Cognitive Psychology **38,** 465–494 (1999)
Article ID cogp.1999.0721, available online at http://www.idealibrary.com on IDEAL®

Phonotactic and Prosodic Effects on Word Segmentation in Infants

Sven L. Mattys and Peter W. Jusczyk

Departments of Psychology and Cognitive Science, Johns Hopkins University

Paul A. Luce

Department of Psychology and Center for Cognitive Science, University at Buffalo

and

James L. Morgan

Department of Cognitive and Linguistic Sciences, Brown University

This research examines the issue of speech segmentation in 9-month-old infants. Two cues known to carry probabilistic information about word boundaries were investigated: Phonotactic regularity and prosodic pattern. The stimuli used in four head turn preference experiments were bisyllabic CVC·CVC nonwords bearing primary stress in either the first or the second syllable (strong/weak vs. weak/strong). Stimuli also differed with respect to the phonotactic nature of their cross-syllabic C·C cluster. Clusters had either a low probability of occurring at a word juncture in fluent speech and a high probability of occurring inside of words (''within-word'' clusters) or a high probability of occurring at a word juncture and a low probability of occurring inside of words (''between-word'' clusters). Our results show that (1) 9-month-olds are sensitive to how phonotactic sequences typically align with word boundaries, (2) altering the stress pattern of the stimuli reverses infants' preference for phonotactic cluster types, (3) the prosodic cue to segmentation is more strongly

This work was supported by Research Grant 15795 from NICHD and Research Scientist Award 01490 from NIMH to P.W.J. and Grant NIH NIDCD 1 R01 DC 0265801-A1 to P.A.L. We thank Mike Vitevitch for his help with the stimulus preparation, Denise Mandel for recording the stimuli, and Jim Sawusch, Bill Hamblen, and Sam Folby for designing the computer software and hardware interfaces for the test procedures. Thanks are also due Ann Marie Jusczyk and Eileen Crowley for help in testing subjects. We also thank Ann Marie Jusczyk, Peter Gordon, Cathy Best, and two anonymous reviewers for their constructive comments on the manuscript.

Correspondence and reprint requests concerning this article should be addressed to Sven Mattys, Department of Psychology, Ames Hall, Johns Hopkins University, 3400 N. Charles Street, Baltimore MD 21218-2686. E-mail: mattys@cogsci.jhu.edu.

0010-0285/99 $30.00

relied upon than the phonotactic cue, and (4) a preference for high-probability between-word phonotactic sequences can be obtained either by placing stress on the second syllable of the stimuli or by inserting a pause between syllables. The implications of these results are discussed in light of an integrated multiple-cue approach to speech segmentation in infancy.

Infants are born with an impressive number of perceptual skills crucial to learning a language (see Jusczyk, 1997, for a review). For instance, they perceive many of the phonetic contrasts that are necessary for sound classification and word discrimination (e.g., Eimas, 1974; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Levitt, Jusczyk, Murray, & Carden, 1988; Morse, 1972; Kuhl, 1979, 1983). Moreover, little experience is necessary to activate these capacities, as demonstrated by young infants' abilities to perceive and discriminate nonnative contrasts (e.g., Aslin, Pisoni, Hennessy, & Perey, 1981; Eimas, 1975; Lasky, Syrdal-Lasky, & Klein, 1975; Polka & Werker, 1994; Streeter, 1976; Trehub, 1976). In addition to these discriminative capacities, infants are also able to cope with sources of variability in the speech signal such as differences in speakers' voices and speaking rates (Jusczyk, Pisoni, & Mullennix, 1992b; Kuhl, 1979, 1983; Kuhl & Miller, 1982).

These language-general capacities found in early infancy soon give way to more language-specific ones. The early perceptual readiness that serves as a springboard for dealing with a wide variety of potentially important contrasts becomes more tailored to the native language during the second half of the first year. As sensitivity to certain nonnative contrasts begins to decline (Werker & Lalonde, 1988; Werker & Tees, 1983, 1984), infants become more attuned to the sound pattern characteristics of words in their native language. The reorganization of infants' perceptual and productive capacities, given experience with their native language, has been documented in a number of ways. With respect to production, sometime between 5 and 10 months of age, infants start to exhibit babbling patterns consistent with the prosodic features of their native language (Levitt, 1993; Levitt & Wang, 1991) and, from 9 months onward, babbling behavior evolves toward native vowels (de Boysson-Bardies, Halle, Sagart, & Durand, 1989; de Boysson-Bardies, Sagart, & Durand, 1984) and consonants (de Boysson-Bardies & Vihman, 1991; de Boysson-Bardies, Vihman, Roug-Hellichius, Durand, Landberg, & Arao, 1992).

As for perception, attunement to the language of their environment is demonstrable in a number of domains. Soon after birth, infants exhibit a very general sensitivity to the rhythmic characteristics of their native language (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988; Nazzi, Bertoncini, & Mehler, 1998). By 6 months, they display the ability to recognize characteristics of typical words in their native language better than those in a foreign language (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993b). Phonetic tuning is reputed to take place around the same age: Six-month-olds have been shown to organize the internal structure of vowel categories around the phonetic dimensions typical of their native language (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994; although see Polka & Bohn, 1996, for a different interpretation of these findings). At 9 months, sensitivity to native prosodic marking of major phrasal units starts to emerge (Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward, & Piwoz, 1992a).

These early perceptual skills constitute an important step toward word learning. The means by which infants begin to extract individual words from the speech input has recently received much attention. Among the potential cues to word boundaries are (1) allophonic variations (Hohne & Jusczyk, 1994; Jusczyk, Hohne, & Bauman, in press-a), (2) distributional regularities (Brent & Cartwright, 1996; Goodsitt, Morgan, & Kuhl, 1993; Saffran, Aslin, & Newport, 1996), (3) prosody (Echols, Crowhurst, & Childers, 1997; Jusczyk, Cutler, & Redanz, 1993a; Jusczyk, Houston, & Newsome, in press-b; Morgan, 1996), and (4) phonotactics (Brent & Cartwright, 1996; Friederici & Wessels, 1993; Myers, Jusczyk, Kemler Nelson, Charles-Luce, Woodward, & Hirsh-Pasek, 1996). The present research focuses on the last two cues: prosody and phonotactics.

It has long been known that young infants are sensitive to suprasegmental distinctions involving rhythm and stress (Chang & Trehub, 1977; Demany, McKenzie, & Vurpillot, 1977; Jusczyk & Thompson, 1978; Morrongiello, 1984; Spring & Dale, 1977; Trehub & Thorpe, 1989). However, recent findings indicate that English-learners, between 6 and 9 months, appear to also know a great deal about the distribution of stress *within the words* of their language (Jusczyk et al., 1993a; Morgan, 1994, 1996; Morgan & Saffran, 1995; Newsome & Jusczyk, 1995; Turk, Jusczyk, & Gerken, 1995). For example, consistent with the fact that most English words receive initial-syllable stress both in adult speech (Cutler & Carter, 1987) and in infant-directed speech (Kelly & Martin, 1994), Jusczyk et al. (1993a) found that 9-month-old infants listened significantly longer to bisyllabic words bearing a strong/weak pattern (e.g., ‘‘butter’’ and ‘‘ardor’’) than to ones bearing a weak/strong pattern (e.g., ‘‘between’’ and ‘‘arouse’’). Thus, by 9 months, English-learners have discovered the predominant stress patterns of words in their native language.

There is growing evidence that such prosodic sensitivity participates in the extraction of words in connected speech, in line with what has been reported for adult English-speakers (e.g., Cutler & Butterfield, 1992; Cutler & Norris, 1988). For example, Echols, Crowhurst, and Childers (1997) observed that, when 9-month-olds were presented with weak–strong–weak nonwords containing a 250-ms silent pause either before or after the strong syllable, the infants listened longest to the stimuli in which the pause occurred before the strong syllable. That is, infants preferred speech configurations in which strong syllables were the onset rather than the offset of a unit.

A further indication that English-learners use stress as a pointer to word onsets is that they have less difficulty extracting words with initial strong syllables from fluent speech than words with initial weak syllables. Jusczyk, Houston, and Newsome (in press-b) found that 7.5-month-old infants had little trouble identifying a familiarized trochaic word (e.g., ''doctor'') embedded in a longer passage but they could not do so with familiarized iambic words (e.g., ''guitar''). Follow-up experiments revealed that infants responded to the entire strong/weak pattern and not only to the strong syllable (e.g., ''dock''). Moreover, when infants were first familiarized with passages that contained an iambic word that was consistently followed by the same weak syllable (e.g., ''guitar is''), and were later tested on the word-straddling trochaic sequence ''taris,'' they listened longer to that sequence (but not to the sole strong syllable ''tar'') than to a matched control sequence. Taken together, these results suggest that English-learning infants begin to segment words by isolating trochaic feet from fluent speech.

During the same period, infants are developing sensitivity to the phonotactics of their language. Phonotactics refer to the constraints on the ordering of segments within and between the words of a language. For example, the sequence [nt] is found within the syllables of many English words, whereas the sequence [mt] is not. Phonotactic regularities constitute a potentially important source of information during word recognition because they inform the listeners about the likelihood that a given phoneme will be adjacent to another given phoneme within and between words (e.g., Auer, 1993; Brown & Hildum, 1956; Eukel, 1980; Greenberg & Jenkins, 1964; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997). Recent research (Vitevitch & Luce, 1999) has demonstrated the role of phonotactics in spoken word processing in adults: When effects of lexical competition are minimized, spoken words and nonwords composed of frequent phonotactic sequences are processed more quickly and accurately than stimuli composed of less frequent sequences.

Early studies of the development of phonotactic awareness focused on children and adolescents (e.g., Brown & Hildum, 1956; Pertz & Bever, 1975). For example, in a study of 3- to 4-year-olds, Messer (1967) found that nonwords consistent with English phonotactics (e.g., [frul]) were judged as possible words and pronounced correctly more often than matched nonwords with atypical sound patterns (e.g., [mrul]). However, even exposure to speech during the first 9 months of life seems to be adequate to develop sensitivity to at least some phonotactic regularities of the language. Jusczyk et al. (1993b) tested infants with lists of English and Dutch words, most of which contained sound sequences that were not phonotactically legal in the other language. All words were produced by an English/Dutch bilingual speaker. When both types of lists were presented to English- and Dutch-learners, 9-month-olds, but not 6-month-olds, listened longer to the lists containing words in their native language. This result suggests that 9-month-olds have developed some sensitivity to phonotactic patterns encountered in their language.

Jusczyk, Luce, and Charles-Luce (1994) went a step further by exploring infants’ sensitivity to phonotactic sequences that occur with different frequencies within words in English. They found that 9-month-olds, but not 6-month-olds, listened significantly longer to monosyllabic nonwords containing high-probability phonotactic sequences (e.g., ‘‘chun’’) than to ones containing low-probability phonotactic sequences (e.g., ‘‘yush’’). Thus, sensitivity to the phonotactic patterns typically found within English words is in place by 9 months of age.

Knowing that sequences of phonemes are typically found in specific positions within words could constitute valuable information for locating word boundaries in the speech stream (Brent & Cartwright, 1996; Cairns, Shillcock, Chater, & Levy, 1997; Church, 1987). For instance, the sequence [br] is generally located at the beginning of a word, whereas the sequence [nt] is typically found at word end. Exploiting this type of regularity during speech processing could contribute to adequate word segmentation and lexical access. Given the continuous nature of connected speech addressed to both adults (Cole & Jakimik, 1980a,b; Klatt, 1980; Liberman & Studdert-Kennedy, 1978) and infants (Aslin, 1993; Mehler, Dupoux, & Segui, 1990; van de Weijer, 1998; Woodward & Aslin, 1990), knowledge of how phonotactic patterns are distributed in the input could be important in isolating words from the speech stream.

Recent computer simulations have demonstrated that the distribution of phonetic segments within and between words is a potentially powerful cue to word segmentation (e.g., Brent & Cartwright, 1996). Some authors even argue that phonotactic sensitivity could bootstrap a series of additional segmentation strategies, such as prosodic segmentation, without explicit use of such concepts as ‘‘syllable’’ or ‘‘strong syllable’’ (Cairns et al., 1997). However, data from studies with infants are limited. In one of their experiments, Myers et al. (1996) noted that the presence of phonotactic cues may have helped English-learning 10.5-month-olds to detect interruptions artificially inserted into weak/strong words. More direct evidence comes from a study by Friederici and Wessels (1993). The authors presented 4.5-, 6-, and 9-month-old Dutch infants with pairs of monosyllables with identical phonetic segments but with different orderings of these segments. Specifically, each pair contained a given cluster of consonants (e.g., [br]) that occurred at stimulus onset in one case (e.g., [bref]) and at stimulus offset in another (e.g., [febr]). The clusters tested were found either at word onset and never at word offset (e.g., [br]) or at word offset and never at word onset (e.g.,[rt]). Nine-month-olds, but not 4.5- and 6-month-olds, showed a listening preference for stimuli containing the critical cluster in a permissible position (e.g., [bref] or [murt]) over the stimuli containing the same cluster in an impermissible position (e.g., [febr] or [rtum]). This result is important because it shows that, by 9 months of age, infants can not only discriminate between phonotactically well-formed and ill-formed word-like stimuli but they can do so based solely on the position of clusters that typically signal word onsets or word offsets.

Previous studies with infants have focused on sensitivity to phonotactic sequences that occur within syllables or words (e.g., Friederici & Wessels, 1993; Jusczyk et al., 1994). Potentially, this sensitivity could arise from infants' storing information about the sound patterns of frequently occurring words. Indeed, 9-month-olds have been shown to engage in some long-term storage of the sound patterns of words that appeared frequently in stories to which they were exposed (Jusczyk & Hohne, 1997). Thus, the source of infants' knowledge of phonotactics may be lexical, deriving from the sound patterns of words that they are beginning to store. However, another possibility is that there is a sublexical source for infants' knowledge of phonotactics (Pitt & McQueen, 1998; Vitevitch & Luce, 1999). Namely, infants may be attuned to the frequency with which certain phonotactic sequences appear between words in the language. Such sensitivity may originate from a sublexical level if the between-word clusters in question do not occur within real words, thus making it unlikely that phonotactic information regarding the probabilities of these clusters is represented in the lexicon. In general, linguistic approaches have not carefully examined phonotactic sequences between words, because, with few exceptions, these sequences do not appear to be subject to significant constraints. Nevertheless, knowledge of the kinds of phonotactic sequences that are more likely to occur between words, rather than within words, could potentially be used as a cue to word boundaries. The present research is a first attempt to explore this possibility. We examined infants' sensitivity to phonotactic sequences that, while legal within words, vary in their probability of occurrence within and between words. We also explored whether infants are sensitive to the correspondence between cross-word phonotactic sequences and other potential markers of word boundaries, such as prosodic regularities.

As discussed above, little empirical work has been devoted to examining the possible role of phonotactics in infants' detection of word boundaries. Our research is a first attempt to demonstrate that infants are differentially sensitive to variations in the frequencies of within- and between-word sequences, a necessary prerequisite to the use of such information in segmenting words in fluent speech. Thus, the goal of the present paper is to (1) further examine the contribution of phonotactic sensitivity to the detection of word boundaries and (2) explore how such phonotactic sensitivity combines with prosodic word-boundary cues in the course of word segmentation. We tested 9-month-old infants because of their apparent sensitivity to both phonotactics and lexical prosody and because it has been suggested that infants at this age may begin to integrate different cues (Lalonde & Werker, 1995; Morgan, 1996).

EXPERIMENT 1

To investigate the role of phonotactic constraints in locating word boundaries, we chose pairs of strong/weak CVC·CVC bisyllabic nonwords (e.g., ‘‘nongkuth’’ [ˈnɔŋ·kʌθ]) that varied in the nature of their cross-syllabic C·C cluster. All the C·C clusters we used were found to occur equally in connected English speech. The only thing that distinguished them was their probability of occurring at a syllable boundary *within* words and *between* words. The stimuli in one condition contained a C·C cluster whose probability was high within words but low between words (the ‘‘within-word’’ condition, e.g., ‘‘nongkuth’’ [ˈnɔŋ·kʌθ]). The stimuli in the other condition contained a matched C·C cluster whose probability was high between words but low within words (the ‘‘between-word’’ condition, e.g., ‘‘nongtuth’’ [ˈnɔŋ·tʌθ] or ‘‘nomkuth’’ [ˈnɔm·kʌθ]).

If 9-month-olds perceive the CVC·CVC stimuli with strong/weak stress patterns as single word-like units, they should prefer the stimuli with the within-word clusters because these are phonotactically better formed than the stimuli with the between-word clusters. The latter contain phonotactic cues that conflict with the CVC·CVC sequences being perceived as a single unit. The phonotactic cues for the between-word stimuli suggest that the two syllables should be treated as separate units. Obviously, our prediction will hold only if the CVC·CVC stimuli are perceived as single units to begin with, that is, before phonotactics are taken into account. However, given that strong/weak stimuli have been shown to be perceived as cohesive units by English-learners (Jusczyk et al., in press-b; Morgan, 1994; Morgan & Saffran, 1995), our predictions have some empirical grounding.

Method

Participants. The participants were 24 infants from monolingual American-English-speaking homes (13 males and 11 females), approximately 9 months of age (mean: 39 weeks, 6 days; range: 37 weeks, 3 days, to 44 weeks, 1 day). Two additional infants were tested but were discarded from the analyses for the following reasons: One stopped looking at the flashing lights and one was suffering from an ear infection.

Stimulus materials. The materials consisted of 21 prerecorded lists of 12 stimuli each. All stimuli were nonsense trochaic CVC·CVC sequences whose second vowel, although unstressed, was fully realized. In order to distinguish these stimuli from those used in subsequent experiments, the present stimuli will all be referred to as SW (for ‘‘strong/weak,’’ an arbitrary label for ‘‘primary stressed/secondary stressed’’). We elected to use full vowels in both syllables in order to preserve vowel quality throughout experiments, whether syllables received stress or not. Even though prosodic segmentation has usually been modeled with full versus reduced syllables, finer stress contrasts have been shown to produce similar effects on infants’ prosodic preferences (Turk et al., 1995).

The stimuli differed in the phonotactic nature of the C·C cluster at their syllabic juncture. Based on the mother’s utterances of the child-directed Bernstein (1982) speech corpus, available in MacWhinney (1991), we chose six C·C clusters that had a low probability of occurring across word boundaries and a higher probability of occurring within words across syllable boundaries (e.g., [ŋ·k]). These ‘‘within-word’’ clusters were matched with sets of C·C clusters

TABLE 1
Cross-Syllabic Consonant Clusters Used in the Stimuli of Experiments 1–4 and Their Mean Frequency of Occurrence

	Type of cluster		
		Between-word	
	Within-word	A	B
	[ŋ·k]	[ŋ·t]	[m·k]
	[f·t]	[f·h]	[v·t]
	[v·n]	[v·m]	[z·n]
	[m·θ]	[m·h]	[n·θ]
	[k·tʃ]	[k·ʃ]	[p·tʃ]
	[ŋ·g]	[ŋ·b]	[n·g]
Between-word frequency:	1.00	19.17	18.83
Cross-syllable within-word frequency:	22.33	0.17	0

Note. The mean frequencies of occurrence come from Bernstein’s (1982) child-directed speech corpus. They correspond to the number of times each given cluster was encountered in the corpus.

that showed the opposite phonotactic probabilities. The ‘‘between-word’’ clusters were relatively frequent across word boundaries of utterances in the corpus but rare (or never found) within words across syllable boundaries. In order to match the phonetic content of the between-word clusters with that of the within-word clusters as closely as possible, we generated two sets of between-word clusters. In one set (‘‘A’’ clusters), the first consonant was identical to that of the matched within-word cluster, while the second consonant differed by one place, manner, or voicing feature (e.g.,[ŋ·t]). In the other set (‘‘B’’ clusters), the second consonant was identical to that of the matched within-word cluster, while the first consonant differed by one place, manner, or voicing feature (e.g., [m·k]). All of the 18 clusters (6 within-word, 6 between-word A, and 6 between-word B), together with their mean occurrence in the Bernstein corpus, are displayed in Table 1. All of the clusters in this study were permissible in English, both within and between words, but differed on their probabilities of occurrence in the two critical positions. Thus, the zero value in Table 1 should be construed as indicative of a low—rather than zero—probability cluster.[1] An analysis of variance run on the high-frequency

---

[1] Although the within-word clusters’ frequencies in the Bernstein corpus are a close approximation of those in the Webster’s dictionary, there are occasional discrepancies as a result of the lexical nature of child-directed speech and of the size of the corpus. For example, while [k·ʃ] is found commonly within words in adult-directed speech (e.g., ‘‘reaction,’’ ‘‘dictionary,’’ ‘‘direction,’’ etc.), such words hardly ever occurred in the child-directed speech sample (‘‘destruction’’ occurred once). Conversely, [k·tʃ], though rare within English words in Webster’s, turned out to be a high-frequency within-word cluster because of numerous occurrences of the words ‘‘picture’’ and ‘‘pictures.’’ Other frequent words in adult-directed speech that contain a cluster listed as between-word are ‘‘somehow,’’ ‘‘enthusiasm,’’ ‘‘capture,’’ etc. These words were not part of the Bernstein corpus. Also, the present cluster counts were obtained using a fairly conservative phonemic transcription system. One pair of clusters, [v·t]–[f·t] may consequently yield slightly different figures depending on the transcription of the string ‘‘have to.’’ If ‘‘have to’’ is assumed to contain the cluster [f·t] (our counts reflect its transcription as [v·t]), the phonotactic probability contrast between the two clusters decreases but nevertheless remains salient and the average frequencies are hardly altered (a change of

means ($M = 22.33$, 19.17, and 18.83) as well as pairwise comparisons did not reveal any significant differences (all Fs < 1).

Bisyllabic CVC·CVC nonsense stimuli were generated from the C·C clusters by adding a CV sequence at the beginning of each cluster and a VC sequence at the end. These CV and VC sequences were chosen so that neither the resulting individual syllables nor the entire bisyllables constituted English words. The mean stimulus duration, measured from a digital waveform display, was 1.05 s in the within-word condition and 1.06 s in the between-word condition (A = 1.04 s and B = 1.09 s). The within-word and between-word stimuli were concatenated into 12-stimulus-long test lists (ISI: 1 s). Each list contained two tokens of each one of the six cluster types for a given condition. The two tokens differed in the segments surrounding the cluster (e.g., ‘‘nangkuth’’ [ˈnɔŋ·kʌθ] and ‘‘chongkudge’’ [ˈtʃoŋ·kʌdʒ]). The 12 stimuli were arranged randomly in the lists, with a different random order for each list. There were 7 lists for the within-word condition and 14 lists for the between-word condition (A and B were each assigned 7 lists). Because of the relatively small set of stimuli that met our selection criteria, 2 of the experimental lists were composed of stimuli that occurred in the other 5 lists. These hybrid lists also contained two tokens of each cluster type. One of the 2 hybrid lists from each condition was used for practice. The same practice lists were used consistently across subjects. The average list duration, resulting from stimulus concatenation, was 23.61 s for the within-word lists and 23.77 s for the between-word lists (A = 23.54 s and B = 24.00 s). A listing of the stimuli, arranged by cluster type, is presented in the Appendix.

The stimuli were recorded with a Shure microphone in a sound-shielded booth by a female native speaker of American English from New York. All stimuli were digitized on a VAX Station Model 3176 computer at a 20-kHz sampling rate via a 12-bit analog-to-digital converter and then concatenated into test lists. Digitized versions of the lists were transferred to a Macintosh Quadra 650 computer for playback during the experiment.

Design. Infants heard the within-word lists and either the between-word A or B lists. They were randomly assigned to one of these two groups. They each heard 2 practice lists and 12 test lists. The order of presentation of the test lists was randomized for each infant, with no more than 3 lists of the same type (within-word or between-word) consecutively. The side of presentation was randomized for each list as well, with no more than 3 lists presented on the same side consecutively. The side of presentation was independent of the stimulus condition.

Apparatus. The computer controlled the presentation of the lists and recorded the observer’s coding of the infant’s response. The audio output for the experiment was generated from the digitized waveforms of the samples. A 12-bit D/A converter was used to recreate the audio signal. The output was fed through anti-aliasing filters and a Kenwood audio amplifier (KA 5700) to one of two 7-in. Advent loudspeakers mounted on the side walls of the testing booth.

The experiment was conducted in a three-sided test booth constructed out of 4 × 6 ft pegboard panels. Except for a small section of preexisting holes in the center panel used for viewing the infant’s head turns, the panels were backed with white cardboard to guard against the possibility that participants might respond to movements behind the panel. The test booth had a red light and a loudspeaker mounted at eye level on each side panel and a green light mounted on the center panel. Directly below the center light, a 5-cm hole accommodated the lens of a video camera employed to record each test session. A white curtain suspended around the top of the booth shielded the infant’s view of the rest of the room. A computer terminal and a response box were located behind the center panel, out of view of the infant. The response box, which was connected to the computer, was equipped with a series of buttons that started and stopped the flashing center and side lights, recorded the direction and duration of head

---

less than three units in the two relevant categories). If anything, the realization of [v·t] as [f·t] in ‘‘have to’’ would have resulted in less extreme differences in cluster probabilities, thereby potentially attenuating our results for the within versus between comparison.

turns, and terminated a trial when the infant looked away for more than 2 s. Information collected about the direction and duration of head turns for each trial was stored in a computer data file. Computer software was responsible for the selection and randomization of the stimuli and for the termination of the test trials. The average listening times for the test lists were calculated by the computer after the completion of each session.

Procedure. A version of the Headturn Preference Procedure (HPP), originally developed by Fernald (1985), was used in the present experiment (see Jusczyk, 1998, for a complete assessment of the procedure and its current implementations). Each infant was held on a caregiver's lap. The caregiver was seated in a chair in the center of the test booth. Each trial began with the blinking of the green light on the center panel. When the infant had oriented in that direction, the light was extinguished and the red light above the loudspeaker on one of the side panels began to flash. When the infant made a headturn of at least 30° in the direction of the loudspeaker, the stimulus for that trial began to play and continued until its completion or until the infant failed to maintain the 30° headturn for 2 consecutive s (e.g., if the infant turned back to the center or the other side or looked at the caregiver, the floor, or the ceiling). If the infant turned briefly away from the target by 30° in any direction, but for less than 2 s, and then looked back again, the time spent looking away was not included in the orientation time. Thus the maximum orientation time for a given trial was the duration of the entire list. The flashing red light remained on for the entire duration of the trial.

Each experimental session began with two practice trials, one on each side. The practice trials were used only to make the infant comfortable with responding to the lights. Responses during practice trials were not recorded. The average loudness level of the lists was 73 ± 2 dB (C) SPL using a Quest (Model 215) should level meter. The test phase began immediately following the two practice trials and consisted of 12 test lists (6 within-word and 6 between-word). An observer hidden behind the center panel looked through the peephole and recorded the direction and duration of the infant's head turns using a response box connected to the computer. The observer was not informed as to which side the lists would be playing from. In addition, both the observer and the infants' caregiver wore foam earplugs and listened to masking music over tight-fitting SONY MDR-V600 headphones. The masker consisted of loud instrumental music, which had been recorded with few silent periods. With such masking, caregivers and observers were made unable to determine the nature of the stimulus on the trial (see Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995, for data on the efficacy of this masking procedure).

Results and Discussion

Mean listening times to within-word and between-word stimuli (collapsed across A and B) were calculated for each of the 24 infants. Across all participants, the average looking times, displayed in Fig. 1, were 11.20 s ($SD = 3.12$) for the within-word stimuli and 9.14 s ($SD = 2.79$) for the between-word stimuli. An analysis of variance performed on these scores indicated a significant difference between the two types of stimuli, $F(1,23) = 9.09$, $p < .007$. That is, the participants showed a clear preference for stimuli containing a high-probability within-word cluster over stimuli containing a high-probability between-word cluster. Importantly, all of the clusters in this experiment, whether they were of the within-word or the between-word type, occurred with the same frequency in the Bernstein (1982) corpus. The two types of clusters differed only in their likelihood of being found at a word boundary or within a word. The present experiment reveals that 9-month-olds were sensitive to just that.

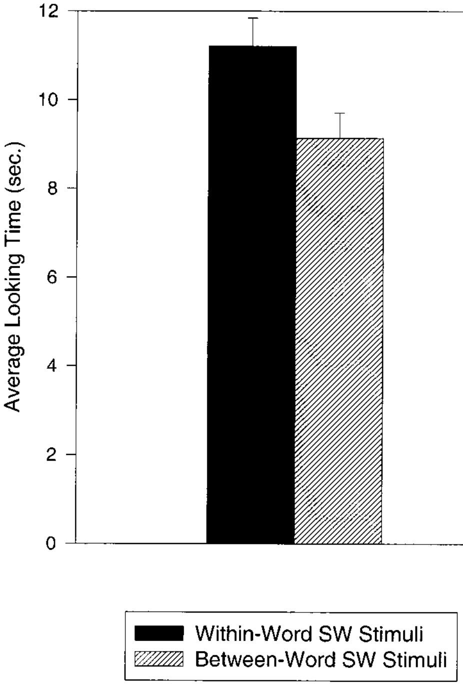

FIG. 1. Average listening times (and standard error bars) for the within-word and between-word strong/weak stimuli (Experiment 1).

As we noted earlier, the difference between the within-word and between-word stimuli is their likelihood of being perceived as one or two units by virtue of their phonotactic configuration. Within-word sequences should promote one-unit perception, whereas between-word sequences should promote two-unit perception. Because the stimuli in this experiment were strong/weak, it is likely that they were represented as single units, that is, a format consistent with the within-word version of the stimuli. Extensive research on the subject has shown not only that infants prefer listing to strong/weak than weak/strong stimuli (Jusczyk et al., 1993a, in press-b; Turk et al., 1995) but also that their perception tends to be more disrupted by the insertion of extraneous noise in the middle of a strong/weak than a weak/strong speech block (Morgan, 1994, 1996; Morgan & Saffran, 1995). These data suggest that English-learners around 8 months of age organize speech sounds into trochaic or strong/weak feet, with strong syllables being treated as word onsets. Therefore, within-word stimuli are better formed than between-word stimuli because only they are consistent with the single-unit nature of the input presentation.

Let us consider what would happen if the same phonotactic sequences occurred in items with a weak/strong stress pattern. Given the tendency of English-learners to treat the occurrence of strong syllables as markers of word onsets (e.g., Jusczyk et al., in press-b), the location of the strong syllable in a weak/strong stimulus should cause this stimulus to be interpreted as consisting of two units and not one. Thus, in weak/strong stimuli, the prosodic cues to word boundary should combine with the phonotactic cues in such a way that within-word sequences might now appear less adequate than between-word sequences. This prediction reflects the fact that, in weak/strong stimuli, prosodic and phonotactic cues would generate conflicting segmentation hypotheses for stimuli with within-word sequences. However, in the case of stimuli with between-word sequences, the prosodic and phonotactic cues would reinforce each other, promoting a two-unit percept. In Experiment 2, we tested this hypothesis using the stimulus lists of Experiment 1, but with primary stress on the second syllable.

EXPERIMENT 2

Method

Participants. The participants were 24 infants from monolingual American-English-speaking homes (13 males and 11 females), approximately 9 months of age (mean: 39 weeks, 1 day; range: 35 weeks, 1 day, to 47 weeks, 3 days). Five additional infants were tested but were discarded from the analyses because of crying (4) and restlessness (1).

Stimulus materials. The CVC·CVC stimuli were the same as in Experiment 1 but were stressed on the second syllable. As in the SW stimuli, the weak syllable of the present WS stimuli bore a full vowel. The speaker was the same as in Experiment 1. The stimulus duration was on average 1.00 s for the stimuli with within-word clusters and 1.07 s for the ones with between-word clusters (A = 1.05 s and B = 1.10 s). The average list duration, resulting from stimulus concatenation, was 23.04 s for the lists of within-word stimuli and 23.89 s for the lists of between-word stimuli (A = 23.59 s and B = 24.19 s). The average loudness level of the lists was 73 ± 2 dB (C) SPL.

Design, apparatus and procedure. These were the same as in the previous experiment.

Results and Discussion

Mean listening times to the within-word and the between-word stimuli (collapsed across A and B) were calculated for each of the 24 infants. Across all participants, the average looking times, displayed in Fig. 2, were 8.18 s ($SD = 2.76$) for the within-word stimuli and 10.25 s ($SD = 3.84$) for the between-word stimuli. An analysis of variance performed on these scores indicated a significant difference between the two types of stimuli, $F(1,23) = 5.55$, $p < .03$. In other words, 9-month-old infants, when presented with WS stimuli, showed a listening preference for those containing clusters usually found at word boundaries, that is, the between-word stimuli. As anticipated, an analysis of variance of the data from the present experiment with those of Experiment 1, with the between-subject factor Stress

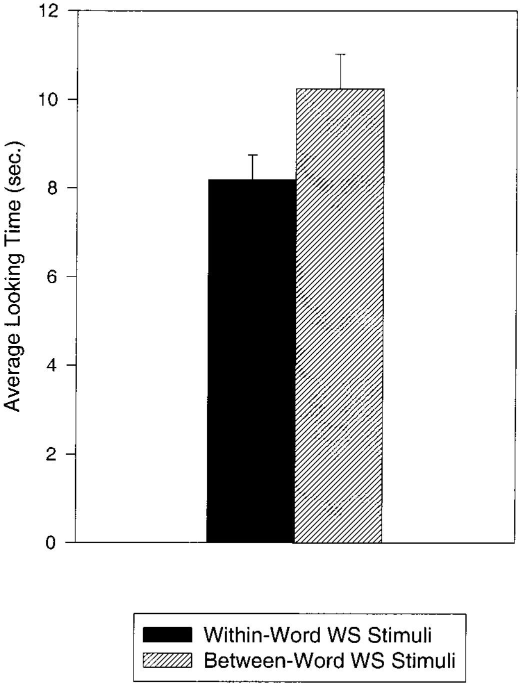

FIG. 2. Average listening times (and standard error bars) for the within-word and between-word weak/strong stimuli (Experiment 2).

Pattern (SW vs. WS) and Phonotactic Structure (within-word vs. between-word), revealed a reliable interaction, $F(1,46) = 13.76$, $p < .001$.

The result of Experiment 2, along with that obtained in Experiment 1, demonstrates that prosody can modulate the effect of phonotactic word-boundary cues to the point of reversing it. The within-word clusters, which were assumed to bring cohesiveness in the strong/weak stimuli, were not the preferred structure for the weak/strong stimuli. Rather, it was the between-word clusters, compatible with the prosody-induced two-unit perception in the weak/strong stimuli, that were preferred for the weak/strong stimuli. This sensitivity to how phonotactic and prosodic cues typically line up is consistent with the hypothesis that infants rely on both sources of knowledge to segment words from the input. This observation fits with the finding that 9-month-olds, but not 6-month-olds, are capable of integrating both rhythmic and distributional information to group the input into processing chunks (Morgan & Saffran, 1995). The present results appear to extend this finding more specifically to phonotactics and to the discovery of word boundaries.

EXPERIMENT 3

In Experiments 1 and 2, the preference for stimuli including high-probability within-words clusters over ones with high-probability between-word clusters, and the reversal of this effect caused by stress change, suggest that 9-month-olds are sensitive to both phonotactic and prosodic cues for word boundary. In particular, the reversal of phonotactic preference from Experiment 1 to Experiment 2 appears to stem from the fact that strong syllables are perceived as word onsets.

If stressing the second syllable of the stimuli generated a perceived word boundary, then simply inserting an explicit boundary between the syllables should also generate a preference for stimuli containing between-word clusters. In other words, the reversal between strong/weak and weak/strong stimuli should also occur between the original strong/weak stimuli and a new version of the strong/weak stimuli in which a pause is inserted at syllable juncture. Specifically, the insertion of a pause between the syllables should make within-word clusters mismatch with the perceived word boundary, hence inverting the preference obtained in the intact SW stimuli.

Method

Participants. The participants were 24 infants from monolingual American-English-speaking homes (12 males and 12 females), approximately 9 months of age (mean: 39 weeks, 0 days; range: 36 weeks, 3 days, to 41 weeks, 5 days). Seven additional infants were tested but were discarded from the analyses for the following reasons: crying (4), failure to look at the flashing lights (2), and parental interference (1).

Stimulus materials. The stimuli were the same as those of Experiment 1. The only difference was that a 500-ms pause was inserted between the two syllables of each stimulus. Using a speech editor, for each stimulus, we located the point corresponding to the end of the first consonant of the critical cluster and the beginning of the second consonant. We inserted 500 ms of silence at that point (e.g., [ˈnɔŋ·kʌθ] became [ˈnɔŋ~kʌθ], where ~ is a 500-ms pause). The duration of the interruption was chosen so that it would be long enough for the two syllables of a stimulus to be perceived as distinct entities but not so long as to interfere with the interstimulus duration.

The average stimulus and list durations were similar to those of Experiment 1 except that each stimulus was 500 ms longer, thus increasing the overall duration of each list by 6 s.

Apparatus and procedure. These were the same as in the previous experiments.

Results and Discussion

Mean listening times to the within-word and between-word stimuli (collapsed across A and B) were calculated for each of the 24 infants. Across all participants, the average looking times, displayed in Fig. 3, were 8.20 s ($SD = 3.22$) for the within-word stimuli and 9.72 s ($SD = 3.48$) for the between-word stimuli. An analysis of variance performed on these scores indicated a significant difference between the two types of lists, $F(1,23) = 7.85$, $p < .02$. When the SW stimuli contained a 500-ms pause at syllable juncture, infants spent more time listening to those containing a high-proba-

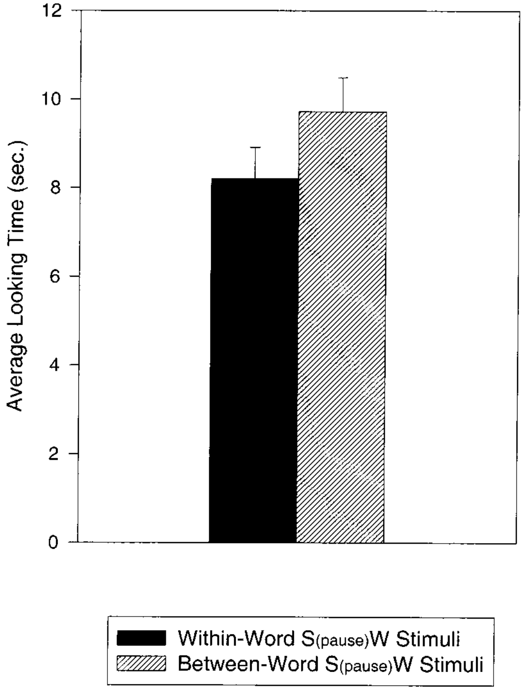

FIG. 3. Average listening times (and standard error bars) for the within-word and between-word strong/weak stimuli when a 500-ms silent pause is inserted between the syllables (Experiment 3).

bility between-word cluster than to those containing a high-probability within-word cluster. Thus, as expected, the insertion of an explicit break between the two syllables of the SW stimuli resulted in a reversal of the phonotactic effect observed in the uninterrupted version of these stimuli in Experiment 1. An analysis of variance combining the data from the present experiment with those from Experiment 1, with the between-subject factor Pause (SW-no-pause vs. SW-pause) and Phonotactic Structure (within-word vs. between-word) revealed a reliable interaction, $F(1,46) = 16.85, p < .001$. Furthermore, a similar analysis performed between the results of Experiments 2 and 3, combining the between-subject factor Type of Boundary Marker (stress vs. pause) and Phonotactic Structure (within-word vs. between-word), showed no reliable interaction, $F(1,46) < 1$. Stress in the second syllable or a cross-syllable interruption had the same effect on phonotactic preference.

To sum up, the presence of an explicit word-boundary marker such as a pause (Experiment 3) and the presence of a word-boundary cue such as stress (Experiment 2) had the same impact on the response to phonotactic constraints: They both generated a preference for high-probability *between-word* phonotactics. This result has two implications. First, it confirms the hypothesis that the listening difference between the within-word and between-word stimuli in Experiment 1 was indeed related to the perception of word boundaries. Had the preference for the within-word stimuli of Experiment 1 been unrelated to their coherence with perceived word structure, merely inserting a pause should not have reversed the effect. Second, the similarity between the results of Experiments 2 and 3 confirms the idea that strong syllables act as word boundary markers. The effect of stress was mimicked closely by that of inserting an explicit pause in the signal, and both manipulations were equally detrimental to the speech continuity induced by a within-word cluster at syllable juncture.

Thus far, the evidence suggests that 9-month-olds are sensitive to both prosodic and phonotactic cues to word boundaries. However, it does not necessarily mean that the two cues have equal contributions. One cue could carry more weight than the other. Therefore, it is possible that, although 9-month-olds are sensitive to both phonotactic and prosodic cues to word boundaries, they may rely more heavily on one of these sources than on the other.

EXPERIMENT 4

To test whether phonotactic or prosodic cues dominated in case of conflict, we presented 9-month-old infants with the within-word WS stimuli of Experiment 2 and the between-word SW stimuli of Experiment 1. These two sets of stimuli were chosen because they both contain conflicting information about potential word boundaries. For the within-word WS stimuli, phonotactics favor one-unit perception and prosody favors two-unit perception. For the between-word SW stimuli, phonotactics favor two-unit perception and prosody favors one-unit perception. From a perceptual point of view, there are possibly three ways to solve this conflict. If prosody is the driving cue, then, regardless of their phonotactics, SW stimuli should be preferred to WS stimuli. Alternatively, if phonotactic cues are primary, then, regardless of their prosody, stimuli with within-word clusters should be preferred. Finally, if phonotactics and prosody have equal weights, no consistent preference should be observed.

Method

Participants. The participants were 24 infants from monolingual American-English-speaking homes (12 males and 12 females), approximately 9 months of age (mean: 39 weeks, 4 days; range: 33 weeks, 2 days, to 42 weeks, 1 day). One additional infant was tested but discarded from the analyses because of restlessness.

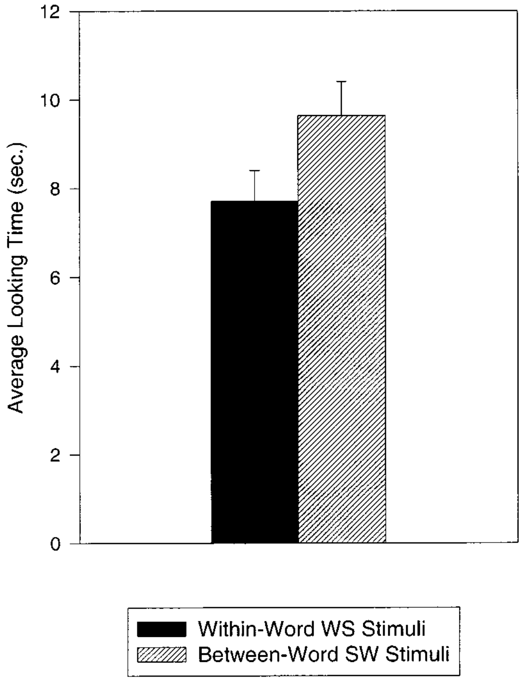

FIG. 4. Average listening times (and standard error bars) for the within-word weak/strong stimuli and the between-word strong/weak stimuli (Experiment 4).

Stimulus materials. The materials consisted of the within-word WS stimulus lists of Experiment 2 and the between-word SW stimulus lists (A and B) of Experiment 1.

Design, apparatus and Procedure. These were the same as in the previous experiments.

Results and Discussion

Mean listening times to the within-word WS stimuli and the between-word SW stimuli (collapsed across A and B) were calculated for each of the 24 infants. Across all participants, the average looking times, displayed in Fig. 4, were 7.70 s ($SD = 3.43$) for the within-word WS stimuli and 9.64 s ($SD = 3.78$) for the between-word SW stimuli. An analysis of variance performed on these scores revealed a significant difference between the two types of list, $F(1,23) = 8.88$, $p < .008$.

This difference in listening times indicates that, when phonotactics and prosody offer conflicting information about word boundaries, prosody overrides phonotactics. In other words, a spoken stimulus bearing the predominant word stress pattern but a low-probability within-word phonemic cluster is preferred to a stimulus bearing a less frequent word stress pattern but a high-probability within-word phonemic cluster. This result replicates and extends earlier reports that the prosodic bias toward trochaic words holds true even when these items are drawn from less common word patterns (Turk et al., 1995). In particular, the preference for strong/weak stimuli was obtained with materials that, if anything, minimized the salience of the prosodic variable (both syllables contained a full vowel) and maximized the phonotactic contrasts (the clusters were chosen to have sharp within- and between-word frequency differences). This indicates that, even in less than optimal conditions, the stress pattern of words constitutes a dominant segmentation cue in 9-month-olds.

The predominance of prosodic cues over phonotactic ones is at odds with a claim recently advanced by Cairns et al. (1997). On the basis of their simulation studies, they argued that prosodic effects could emerge from the discovery of phonotactic regularity. According to their simulations, segmentation attempts resulting from the assimilation of phonotactic regularities yielded performances ‘‘skewed towards successful detection of strong-initial words to a striking degree’’ (Cairns et al., 1997, p. 138). However, had prosody evolved as a byproduct of phonotactic computation for these English-learning 9-month-olds, a conflict between the two cues would likely have caused the phonotactic segmentation cue to override the prosodic one. Instead, the present results are best accommodated by models in which prosody is an initial cue yielding a coarse first pass at word boundaries that is subsequently supplemented with additional cues such as phonotactic and allophonic constraints (e.g., Jusczyk, Houston, & Goodman, 1998; Myers et al., 1996).

GENERAL DISCUSSION

The present study had two aims: First, we attempted to determine if infants are sensitive to probabilistic phonotactic patterns that may signal the presence of word boundaries. In particular, we examined whether infants are differentially sensitive to sequences of consonants that are typical between words compared to those sequences that are typical within words. Demonstrating such a sensitivity is a crucial first step in evaluating the hypothesis that infants exploit phonotactics in detecting word boundaries in fluent speech (see Cairns et al., 1997). Second, we were interested in how phonotactic and prosodic cues interact and which of the two sets of cues dominate when placed in conflict. The present study is unique in its combined focus on the potential utility of probabilistic phonotactics (i.e., legal sequences varying in frequency) and prosody in the segmentation of fluent speech by infants.

In Experiment 1, infants were found to listen significantly longer to stimuli with strong/weak stress patterns when these contained a cluster that occurs frequently within words in English. This finding suggests that English-learning 9-month-olds have discovered that certain sound sequences typically occur at word boundaries, while others are more likely to be found within words.

The goal of Experiment 2 was to investigate how the phonotactic cue for segmentation interacts with the tendency in 9-month-old infants to perceive strong syllables as word onsets. Experiment 2 was identical to Experiment 1 with the exception that the stimuli of Experiment 2 received primary stress in their second syllable. When tested for their listening preference for these weak/strong stimuli, 9-month-olds listened longer to the ones containing a high-frequency between-word cluster.

The hypothesis that infants responded to the word-boundary information contained in the phonotactic clusters was confirmed in Experiment 3, where the preference for within-word clusters in strong/weak stimuli was turned around in favor of between-word clusters as a result of inserting a 500-ms pause between the syllables. The ‘‘two-unit’’ structure of these stimuli was more consistent with the between-word than the within-word phonotactic sequences. Thus, the findings support the view that infants use phonotactic regularities to infer potential word boundaries.

Together, the results of the first three experiments indicate that, at 9 months, infants can rely on both phonotactic and prosodic regularities to compute the likely position of word boundaries. In particular, the occurrence of a strong syllable is associated with a preceding high-probability between-word phonotactic sequence, whereas an unstressed syllable is associated with a high-probability within-word phonotactic sequence.

Finally, Experiment 4 explored how English-learning 9-month-olds respond to conflicting phonotactic and prosodic cues to word boundaries. The results indicated that, although both phonotactics and prosody were used by the 9-month-olds, prosody was the predominant cue. Infants listened longer to strong/weak stimuli that violated phonotactic cohesion than to weak/strong stimuli that did not. Thus, at this age, items bearing the predominant stress pattern but less typical phonotactics won out over items with well-formed phonotactics but less typical prosody.

The present results show that, by 9 months of age, English learners have received enough exposure to their native language to discover that certain sequences of sounds typically occur between words and others within words. Critically, the results indicate that this effect holds true when the absolute frequency of occurrence, computed from a corpus of child-directed fluent speech (Bernstein, 1982), is controlled across all cluster types. The fact that infants are sensitive to between-word sequences suggests a sublexical, rather than purely lexical, basis for the effects of probabilistic phonotactics. If infants were merely drawing on information from stored sound patterns of isolated words, it is unlikely that they would demonstrate the sensitivity to between-word phonotactic probabilities observed in the present study. This argument does not preclude the possibility that the between-word effects we observed might emanate from representations of multiple word sequences stored in memory. However, the present data appear to conflict with an account of phonotactic knowledge that is based strictly on representations of isolated lexical items.

It is worth noting that a slightly different way to interpret the results observed in this study would be to posit that our data were obtained based solely on sensitivity to within-word clusters. According to this account, the 9-month-olds' patterns of preference would have emerged from their sensitivity to which phoneme sequences constitute an acceptable word rather than from their knowledge of how such sequences align with word boundaries. Although we cannot entirely discount the word well-formedness hypothesis, the design of the present experiments and the choice of the test clusters make this possibility unlikely. Sensitivity to only within-word cluster frequency would imply that the preference for between-word stimuli provoked by the insertion of a cross-syllable pause (Experiment 3) was caused by infants' knowing which phonemes are typically encountered at the end of monosyllables and which ones are encountered at the beginning. Using the stimulus set [ˈnɔŋ~k,θ] versus [ˈnɔŋ~tʌθ] and [ˈnɔm~kʌθ] as an example, infants would have had to base their preference on the knowledge that /m/ is a more plausible word-final phoneme than /ŋ/ and /t/ a more plausible word-initial phoneme than /k/. This possibility is improbable not only because it presupposes that infants have an extensive knowledge of words at 9 months but also, and more critically, because there is no obvious trace of such a distribution of consonants among the words of either the Bernstein corpus or the Webster's dictionary. Thus, our results, as a whole, are better accounted for by infants' sensitivity to biphone cooccurrences as they relate to word boundaries than by infants' sensitivity to segment position in words.[2]

The data also show that infants not only can draw on phonotactic regularities to detect potential word boundaries but also combine this sensitivity with their preference for initial-stress words. When a strong syllable is encountered in the signal, infants seem to treat it as a potential marker of a word onset. Hence, they tend to favor between-word phonotactic clusters at the onsets of strong syllables, but within-word clusters at the offsets of strong syllables when these are followed by weak syllables. Note that, contrary to conventional views of prosodic segmentation, which are based on a contrast between full and reduced syllables (e.g., Cutler & Norris, 1988; Jusczyk et al., 1993a; but see Turk et al., 1995), the prosodic effects in the present study involved an opposition between syllables that received primary stress and

---

[2] Similarly, we did not find evidence that infants paid more attention to either the final or the initial consonant of the monosyllabic components of our bisyllabic stimuli, which would have run against the hypothesis that infants are sensitive to the frequency of consonant *sequences.* When between-word stimuli were preferred over within-word stimuli (Experiments 2–4), the A and B between-word stimuli—which departed from the within-word stimuli by either the first or the second consonant—were not statistically different from each other (all interactions showed ps $>$ .20).

ones that could potentially receive secondary stress. This is because the latter were produced with a full rather than a reduced vowel. Thus, the current findings suggest not only that the 9-month-olds' perceptual sensitivity goes beyond simple metrical (i.e., full vs. reduced) distinctions in the signal but also that the fine-grained difference between primary and secondary stress enters in the computation of word boundaries (this idea is also present in recent work on adults by Mattys, in press; Mattys & Samuel, submitted for publication; Vroomen & de Gelder, submitted for publication).

The interaction between phonotactic structure and prosody found in Experiments 1–3 clearly demonstrates that, at 9 months, both cues are at work. This result is consistent with earlier findings showing that the age under study may be a critical age for cue integration (Lalonde & Werker, 1995; Morgan & Saffran, 1995). Thus, early on, infants seem to have the capacity to combine their sensitivity to probabilistic information from various sources, which has been shown to be a powerful problem-solving strategy in humans and animals (Kelly & Martin, 1994).

As noted by others, (e.g., Brent & Cartwright, 1996; Jusczyk et al., 1998; Morgan & Saffran, 1995; Myers et al., 1996), the integration of multiple segmentation cues is a necessary step in the course of language acquisition. No single cue could possibly account for the correct parsing of all English words because each cue fails in certain situations. For example, potential allophonic markers of word boundaries sometimes occur within words. Similarly, reliance on the occurrence of strong syllables to mark word onsets will result in missegmentations of words beginning with a weak syllable. Thus, to a certain extent, learning how to segment words is not so much the process of discovering word-boundary cues in the signal as it is the process of discovering how to integrate these cues successfully. Our results, along with previous findings (e.g., Lalonde & Werker, 1995; Morgan & Saffran, 1995), suggest that this skill is already in place by 9 months of age.

The theoretical relevance of a multiple-cue integration approach to segmentation has recently been underscored by Christiansen, Allen, and Seidenberg (1998) by means of a connectionist network. They trained their network on a corpus of child-directed speech in which the phonological, utterance boundary, and stress information had been explicitly marked. Considered individually, none of these sources of information provided reliable cues to word boundaries. However, when the sources were combined, the overall segmentation result was substantially improved. Christiansen et al. concluded that the integration of partial segmentation cues was more than just the sum of their individual contributions: The integration of cues was a cue in itself.

Despite the suggestion that segmentation greatly benefits from the conjunction of multiple cues, individual cues can sometimes provide relatively satisfying segmentation results as well (e.g., Brent & Cartwright, 1996; Cairns et al., 1997; Cutler & Carter, 1987). For instance, based on results with a neural network trained on a corpus of conversational English, Cairns et al. contend that sensitivity to the distribution of the phonotactic patterns of spoken English not only suffices to locate word boundaries, but also constitutes a viable method for bootstrapping other segmentation cues such as prosody. However, the present findings with English-learning infants cast doubt on the view that prosodic cues emerge from knowledge of phonotactics. Had a prosodic segmentation strategy been derived from phonotactics, it is not clear why the weak/strong stimuli in Experiment 2 should have reversed the direction of the phonotactic preference observed for the strong/weak stimuli of Experiment 1. Moreover, consider how infants responded when the two cues provided conflicting information about potential word boundaries in Experiment 4. If the use of stressed syllables to segment speech was derived from the computation of phonotactic regularities, then the situation in Experiment 4 should have favored a high-probability phonotactic choice and not a prosodic one. The pattern of results showed otherwise: The infants favored standard prosody over standard phonotactics.

However, our claims about the implications of our results for the developmental course of word boundary cues integration must be tempered with some caution. The present data only allow for speculations about how prosody and phonotactics combine before and after 9 months. For example, one could still argue that phonotactic sensitivity provides the grounds for prosodic sensitivity at an earlier age and that prosody rapidly takes over as the dominant cue. Nevertheless, independent studies of phonotactic and prosodic sensitivity in younger infants reveal that neither cue has significantly emerged by 6 months (Jusczyk et al., 1993a,b, 1994) and that prosodic sensitivity is in place at 7.5 months (Jusczyk et al., in press-b; Newsome & Jusczyk, 1995). Thus, for phonotactic sensitivity to be the initial source of word segmentation abilities, it would have to emerge between 6 and 7 months, then initiate prosodic sensitivity, and recede to a secondary cue by 9 months. Although possible, such a scenario seems unlikely to us.

A more plausible account of the present data and of former research is one in which the prosodic cues constitute a first-pass strategy that enables English-learners to begin to segment content words from fluent speech. Given the preponderance of English words with strong initial syllables, infants would succeed in recovering many words. Yet, other words would be missed or incorrectly segmented. Thus, clearly, English learners need to supplement stress cues with other potential cues to word boundaries. However, what an initial, prosodically based segmentation strategy provides is a way to ‘‘divide and conquer’’ the signal (Jusczyk, in press). Specifically, this coarse pass at word segmentation breaks the input into smaller chunks and gives the learners more opportunities to observe how potential phonotactic and allophonic cues are distributed within such chunks. For example, it may reveal which kinds of elements or sequences are typically found at the onsets or offsets of chunks. The discovery of such regularities could provide learners with other possible cues to the onsets and offsets of words in fluent speech. In the end, word segmentation would be multiply determined from these different sources of information. How and when infants learn to weight and integrate these different sources of information has yet to be established. The relationship between multiple cues presumably undergoes important remodeling as the infant's lexicon becomes larger and the perceptual and mnemonic strategies more sophisticated. Ultimately, as language learners' lexicons grow larger, the integration of these various cues could take the form of a Possible Word Constraint (PWC), as Norris, McQueen, Cutler, and Butterfield (1997) have recently suggested. The PWC holds that input should be segmented so as to produce strings of possible words. Thus, any parses that leave portions that are impossible words are disfavored. Future research could benefit from testing older infants and exploring how infants resolve conflicts among the potential cues, when the potential parses leave possible words and when they leave impossible words.

In conclusion, the data reported here provide information about how two categories of cues to English word boundaries, phonotactics and prosody, are used and combined. By 9 months, English-learners seem to have discovered how sequences are typically distributed both between and within words. They also show some sensitivity to how phonotactic sequences typically align with prosodic cues to word boundaries. The two cues in combination offer a potentially reliable segmentation tool and may constitute the conduit for increasingly more complex and efficient parsing strategies.

APPENDIX: Stimuli Used in Experiments 1–4

		Between-word Stimuli			
Within-word Stimuli		A		B	
ŋ·k	nɔŋ·kʌθ	**ŋ·t**	nɔŋ·tʌθ	**m·k**	nɔm·kʌθ
	tʃoŋ·kʌdʒ		tʃoŋ·tʌdʒ		tʃom·kʌdʒ
	dʒɑʊŋ·kʌθ		dʒɑʊŋ·tʌθ		dʒɑʊm·kʌθ
	fɑɪŋ·kʌv		fɑɪŋ·tʌv		fɑɪm·kʌv
	vʌŋ·kʌdʒ		vʌŋ·tʌdʒ		vʌm·kʌdʒ
	pɔɪŋ·kʌv		pɔɪŋ·tʌv		pɔɪm·kʌv
	bəʊŋ·kʌt		bəʊŋ·tʌt		bəʊm·kʌt
	gʊŋ·kʌv		gʊŋ·tʌv		gʊm·kʌv
	suŋ·kʌdʒ		suŋ·tʌdʒ		sum·kʌdʒ
	zɜŋ·kʌθ		zɜŋ·tʌθ		zɜm·kʌθ
f·t	tʃef·tʌs	**f·h**	tʃef·hʌs	**v·t**	tʃev·tʌs
	dʒeɪf·tʌz		dʒeɪf·hʌz		dʒeɪv·tʌz
	θæf·tʌdʒ		θæf·hʌdʒ		θæv·tʌdʒ
	mof·tʌθ		mof·hʌθ		mov·tʌθ
	kɑʊf·tʌp		kɑʊf·hʌp		kɑʊv·tʌp
	dɔɪf·tʌs		dɔɪf·hʌs		dɔɪv·tʌs
	sʊf·tʌz		sʊf·hʌz		sʊv·tʌz
	zuf·tʌdʒ		zuf·hʌdʒ		zuv·tʌdʒ
	ʃɜf·tʌθ		ʃɜf·hʌθ		ʃɜv·tʌθ
	tʃef·tædʒ		tʃef·hædʒ		tʃev·tædʒ
v·n	dʒev·nɔɪf	**v·m**	dʒev·mɔɪf	**z·n**	dʒez·nɔɪf
	θeɪv·nʌv		θeɪv·mʌv		θeɪz·nʌv
	meɪv·nɔɪn		meɪv·mɔɪn		meɪz·nɔɪn
	tɑʊv·nɔɪp		tɑʊv·mɔɪp		tɑʊz·nɔɪp
	kɑɪv·nʌz		kɑɪv·mʌz		kɑɪz·nʌz
	dɔv·nʌθ		dɔv·mʌθ		dɔz·nʌθ
	fɔv·nɔɪm		fɔv·mɔɪm		fɔz·nɔɪm
	θɔɪv·nʌp		θɔɪv·mʌp		θɔɪz·nʌp
	mʊv·nʌz		mʊv·mʌz		mʊz·nʌz
	nʊv·nɔɪʃ		nʊv·mɔɪʃ		nʊz·nɔɪʃ
m·θ	gim·θʌdʒ	**m·h**	gim·hʌdʒ	**n·θ**	gin·θʌdʒ
	zim·θɔɪf		zim·hɔɪf		zin·θɔɪf
	ʃɑʊm·θeɪv		ʃɑʊm·heɪv		ʃɑʊn·θeɪv
	tʃɑʊm·θʌp		tʃɑʊm·hʌp		tʃɑʊn·θʌp
	dʒɑʊm·θɔɪt		dʒɑʊm·hɔɪt		dʒɑʊn·θɔɪt
	fɔɪm·θʌs		fɔɪm·hʌs		fɔɪn·θʌs
	vɔɪm·θɔɪz		vɔɪm·hɔɪz		vɔɪn·θɔɪz
	pʊm·θʌdʒ		pʊm·hʌdʒ		pʊn·θʌdʒ
	tʊm·θɔɪf		tʊm·hɔɪf		tʊn·θɔɪf
	ʃum·θʌv		ʃum·hʌv		ʃun·θʌv

APPENDIX: Continued

		Between-word Stimuli			
Within-word Stimuli		A		B	
k·tʃ	gɪk·tʃeɪt	**k·ʃ**	gɪk·ʃeɪt	**p·tʃ**	gɪp·tʃeɪt
	zek·tʃɔɪb		zek·ʃɔɪb		zep·tʃɔɪb
	fæk·tʃeɪg		fæk·ʃeɪg		fæp·tʃeɪg
	θok·tʃʌs		θok·ʃʌs		θop·tʃʌs
	vɑʊk·tʃʌz		vɑʊk·ʃʌz		vɑʊp·tʃʌz
	nɔk·tʃeɪdʒ		nɔk·ʃeɪdʒ		nɔp·tʃeɪdʒ
	pɔɪk·tʃɔɪf		pɔɪk·ʃɔɪf		pɔɪp·tʃɔɪf
	tʊk·tʃeɪθ		tʊk·ʃeɪθ		tʊp·tʃeɪθ
	dʊk·tʃəʊv		dʊk·ʃəʊv		dʊp·tʃəʊv
	zuk·tʃeɪt		zuk·ʃeɪt		zup·tʃeɪt
ŋ·g	θɑʊŋ·gʌp	**ŋ·b**	θɑʊŋ·bʌp	**n·g**	θɑʊn·gʌp
	tʃɑʊŋ·gɔdʒ		tʃɑʊŋ·bɔdʒ		tʃɑʊn·gɔdʒ
	dʒɑɪŋ·gʌθ		dʒɑɪŋ·bʌθ		dʒɑɪn·gʌθ
	θɑɪŋ·gɔv		θɑɪŋ·bɔv		θɑɪn·gɔv
	vɔɪŋ·gʌp		vɔɪŋ·bʌp		vɔɪn·gʌp
	mɔɪŋ·gɔθ		mɔɪŋ·bɔθ		mɔɪn·gɔθ
	nʊŋ·gʌtʃ		nʊŋ·bʌtʃ		nʊn·gʌtʃ
	pʊŋ·gɔdʒ		pʊŋ·bɔdʒ		pʊn·gɔdʒ
	zuŋ·gʌθ		zuŋ·bʌθ		zun·gʌθ
	ʃuŋ·gɔv		ʃuŋ·bɔv		ʃun·gɔv

Note. For clarity, the stimuli are displayed according to cluster type. The test lists were assembled by grouping two stimuli of each relevant cluster type into 12-stimulus-long lists (see *Stimulus materials* under Experiment 1 for further details).

REFERENCES

Aslin, R. N. (1993). Segmentation of fluent speech into words: Learning models and the role of maternal input. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life.* Dordrecht: Kluwer.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development,* **52,** 1135–1145.

Auer, E. T. (1993). *Dynamic processing in spoken word recognition: The influence of paradigmatic and syntagmatic states.* Buffalo, NY: Univ. at Buffalo. [unpublished doctoral dissertation]

Bernstein, N. (1982). Acoustic study of mothers' speech to language-learning children: An analysis of vowel articulatory characteristics. *Dissertation Abstracts International,* **43-**04B. [No. AAI8220909]

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition,* **61,** 93–125.

Brown, R. W. & Hildum, D. C. (1956). Expectancy and the perception of syllables. *Language,* **32,** 411–419.

Cairns, P., Shillcock, R., Chater, N., & Levy, J. (1997). Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology,* **33,** 111–153.

Chang, H. W., & Trehub, S. E. (1977). Infants' perception of temporal grouping in auditory patterns. *Child Development,* **48,** 1666–1670.

Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes,* **13,** 221–268.

Church, K. (1987). Phonological parsing and lexical retrieval. *Cognition,* **25,** 53–69.

Cole, R. A., & Jakimik, J. (1980a). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech.* Hillsdale, NJ: Erlbaum.

Cole, R. A., & Jakimik, J. (1980b). How are syllables used to recognize words? *Journal of the Acoustical Society of America,* **67,** 965–970.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language,* **31,** 218–236.

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., & Norris, D. G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance,* **14,** 113–121.

de Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A cross-linguistic investigation of vowel formats in babbling. *Journal of Child Language,* **16,** 1–17.

de Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language,* **11,** 1–15.

de Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language,* **67,** 297–319.

de Boysson-Bardies, B., Vihman, M. M., Roug-Hellichius, L., Durand, C., Landberg, I., & Arao, F. (1992). Material evidence of selection from the target language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications.* Timonium, MD: York Press.

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature,* **266,** 718–719.

Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). Perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language,* **36,** 202–225.

Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception and Psychophysics,* **16,** 513–521.

Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-1] distinction by young infants. *Perception and Psychophysics,* **18,** 341–347.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science,* **171,** 303–306.

Eukel, B. (1980). Phonotactic basis for word frequency effects: Implications for lexical distance metrics. *Journal of the Acoustical Society of America,* **68,** s33.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development,* **8,** 181–195.

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception and Psychophysics,* **54,** 287–295.

Goodsitt, J. V., Morgan, J. L., & Kuhl, P. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language,* **20,** 229–252.

Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word,* **20,** 157–177.

Hohne, E. A., & Jusczyk, P. W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics,* **56,** 613–623.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.

Jusczyk, P. W. (1998). Using the headturn preference procedure to study language. In C. Rovee-Collier, L. P. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12). Stamford, CT: Ablex.

Jusczyk, P. W. (in press). Bootstrapping from the signal: Some further directions. In J. Weissenborn & B. Hoehle (Eds.), *How to get into language: Approaches to bootstrapping in early language development.* Amsterdam: Benjamins.

Jusczyk, P. W., Cutler, A., & Redanz, N. (1993a). Preference for the predominant stress patterns of English words. *Child Development,* **64,** 675–687.

Jusczyk, P. W., Friederici, A. D., Wessels, J., Svenkerud, V. Y., & Jusczyk, A. M. (1993b). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language,* **32,** 402–420.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992a). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology,* **24,** 252–293.

Jusczyk, P. W., & Hohne, E. A. (1997). Infants' memory for spoken words. *Science,* **27,** 1984–1986.

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (in press-a). An investigation of infants' use of allophonic cues to word boundaries in fluent speech contexts. *Perception and Psychophysics.*

Jusczyk, P. W., Houston, D., Goodman, M. (1998). Speech perception during the first year. In A. Slater (Ed.), *Perceptual development: Visual, auditory and language perception.* London: UCL Press.

Jusczyk, P. W., Houston, D., & Newsome, M. (in press-b). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology.*

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language,* **33,** 630–645.

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992b). Some consequences of stimulus variability on speech processing by 2-month old infants. *Cognition,* **43,** 253–291.

Jusczyk, P. W., & Thompson, E. J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. *Perception and Psychophysics,* **23,** 105–109.

Kelly, M. H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition and language. *Lingua,* **92,** 105–140.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development,* **18,** 111–116.

Klatt, D. H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R. A. Cole (Ed.), *Perception and production of fluent speech.* Hillsdale, NJ: Erlbaum.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America,* **66,** 1668–1679.

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development,* **5,** 263–285.

Kuhl, P. K. (1991). Human adults and human infants show a ‘‘perceptual magnet effect’’ for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics,* **50,** 93–107.

Kuhl, P. K., & Miller, J. D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. *Perception & Psychophysics,* **31,** 279–292.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science,* **255,** 606–608.

Lalonde, C. E., & Werker, J. F. (1995). Cognitive influences on cross-language speech perception in infants. *Infant Behavior and Development,* **18,** 459–475.

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology,* **20,** 215–225.

Levitt, A. G. (1993). The acquisition of prosody: Evidence from French- and English-learning infants. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Development neurocognition: Speech and face processing in the first year of life.* Dordrecht: Kluwer.

Levitt, A. G., Jusczyk, P. W., Murray, J., & Carden, G. (1988). The perception of place of articulation contrasts in voiced and voiceless fricatives by two-month-old infants. *Journal of Experimental Psychology: Human Perception and Performance,* **14,** 361–368.

Levitt, A. G., & Wang, Q. (1991). Evidence for language-specific rhythm influences in the reduplicative babbling of French- and English-learning infants. *Language and Speech,* **34,** 235–249.

Liberman, A. M., & Studdert-Kennedy, M. (1978). Phonetic perception. In R. Held, H. W. Leibowicz, & H. L. Teuber (Eds.), *Handbook of sensory physiology.* Berlin: Springer-Verlag.

MacWhinney, B. (1991). *The CHILDES project: Tools for analyzing talk.* Hillsdale, NJ: Erlbaum.

Mattys, S. L. (in press). The perception of primary and secondary stress in English. *Perception and Psychophysics.*

Mattys, S. L., & Samuel, A. H. Implications of stress pattern differences in spoken word recognition. [submitted for publication]

Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing.* Hillsdale, NJ: Erlbaum.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition,* **29,** 144–178.

Messer, S. (1967). Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior,* **6,** 609–613.

Morgan, J. L. (1994). Converging measures of speech segmentation in prelingual infants. *Infant Behavior and Development,* **17,** 387–400.

Morgan, J. L. (1996). A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language,* **35,** 666–688.

Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development,* **66,** 911–936.

Morrongiello, B. A. (1984). Auditory temporal pattern perception in 6- and 12-month-old infants. *Developmental Psychology,* **20,** 441–448.

Morse, P. A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology,* **13,** 477–492.

Myers, J., Jusczyk, P. W., Kemler Nelson, D. G., Charles-Luce, J., Woodward, A. L., & Hirsh-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language,* **23,** 1–30.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance,* **24,** 756–766.

Newsome, M., & Jusczyk, P. W. (1995). Do infants use stress as a cue for segmenting fluent speech? In D. McLaughlin & S. McEwen (Eds.), *19*th Annual Boston University Conference on Language Development, 2. Somerville, MA: Cascadilla Press.

Norris, D. G., McQueen, J., Cutler, A., & Butterfield, S. (1997). The possible word constraint in the segmentation of continuous speech. *Cognitive Psychology,* **34,** 191–243.

Pertz, D. L., & Bever, T. G. (1975). Sensitivity to phonological universals in children and adolescents. *Language,* **51,** 149–162.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language,* **39,** 347–370.

Polka, L., & Bohn, O.-S. (1996). Cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the American Society of America,* **100,** 577–592.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance,* **20,** 421–435.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science,* **274,** 1926–1928.

Spring, D. R., & Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research,* **20,** 224–232.

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature,* **259,** 39–41.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development,* **47,** 466–472.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structures. *Canadian Journal of Psychology,* **43,** 217–229.

Turk, A. E., Jusczyk, P. W., & Gerken, L. A. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech,* **38,** 143–158.

van de Weijer, J. (1998). *Language input for word discovery.* MPI Series Monographs, 9.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science,* **9,** 325–329.

Vitevitch, M., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language,* **40,** 374–408.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech,* **40,** 47–62.

Vroomen, J., & de Gelder, D. Stress is a cue to word onset: Stress based speech segmentation. [submitted for publication]

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental changes. *Developmental Psychology,* **24,** 672–683.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology,* **37,** 278–286.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development,* **7,** 49–63.

Woodward, J. Z., & Aslin, R. N. (1990). *Segmentation cues in maternal speech to infants.* Paper presented at the meeting of the International Conference on Infant Studies, Montreal, Québec.

Accepted February 1, 1999