

2. Experiment 1: stress and the representational landscape

Infants' sensitivity to the stress related properties of language (Jusczyk & Thompson, 1978) suggests that exposure to syllables that differ in their acoustic properties (i.e. for lexical stress the change in duration, amplitude and pitch) may result in differing perceptions of these syllable types. We propose that infants' perceptual differentiation of stressed and unstressed syllables results in a *representational* differentiation of the two types of syllables. This means that the same syllable will be represented differently depending on whether it is stressed or unstressed. Lexical stress thus changes the representational landscape over which the infants carry out their distributional analysis. Here, we employ a corpus analysis to demonstrate how this can facilitate the task of speech segmentation. Specifically, we investigate whether the hypothesized representational

differentiation between stressed and unstressed syllables leads to better segmentation performance when utilizing transitional probabilities.

Research has demonstrated that infants are able to use the transitional probability information in the language as a possible cue to segment speech (Aslin, Saffran, & Newport, 1998; Saffran, et al., 1996). Transitional probabilities refer to the sequential co-occurrences of sounds. For example, in the sequence ‘pretty baby’, the transitional probability is higher for the syllable sequence (bay.bi) since it forms a word than for (ti#bay) which crosses a boundary. Saffran, et al. (1996) examined whether infants could use such statistical information to segment continuous speech. They created an artificial language of trisyllabic nonwords. The statistical information available to the infants was the transitional probabilities between successive syllables (1.0 word internally = ‘words’ and .33 across word boundaries = ‘part-words’). After a short exposure to the language (2 min), they tested 8-month old infants on their ability to distinguish ‘words’ from part-words. The results showed that even when transitional probabilities are the only type of information present and frequency of co-occurrence of syllables is controlled (Aslin, et al., 1998), infants were able to successfully differentiate words from part-words.

In our corpus analysis, we explore how introducing stressed syllables changes the way in which transitional probabilities are calculated. If stressed syllables are treated differently from their unstressed counterparts of equal vowel quality, then transitional probabilities will be calculated differently as well.

2.1. Method

2.1.1. Materials

We used a version of the Korman (1984) corpus that Christiansen, Allen, and Seidenberg (1998) had transformed into a phonologically transcribed corpus with indications of lexical stress. The Korman corpus contains speech directed to British infants between the ages of six and sixteen weeks. Ninety percent of the infant-directed utterances were used to train Christiansen et al.’s neural network models, and we used this portion of the modified corpus for our analyses.¹ Whereas Christiansen et al. used single-phoneme representations we employ whole-syllable representations to simplify our analysis.

All the disyllabic words were first extracted from the corpus, yielding 2952 word tokens distributed over 255 different word types. For each disyllabic word we pulled out two associated disyllabic nonwords. One consisted of the last syllable of the previous word (which could be a monosyllabic word) and the first syllable of the target word, and the other of the second syllable of the target word and the first syllable of the following word (which could be a monosyllabic word). For example, for the target word *sleepy* /slipI/ in the utterance ‘Are you a sleepy head?’ /A ju eI slipI hed/ we would pull out the disyllables /eIsli/ and /pIhed/. We did not extract disyllabic nonwords that straddled an utterance

¹ Christiansen et al. (1998) represented function words as having primary stress based on early evidence suggesting that there is little stress differentiation of content and function words in child-directed speech (Bernstein-Ratner, 1986). More recently, however, Shi, Werker & Morgan (1999) have found evidence in support of such differentiation. For the purpose of our analyses we have therefore chosen to represent function words as having no stress.

boundary as they are not likely to be perceived as a unit. Word tokens occurring as single-word utterances thus had no corresponding nonwords and were therefore removed from further analyses. Three word types—*gracious*, *presto*, and *okay*—only occurred as single-word utterances and could consequently not be included in our analyses. Six additional word types—*upstairs*, *inside*, *outside*, *downstairs*, *hello*, and *seaside*—were omitted because they incorrectly had been assigned primary stress on both syllables according to the MRC Psycholinguistic Database available from the Oxford Text Archive.² The remaining corpus comprised a total of 2407 disyllabic word tokens representing 246 word types. For all words, we randomly selected one of the two associated disyllabic nonwords for a pairwise comparison with the target word. Two versions of each word-nonword pair were created. In one version, the *stress condition*, lexical stress was encoded by appending the level of stress, primary, secondary or none, as a number, 2, 1, or 0, respectively, to the representation of a syllable (e.g. /sli-pI/ → /sli2-pI0/). This allows for differences in the representation of stressed and unstressed syllables that consist of the same phonemes. In the second version, the *no-stress condition*, no indication of stress was included in the syllable representations. Whereas, there were 814 unique syllables in the no-stress condition, the addition of stress level to the syllable representations increased the number of syllables by 7.49% (61 syllables) to a total of 875 unique syllables in the stress condition. If differential representation of stressed and unstressed syllables facilitates segmentation, we would expect better segmentation of words from nonwords in the stress condition compared with the no-stress condition.

2.1.2. Procedure

Our hypothesis suggests that lexical stress changes the basic representational landscape over which infants carry out their statistical analyses in early speech segmentation, and that this in turn facilitates segmentation of speech. Previous research has suggested that infants pay attention the transitional probabilities between syllables occurring in the speech stream (Saffran, 2001; Saffran et al., 1996). We therefore chose to use transitional probabilities as the dependent measure for our analyses. The transitional probability between two syllables, *X* and *Y*, is calculated as the frequency with which *X* follows *Y* in the language divided by the frequency of *X*, that is

$$TP = \frac{Freq(XY)}{Freq(X)}$$

A high transitional probability indicates that *X* is often followed by *Y*. The transitional probabilities between syllables *within* a multi-syllabic word tend to be high because relatively few combinations of syllables are used to make up words. On the other hand, the transitional probabilities for syllables that occur *between* words tend to be low because practically any syllable can begin a new word. Thus, the transitional probabilities between

² The pronunciations for these words as found in the British Dictionary do not assign main stress on both syllables. Given this conflicting information, we decided to leave the eight words out of our analyses. Moreover, pilot work in which the words were included with both syllables stressed yielded the same results as reported here.

the syllables in the nonwords from our corpus should be low relative to the transitional probabilities between syllables within the disyllabic target words.

Transitional probability information is likely to be used by a learner to inform the process of deciding which syllables form coherent units in the speech stream. A child faced with the task of segmenting fluent speech is presented with a continuous string of word tokens between which boundaries have to be found. The main focus of our analyses is therefore on comparisons at the word token level. However, for completeness we also include results from word type analyses in which a word token and its associated nonword is randomly chosen to represent a given word type. In both cases, we would expect that words would have higher transitional probabilities than nonwords, and that the difference between word and nonword transitional probabilities will be greater for the stress condition than the no-stress condition.

2.2. Results and discussion

The first analysis investigated whether the addition of lexical stress to the syllable representations promoted better segmentation performance. A pair-wise comparison between the disyllabic words in the two conditions showed that the addition of stress resulted in a significantly higher transitional-probability mean for the stress condition (tokens: $t(4812)=8.74$, $P<.0001$; types: $t(490)=2.23$, $P<.03$)—see Table 1. There was no difference between transitional-probability means for nonwords in the two conditions (tokens: $t(4812)=0.18$, $P>.88$; types: $t(490)=0.88$, $P>.38$). When distinguishing between words and nonwords using transitional probabilities it is the relative difference between the two that is important. The larger the difference in transitional probabilities between a word and its associated nonwords, the easier it will be to distinguish actual words from nonwords that straddle a word boundary (as is the case with our nonwords). We therefore computed the differences between transitional probabilities for the target words and their associated nonwords. This difference was significantly higher for word-nonword token comparisons in the stress condition compared to the nonstress condition ($t(4812)=7.12$, $P<.0001$; types, $t(490)=2.21$, $P<.03$). The word–nonword differences were significantly different from chance both in the stress condition (tokens: $t(2406)=72.01$, $P<.0001$; types: $t(245)=16.71$, $P<.0001$) and the no-stress condition (tokens: $t(2406)=60.94$, $P<.0001$; types: $t(245)=12.47$, $P<.0001$).

Our analysis has revealed two key results. First, even without a separate representation of stress, transitional probabilities provide useful information for word segmentation as evidenced by the reliable word-nonword difference in the no-stress condition. Second, the addition of stress in the syllable representations significantly improved segmentation of

Table 1
Transitional probability means for words and nonwords in the two stress conditions

Condition	Words		Nonwords		Word-nonword difference	
	Tokens	Types	Token	Types	Tokens	Types
Stress	0.73	0.63	0.15	0.10	0.58	0.53
No-stress	0.64	0.55	0.15	0.12	0.49	0.43

Table 2

Transitional-probability means for words and nonwords from the stress condition as a function of stress pattern

Stress pattern	Words		Nonwords		Word-nonword difference	
	Tokens	Types	Token	Types	Tokens	Types
Trochaic	0.77	0.70	0.15	0.09	0.62	0.61
Iambic	0.29	0.28	0.13	0.16	0.16	0.12

N = 206 for trochaic tokens; *N* = 2,170 for trochaic types; *N* = 237 for iambic tokens; *N* = 40 for iambic types.

words in the stress condition in comparison with the no-stress condition. This result thus confirms our hypothesis that lexical stress benefits the learner by changing the representational landscape in such way as to provide more information that the learner can use in the task of segmenting speech; specifically, words are more easily distinguished from nonwords in terms of transitional probabilities between syllables.

In a second analysis, we investigated whether the trochaic stress pattern predominant in English provides any advantage over the iambic stress patterns when segmenting speech using transitional probabilities. Table 2 shows the transitional-probability means for words and associated nonwords for the disyllabic items in the stress condition as a function of stress pattern. The trochaic words yielded significantly higher transitional means than the iambic words (tokens: $t(2405) = 24.57$, $P < .0001$; types: $t(244) = 6.41$, $P < .0001$). For the nonwords, the iambic condition resulted in marginally lower transitional-probability means (tokens: $t(2405) = 1.78$, $P = .075$; types: $t(244) = 1.77$, $P = .078$). Most importantly, the word-nonword difference was significantly larger for the trochaic words in comparison with the iambic words (type: $t(2405) = 18.49$, $P < .0001$; types: $t(244) = 6.15$, $P < .0001$). These word-nonword differences were significantly different from chance in the trochaic condition (tokens: $t(2169) = 77.42$, $P < .0001$; types: $t(205) = 18.45$, $P < .0001$) and partially for the iambic condition (tokens: $t(236) = 8.13$, $P < .0001$; types: $t(39) = 1.85$, $P = .072$). Thus, the transitional probability distribution of syllables is such that trochaic words should be much less likely to be confused with nonwords than iambic words. These results indicate that a system without any built-in bias towards trochaic stress nevertheless benefits from the existence of the abundance of such stress patterns in languages like English.

Previous research has shown that transitional probabilities between individual phonemes provide a useful source of information for the detection of word boundaries both in a small artificial corpus (Elman, 1990) and in a large corpus of child-directed speech (Hockema, 2004). However, these computational studies did not determine how useful this information would be in terms of separating actual words from nonwords. Studies with infants (Aslin et al., 1998; Saffran et al., 1996), on the other hand, have indicated how transitional probabilities between syllables may be used to distinguish between word and nonwords in simple artificial languages. Our corpus analyses build on this previous work, demonstrating that transitional probabilities between syllables provide reliable information for distinguishing between words and nonwords in actual child-directed speech. The analyses also show that transitional probabilities allow for even better performance when, as we have hypothesized here, stressed and unstressed syllables are represented differently. The presence of salient acoustic information in the form of lexical

stress alters the landscape of the input over which statistical analyses are carried out, such that simple distributional learning devices end up finding words with the predominant stress pattern (trochaic for English) easier to segment. The following experiments test key predictions of this scenario: Experiment 2–4 tests whether young infants learning English indeed show a bias for segmenting stress-initial words from continuous speech *in the absence of other cues to word boundaries*. Experiment 5 tests whether infants indeed encode stress as part of their proto-lexical representations.

3. Experiment 2: stress and segmentation

The corpus analyses in Experiment 1 suggest that the predominant stress pattern of a language is a useful cue for word segmentation. Given the exposure to trochaic patterns and the likelihood that trochaic patterns are more reliably parsed from the speech stream based on transitional probabilities, we hypothesize that infants will attend to stress information and use this to calculate transitional probabilities and segment an artificial language. To test this hypothesis, we conducted an experiment that examines whether stress, independent of other cues, influences in a consistent manner the way infants segment word-like units from continuous speech. If infants recognize sequences based on the position of stress in a continuous speech stream, then it is a salient property of the speech stream that can be utilized successfully for word segmentation.

In addition to a general effect of stress on segmentation, we also investigate whether infants demonstrate the preferred parsing strategy for their language (English). Cutler and Norris (1988) found that English speakers associate a strong syllable with the beginning of a word (Metrical Segmentation Strategy (MSS)). Thus, if infants demonstrate a trochaic bias, following patterns in the language input (Cutler & Carter, 1987), then this will confirm that infants have analyzed the input and extended this ‘knowledge’ to new information in the signal. Thus, not only would infants be statistically analyzing the speech stream online, but they would also be transferring the ‘learned’ information to new situations, which suggests that some type of statistical generalization about the language has been formed.

We tested these predictions using an artificial language learning paradigm. Two groups of 9-month-old infants were presented with a 2-min continuous stream of CV syllables. Crucially, no segmentation cues other than stress were present in the familiarization stimuli (i.e. there were no cues from pauses or consonant and vowel co-occurrence information). The goal of this experiment was to establish whether stress leads to a preference for segmenting syllable sequences as SWW versus WSW or WWS. The use of trisyllabic sequences in the testing phase of the experiment allowed us to explore whether the putative trochaic segmentation strategy results in extracting only SW sequences, or whether it results in larger units that begin with a stressed syllable (Cutler & Norris, 1988). If the former is true, infants should differentiate both SWW and WSW sequences from WWS sequences, since both form a trochaic grouping ((SW)W and W(SW)). Otherwise, WSW and WWS sequences should pattern together. While this will ultimately result in some mis-parses of the speech stream, overall it would be a successful strategy for early

word segmentation. Of course, if infants are only pulling out the stressed syllable, then all the test items should be treated in a similar fashion.

3.1. Method

3.1.1. Participants

Twenty-four English-learning infants, twelve per group, with a mean age of 9 months 10 days ($SD = 7$ days) were recruited through brochures that parents filled out and returned to the Mintz Language Development Lab at the University of Southern California. All participants in the study received a tee shirt. Each infant was randomly assigned to one of two testing groups. Both groups heard the same familiarization sequence, but one group was then tested on initial stress parses (initial group (IG)) and the other on final stress parses (final group (FG)). Both groups were also presented with test sequences corresponding to a medial stress parse and to trisyllabic sequences that never occurred during familiarization (i.e. control sequences that ensure that infants were in fact learning the order of the syllables in the speech stream). Nine additional infants were tested but excluded from further analyses for the following reasons: fussiness or crying (6), and equipment failure (3). Parental consent was obtained for all infants prior to testing in accordance with the University of Southern California human subjects review committee and the principles of ethical treatment established by the American Psychological Association.

3.1.2. Stimuli

The familiarization stimuli consisted of a sequence of syllables with certain controlled properties. First, the initial syllable and every third syllable thereafter in the familiarization string was stressed. We created tri-syllabic strings of syllables that we would treat as ‘words’ during the test phase. However, while we decided what would constitute a ‘word’, there was no information in the familiarization stream to treat these units as anything other than independent syllables. Nonetheless, we assumed that if the stressed syllable was used as a cue to positing a boundary, then the units between stressed syllables would be considered part of the ‘word’.³

It is not possible for a human talker to produce the entire familiarization string in one breath and still maintain the desired prosodic contour from beginning to end. Therefore, the familiarization string was constructed from smaller units. All source items were naturally produced by a female native speaker of English (SC) and recorded on a Marantz cassette recorder with a close talking microphone. To create the naturally produced string, the larger familiarization sequence was divided into three syllable CVCVCV strings with word level stress (prominence of a syllable within the word). In order to preserve a natural sounding contour, each sequence was produced within a 5-syllable window corresponding to the prosodic pattern of the English word ‘unbeLIEvable’ (stress is denoted by capital letters).

³ Any reference to ‘word’ in these experiments is simply for convenience. By ‘word’ we mean parsed sequences of syllables treated as a unit.

The crucial three syllable string was then cut out from the center of the 5 syllable frame at zero crossing of the waveform and spliced together with the other items, resulting in the familiarization string. For example, 'kugabigamu' was produced with stress on the syllable 'bi'. 'gabiga' was then cut from the utterance. The sequence 'gamuNEPoku' was produced with stress on the 'NE' syllable and 'munepo' was cut from the string. The two sequences were spliced together to create the string 'gabigamunepo'. This process was used to create nine tri-syllabic sequences, which were then concatenated into a 27-syllable sequence. This syllable sequence (indicated in bold) was then repeated 24 times to create a 2 min long familiarization string of which a part is shown in (1) with the stressed syllables denoted by capital letters.

(1)

...gabigamunepokUTANedokULEPOGADoneMUTaledobitAPomubILEkugabigamunepoku
TANedokULEPOGADoneMUTaledobitAPomubILEkug**ABIGAMUNEPokUTANedokULEPOGA**
donemutaledobitAPomubILEkugABIGAMUNepokUTANedokULEPOGADoneMUTaledobita
PomubILEkugABIGAMUNepokUTANedokULEPOGADoneMUTaledobita PomubILEku....

All infants were presented with exactly the same familiarization stimuli, namely the continuous 2-min loop above.

Immediately following the familiarization phase, the infants were presented with a set of test items (see Table 3). The initial stress group (IG) was tested on three types of stimuli: items corresponding to initial stress familiarization sequences (DOBITA), items corresponding to medially stressed sequences (LEDOBI), and control items (syllable sequences that never co-occurred during familiarization (BIPOTA)). The final stress group (FG) was also tested on three types of stimuli: items corresponding to final stress familiarization sequences (TALEDO), items corresponding to medial stress familiarization sequences (LEDOBI), and control items (syllable sequences that never co-occurred during familiarization (BIPOTA)). All test items were presented *without a stress contour*, however, the items correspond to groupings based on the position of the stressed syllable in the familiarization string.

To construct the test items, new recordings of the relevant syllables were created. Each syllable was produced in a CVC frame (in the case of the first two syllables) or a CV frame (for the final syllable). For example, 'BIPOTA' was produced as 'BIP' 'POT' and 'TA' syllable sequences. This retained the coarticulation information of the following consonant on the vowel. We ensured that all test items were appropriately coarticulated within and across syllable boundaries (Curtin, Mintz, & Byrd, 2001). Each CV syllable was then cut from the CVC frame at the waveform zero crossing and spliced together with the following onset

Table 3
Test items used in Experiments 2 and 3

Groupings based on:			
Initial stress	Medial stress	Final stress	Control (never co-occurred)
DOBITA	LEDOBI	TALEDO	BIPOTA
NEPOKU	MUNEPO	GAMUNE	LEMUDO
POMUBI	TAPOMU	BITAPO	GANEKU
KULEPO	DOKULE	NEDOKU	KUBINE

consonant. For example, the ‘p’ was cut from Bip, and the ‘t’ was cut from POT. The remaining CV syllables were then all spliced together to create BIPOTA. The rationale for splicing off the coda and then splicing the CV with the following onset consonant was to maintain the onset production of the syllables since there are no codas in the familiarization phase. Moreover, since the items were not presented in the test phase with stressed syllables (i.e. one prominent syllable relative to the other syllables in the string), it was necessary to produce monosyllabic utterances in order to maintain equal prominence across syllables.

In order to ensure proper stop duration during the splicing, the stop closure length was evaluated based on the second syllable of production of a stop in the familiarization recordings. That is, the closure length was determined based on the pre-tonic occurrence of the stop. The splicing together of syllables for the familiar test items results in strings that correspond to sequences in the familiarization string. For example, the string ‘pomubi’ corresponds to a stress-initial parse of the familiarization string:.

...gabigamUNEPokUTANedOKulePOGadoneMUTaleDObitaPOMubiLEku...

Again, however, the test items do not have any SW stress contours. This technique was used to create all of the test sequences for both groups of infants.

The medially stressed items in each experimental group can be viewed as another type of control item: They allow for a controlled test of whether word segmentation is biased towards either stress initial or stress final parses. Importantly, the medial stress control items form a continuous string in the familiarization string of syllables, just as the stress-initial and stress-final test items do. The second set of control words consisted of syllables that, while occurring in the familiarization phase, never co-occurred. This ensured that infants are in fact encoding syllable transition regularities.

3.1.3. Procedure

A preferential listening procedure adapted from Jusczyk and Aslin (1995) was used. The experiment was conducted in a sound-attenuated room with the experimenter situated outside the room observing the infant’s looking behavior on a video monitor and coding the infant’s looking behavior using a keyboard connected to a computer. The infant sat on the caregiver’s lap. The caregiver listened to music over headphones, in order to mask the stimuli.

Situated directly in front of the infant was a video camera and a central light. There were lights on either side of the infant with loudspeakers directly above. During the familiarization phase of the experiment, infants were presented with a 2-min continuous sequence. The familiarization loop was played continuously as a procedure with the lights repeated. However, only the lights, not the familiarization stimuli, were contingent on the infant’s head-turn behavior during this phase of the experiment (Mintz, 1996; Saffran et al., 1996) allowing for the uninterrupted presentation of the familiarization loop. The light procedure during familiarization worked as follows. The center light flashed to orient the child’s gaze to the front of the room. Once the child looked at the central light, the experimenter pressed a key attached to a button-box initiating the flashing of a randomly chosen side-light. As the infant looked towards the flashing side-light the experimenter pressed the appropriate key to indicate a head-turn. The light flashed until the infant

looked away for 2 consecutive seconds at which point it extinguished and the center light began flashing again. The purpose for using lights during this phase of the experiment was to familiarize infants to the contingency between their looking behavior and the light activity, which is important in the testing phase.

Following the familiarization phase was a contingency phase consisting of four trials using pure tones in order to provide a correlation between lights, sounds, and head-turns. The procedure with the lights was identical to that of the familiarization phase; however, now the sounds were contingent on the infants' looking behavior towards a randomly chosen side-light. Following the contingency phase, the test phase consisting of twelve test items was presented using the same procedure. A test trial was terminated if the infant looked away from the appropriate side-light for two consecutive seconds or if the trial timed out after a maximum of fifteen seconds. All items were presented in a randomized order. The computer recorded orientation times for each of the test trials.

3.2. Results and discussion

The infants in the initial stress group (IG) looked significantly longer to the Control items ($M=13.69$, $SE=1.27$) than to the items corresponding to an Initial Stress parse ($M=10.30$, $SE=1.14$) of the familiarization string ($t(11)=2.52$, $P=.02$). There was no difference in looking times for the Control versus Medial Stress ($M=12.38$, $SE=1.55$) sequences ($t(11)=-1.1$, $P=.29$). This difference was also significant according to the Wilcoxon Matched-Pairs Signed-Ranks Test (2-tailed $P=.03$). Of the infants in this group, nine of the twelve infants had significantly longer looking times for the Control sequences than the Initial sequences.⁴ The Medial-Control matched pairs test was not significant ($P=.209$) (Fig. 1).

The IG infants show the least amount of interest in the initial stress sequence because this pattern is very familiar to them in comparison to the control sequence. As predicted, the IG infants appear to correlate a strong syllable with the beginning of a sequence.

Mean looking times for Finally Stressed, Medially Stressed, and Control items were 10.98 s ($SE 1.22$), 9.35 ($SE 1.23$), and 10.07 ($SE .81$) respectively, for the second group of infants who were tested on sequences that correlated with a final stress grouping during familiarization (FG). The FG demonstrated no differences in looking times for any of the items (Control, Final $P=.621$, Medial, Final $P=.474$). There were no significant matched-pairs according to the Wilcoxon Signed-Ranked Test (Control, Final $P=.99$, Medial, Final $P=.373$) (Fig. 2).

The data suggest that stress influences what infants perceive as salient units in the speech stream—perhaps for the reasons indicated in the corpus analyses: this way of

⁴ The novelty preference is consistent with research on statistical learning (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999). This preference can also be explained by examining the stimuli. The number of exposures during familiarization affects the direction of preference (Hunter & Ames, 1988). The greater amount of exposure (a set 2 min during familiarization) the more likely a novelty result (Aslin, 1999; 2000). Moreover, during the test phase, the infants were presented with two sequences that occurred in the familiarization (for example, 'dobita' and 'ledobi' were presented to the IG infants) and one sequence that did not (for example, 'bipota'), which could also contribute to greater interest in novel stimuli.

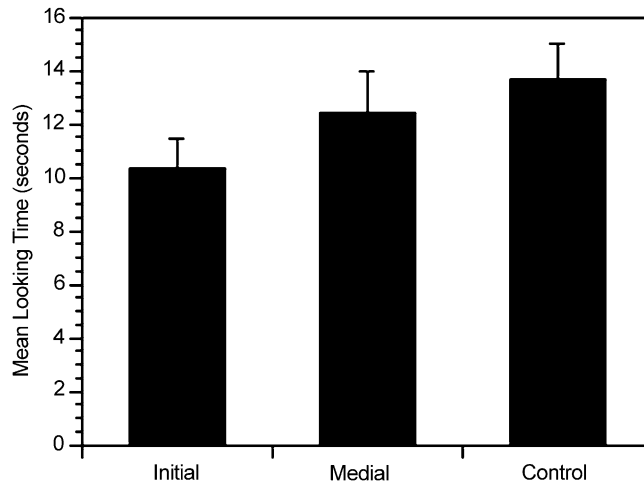


Fig. 1. 9-month-olds in the Initial-Stress Group.

carving up the speech stream is the most informative and will have been statistically reinforced. It is not the case, however, that infants are only extracting strong–weak patterns. As we saw, infants are not simply ignoring the initial syllable of a trisyllabic string that corresponds to a weak–strong–weak pattern. If this were the case, the medial stress items would pattern with the initial stress items when compared to the control items, which was not found. Infants appear to correlate stress syllables to the beginning of a sequence suggesting that they are using an initial, strong syllable strategy (Cutler & Norris, 1988) in conjunction with the trochaic bias they have developed over their first nine months of exposure to English. Importantly, it is both of these strategies that are

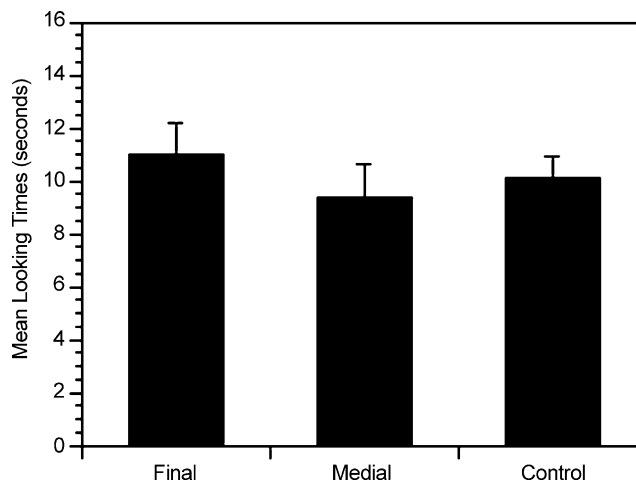


Fig. 2. 9-month-olds in the Final-Stress Group.

driving the segmentation. If it were simply a trochaic bias, then similar patterns should have resulted with the Medial-Stress items since they too conform to the preferred trochaic pattern. Since infants demonstrate a sensitivity to stress and show a trochaic bias by 9-months, we next asked whether or not this preference is exhibited in younger infants. To this end, we tested 7-month-old infants using the same procedure and stimuli.

4. Experiment 3: stress and segmentation with 7-month-olds

By 9-months infants can use an initial, strong syllable strategy to parse sequences in the familiarization stream. Given infants' strong sensitivity to stress throughout their first year, we explored whether this cue to word boundaries can be utilized by younger infants at the initial stages of word segmentation.

4.1. Method

4.1.1. Participants

Thirty-eight infants were recruited using the same methods as Experiment 2, and all the same guidelines for ethical treatment were followed. Fourteen infants were excluded from the study due to the following reasons: fussiness or crying (9), equipment failure (3), infant fell asleep (1), not exposed predominantly to English (1). The remaining twenty-four participants were English-learning infants, twelve per group, with a mean age of 7 months, eight days ($SD=6$ days). The infants were randomly assigned to either the initial-stress or the final-stress groups.

4.1.2. Stimuli

The same stimuli as Experiment 2 was used for both the familiarization, contingency and testing phases of this experiment.

4.1.3. Procedure

The same procedure as in Experiment 2 was used.

4.2. Results and discussion

The results of the 7-month-olds were similar to those of the 9-month-olds. There was a novelty preference for the Control items ($M=13.03$, $SE=1.39$), which had significantly longer looking times than the sequences that corresponded to Initially Stressed items ($M=8.79$, $SE=1.13$) in the familiarization ($t(11)=3.85$, $P<.01$) (Fig. 3). There was no significant difference in the mean looking times between the Control and Medial items ($M=10.37$, $SE=1.8$). According to the Wilcoxon Matched-Pairs Signed-Ranks Test, this difference was also significant (2-tailed $P<.001$). In this group, ten of the twelve infants had significantly longer looking times for the Control sequences than the Initial sequences. No other matched-pairs were significant (Control, Medial $P=.1579$).

The IG infants appear to use an initial parsing strategy; that is, interpreting the strong syllable as a word-initial boundary. The stress-initial sequence is a 'better' trochaic pattern

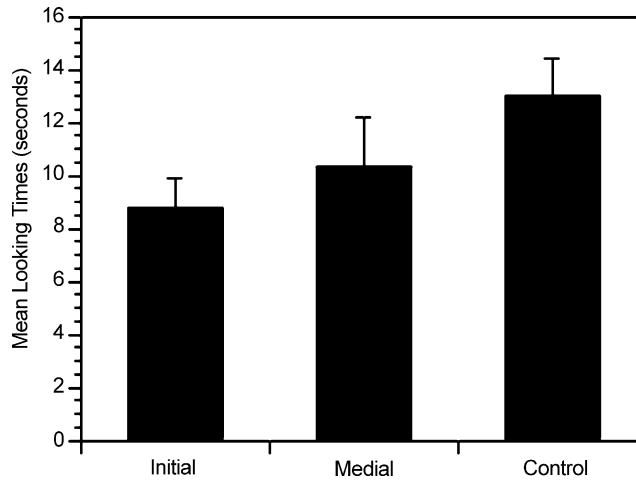


Fig. 3. Results of 7-month-olds in the initial-stress group.

(SWW) than the medial stress sequence (WSW). The results suggest an stress-initial parsing strategy, since the sequence of syllables corresponding to medial stressed sequences are not treated the same as the ones corresponding to initial stress, even though these items also conform to a trochaic pattern.

The results are quite different for the FG infants, for whom the test sequences correspond to final and medial stress sequences in the familiarization string (Fig. 4).

The means for the Final, Medial, and Control test sequences were 9.69 (SE .09), 10.71 (SE 1.38), and 9.8 (SE .87) respectively. There was no significant difference in looking times for any of the test items according to a paired-samples *t*-test for Final vs. Control

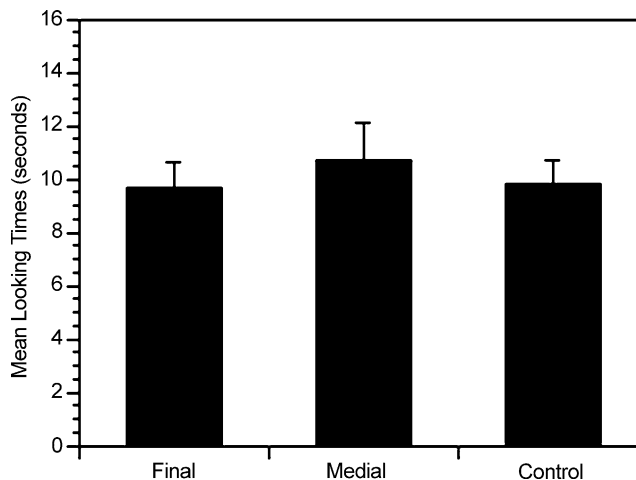


Fig. 4. Results of 7-Month-olds in the final-stress group.

sequences ($t(11) = .143$, $P = .89$) or for Medial vs. Control ($t(11) = .829$, $P = .4245$). Wilcoxon Matched-Pairs Signed-Ranks Test demonstrated that none of the matched pairs were significantly different from each other (Control, Final $P = .806$, Control, Medial $P = .7532$). It appears that the infants in this group are treating all of these items equally. Thus, when presented with items corresponding to medial and final stress, the infants show no preference for either.

In Experiments 2 and 3, we assume that it is the stressed syllable that is determining the parse of the familiarization string. The familiarization string was constructed such that stress was the only cue to a potential word boundary, and the introduction of stress potentially changed how the transitional probabilities of the string were calculated. However, it is conceivable that infants were paying attention to a property in the string of which we are unaware. To ensure that our results were indeed relevant to the stress patterns, and not to some idiosyncratic property of the phoneme or syllable sequences, we further tested a group of 7-month-old infants on the same familiarization string sequences and test items as in Experiments 2 and 3; however, rather than having every third syllable stressed, there was no prosodic contour in the familiarization string.

5. Experiment 4: segmentation without stress

To ensure that stress was in fact the salient cue used by the infant to segment the familiarization stream, we created a new familiarization string in which all the syllables had the same stress. In other words, all syllables were equally emphasized.

5.1. Method

5.1.1. Participants

English infants were recruited through Janet Werker's Infant Studies Centre in Vancouver, Canada. One group of sixteen 7-month-old infants completed the study. Three infants were excluded due to equipment error. All infants received a tee shirt and an infant scientist degree for participating in the study. All local guidelines of the human subjects review committee and the principles of ethical treatment established by the American and Canadian Psychological Associations were followed.

5.1.2. Materials

For the familiarization phase, we used the same syllable sequences; however, we created the string in the same way as the test items used in Experiments 2 and 3. As in the construction of the test items in the previous experiments, each syllable was produced in a CVC frame ('BIG' 'GAM' 'MUN', etc.). Each CV syllable was then cut from the CVC frame at the waveform zero crossing and spliced together with the following onset consonant. By using the monosyllabic utterances, we were able to maintain equal prominence across syllables, and by producing the CV syllables within the larger CVC frame we were able to maintain the appropriate coarticulatory information across syllables as was the case in Experiment 2 and 3.

All infants were presented with the same test items as in the previous experiments, so that all infants heard sequences that corresponded in the earlier experiments to initial-, medial-, and final-stress groupings of the syllables, as well as control items.

5.1.3. Procedure

We used the preferential looking procedure adapted by Maye, Werker, and Gerken (2002). The experiment was conducted in a sound-attenuated room where infants were seated on their parent's lap, in front of a television monitor and a speaker. The monitor and speaker were connected to a computer in an adjoining control room. The presentation of auditory and visual stimuli was controlled by Habit, the program developed by L. B. Cohen (University of Texas, Austin). Infants' gaze was monitored over closed-circuit TV in the control room via a video camera positioned below the television monitor, and the entire experiment was videotaped to check reliability of looking-time measurements.

In the familiarization phase, infants first saw a salient visual stimulus (flashing ball) to draw their attention to the screen. Once the infant was oriented to the screen, the experimenter began the familiarization phase, which was not contingent on the infants looking behavior, for two continuous minutes, during which an unrelated picture was displayed (a field of flowers). Following familiarization, the infant-controlled test phase began. During the test phase, if the infant looked away from the monitor for more than two consecutive seconds, the trial was terminated. There was also a maximum looking time of 15 s at which point if the infant had not looked away for two consecutive seconds, the trial ended. On each trial the flashing ball was presented to draw the infant's attention to the television monitor. Once the infant oriented towards the screen, the experimenter began a trial where the auditory stimulus was presented paired with an unrelated visual stimulus (a black-and-white checkerboard pattern). Throughout the experiment the parent listened to music through headphones, to ensure that they would not inadvertently influence their infant's looking behavior. Looking times to the auditory/visual stimulus pairing were recorded on the computer and measured.

5.2. Results and discussion

As there was no difference in looking behavior for the initial, medial and final test sequences, we collapsed all items into a signal category (word). We compared looking behavior of *t* word sequences and control sequences and found no difference ($t(15) = .151$, $P = .882$) (Fig. 5).

In the absence of stress, infants did not systematically parse the familiarization stream into units that were captured by any of our test categories, and as a result they did not prefer one type of test stimulus over another. This confirms that in the previous experiments, infants parsed the familiarization stream on the basis of stress.

The question now turns to whether or not stress simply denotes a potential word boundary or whether the difference between stressed syllables and their unstressed counterparts exists in the representation of the segmented sequences. In other words, is stress simply a tool for segmentation or is it encoded in the representation of these potential 'words'? We hypothesize based on the corpus analysis results in Experiment 1 that stressed and unstressed syllables which are made up of the same segmental units are

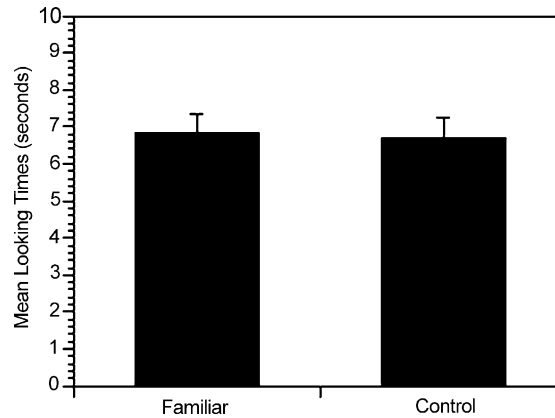


Fig. 5. Looking times for initial, medial and final items (word) and the control items.

treated differently by the statistical learning mechanism, thus changing the representational landscape of the input. We now test whether or not this information is encoded in the representation. Having obtained similar results with 9- and 7-month-old infants in the previous experiments, the final experiment is conducted with 7-month-olds.

6. Experiment 5: representations of segmented sequences

Experiments 2–4 provide evidence that stress is a salient property of the speech signal that infants use to posit initial ‘word’ boundaries. We now ask about the representations of the segmented units. Specifically, do infants encode the stress pattern of the representation of the parsed sequence, or do they only encode the segmental information?

Using a modified version of the Headturn Preference Procedure (Kemler Nelson et al., 1995), Jusczyk and Aslin (1995) found evidence for a ‘word-like’ status of parsed items. They familiarized infants with two different tokens of monosyllabic words spoken in isolation on alternating 30-second trials. They then tested infants on fluent speech passages with either familiar or novel words. The results showed that 7.5-month olds, but not 6-month-olds, listened significantly longer to the passages containing the familiar words, suggesting that by 7.5-months, infants are able to recognize a word in fluent speech after hearing it previously in isolation. Moreover, when they reversed the experiment and exposed the same aged infants to the passages and then tested them on the words in isolation, they found that the infants listened longer to the targeted words heard in the familiarization passage than to novel words suggesting infants can recognize sound patterns in fluent speech.

Recently, Saffran (2001) found support for the ‘word-like’ status of segmented sequences defined by transitional probabilities. In that experiment, a familiarization phase was combined with a test phase in which ‘word’ and ‘part-word’ targets were

placed in either English or nonsense sentence frames. The infants listened longer to words over part-words in English sentences, and no preference was found in nonsense passages. This suggests that units parsed during segmentation have a privileged status relevant to the target language. One goal of our study is to further examine the content of segmented units and, in particular, whether infants might represent stressed syllables as unique entities rather than as members of categories to which their unstressed counterparts belong.

This experiment utilizes the same artificial language learning paradigm as Experiments 2 and 3. Infants were first familiarized with the same continuous stream of trisyllabic nonsense sequences as the previous experiments. The testing phase procedure was adapted from Saffran (2001). Infants were tested on real English sentences that contained a target word either from the familiarization stream, or a control item. This variable (Familiar vs. Control) was crossed with a variable that determined whether the target conformed to the preferred initial stress pattern (SWW), or whether the stress pattern was altered. The rationale for using a common English sentence was to place the infant into a type of ‘language mode’, thereby promoting the parsed sequence to a ‘protoword’ (Saffran, 2001). Recall that we are ultimately trying to establish how newly parsed sequences are represented by the infant. This study differs from the previous experiments in that in the test phase we compare initially stressed sequences, ‘DObita’ (the infants’ preferred parse), to segmentally identical sequences with the stress shifted to the medial syllable, ‘doBIta’. In the previous experiments there was no stress differences in the test items, as all syllables were equally emphasized, ‘DOBITA’.

If stress is part of the lexical representation, then a ‘word’ conforming to the stress pattern of the familiarization string will be treated differently than a sequence that does not conform to this pattern. For example, infants should demonstrate different listening behaviors for the familiar, correct stress pattern ‘dobita’ than the incorrect stress pattern in ‘dobita’ sequences. The sentences containing an incorrectly stressed syllable should pattern like the completely unfamiliar control sentences.

6.1. Method

6.1.1. Participants

One group of sixteen English infants with a mean age of 7-months, 2 days ($SD=8$ days) participated in this experiment. All infants received the same familiarization and test stimuli. Eight additional infants were tested but excluded from analysis for the following reasons: fussiness or crying (5), and not predominantly exposed to English (3). All infants were recruited as in Experiments 2 and 3 at the University of Southern California, and all procedures conformed to the local human subjects review committee and the American Psychological Association guidelines.

6.1.2. Materials

The same familiarization stream as in Experiments 2 and 3 was used. The test stimuli consisted naturally produced English sentences in which the final word either was either taken from the familiarization string (corresponding to a stress-initial parse sequence), or was a novel sequence (i.e. consisted of syllable sequences that never co-occurred in

Table 4
Test frames and targets for Experiment 5

TEST FRAME	TARGETS CONDITION	INITIAL STRESS	MEDIAL STRESS
I like your <i>X</i>	'WORD'	dobita	dobita
I like your <i>X</i>	CONTROL	potaga	potaga

the familiarization string).⁵ Placing the target item in an actual English sentence (as opposed to isolated presentation) arguably increases the likelihood that the test item will be treated as a potential lexical item (Saffran, 2001), and thus allows us to test whether the stressed syllable is encoded as part of the lexical representation. Each sentence was recorded as a whole, in order to maintain a natural sounding overall intonational pattern for the sentences, as well as the appropriate coarticulatory cues. The test sentences and target words are in Table 4.

All sentences were recorded by an adult, female, native speaker of English (SC). It was necessary to ensure that there was no pre-pausal lengthening of the last syllable, since length/duration is a cue for stress. The sentences were recorded in a longer frame (e.g. '*I like your dobita option*') and then the final word was spliced. The speaker naturally produced a glottal stop before the vowel final word. This provided a natural break between the vowels and a mark for splicing off the final word.

One hundred milliseconds of silence was added to the beginning of the sentences and up to 400 ms. of silence was added to the end depending on the overall length of the sentence. This ensured that each sentence was 1.5 s in length.

6.1.3. Procedure

The testing procedure was the same as Experiments 2 and 3. All infants heard all the test sentences. Each individual test sentence was repeated until the infant looked away for 2 consecutive seconds or a maximum playtime of 15 repetitions (33 s) was reached for the trial.

6.2. Results and discussion

The results demonstrate that infants are sensitive to the stress pattern of the sequence extracted from the familiarization based on their preferred parsing strategy. Mean looking times for Initially Stressed Words and Medially Stressed Words were 15.4 s (SE 1.61) and 9.2 (SE .81) respectively. Infants listened significantly longer to Initially Stressed versus Medially Stressed familiar words (paired $t(15)=3.56$, $P < .01$). When compared to the control items, infants also demonstrated a familiarity preference for the Initially Stressed Word sequences (Initially Stressed Word vs. Initially Stressed Control $t(15)=2.57$, $P < .05$; Initially Stressed Word vs. Medially Stressed Control $t(15)=3.87$, $P < .01$). Mean looking times for the Initial Stressed Control and Medial Stressed Control were 11.2

⁵ It should be noted that the nonsense word always corresponded to a noun position in the sentence. Since nouns are overwhelmingly initially stressed in English (Cutler & Carter, 1987), this might lead to a preference for initially stressed strings. However, as an initial experiment, maintaining a constant position within the sentence reduces overall variability and also the possibility of a parsing bias based on the distributional properties of the syntactic position.

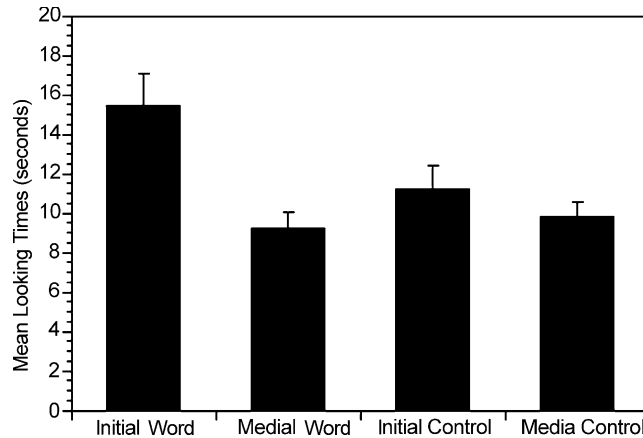


Fig. 6. Results of the listening preference for stressed items.

(SE 1.17) and 9.8 (SE .71) respectively. There was no significant listening preference for the Initial Stressed Control vs. the Medial Stressed Control items ($t(15) = 1.301$, $P = .21$), nor was there any significant listening preferences demonstrated between the Medially Stressed Words and Medially Stressed Controls ($t(15) = -.578$, $P = .57$).

Fourteen of the sixteen infants listened longer to the Initially Stressed Word sequences than to the Medially Stressed Word sequences. This difference was also significant by a Wilcoxon Matched Pairs Signed Ranks Test (2-tailed $P < .01$). Eleven of the sixteen infants had significantly longer looking times for the Initially Stressed Word sequences than the Initially Stressed Control sequences (2-tailed $P < .05$). Moreover, twelve infants significantly preferred the Initially Stressed Word sequences to the Medially Stressed Control sequences ($P = < .01$). No other matched-pairs were significant (Medial Stress Word, Initial Stress Control, $P = .162$, Medial Stress Word, Medial Stress Control, $P = .534$, Medial Stress Control, Initial Stress Control, $P = .162$) (Fig. 6).

The infants prefer the Initial Word sequence over the Medial Word sequence as well as over the Control SWW and WSW Sequences.⁶ The results support the initial parsing strategy demonstrated in Experiment 2. Moreover, the results suggest that stress information is retained in the proto-lexical representation, and not discarded after the sequence was parsed out of the continuous speech stream.⁷

⁶ The resulting familiarity preference is probably due to (1) task demands: test materials in Experiment 2 were situated in familiar sounding English sentences, and infants had to parse test sentences (Hunter & Ames, 1988), and (2) prior knowledge: infants prefer to listen to materials that are consistent with or match their native language (Jusczyk, 1997; Saffran, 2001).

⁷ It is possible that infants have a natural preference for sequences beginning with a 'd'-initial stressed syllables over sequences beginning with a 'p' sound, and stress medial 'd'-words. If it is the case that there is simply a preference for one type of word form over another, then in a preference task without familiarization, we should see a difference in looking times for the two types of words. We tested 8 7-month-olds and found that infants did not prefer the stress-initial 'd' words over any of the other test frames ('d'-initial vs. 'd'-medial $t(7) = -.44$, $P = .67$; 'd'-initial vs. 'p'-medial $t(7) = -1.44$, $P = .19$). Interestingly, infants demonstrated longer listening times for the 'p'-initial word over the 'd'-initial sequence ($t(7) = -2.231$, $P = .06$).

7. General discussion

Infants' bias for attending to stressed syllables provides a valuable tool for extracting word units (Gleitman & Wanner, 1982). This together with a language specific parsing preference, forms the initial representation for segmented units. We hypothesized that infants represent syllables with stress differently from syllables with no stress, and set out to determine whether this type of change to the representational landscape would result in quantifiable segmentation advantages. We also investigated whether infants are sensitive to such information even in the absence of other cues. Related to these two questions, we asked whether the distributional properties of lexical stress that are available to the infant are merely used as cues for segmentation, or whether they are encoded as part of the representation of a potential word form.

Our findings suggest that if a stressed syllable is represented differently from an unstressed syllable containing identical segments, then the input is significantly more informative for positing word boundaries. This information can then be quite useful when using transitional probabilities for segmentation since an occurrence of a stressed syllable will be counted differently from the occurrence of an unstressed syllable with the same segmental content. This changes the shape of the representational landscape such that segmentation using transitional probabilities results in more reliable outcomes, in particular with respect to words that have trochaic stress patterns.

We also found that both 9- and 7-month old infants could successfully parse the stream and preferred sequences in the test phase that corresponded to an initially stressed syllable from the familiarization phase (even though all of the sequences in the test phase had equally emphasized syllables). These results confirm the hypothesis that stress is a salient cue used by the infant to parse sequences from the continuous stream of speech (Jusczyk et al., 1999). In addition, these results demonstrate that after several months of exposure to the input language, the predominant pattern in the language can be used as a cue to parse novel utterances. While we have not explicitly tested this result with a nontrochaic language, based on the results of Polka et al. (2002) for Canadian French, an iambic language, we would predict that infants exposed to French would parse the speech stream following a WWS pattern, and they would prefer items that conform to this pattern (*dobit_A*) to ones where the pattern is altered (*dobit_a*). This suggests that information about the statistical pattern of the language is learned by the infant and used to generalize to new situations.

A possible alternative explanation for the stress-initial bias demonstrated in our studies may be found in terms of the way in which transitional probabilities is calculated.⁸ For example, there are two ways in which we can calculate the transitional probabilities across the syllables in the familiarization stream from Experiments 2–5. If stressed and unstressed syllables are encoded in the same way, then the within-‘word’ probabilities (probabilities of syllables co-occurring together within a word) are the same across all items and the transitional probabilities across potential ‘word’ boundaries are equal. The transitional probabilities for an initial parse (*dobit_a*), a medial parse (*ledobi*) and a final

⁸ We would like to thank an anonymous reviewer for suggesting this alternative explanation.

parse (taledo) are all .33. If however, the stressed and unstressed syllables are treated as different, then the transitional probabilities across items in the familiarization stream change. The initial bigram transitional probability of an initial parse (DObita) is 1.00, for a medial parse (leDObi) it is .50, and for a final parse (taleDO) it is .50. By treating the stressed and unstressed syllables as unique, there is a higher initial bigram transitional probability for an initial parse of the stream. If our infants are indeed segmenting based on stress-modulated transitional probabilities, then their stress-initial bias might not be due to a trochaic bias resulting from exposure to English. Instead, it could be the result of a general bias toward high transitional probabilities in the initial part of a sequence. One way to distinguish between the two explanations would be to run the same experiments as presented here with infants who are primarily exposed to an iambic language (such as, Canadian French). If high initial bigram transitional probabilities are influencing parsing, then we would expect to obtain the same results as reported here; that is, a preference for an initial parse. However, if our results are due to an English-specific trochaic bias resulting from language experience—as we have argued here—then we would predict that learners of an iambic language would show the opposite pattern of responses; that is, a preference for final-parse strings. The experience-based iambic bias found in infants exposed to Canadian French (Polka et al., 2002) provides indirect support of the latter prediction.

Our results suggest that stress might play multiple roles, including aiding in segmentation, and modulating transitional probabilities. Additionally, if stress is changing the representational landscape over which transitional probabilities are calculated, then we need to examine how the nontypical stress patterns in English are learned. Recent work by Thiessen and Saffran (2004) sheds light on this question. They presented infants first with words separated by pauses that conformed to either a trochaic or iambic stress pattern. They then had a word segmentation phase followed by a test phase. Their results demonstrated that familiarity with a particular pattern influenced segmentation. In other words, infants could learn the less frequent iambic pattern.

Our studies not only examined how English input might lead to a stress initial parsing strategy, but also probed whether the information in the signal is encoded as part of the representation of the sequences excised from the speech stream. We found that infants demonstrated significantly different behaviors for sequences that contained identical segments but differing stress patterns (DObita versus doBIta) when these were placed in natural language contexts. Of interest, in our Experiments 2&3, the test items were not exact matches of the familiarization stream. Rather in the test items all the syllables were all equally emphasized—produced in isolation and hence stressed. However, in Experiment 5, a switch in stress resulted in the item being treated as novel. If we think of the representation of the parsed sequences as attracting the best match, then the items in Experiments 2&3 best match the stress initial parse (DObita encoded and DOBITA recognized versus LEDOBI or TALEDO). In the case of Experiment 5, the infant encoded DObita and preferred DObita to doBIta as it is a better match. These results suggest that a ‘best match’ to the information encoded is used in the recognition of familiar forms.

Moreover, the results from Experiment 5 further confirms our hypothesis that stressed syllables are treated as unique with respect to their segmentally identical but unstressed counterparts. In other words, the stress pattern of the parsed sequence was encoded. Recall

that some theories of spoken word recognition posit relatively abstract representations of lexical items, in which stress and unstressed syllables are not differentiated (Lahiri & Marslen-Wilson, 1991; McClelland & Elman, 1986; Studdert-Kennedy, 1976). As such, spoken word recognition is achieved with respect to abstract, underspecified phonological representations in which all predictable and nondistinct information is absent (Lahiri & Marslen-Wilson, 1991). However, if it is the case that stress syllables are encoded in the representation of parsed sequences, then these results lend support for exemplar representations, which encode details of the speech input—including stress—(Goldinger, 1992, 1996, 1998) rather than abstract representations that require stress to be computed on-line (Halle, 1998).

While our findings suggest that infants encode stress in segmented units, findings reported by Vihman, Nakai & DePaolis (see Vihman, 2000) suggest that infants might not always draw on this information. They tested how altering the stress pattern of a familiar word would affect listening by 11-month old English learning infants by using unaltered (correctly stressed) familiar words and compared them to altered familiar words (incorrectly stressed, trochaic → iambic and iambic → trochaic). The results demonstrated no significant listening difference. Vihman et al. claim that at 11-months of age, segmental patterning plays a more important role in infants' mental representation of words than does prosodic patterning. One possible explanation for these divergent results is that differential access to information may be dependent on the developmental level (Werker & Curtin, *in press*). In other words, what is important or attended to most might be different at different points in development, and for different tasks.

Support for this explanation comes from a number of studies that examined the importance of information at different ages. In the case of segmentation, when pitted against other cues, stress overrides transitional probabilities in 8-month olds (Johnson & Jusczyk, 2001), but transitional probabilities override stress when infants are 6-months (Saffran & Thiessen, 2003). These results imply that cues are weighted differently over time. In illustration, to determine whether infants grouped test syllables into a unit based on a specific rhythmic pattern (Morgan & Saffran, 1995) tested infants on their latency to detect a buzz that was inserted either at a position that preserved or violated the grouping. At nine months infants detect the buzz if the rhythm and sequential syllabic information both correspond to that which they had learned (e.g. treat GOKa as a unit, whether in a tiGOKa, deGOKa, GOKati, GOKade context). At six months infants only demonstrate this when the rhythmic pattern remains the same, in spite of whether the syllable sequences change or not. These results suggest that only the older infants were able to pick up and integrate both sets of cues in the short training phase.

We argue that infants are encoding all the information and that the stress information is available to infants at 7 and 9 months of age, and that this information continues to be encoded throughout development. Curtin and Werker (2004) demonstrate in an associative word-learning task that 12-month-old infants encode stress information. The infants noticed a switch not only when the stress pattern remains constant but the segments change, but also when the stress pattern of newly learned words is altered. This suggests that stress and segmental information are encoded in the representation of a word form, as was the case with the Morgan and Saffran (1995) results with older infants.

Our results suggest a picture of early speech segmentation in which stress information in the speech signal both plays a role in segmenting the speech stream while also becoming part of the representation of speech sequences. The connectionist speech segmentation model by Christiansen et al. (1998) is consistent with this picture. Their Simple Recurrent Network (Elman, 1990) model was trained to integrate sets of phonetic features with information about lexical stress and utterance boundary information derived from a corpus of child-directed speech. The network was trained to predict the appropriate values of these three cues for the next segment or indicate the presence of an utterance boundary. After training, the network was able to integrate the input such that it would activate a boundary unit not only at utterance boundaries, but also at word boundaries inside utterances. The network was thus able to generalize patterns of cue information that occurred at the end of utterances to cases where the same patterns occurred within an utterance. This model performed well on the word segmentation task while capturing additional aspects of infant segmentation, such as the bias toward the dominant trochaic stress pattern in English, the ability to distinguish between phonotactically legal and illegal novel words, and having segmentation errors being constrained by English phonotactics. Of particular importance for our purposes is that lexical stress both shaped and informed learning, and was an intricate part of not only the model's ability to segment words but also of its representation of speech sequences.

In summary, the input can provide the language learner with information about the predominant patterns of the target language. This information is used to begin organizing the speech input. At the same time, the information is encoded in early representations and generalizations about the language can begin to emerge. This is demonstrated when the infant is presented with novel sequences and demonstrates biases that are the result of a few months of exposure to the ambient language. Thus, lexical stress permeates the infants' emerging speech processing abilities by changing the representational landscape over which they do statistical learning, laying the foundation for later language learning.

Acknowledgements

This behavioral research was supported by grants from the Social Sciences and Humanities Research Council of Canada (SSHRC) (doctoral #752-98-0283, post-doctoral #756-2001-43) awarded to Suzanne Curtin, and a Zumberge Faculty Research and Innovation Fund (USC) and an Equipment Grant from Intel Corporation award to Toben Mintz. The corpus analyses were supported by a Human Frontiers Science Program Grant (RGP0177/2001-B) awarded to Morten H. Christiansen. We would like to thank Janet Werker for helping with the control experiments and for the use of her lab for these studies, and Dani Byrd for help constructing the stimuli. We would like to thank Rachel Walker, Maryellen MacDonald and Mark Seidenberg for their helpful comments. We would also like to thank Jacques Mehler and two anonymous reviewers for their constructive and insightful comments. Thanks to the parents and infants for participating in these studies.

References

- Aslin, R. N. (1999). Utterance-final bias in word recognition by 8-month-olds. *Biennial meeting of the society for research of child development, Albuquerque, NM*.
- Aslin, R. N. (2000). Interpretation of infant listening times using the headturn preference technique. *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(4), 321–324.
- Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother–child speech. *Journal of Phonetics*, 14, 303–309.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row xiv, p. 470. Reprinted 1991, Boston: MIT Press.
- Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 2–3.
- Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 20, 521–533.
- Curtin, S., Mintz, T. H., & Byrd, D. (2001). Coarticulatory cues enhance infants' recognition of syllable sequences in speech. In A. H. J. Do, L. Dominguez, & A. Johansen (Eds.), *BUCLD 25: Proceedings of the 25th annual Boston University conference on language development* (pp. 190–201). Somerville, MA: Cascadilla Press.
- Curtin, S., & Werker, J. F. (2004). Patterns of word-object associations. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *BUCLD 28: Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 120–128). Somerville, MA: Cascadilla Press.
- Cutler, A. (1979). Errors of stress and intonation. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the Tongue, Ear, Pen, and Hand* (pp. 167–180). New York: Academic Press.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133–142.
- Cutler, A., & Isard, S. (1980). The production of prosody. In B. Butterworth, *Language production: Speech and talk* (Vol. 1) (pp. 229–269). London: Academic Press.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113–121.
- 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.
- Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner, & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (p. 348). Cambridge, England: Cambridge University Press.
- Goldinger, S. D. (1992). *Words and voices: Implicit and explicit memory for spoken words (Research on speech perception Tech. Rep. No 7)*. Bloomington, IN: Indiana University Press.
- Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166–1183.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105-2, 251–279.
- Halle, M. (1998). The stress of English words: 1968–1998. *Linguistic Inquiry*, 29(4), 539–568.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago.
- Hockema, S. (2004). Finding words in speech: An investigation of American English. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *BUCLD 28: Proceedings of the 28th annual Boston University conference on language development* (pp. 244–255). Somerville, MA: Cascadilla Press.
- 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.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-Month-Olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44(4), 548–567.

- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: The Massachusetts Institute of Technology 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, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675–687.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3–4), 159–207.
- Jusczyk, P. W., & Thompson, E. J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. *Perception & Psychophysics*, 23, 105–109.
- Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The headturn preference procedure in for testing auditory perception. *Infant Behavior and Development*, 18, 1111–1116.
- Korman, M. (1984). Adaptive aspects of maternal vocalizations in differing contexts at ten weeks. *First Language*, 5, 44–45.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the mental lexicon. *Cognition*, 38, 245–294.
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38(4), 465–494.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1–86.
- Mehler, J., & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M. S. E. Gazzaniga (Ed.), (pp. 943–954). Cambridge, MA: The MIT Press.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Mintz, T. H. (1996). *The roles of linguistic input and innate mechanisms in children's acquisition of grammatical categories*. Doctoral dissertation, University of Rochester.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, 66, 911–936.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756–766.
- Polka, L., Sundara, M., & Blue, S. (2002). The role of language experience in word segmentation: A comparison of English, French, and bilingual infants. *The 143rd Meeting of the Acoustical Society of America: Special Session in Memory of Peter Jusczyk*, Pittsburgh, PA.
- Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, 81, 149–169.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.
- 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(1), 27–52.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621.
- Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, 39, 484–494.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2), B11–B21.
- Studdert-Kennedy, M. (1976). Speech perception. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 243–293). New York: Academic Press.
- Svartvik, J., Quirk, R. (1980). *A Corpus of English Conversation*. Lund., CWK Gleerup.

- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39(4), 706–716.
- Thiessen, E. D., & Saffran, J. R. (2004). Infants' acquisition of stress-based word segmentation strategies. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *BUCLD 28: Proceedings of the 28th annual Boston University conference on language development* (pp. 608–619). Sommerville, MA: Cascadia Press.
- Vihman, M. (2000). *From prosodic response to segmental patterning: The origins of representation for speech*. MS. University of Wales, Bangor.
- Werker, J. F., Curtin, S. (in press). *PRIMIR: A Developmental Framework of Infant Speech Processing, Language Learning and Development*.
- Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75, 1866–1878.
- Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50, 509–535.