Developmental Psychology
2003, Vol. 39, No. 4, 706–716

0012-1649/03/$12.00 DOI: 10.1037/0012-1649.39.4.706

When Cues Collide: Use of Stress and Statistical Cues to Word Boundaries by 7- to 9-Month-Old Infants

Erik D. Thiessen and Jenny R. Saffran
University of Wisconsin—Madison

Prior research suggests that stress cues are particularly important for English-hearing infants' detection of word boundaries. It is unclear, though, how infants learn to attend to stress as a cue to word segmentation. This series of experiments was designed to explore infants' attention to conflicting cues at different ages. Experiment 1 replicated previous findings: When stress and statistical cues indicated different word boundaries, 9-month-old infants used syllable stress as a cue to segmentation while ignoring statistical cues. However, in Experiment 2, 7-month-old infants attended more to statistical cues than to stress cues. These results raise the possibility that infants use their statistical learning abilities to locate words in speech and use those words to discover the regular pattern of stress cues in English. Infants at different ages may deploy different segmentation strategies as a function of their current linguistic experience.

To achieve mastery of their native language, infants must identify and learn words. Identifying words in an unfamiliar language is no simple task. Unlike the white spaces that mark the boundaries between words in a written text, speakers do not consistently place silent pauses between words when speaking (e.g., Cole & Jakimik, 1980). Further, because of the variation between languages, there is no infallible acoustic cue to word boundaries present in all languages (e.g., Cutler & Carter, 1987). Cross-linguistic variation suggests that infants must embark on the task of learning words equipped with one or more highly adaptable learning strategies. Despite the difficulty of identifying and learning words, infants begin to accomplish this task early in their lives; by 7.5 months of age, infants presented with a repetitive stream of speech are able to recognize some of the words that made up that stream when they are later presented in isolation (Jusczyk & Aslin, 1995).

One ability that infants may bring to bear on word segmentation is the statistical learning mechanism demonstrated by Saffran and colleagues (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996). In their experiments, 8-month-old infants were exposed to artificial languages consisting of trisyllabic nonsense words. These languages were synthesized so that there were no acoustic cues to word boundaries and no silences between syllables. However, infants were able to distinguish between words and nonwords on the basis of the level of statistical coherence between their syllables. Infants performed successfully on this task even when the nonwords were "part-words" consisting of the last syllable of one word and the first two syllables of another word. Aslin et al. (1998) later showed that this discrimination was not based simply on frequency (words occurred more frequently than part-words in their original experiment). Instead, infants discriminated on the basis of the differences in transitional probabilities—the odds that one syllable would follow another—between words and part-words (transitional probabilities are higher between syllables that are part of the same word because these syllables consistently occur together).

Much of the research assessing infants' ability to use statistical information to segment speech has, understandably, been carried out using speech that is stripped of all but statistical information (though see Goodsitt, Morgan, & Kuhl, 1993; Morgan & Saffran, 1995). This stripping has been a necessary step in the effort to establish that infants can segment speech on the basis of statistical cues alone. However, these experiments lack a certain degree of ecological validity. In an infant's natural environment, the speech stream contains multiple redundant cues to word boundaries, including not only statistical information but also stress cues, phonotactic rules, and allophonic cues (Jusczyk, 1999; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Hohne, & Bauman, 1999). Thus, the question of how statistical cues interact with language-specific acoustic cues to word boundaries has been left largely unresolved.

The acoustic cue that appears to have the most widespread and powerful influence on word segmentation in English is lexical stress. English is a predominantly *trochaic* language, meaning that words in English are usually stressed on their first syllables. In a corpus study, Cutler and Carter (1987) found that approximately 90% of English content words are stressed on their first syllables. Adult speakers of English appear to be biased to treat stressed syllables as a cue to word onset (Cutler & Norris, 1988), which suggests that adults use a metrical segmentation strategy when listening, assuming that a strong syllable signals the onset of a new word.

Stress cues may also play a role in infant word segmentation. Infants are clearly able to detect stress in fluent speech. Mehler et

Erik D. Thiessen and Jenny R. Saffran, Department of Psychology and Waisman Center, University of Wisconsin—Madison.

This research was supported by the Waisman Center, University of Wisconsin—Madison, and by National Institutes of Health Grant R01HD37466 and National Science Foundation Grant BCS-9983630.

We thank Susan Bunton, Michelle Loman, Joanna Miller, Aimee Niebuhr, Kari Reeck, Becky Seibel, and, especially, Rachel Robertson, Diana Wilson, and Kim Zinski for their assistance in conducting this research; the parents of the infants for their time and generosity; and Keith Kluender and Michael Kaschak for helpful suggestions.

Correspondence concerning this article should be addressed to Erik D. Thiessen, Department of Psychology, University of Wisconsin—Madison, Madison, Wisconsin 53706. E-mail: edthiessen@facstaff.wisc.edu

al. (1988) demonstrated that infants could use prosodic patterns to discriminate between their own language and a foreign language by 4 days of age. Jusczyk and Thompson (1978) showed that infants as young as 2 months dishabituate to a change in stress contour. By 9 months, English-hearing infants show a bias toward *trochees*, preferring to listen to words with a strong/weak stress pattern, as opposed to *iambs*, words with a weak/strong pattern (Jusczyk, Cutler, & Redanz, 1993). More recent research strongly suggests that infants not only are aware of the stress patterns in their native language but use stress cues to locate word boundaries in much the same ways that adults do. Like adults, English-hearing infants treat trochaic syllable sequences as more coherent than iambic sequences by 9 months (Echols, Crowhurst, & Childers, 1997; but see Morgan & Saffran, 1995) and are better at recognizing trochaic words heard in fluent speech than they are at recognizing iambic words by 7.5 months (Jusczyk, Houston, & Newsome, 1999).

Jusczyk, Houston, and Newsome (1999) have suggested that infants, like adults, use a metrical segmentation strategy, attending to strong/weak syllable sequences and extracting them from fluent speech. For example, when 7.5-month-old infants are presented with a passage that contains several repetitions of an iambic word consistently followed by the same unstressed monosyllabic word (such as the phrase "guitar is"), they will listen preferentially to the strong/weak sequence (e.g., "taris"). This suggests that they have incorrectly segmented the stressed (second) syllable of the iamb as a word onset and appended the monosyllable to it to create the preferred trochaic pattern. On the basis of these and other findings, Jusczyk (1999) has argued that rhythm is among the first cues to word boundaries available to infants.

The hypothesis that stress is the earliest cue used for word segmentation presents something of a chicken-and-egg problem. How can infants discover that words in English tend to be stressed on their first syllables if they are not familiar with any words? Infants must segment at least a few words in order to notice that words tend to be stressed on the first syllable. But how can they learn any words if stress patterns are their first cues to word boundaries? If stress is the first cue that infants use to segment words, then they cannot be learning about the stress pattern of English from words that they have segmented out of fluent speech. There are three possible ways around this problem:

1. Infants do not need to be familiar with any words to discover the stress pattern of English. For example, the trochaic bias could be innate. However, this is unlikely, given that many languages do not use trochaic word stress (Hyman, 1977).

2. Infants learn about English stress patterns from words heard in isolation, rather than words heard in fluent speech. Johnson and Jusczyk (2001) suggested that this is the strategy infants use to learn about the stress pattern of English.

3. Before infants come to depend primarily on stress cues for word segmentation, they use another strategy that allows them to isolate words from fluent speech and detect the regular rhythmic pattern of English from the example of these early words. For example, infants might be using statistical cues to segment their first words.

The second and third possibilities are the two with the most potential to allow young language learners to succeed in diverse linguistic environments. However, they make very different predictions about whether infants will attend to stress or statistical cues in a word segmentation task at an age when they are just beginning to segment words from fluent speech. If infants can learn about the predominant word stress of English from words heard in isolation, there is no need for statistical information to play a central role in infants' early word segmentation. This hypothesis suggests that the younger an infant is, the more strongly he or she will rely on stress cues to segment words from fluent speech. In this view, statistical cues play only a secondary role, being used when stress cues prove unreliable—as when listening to an unfamiliar language with a different stress pattern.

The suggestion that infants learn about the predominant stress of English from words that they hear in isolation, while plausible, is not without its difficulties. It is plausible because the diminutive terms that infants often hear in isolation adhere to the standard pattern of English intonation, for example "doggy," "kitty," and "birdy" (Jusczyk, Houston, & Newsome, 1999). Also, English names are likely to begin with a stressed syllable (Cutler, McQueen, & Robinson, 1990). Further, infants may learn words heard in isolation more easily than words heard in fluent speech (Brent & Siskind, 2001). However, on the occasions when infants hear a word in isolation, there is no guarantee they will treat it as a single word. An infant could represent a two-syllable utterance as two separate words, rather than one bisyllabic word. So infants must not only hear words in isolation to learn about the predominant stress pattern of English but must also treat these words as single units to discover the language's predominant rhythmic pattern.

Furthermore, although many of the utterances infants hear will begin with a stressed syllable, many will not. Sentences (at least those spoken to adults) are more likely to begin with an unstressed than a stressed syllable, because most determiners in English are unstressed, and many sentences begin with determiners (Cutler & Carter, 1987). Infant-directed speech may contain a higher ratio of utterances that begin with strong syllables owing to sentence-initial uses of the child's name or "wh-questions." Nevertheless, it is possible that infants could have trouble learning that a strong/weak rhythm is the preferred word stress of English. Even though most *words* that an infant hears begin with a stressed syllable, many *utterances* that an infant hears will begin with an unstressed syllable. Unless an infant has some way of separating single-word utterances from those that consist of more than one word, it could be difficult to learn that words ought to begin with a stressed syllable. Attending to other cues in early word segmentation, rather than stress cues alone, may help infants avoid this problem.

If stress is indeed the earliest cue that infants use to segment words, then infants ought to favor stress cues over statistical cues at the youngest age at which they can segment words from fluent speech. Obtaining results consistent with this hypothesis, Jusczyk, Houston, and Newsome (1999) found that 7.5-month-old infants used stress as a cue to word segmentation. These results are particularly significant because 7.5 months is the youngest age at which infants appear to be able to segment words from fluent speech (Jusczyk & Aslin, 1995). It is important to note, however, that stress cues were not placed in conflict with other cues in the Jusczyk, Houston, and Newsome experiments. Instead, their results suggest that when there are several candidate words with equally good internal statistics (e.g., the trochee "kingdom" vs. the iamb “guitar” vs. the nonword trochee “taris” that occurs when “guitar” is consistently followed by “is”), infants will segment the speech in a trochaic manner (e.g., segmenting “kingdom” and “taris” but not “guitar”).

Therefore, it is as yet unclear that stress is the earliest cue that infants use for word segmentation. It is possible that stress cues do not assume primary importance until after infants have been segmenting words from fluent speech for some time. If this is the case, there should be a young age at which infants will favor some other cue to word boundaries—possibly statistical information—over stress cues. Of course, it is quite possible that statistical cues are not the only cues that infants could use to learn about the relation between stress and word position. Infants may use a consortium of cues to learn about stress (e.g., words heard in isolation). Though this series of experiments focuses only on two cues—statistical information and stress cues—others should be explored as well in future research.

To distinguish between these two accounts of infants’ earliest word segmentation, it is necessary to give a word segmentation task to as young a group of infants as possible, preferably at an age when they have just begun to segment words from fluent speech. If stress is the first cue that infants use to segment words from fluent speech, then infants ought to weight this cue heavily at the youngest age that they demonstrate word segmentation. If, however, statistical information, or some other type of information, is the first cue that infants use to segment words from fluent speech, rhythmic cues should play a less important role in early word segmentation. Jusczyk and Aslin (1995) found that whereas 7.5-month-old infants were able to successfully segment words from fluent speech, 6-month-old infants were not, which suggests that infants between 6 and 7 months of age constitute a good candidate age group in which to explore these divergent predictions.

This age group is significantly younger than any group previously found to segment words from fluent speech (e.g., Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999). Past research with slightly older infants has suggested that infants will use stress as a cue to word segmentation when statistical cues are either absent or ambiguous; however, these cues have never been placed in conflict in studying a group of infants so young. The youngest age at which infants have been found to explicitly favor stress cues over statistical cues is 8 months, nearly 6 weeks older than this 6- to 7-month age group (Johnson & Jusczyk, 2001). Therefore, in order to gain a better understanding of how young infants weight these two cues, we need a word segmentation paradigm in which stress and statistical cues can be placed in conflict.

The statistical learning paradigm used by Saffran et al. (1996) provides an appropriate word segmentation task. Because the languages used in Saffran et al.’s experiments were synthesized, they provided no other cues to word boundaries except the statistical relations between syllables. Stress features can be edited into these stimuli, allowing for a direct comparison of the importance of stress cues and statistical cues to young word learners. Such a procedure would allow us to discern whether young infants favor stress cues to segmentation over statistical cues, as older infants do.

Although the age group of most interest here was 6- to 7-month-old infants, there was a difficulty that had to be overcome before their performance could be examined. It was necessary to establish that the synthesized stress cues being used in these experiments were reasonable approximations of natural stress cues. To do so, we needed to first compare infants’ responses to our synthesized speech with their response to natural stressed speech in a situation in which that response was predictable. Previous research using similar procedures but with naturally produced stimuli suggested that at 9 months of age, infants should mis-segment iambically stressed words in fluent speech (Echols et al., 1997; Johnson & Jusczyk, 2001). By observing how infants at this age reacted to the synthesized stimuli in Experiment 1, we could be more confident that infants would treat synthesized stress cues in the way that they treat natural stress cues at an age for which there was a clearly predicted response: Infants should favor stress cues to word boundaries over statistical cues. At a younger age, that may not be the case, an issue to which we return in Experiment 2.

Experiment 1

The purpose of this experiment was to determine whether 9-month-olds would preferentially attend to stress cues over statistical cues in a word segmentation task using synthesized speech, as predicted by research using natural speech (Johnson & Jusczyk, 2001). If infants at this age follow a metrical segmentation strategy rather than a statistical segmentation strategy, they should treat stressed syllables in fluent speech as word onsets regardless of the statistical properties of the input. If they instead favor statistical cues, infants should be able to correctly locate word onsets, whether they are stressed or unstressed, as long as statistical information identifying those onsets is consistently available. Presenting infants with fluent speech that is entirely iambic can test these contrasting predictions. That is, if an infant hears the iambically stressed two-word string “diTI#buGO” (with no pauses between the words), will they correctly treat “diTI” as a word, or will they incorrectly identify “TIbu” as a word? If infants depend mainly on stress cues, they should mis-segment words in iambically stressed speech, because the second syllable should erroneously be taken as a word onset. However, if statistical information is a more powerful indicator of word boundaries than are stress cues, 9-month-old infants should be able to segment the words in the speech stream correctly, as they are able to do in the absence of stress cues (Saffran et al., 1996).

In previous experiments using the headturn preference procedure, it has often been difficult to predict infants’ direction of preference a priori (e.g., Aslin, 2000). In some experiments, infants act as though they are bored with the words that they have segmented from speech and listen longer to novel words during test trials (e.g., Aslin et al., 1998; Echols et al., 1997; Saffran et al., 1996). In other experiments, infants prefer to listen longer to the familiar stimuli (the words they segmented from speech) during test trials (e.g., Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999). Because this experiment pits two segmentation cues against each other, infants might segment either words or part-words from fluent speech depending on the cue to which they choose to attend. Therefore, these results will be uninterpretable unless infants’ baseline direction of preference after learning is known. For example, imagine that the infants, after being familiarized with an iambic language, show a preference for part-words. Does this suggest that infants have incorrectly segmented the language and are showing a familiarity preference for the items that they have segmented? Or does this suggest that they have correctly segmented the language and are showing a novelty preference for the items that they (correctly) failed to segment from fluent speech? Either interpretation is plausible. Without knowing whether infants are expected to show a familiarity or a novelty preference given these materials, it is impossible to interpret a preference for words or part-words.

In order to resolve this difficulty, we included a separate group of infants exposed to the same words spoken with trochaic stress cues, for example, "DIti#BUgo#DApu#DObi." In a language with trochaic words, both stress cues and statistical cues point to the *same* word boundaries. Therefore, we can be certain which items infants are segmenting from fluent speech; because there are no conflicting cues in such a language, infants should successfully segment the words from fluent speech. If infants listen longer to part-words after hearing the trochaic language, we can say that infants at this age will exhibit a novelty preference given this task. If they instead listen longer to words, we can say that infants at this age will exhibit a familiarity preference in this task. With this baseline information available for comparison, it becomes possible to know whether infants correctly or incorrectly segment the speech stream after listening to the iambic language, in which stress cues are pitted against statistical cues.

Method

Participants. Participants were 31 infants ranging in age from 8.5 months to 9.0 months. The average age of the participants was 8.73 months. To obtain the 31 infants for this experiment, it was necessary to test 45. The other 14 were excluded for the following reasons: experimental error (3), refusing to look at sidelights (2), failure to complete at least 8 of 12 trials within 15 min (2), looking times of less than 3 s, on average, to either sidelight (2), crying (2), and parental interference (1). An additional subject (from the trochaic condition) was excluded for a looking-time preference greater than two standard deviations from the mean. A second infant (also from the trochaic condition), with the greatest preference in the opposite direction, was excluded to lessen the chance of bias. No other infants exhibited preferences greater than two standard deviations from the mean. Of the 31 participants included in the data analysis, 15 were randomly assigned to the trochaic condition, and 16 were randomly assigned to the iambic condition.

Stimuli: Acoustics. All stimuli were generated by the MacInTalk speech synthesizer running on a PowerMac 5300 (Apple Computer, Inc., Cupertino, CA). Two artificial languages were synthesized for use during familiarization, one iambic and one trochaic, each consisting of the same four bisyllabic nonsense words: *dapu*, *dobi*, *bugo*, and *diti*. Each language consisted of the same words spoken in the same randomized order. There were no pauses between words, and all syllables were fully coarticulated. The synthesizer produced syllables with a monotonic F0 (fundamental frequency) of 200 Hz. We used synthesized speech because it allowed better control over the acoustics of stress. Pitch contour can have unnatural variation when it occurs across spliced natural utterances. Therefore, even though synthesized speech is less like what an infant hears in the natural environment, it allows tighter control over the parameters of stress.

Both familiarization languages were 2 min 20 s in duration. Stress was synthesized by altering three parameters of the stimuli: vowel length, amplitude, and pitch (Lieberman, 1960). In the iambic language, each word was stressed on the second syllable, and in the trochaic language, each word was stressed on the first syllable. Thus, each infant heard a stream of speech like one of the following:

Iambic: "daPU#buGO#diTI#doBI"
Trochaic: "DApu#BUgo#DIti#DObi."

According to Crystal and House (1987), the ratio between stressed and unstressed syllable duration (both short and long vowels) is approximately 2:1. Most of the syllable lengthening that is due to stress occurs on the vowels. Although the consonants in stressed syllables do lengthen, the ratio between stressed and unstressed consonants is not as large; it is approximately 1.3:1 (Crystal & House, 1987). Because consonants are much shorter than vowels to begin with, the absolute effect of stress lengthening is considerably smaller for consonants than for vowels.

Vowels were lengthened to match Crystal and House's (1987) estimates of the ratio of stressed vowels to unstressed vowels (range = 1.8:1 to 2:1; M = 1.87:1). Consonants were not lengthened, because consonant lengthening could result in the extension of formant transitions, which could, in turn, make the consonants themselves more difficult to recognize. Thus, the stressed syllables were, on average, 310 ms long, and the unstressed syllables were all approximately 185 ms long. This ratio (1.67:1) is a close match with Crystal and House's (1990) report of the ratio of stressed consonant–vowel (CV) syllables to unstressed CV syllables in a fast talker (1.85:1).

Amplitude and fundamental frequency also increase in stressed syllables. Stressed syllables can have a peak amplitude between 4 and 8 dB higher than unstressed syllables (Bernstein-Ratner & Pye, 1984; Schwartz, Petinou, Goffman, Lazowski, & Cartusciello, 1995); the stressed syllables in this experiment were 4 dB louder than their unstressed counterparts. F0 values were based on the pitch contours of an adult female native speaker of English. Average pitch peak values varied from 255 to 270 Hz, depending on the vowel. The pitch contour, resynthesized in stressed syllables using Kay Elemetrics' Computer Speech Lab (Kay Elemetrics Corp., Lincoln Park, NJ), was somewhat different depending on whether the syllable began with a voiced or an unvoiced consonant. In the case of a voiced consonant, the pitch contour was a roughly inverted parabola, whereas in the case of voiceless consonants, the pitch contour described a falling plateau. This is because during the onset of a voiceless consonant, there are no glottal pulses, and thus no F0 contour (Stevens, 1998). Therefore, when voicing began in the syllable after a voiceless consonant, the value of the F0 was roughly where it would have been had the consonant been voiced, rather than starting at the ambient value (200 Hz). The beginning part of the pitch contour parabola was, in these cases, simply cut off. Unstressed syllables remained at the monotonic pitch of 200 Hz at which they were synthesized.

Because English-hearing infants prefer to listen longer to trochees than to iambs (Jusczyk, Cutler, & Redanz, 1993), all items were presented with neutral stress during test trials. This mismatch between the articulations of the test items from the familiarization session to the test trials likely made the task more difficult for infants (Hunter & Ames, 1988). However, this design feature was necessary because if infants had heard the test items articulated in the same way during both the test trials and the familiarization period (i.e., with stress cues), they would likely have preferred to listen to the trochees whether or not they had segmented these items from speech, and this would have made interpretation difficult.

Stimuli: Statistical structure. Both the trochaic and the iambic languages contained identical statistical cues to word boundaries. The transitional probability between syllables within a word was 1.0, whereas the transitional probabilities across word boundaries ranged between 0.2 and 0.4. However, the relationships between stress cues and statistical cues were critically different in the two languages. In the trochaic language, both stress cues and statistical cues indicated the same word boundaries. As such, infants should have had no difficulty correctly segmenting words when listening to the trochaic language. But in the iambic language, stress cues and statistical cues conflicted. If infants attended primarily to statistical cues, it was possible to segment the language correctly. However, if infants attended to stress cues, they would likely mis-segment the words in this language. Because these words were stressed on their second syllables, infants using a metrical segmentation strategy would treat the second syllable of a word as the word onset. Therefore, they would assume that a part-word like "PUbu" (actually the second syllable of "daPU" and the first syllable of "buGO") was a word, as opposed to a syllable sequence that crossed word boundaries.

An interpretational difficulty that sometimes arises when testing infants' discrimination between part-words and whole words is that words occur more frequently than part-words. This is because after a word such as *dapu*, any of the other words in the language (here, *bugo*, *diti*, or *dobi*) can occur. Each of them will occur, on average, one third of the time after *dapu*. Therefore, the part-word formed across the boundaries of words occurs only one third as often as whole words.

Differences in test item frequency present a problem for an experiment such as the current one, because they add another factor that could influence infant preference. In order to remove this potential confound, we constructed the two familiarization languages (iambic and trochaic) so that two of the words—*dapu* and *dobi*—occurred twice as often (90 times) as the other two—*diti* and *bugo*—which occurred only 45 times each. Therefore, the part-words formed across the boundaries of the two frequent words (*dapu* and *dobi*) occurred 45 times each, just as often as the infrequent whole words (see Aslin et al., 1998, for further discussion of the statistical features of this type of language). The two infrequent words (*diti* and *bugo*) and the two part-words formed across the boundaries of the frequent words (*pudo* and *bida*) served as the test items in these experiments.

Procedure. Infants were tested individually in a double-walled sound-attenuated room while seated on a parent's lap. An experimenter outside the booth observed the infants' looking behavior on a video monitor connected to an infrared camera inside the room and coded the direction of the infants' gaze on-line. The parent inside the room listened to masking music to eliminate bias, and the observer was similarly unable to hear the stimulus being played to the infant.

At the beginning of the familiarization phase, a light in the center of the wall facing the infant began to flash, directing the infant's gaze forward. Simultaneously, one of the two languages (either iambic or trochaic; each infant heard only one language) began to play from the speakers beneath the two sidelights—one light and speaker on each sidewall—in the room. The familiarization phase lasted 2 min 20 s.

Immediately after familiarization, 12 test trials were presented. All infants heard the same test trials regardless of familiarization condition. Six of these trials were part-word trials, and six were word trials. Each test item occurred on three trials during the testing session. A test trial began with the blinking light at the center of the wall facing the infant drawing the infant's gaze forward. When the observer signaled the computer that the infant had fixated on the center light, one of the sidelights began to flash, and the center light simultaneously stopped. As soon as the infant made a head turn of at least 30° in the direction of the flashing sidelight, the experimenter signaled the computer, and one of the test items was presented from the speaker beneath the flashing light. Test items were presented in random order, with six trials (three words and three part-words) presented from each side speaker. The test item continued to play for as long as the infant continued to look at the flashing sidelight. When the infant looked away for more than 2 s, the test item stopped playing, and the center light began to blink again. This procedure was repeated until the infant had completed all 12 test trials.

Results and Discussion

First, we compared listening times to words and part-words for infants exposed to the trochaic language. As shown in Figure 1, infants listened to words for 7.03 s ($SE = 0.34$) during the test trials and to part-words for 6.43 s ($SE = 0.36$). Twelve of the 15 infants listened longer to words than to part-words during the test trials after exposure to the trochaic language. A paired *t* test (all *t* tests reported are two-tailed) indicated that the difference in looking times between words and part-words was significant, $t(14) = 2.47$, $p < .05$.

Second, we compared listening times to words and part-words for infants exposed to the iambic language. As shown in Figure 1, these infants listened to words for 7.73 s ($SE = 0.51$) and to part-words for 8.92 s ($SE = 0.59$). Thirteen of the 16 infants listened longer to part-words than to words during the test trials after exposure to the iambic language. A paired *t* test indicated that

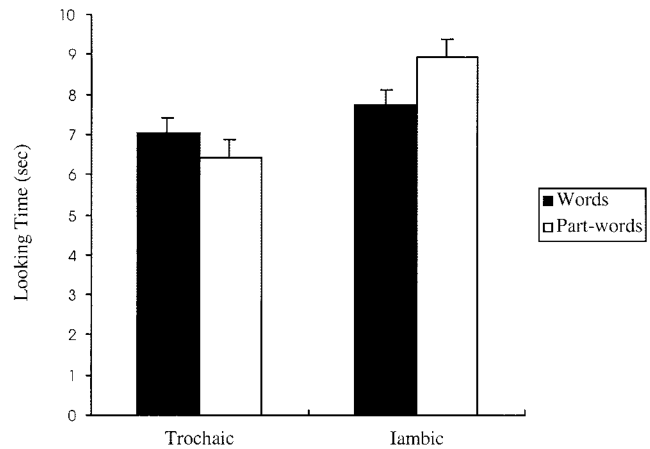

Figure 1. Nine-month-old infants' looking times to words and part-words in the trochaic and iambic conditions of Experiment 1. Vertical lines depict standard errors of the means.

the difference in looking times between words and part-words was significant, $t(15) = 3.07$, $p < .05$.

After listening to the trochaic language, infants listened longer to words than to part-words. In contrast, infants listened longer to part-words than to words after listening to the iambic language. To assess whether this difference in direction of preference was significant, we performed a 2 × 2 analysis of variance (ANOVA; Item × Condition). The main effect for item (word vs. part-word) was not significant, $F(1, 58) = 0.41$, $p = .53$. There was a significant main effect for condition (trochaic vs. iambic), $F(1, 58) = 12.53$, $p < .05$, that was due to the fact that infants listened longer to test items after the iambic familiarization than after the trochaic familiarization. Most important, the significant interaction between item and condition, $F(1, 58) = 4.01$, $p < .05$, indicates that infants exhibited a different direction of preference in the two conditions (infants in the trochaic condition listened longer to words, whereas infants in the iambic condition listened longer to part-words).

Nine-month-old infants showed a familiarity preference for words over part-words after listening to the trochaic language. This finding establishes that infants are able to perform a word segmentation task even when the language they must segment contains synthesized stress cues. However, infants showed a different pattern of responses after listening to the iambic language, listening longer to the part-words. Assuming that infants continue to show the same direction of preference as they did after hearing the trochaic language—listening longer to the items that they have segmented as "words"—these results indicate that infants in the iambic language condition are incorrectly treating part-words as words. Such a pattern of mis-segmentation would only occur if infants were being consistently misled by the stress cues of the iambic language. Recall that infants in the iambic condition heard syllable strings such as "diTI#buGO#doBI#daPU" and were tested on part-words such as *tibu*. Although this test item actually crosses a word boundary, its first syllable was stressed when infants heard it during the familiarization period. Thus, the finding that infants listened longer to part-words after hearing the iambic language suggests that they were segmenting part-words as words. At 9 months of age, when stress and statistical cues conflict, infants weight stress cues more heavily than statistical cues. For this age group, stress cues are a more powerful index of word boundaries.

This finding replicates Johnson and Jusczyk's (2001) results with 8-month-old infants: When stress cues and statistical cues conflict, infants at this age attend preferentially to stress cues. Interestingly enough, participants in our experiment showed a familiarity preference, whereas participants in Johnson and Jusczyk's experiment showed a novelty preference. In order to better interpret this difference, it is necessary to have an understanding of the processes that drive infant preferential behavior. According to Hunter and Ames's (1988) model of infant preferential responses, infant preferences progress on a continuum from familiarity to novelty. Early in the processing of a stimulus, infants are likely to show a familiarity preference. As they habituate, they become less interested in the stimulus and will show a novelty preference. A number of factors affect the amount of time an infant requires to habituate: individual processing differences, age (older infants habituate more quickly), and task difficulty (which can be defined by a number of parameters, such as the amount of time an infant is exposed to the stimulus). As a general rule, the more difficult a task is, the longer an infant requires to habituate and the more likely an infant is to show a familiarity preference.

There are several procedural differences between our experiment and Johnson and Jusczyk's (2001). Although the words were longer in Johnson and Jusczyk's experiment, a factor that may have made their task more difficult, a number of other factors likely make our learning task more difficult. Johnson and Jusczyk's stimuli were articulated somewhat more slowly than ours. In addition, their materials contained less of a mismatch between familiarization and test items than ours did, because their stressed syllables were acoustically more similar to their unstressed syllables than was the case in the current experiment (our stressed syllables were 125 ms longer and 70 Hz higher pitched than our unstressed syllables, whereas Johnson and Jusczyk's stressed syllables were 40 ms longer and 20 Hz higher pitched than their unstressed syllables). Words in our stimuli occurred such that our test items were matched for frequency, whereas in the Johnson and Jusczyk experiment the test words occurred three times more often than part-words. Finally, Johnson and Jusczyk used natural speech. All of these differences may have led Johnson and Jusczyk's task to be somewhat easier than ours and thus more likely to elicit a novelty preference. Regardless of the procedural differences between the two experiments, the most important point is that this experiment provides a confirmation of Johnson and Jusczyk's finding that 8-month-old infants depend more on stress cues than statistical cues to word segmentation. When stress cues were concordant with statistical cues, 8.5- to 9.0-month-old infants were able to segment words from fluent speech correctly. But when the stress cues and statistical cues were discordant, infants incorrectly segmented part-words from speech.

It is important to note that such an interpretation depends on the assumption that infants show the same direction of preference after listening to both the trochaic and iambic languages. Is it possible that infants correctly segmented the words in both languages but simply showed a different direction of preference after exposure to each, with a familiarity preference after hearing the trochaic language and a novelty preference after listening to the iambic language?

That explanation, although it could serve to explain the findings in this particular experiment, is inconsistent with what is known about infant preferential responses. Recall that infants are more likely to show a familiarity preference when the task is difficult (Hunter & Ames, 1988). Of the two word segmentation tasks infants were faced with in Experiment 1, the iambic condition task was the more difficult. In the trochaic language, there were two cues to help infants locate word boundaries, statistics and word stress. Because these two cues conflicted in the iambic language, it ought to have been more difficult to segment. Even if infants completely ignored statistical cues and attended only to stress, the iambic language should not have been any easier to segment than the trochaic language, because the trochaic language also contained stress cues. Therefore, if there were any language likely to elicit a novelty preference, it would be the trochaic language, because infants should find that language easier to segment.

We can safely assume, then, that if infants show a familiarity preference after segmenting the easier trochaic language, they are still showing a familiarity preference after segmenting the more difficult iambic language. Their preference for different test items arises from the fact that they are segmenting the iambic language differently than they are segmenting the trochaic language. The fact that the only difference between these two languages is the placement of stressed syllables suggests that infants are relying on stress cues for word segmentation. Our results are consistent with evidence that 9-month-old infants know a great deal about the prosodic regularities of their language (e.g., Jusczyk, Cutler, & Redanz, 1993) and are also consistent with word segmentation data reported by Johnson and Jusczyk (2001) with different stimuli. The consistency between our results and those of previous studies that used natural speech suggests that even though our language was synthesized, infants treated it in much the same way that they treat natural speech. The suggestion that infants will treat synthesized speech similarly to the way they treat natural speech is also consistent with the results of Saffran (2001): Infants given synthetic speech stimuli responded to the words that they segmented from such speech as potential English words, ready to be integrated into the native language.

These findings leave a question unanswered: How do infants learn that words in English tend to be stressed on their first syllables? One possibility is that infants first segment words from the speech stream through the use of statistical cues. Once they have segmented a sufficient number of words to begin to detect prosodic regularities, they would notice that words tend to be stressed on the first syllable. This would allow infants to use stress cues and eventually to attend to them more than statistical cues, as do 9-month-olds.

If the hypothesis that infants first attempt to segment words using statistical cues is correct, there should be an age younger than 9 months at which infants weight statistical cues to word boundaries more heavily than stress cues. If infants are not aware that stress is a cue to word boundaries, or if they are aware that stress is correlated with word boundaries, but do not think it is as important a cue as statistical information, they should segment the same words from the iambic language as from the trochaic language. Experiment 2 was intended to determine whether infants between 6.5 and 7.0 months of age segment words from fluent speech by weighting statistical cues more heavily than stress cues. In this experiment, infants were exposed to the same languages as in Experiment 1.

Experiment 2

Method

Participants. Participants were 30 infants ranging in age from 6.5 months to 7.0 months. The average age of the participants was 6.86

months. To obtain the 30 infants for this experiment, it was necessary to test 44. The other 14 were excluded for the following reasons: crying (5), refusal to look to the sidelights (2), failure to complete at least 8 of 12 trials within 15 min (2), missing more than one trial of the same test item (2), ear infections (1), experimental error (1), and looking for less than 3 s, on average, to one of the sides of the room (1). No infants exhibited listening preferences greater than two standard deviations from the mean. Of the 30 participants who were included in the data analysis, 15 were randomly assigned to the trochaic condition, and 15 were randomly assigned to the iambic condition.

Stimuli. The stimuli were identical to those used in Experiment 1.

Procedure. The procedure was identical to that used in Experiment 1.

Results and Discussion

First, we compared listening times to words and to part-words for infants exposed to the trochaic language. As shown in Figure 2, these infants listened to words for 5.9 s ($SE = 0.60$) during the test trials and to part-words for 7.4 s ($SE = 0.78$). Twelve of the 15 infants listened longer to part-words than to words during the test trials after exposure to the trochaic language. A paired t test indicated that the difference in looking times between words and part-words was significant, $t(14) = 2.97$, $p < .05$.

Second, we compared listening times to words and to part-words for infants exposed to the iambic language. As shown in Figure 2, these infants listened to words for 7.22 s ($SE = 0.38$) during the test trials and to part-words for 8.89 s ($SE = 0.52$). Thirteen of the 15 infants listened longer to part-words than to words during the test trials after exposure to the iambic language. A paired t test indicated that the difference in looking times between words and part-words was significant, $t(14) = 3.83$, $p < .05$.

After listening to the trochaic language, infants listened longer to part-words than to words. Similarly, after listening to the iambic language, infants listened longer to part-words than to words. In order to ask whether infants were showing different patterns of preference across the two conditions, we conducted a 2 × 2 (Item × Condition) ANOVA. There was a significant main effect of item (words vs. part-words), $F(1, 56) = 7.46$, $p < .05$, that was due to the fact that, across conditions, infants listened longer to part-words than to words. There was also a significant main effect of condition, $F(1, 56) = 5.49$, $p < .05$, because infants listened longer to test items after the iambic familiarization than after the trochaic familiarization. However, unlike the pattern of results for the 9-month-olds, the interaction between item and condition was not significant, $F(1, 56) = 0.01$, $p = .91$. These results suggest that infants were segmenting the same items from the iambic language as they were from the trochaic language, as would be expected if 7-month-old infants were weighting statistical cues for word segmentation more heavily than stress cues. Because the statistical cues to word boundaries were identical in both the trochaic and the iambic languages, infants should have segmented the same items across conditions if they were attending primarily to statistical cues.

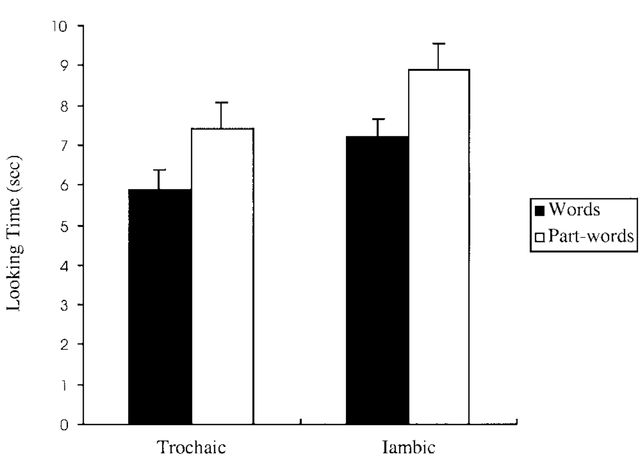

Figure 2. Seven-month-old infants' looking times to words and part-words in the trochaic and iambic conditions of Experiment 2. Vertical lines depict standard errors of the means.

Cross-age comparison. At these two different ages, 6.5–7.0 months and 8.5–9.0 months, infants show very different patterns of listening behavior. At 9 months, after listening to the iambic language, infants show a direction of preference opposite to the one they show after listening to the trochaic language. But at 7 months, after listening to both languages, infants show identical preferences. This pattern of results suggests that the younger infants are achieving the same segmentation results after listening to both the iambic and the trochaic languages. In order to be certain that infants were performing differently in these two experiments, we conducted a cross-age ANOVA on the preference scores of infants in Experiments 1 and 2. We generated preference scores by subtracting each infant's average looking time to part-words from his or her average looking time to words, so that an infant who listened to words for 4.8 s and to part-words for 4.1 s would have a preference score of 0.7 s. As shown in Figure 3, 9-month-olds had, on average, a positive preference score after listening to the trochaic language and a negative preference score after listening to the iambic language. Conversely, 7-month-old infants had a negative preference score after listening to both languages. A 2 × 2 (Age × Language) ANOVA was performed to determine whether these preference patterns across age were significantly different. There was a significant main effect for age, $F(1, 57) = 10.38$, $p < .05$. There was also a significant main effect for language, $F(1, 57) = 5.41$, $p < .05$. Most important, there was a significant interaction between condition and age, $F(1, 57) = 4.08$, $p < .05$. This indicates that the pattern of preference of 7-month-old infants (negative preference scores after hearing either language) was significantly different than the pattern of preference of 9-month-old infants (positive preference scores in the trochaic condition, negative preference scores in the iambic condition).

The results of Experiments 1 and 2 suggest that 6.5- to 7.0-month-old infants were weighting statistical cues more heavily than stress cues—though it is quite possible that they could have used stress as a cue to word segmentation if it did not conflict with statistics—whereas 8.5- to 9.0-month-old infants attended more to stress as a cue to word boundaries. But in order to interpret the results of these experiments with more certainty, we need a better understanding of the basic processes driving infant direction of preference. After exposure to the trochaic language (in which infants should be segmenting the correct words), 9-month-olds showed a familiarity preference, whereas 7-month-olds showed a novelty preference. This finding is somewhat surprising. In general, when two groups of infants at different ages are exposed to the same set of stimuli, one expects the older infants to find the task easier, and thus it should be the older infants who show a novelty preference (Hunter & Ames, 1988). In this set of experi-

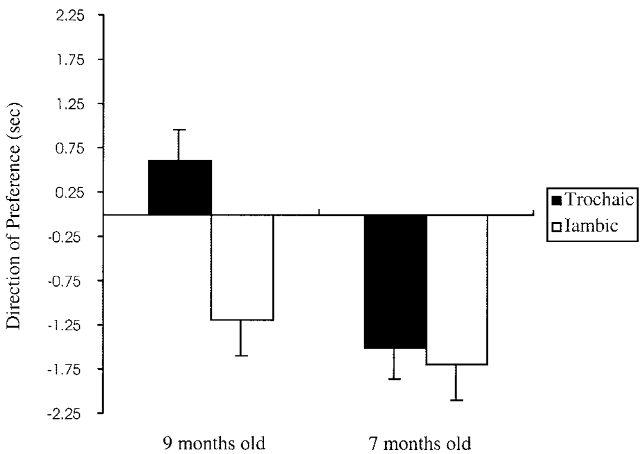

Figure 3. Direction-of-preference scores for 7- and 9-month-old infants for both iambic and trochaic languages. Vertical lines depict standard errors of the means.

ments, precisely the reverse result was obtained, which suggests that the older infants found the task more difficult.

One possible explanation for this counterintuitive finding is that although the infants were familiarized with nonsense words that have a naturalistic stress pattern, they were tested on words with a monotonic pitch contour. If 9-month-olds are more attentive to stress cues than are 7-month-olds, the mismatch between familiarization items and test items might have been more salient to the older infants. Because a mismatch between test items and familiarization items is one of the factors that determine task difficulty (Hunter & Ames, 1988), this could explain why 9-month-olds appeared to find the segmentation task more difficult than the 7-month-olds. Their attentiveness to stress cues could have made it more difficult for them to match the words they had segmented with the test items.

If this hypothesis is correct, 9-month-olds should show a novelty preference when the mismatch between familiarization items and test items is removed. Experiment 3 was designed to test the mismatch hypothesis by presenting 9-month-old infants with the same language that was used in Experiments 1 and 2 but with it synthesized in a monotone. This removed the mismatch between familiarization items and test items. If the mismatch was responsible for 9-month-olds showing a familiarity preference in Experiment 1, they should show a novelty preference in Experiment 3.

Experiment 3

Method

Participants. Participants were 13 infants ranging in age from 8.5 months to 9.0 months. The average age of the participants was 8.83 months. To obtain the 13 infants for this experiment, it was necessary to test 26.[1] The other 13 were excluded for the following reasons: crying (5), looking times of less than 3 s to one side (4), and failing to complete at least two trials of all four test items (4). No infants were excluded because of listening preferences that were greater than two standard deviations from the mean.

Stimuli. The familiarization language had the same word order and statistical structure as the languages used in Experiments 1 and 2. However, no stress cues were present. All words were synthesized monotonically with the MacInTalk speech synthesis program. Average syllable length was 223.8 ms. The duration of the language was 2 min. The test items were identical to those used in Experiments 1 and 2.

Procedure. The procedure was identical to that used in Experiments 1 and 2.

Results and Discussion

As shown in Figure 4, infants listened to words for 5.45 s ($SE = 0.44$) during the test trials and to part-words for 6.35 s ($SE = 0.36$). Eleven of the 13 infants listened longer to part-words than to words during the test trials after exposure to the familiarization language. A paired t test indicated that the difference in looking times between words and part-words was significant, $t(12) = 3.37$, $p < .05$.

After familiarization with a language that contained no stress cues (i.e., one that closely matched the test items used in these experiments), 9-month-old infants showed a novelty preference. In contrast, 9-month-old infants showed a familiarity preference when exposed to a language that contained stress cues. We performed a 2 × 2 (Experiment × Item) ANOVA that included infants in Experiment 3 and infants in the trochaic condition of Experiment 1 to determine whether this change in direction of preference was significant. There was a significant main effect of experiment, $F(1, 52) = 4.81$, $p < .05$, because infants in Experiment 1 had longer overall listening times than infants in Experiment 3. There was no significant main effect of item, $F(1, 52) = 1.41$, $p = .71$. Most important, there was a significant interaction between experiment and item, $F(1, 52) = 4.05$, $p < .05$. The interaction indicates that infants in the two experiments exhibited different patterns of behavior. After exposure to the trochaic language in Experiment 1, infants preferred to listen to the words, whereas after exposure to the monotonic familiarization language of Experiment 3, infants preferred to listen to the part-words.

As predicted, 9-month-old infants in this experiment—with no stressed syllables present in either the familiarization or test stimuli—showed a novelty preference, just as 7-month-old infants in Experiment 2 did. This is consistent with the hypothesis that the 9-month-olds' familiarity preference in Experiment 1 was due to difficulty in making the match between the stressed familiarization syllables and the monotonic test syllables: Once that mismatch was removed by making the familiarization syllables monotonic, 9-month-old infants showed a novelty preference.

An interesting pattern emerged when we compared the data from Experiments 1, 2, and 3. In both Experiments 1 and 2, infants listened longer to test items (averaged across words and part-words) after hearing the iambic language than after hearing the trochaic language. And at 9 months, infants had longer overall looking times after hearing the trochaic language (Experiment 1) than after hearing the monotonic language (Experiment 3). Why might infants show different overall levels of interest in the test stimuli after each of the three types of familiarization experiences we used? One possible explanation hinges on the differing levels of discontinuity between the three familiarization languages and the test items, which were the same for all three experiments. In the monotonic language, there were no stress cues, and only one

[1] The attrition rate in this experiment was higher than in the previous two. This may have been because these stimuli were less like natural speech and thus less interesting.

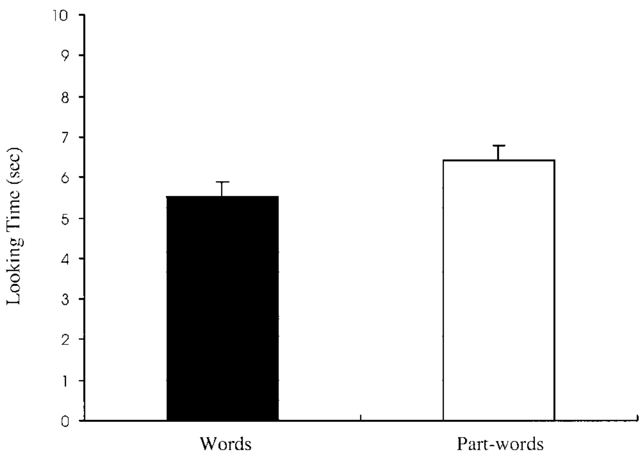

Figure 4. Nine-month-old infants' looking time to words and part-words in Experiment 3. Vertical lines depict standard errors of the means.

set of word boundaries was indicated. In the trochaic language, stress cues were present, but still only one set of word boundaries was indicated (by both stress and statistical cues). Stress cues were also present in the iambic language, but in this case two separate sets of word boundaries were indicated; stress and statistical cues highlighted a different set of syllables as word onsets. In comparison, the test items never contained stress cues and always indicated only one set of word onsets (via statistical information). As such, the monotonic language was most similar to the test items, the trochaic language was moderately novel, and the iambic language was maximally dissimilar to the test items.

It is possible that the levels of discontinuity between the three familiarization languages and the test items account for the differences in overall looking time during the test period across the three experiments.[2] Infants appeared to be listening longer when the test items were the least similar to what they had heard during familiarization. This does not mean that infants at all ages were showing an overall novelty preference, given that infants in Experiment 1 showed a familiarity preference for words over part-words. Instead, this pattern of results suggests that the characteristics of the familiarization language affected the degree to which infants were interested in the test items. If this is the case, then even the 7-month-olds must have considered the iambic familiarization to be less similar to the test items than was the trochaic familiarization. For this to be so, the infants must have recognized that stress cues indicated different word boundaries than did statistical cues in the iambic language.

If this explanation of infants' overall looking times is correct, then even 7-month-olds must have been weakly aware of the correlation between stress and word onsets, even though they did not follow the stress cues when they were placed in conflict with the statistical cues in Experiment 2. That is, it is conceivable that infants would have used stress as a cue to segmentation if statistical cues had not conflicted with stress cues. The possibility that even young infants are sensitive to stress, even if they do not use prosodic markers of word boundaries like older infants, is consistent with other findings concerning the awareness of prosodic contours in young infants (e.g., Jusczyk & Thompson, 1978). In addition, younger infants do not rely exclusively on trochees as a cue to word segmentation, even though they are likely to be more familiar with trochees than iambs (e.g., Morgan, 1996; Morgan & Saffran, 1995). Our results, along with others in the literature, support the hypothesis that there is a developmental progression between the time when infants are first aware of the use of stress as a cue to word boundaries and the time when they learn to rely primarily on trochaic stress in English.

General Discussion

This series of experiments has, first of all, shown that infants are able to segment words from fluent speech at a younger age than previous reports have indicated. Using a somewhat different procedure, Jusczyk and Aslin (1995) found that although 7.5-month-olds were able to segment words from fluent speech, 6-month-olds showed no such ability. The results of Experiment 2 demonstrate that 6.5- to 7-month-old infants are able to segment words from fluent speech. Perhaps infants develop the ability to segment words from fluent speech between 6 and 6.5 months. Alternatively, the segmentation task in this series of experiments may have been easier for the infants than the tasks used in other experiments. The experimental methods used in these experiments expose infants to far fewer words, with far more repetition, than does the Jusczyk and Aslin procedure. If the difference between the success and failure of the young infants in each experiment is due more to procedural differences than to developmental differences, it becomes difficult to say at what age infants actually acquire the ability to segment words from fluent speech. Whatever the reasons for our participants' success, this is, to our knowledge, the first evidence of word segmentation by infants younger than 7.5 months of age.

With regard to the hypothesis under investigation, our results suggest that young (7-month-old) infants are less reliant on stress cues to word segmentation than are their older (9-month-old) counterparts. At 9 months of age, infants are misled by stress cues in fluent iambic speech, treating stressed syllables as word onsets even though those stressed syllables are actually the second syllables of words. This contrasts sharply with the pattern of results shown by the younger infants in this series of experiments. At 7 months of age, infants segment the same words from the iambic language as they do from the trochaic language. These results suggest that they are using statistical cues (which are invariant across the two languages) to segment these words from fluent speech, unlike 9-month-old infants. These results do not imply that infants at 6.5–7 months are unaware of stress, or even of its role as a cue to word boundaries (e.g., they might attend to stress cues in a language with no statistical information). Indeed, there are suggestions in the data that 7-month-old infants are aware that iambic stress is unusual; this may explain why infants at this age listen longer to test items after hearing the iambic language. Younger infants may weigh stress cues less heavily than older infants, a possibility echoing other research suggesting that 6- to 7-month-old infants are less sensitive to English-typical stress cues than are older infants (e.g., Echols et al., 1997; Jusczyk, Cutler, & Redanz, 1993; Morgan, 1996).

Does this mean that 7-month-old infants are better at segmenting words from fluent speech than are 9-month-old infants? Although the younger infants are not misled by stress, unlike the

[2] We thank an anonymous reviewer for this suggestion.

older infants, this does not mean that infants are becoming less able to accomplish a task (word segmentation) in the same way that they become less able to discriminate between phonemes not used in their language (Werker & Tees, 1984). Segmenting words from fluent speech using stress cues might be quite an adaptive strategy. An infant needs to hear a syllable only once to decide whether or not it is stressed. Therefore, hearing one instance of a trochaically stressed word is, in theory, all an infant needs to isolate it from fluent speech. In contrast, an infant needs to hear several instances of a word to learn that its syllables display some statistical coherence. A few exposures are not sufficient to build up a reliable statistical representation of a word. So despite the fact that using stress cues to segment words may lead infants to mis-segment words more often than does using statistical cues, it could be the easier of the two cues to employ, particularly given the complex statistical cues present in natural languages.

Why then did the young infants in this experiment not rely more on stress cues to segment words if stress cues are such an easy and robust cue to use? There are several possibilities. The current results indicate that younger infants are less attentive to trochaic stress as a cue to word segmentation than are older infants. There are a number of results in the literature that point to the same conclusion (e.g., Echols et al., 1997; Jusczyk, Cutler, & Redanz, 1993). Even after infants begin to use stress as a cue to syllable coherence, they do not attend solely to the trochaic pattern of English. Morgan (1996) found that 9-month-old infants were more sensitive to trochaic rhythm as a cue to syllable coherence than they were to iambic rhythm, whereas 6-month-old infants seemed to be equally attentive to both kinds of rhythm as a cue to syllable coherence. This finding suggests that even after infants begin to attend to stress in relation to word segmentation, they require experience with language before they show the metrical segmentation strategy adults and older infants use.

It may be that awareness of stress cues and awareness of statistical cues develop in parallel, with infants attending to each preferentially at different points in time. Infants may use statistical segmentation as a primary cue early in segmentation and attend more to stress cues later, after they have had sufficient exposure to words in isolation to learn about the predominantly trochaic pattern of English words. If this is the case, then the main milestone in the development of infants' word segmentation is the integration of these cues, so that infants can take advantage of multiple, partly redundant cues instead of fixating primarily on one. Morgan and Saffran (1995) demonstrated that older infants are more likely to attend to multiple cues to word segmentation and to attempt to integrate them than are younger (6-month-old) infants. Similarly, Johnson and Jusczyk (2001) found that 11-month-old infants were not misled by stress cues like the 8-month-old infants in their experiment were. Older infants have apparently learned that stress cues are not infallible indices of word boundaries and will attempt to integrate multiple cues to word segmentation.

Another possible explanation for the current results is that the younger infants were treating the task differently than the older infants. Younger infants may treat the task solely as a statistical learning task, attending to test items based only on their statistical coherence (or lack thereof). For a sequential statistical learning task, stress is simply not relevant. The older infants, on the other hand, may have been treating the familiarization stimulus as a language and trying to extract words from it, not just areas of statistical coherence. For this task, stress is more relevant, causing infants to mis-segment the iambic language. This account relies on the assumption that statistical learning is a domain-general process that is present very early in infancy (Kirkham, Slemmer, & Johnson, 2002) and that may be used on language-like stimuli long before infants appear to have any kind of linguistic knowledge.[3] However, there is nothing in the current data to confirm or deny the presence of different stages of linguistic awareness, on which this account also relies.

A final explanation consistent with the results of these experiments is that statistical segmentation helps infants to acquire the vocabulary necessary to discover the regularity of word stress in English. Although our experiments did not directly show learning of stress cues from statistical segmentation, the results do suggest that infants do not begin to attend primarily to stress cues until after they have gone through a period in which they rely more heavily on statistical information. Infants may use their statistical learning abilities to segment their first words from fluent speech, which—along with any words they learn from exposure in isolation—permits them to acquire a large enough vocabulary to learn that stress is highly correlated with word onsets in English.

On this account, general awareness of stress as a cue to syllabic coherence develops sometime between birth and 6 months (Morgan, 1996; Morgan & Saffran, 1995). At this early stage, infants are attentive to stress but have not yet discovered that stress predicts word onsets. By 7.5 months, infants are taking advantage of the correlation between stress and word onsets and are highly attentive to trochaic stress as a cue to word boundaries—although it is unclear what infants would do if stress and statistical information conflicted at this age (Jusczyk, Houston, & Newsome, 1999). By 8 months, infants are attending primarily to stress as a cue to word segmentation, treating stressed syllables as word onsets even when that information conflicts with statistical information (Johnson & Jusczyk, 2001). Around 11 months, infants become less reliant on stress cues (possibly because they have become aware of iambic words) and once again favor statistical cues over stress cues (Johnson & Jusczyk, 2001). One of the milestones of infant word segmentation is the movement from reliance on a single cue toward attention to multiple cues (e.g., Morgan & Saffran, 1995); by 10.5 months, infants may have discovered that stress cues correlate less well with multiple other cues to word segmentation than do statistical cues, and they thus favor statistical cues when the two are placed in conflict.

The current results suggest that 6.5- to 7.0-month-old infants are in a state of transition toward primarily attending to stress cues, and the results are consistent with the hypothesis that statistical learning may play a role in that transition. Though 6.5- to 7-month-old infants may have some weak knowledge of stress cues, phonology does not yet trump statistics. However, it is quite conceivable that infants would use stress as a cue to word segmentation if it were not opposed by statistical cues. The developmental timetable suggested by these and other results is rapid but not outside the realm of the changes in speech perception and word segmentation ability that occur in infancy (e.g., Jusczyk, 1997; Polka & Werker, 1994). If this is the case, then infants are rapidly acquiring phonological knowledge, aided by their use of statistical learning mechanisms, to generate a new strategy for segmentation—the use of stress cues. The output of one type of learning thus provides

---

[3] We thank an anonymous reviewer for this suggestion.

materials that afford the development of novel learning mechanisms, leading infants to flexibly adapt their learning strategies to the nature of the input.

More generally, English-hearing infants (and presumably infants hearing other languages) have become quite proficient and sophisticated language listeners by the end of their 1st year. They are familiar with their native language's predominant stress pattern (Jusczyk, Cutler, & Redanz, 1993). They can distinguish between allophones that occur across word boundaries and those that occur between syllables within a single word (Jusczyk, Hohne, & Bauman, 1999). They can even distinguish between consonant clusters that are likely to occur within a word as opposed to those that tend to occur across words (Mattys, Jusczyk, Luce, & Morgan, 1999). In short, their knowledge of the acoustic structure of their native language is much more sophisticated than one would suspect from a superficial investigation of infants' language competence. Learning all of these regular acoustic markers should require, first, some knowledge of word boundaries. It may be that statistical learning provides infants with their first window into the acoustic regularities of English and plays a broader role in infant language learning than has previously been suspected.

References

Aslin, R. N. (2000, July). *Interpretation of infant listening times using the headturn preference technique.* Paper presented at the International Conference on Infant Studies, Brighton, England.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9,* 321–324.

Bernstein-Ratner, N., & Pye, C. (1984). Higher pitch in BT is not universal: Acoustic evidence from Quiche Mayan. *Journal of Child Language, 11,* 515–522.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81,* B33–B44.

Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.

Crystal, T. H., & House, A. S. (1987). Segmental duration in connected-speech syllables: Syllabic stress. *Journal of the Acoustical Society of America, 83,* 1574–1585.

Crystal, T. H., & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America, 88,* 101–112.

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., McQueen, J., & Robinson K. (1990). Elizabeth and John: Sound patterns of men's and women's names. *Journal of Linguistics, 26,* 471–482.

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.

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.

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

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.

Hyman, L. M. (1977). On the nature of linguistic stress. In L. M. Hyman (Ed.), *Studies in stress and accent* (Southern California Occasional Papers in Linguistics No. 4, pp. 37–82). Los Angeles: University of Southern California.

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,* 548–567.

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

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Science, 3,* 323–328.

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. (1993). Preference for the predominant stress pattern of English words. *Child Development, 64,* 675–687.

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

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics, 61,* 1465–1476.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39,* 159–207.

Jusczyk, P. W., & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception & Psychophysics, 23,* 105–109.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83,* B35–B42.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America, 32,* 451–454.

Mattys, S. L., Juszczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38,* 465–494.

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.

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.

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. (2001). Words in a sea of sound: The output of 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,* 1926–1928.

Schwartz, R. G., Petinou, K., Goffman, L., Lazowski, G., & Cartusciello, C. (1995). Young children's production of syllable stress: An acoustic analysis. *Journal of the Acoustical Society of America, 99,* 3192–3200.

Stevens, K. N. (1998). *Acoustic phonetics.* Cambridge, MA: MIT Press.

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

Received April 29, 2002
Revision received January 23, 2003
Accepted February 4, 2003 ■