

Infants' sensitivity to allophonic cues for word segmentation

PETER W. JUSCZYK

Johns Hopkins University, Baltimore, Maryland

ELIZABETH A. HOHNE

AT&T Labs, Holmdel, New Jersey

and

ANGELA BAUMAN

San Diego State University, San Diego, California

A series of four experiments was conducted to determine whether English-learning infants can use allophonic cues to word boundaries to segment words from fluent speech. Infants were familiarized with a pair of two-syllable items, such as *nitrates* and *night rates* and then were tested on their ability to detect these same words in fluent speech passages. The presence of allophonic cues to word boundaries did not help 9-month-olds to distinguish one of the familiarized words from an acoustically similar foil. Infants familiarized with *nitrates* were just as likely to listen to a passage about *night rates* as they were to listen to one about *nitrates*. Nevertheless, when the passages contained distributional cues that favored the extraction of the familiarized targets, 9-month-olds were able to segment these items from fluent speech. By the age of 10.5 months, infants were able to rely solely on allophonic cues to locate the familiarized target words in passages. We consider what implications these findings have for understanding how word segmentation skills develop.

For fluent speakers of a language, the task of segmenting speech into words seems relatively easy. Only under rather special circumstances (e.g., decoding unfamiliar names on the radio, listening to speech in a foreign language, etc.) are most listeners aware of the potential difficulties involved in speech segmentation. Yet the difficulties posed by word segmentation are well known to those involved in devising automatic speech recognition devices. In conversational speech, the acoustic shapes of words are distorted by the nature of surrounding words (Lieberman & Studdert-Kennedy, 1978; Mills, 1980). Moreover, the boundaries of words are often not clearly marked in the speech stream (Cole & Jakimik, 1978, 1980; Klatt, 1979, 1989)—a fact that poses great difficulty for accurate machine recognition of words (Bernstein & Franco, 1996; Marcus, 1984; Reddy, 1976; Waibel, 1986). Yet, judging by the pace at which infants acquire a native

language vocabulary (Bates et al., 1994), language learners show considerable mastery of word segmentation skills before their second birthdays.

Indeed, some investigations indicate that infants begin to display some limited word segmentation abilities as young as 7.5 months of age (Echols, Crowhurst, & Childers, 1997; Jusczyk, 1996; Jusczyk & Aslin, 1995; Newsome & Jusczyk, 1995; Saffran, Aslin, & Newport, 1996). For example, Jusczyk and Aslin first demonstrated that 7.5-month-old English-learners detect the occurrence of repeated words in fluent speech passages. In one experiment, they familiarized infants with a pair of words (e.g., *feet* and *bike*) that were repeated in citation form. Then, the infants heard four different six-sentence passages. Two of these passages included one of the familiarized words in each sentence; the other two were comparable but included two other words that the infants had not heard during the familiarization period. The infants listened significantly longer to the passages containing the familiarized words. In a subsequent experiment, Jusczyk and Aslin used two of the passages during the familiarization phase and tested the infants on repetitions of the words in citation form. Once again, the same pattern of results ensued: Infants listened significantly longer to words they had heard during familiarization. This latter result suggests that, during familiarization, the infants were able to extract the repeated target words from their surrounding sentential contexts. Thus, they came to recognize the sound patterns of these items even when they were

The research reported here was supported by Research Grant 15795 from NICHD to P.W.J. We thank Ann Marie Jusczyk for helpful comments on a previous version of this manuscript. Jan Charles-Luce provided us with invaluable information regarding the acoustic measurements that are reported in the paper. In addition, we are grateful to Ann Marie Jusczyk, Nancy Redanz, Amy Gambon-Dennis, Aileen Warden, and Debra Dombrowski for assistance in recruiting and testing infants. Correspondence concerning this article should be addressed to P. W. Jusczyk, Ames Hall, Department of Psychology, Johns Hopkins University, Baltimore, MD 21218 (e-mail: jusczyk@jhu.edu).

—Accepted by previous editor, Myron L. Braunstein

presented only in complete sentences. In contrast to the 7.5-month-olds, 6-month-olds gave no evidence of detecting the occurrence of familiarized words in fluent speech.

How do infants begin to segment fluent speech into words? A number of different suggestions have been offered regarding cues to the possible location of word boundaries in speech. For example, Cutler and her colleagues (Cutler, 1976, 1990; Cutler & Butterfield, 1992; Cutler & Carter, 1987; Cutler & Norris, 1988) have pointed out that a high proportion of content words in English conversational speech have an initial stressed syllable. Hence, listeners could use what Cutler and Norris termed a *metrical segmentation strategy* (MSS), whereby stressed syllables are identified with the onsets of new words in fluent speech. Moreover, evidence from studies with English-learners suggests that infants may use stress cues in segmenting fluent speech. Jusczyk, Cutler, and Redanz (1993) noted a developing sensitivity to words with the predominant English stress pattern. Specifically, 9-month-old, but not 6-month-old, English-learners showed significant listening preferences for words with strong/weak, as opposed to weak/strong, stress patterns. More recently, Newsome and Jusczyk (1995) used the same procedure as that in Jusczyk and Aslin (1995) to test 7.5-month-olds' abilities to detect bisyllabic words in fluent speech contexts. Consistent with the predictions of the MSS, they found that the infants were able to detect the occurrence of bisyllabic words with strong/weak stress (e.g., *hamlet* and *kingdom*), but not ones with weak/strong stress (e.g., *guitar* and *device*). Similarly, Echols et al. (1997) and Morgan and Saffran (1995) have reported processing advantages for strong/weak over weak/strong stress patterns in their studies with English-learning 9-month-olds.

In addition to stress-based cues to word boundaries, it has been suggested (Saffran, Aslin, & Newport, 1996) that infants might use information about statistical regularities in the input to segment words from fluent speech. Saffran, Aslin, and Newport exposed 8-month-olds to 2 min of a continuous stream of speech, consisting of four trisyllabic nonsense words, repeating in a random order. These nonsense words were made from concatenations of isolated syllables. Unlike real English words, all three syllables were equally stressed. The only cues to word boundaries in this continuous stream were the transitional probabilities between syllable pairs. Specifically, transitional probabilities were higher for syllable pairs within words than between words. During the test phase, the infants were found to distinguish trisyllables that conformed to words during the familiarization phase from trisyllables that spanned word boundaries.

Closely related to the view that infants use statistical regularities to detect word boundaries is the notion that learners take advantage of distributional properties in the input (Brent & Cartwright, 1996; Suomi, 1993). For example, an infant who learns some words spoken as isolated utterances might then match a stored representation of their sound patterns to fluent speech, thus breaking

the input into a known and one or more unknown strings. These unknown strings might then be stored as unanalyzed units. For instance, if *cat* is known, it can be matched in the utterance *See the cat*, with the result that *See the* would be an unanalyzed unit. Eventually, the learner will hear *see* and *the* in other distributional contexts, and the contrast between these and the stored sequence of *see the* lead to the decomposition of the unit into the lexical items *see* and *the*. Although this approach has not yet been directly tested with infants, Brent and Cartwright have described a successful computer simulation model for word segmentation that is based on distributional cues.

The focus of the present investigation is on another potential source of information about word boundaries in fluent speech. It has been noted that attention to the particular contexts in which variants (or allophones) of the same phoneme appear can provide cues about the locus of word boundaries (Bolinger & Gerstman, 1957; Church, 1987a; Lehiste, 1960; Umeda & Coker, 1974). Church (1987b) observes that the allophone of /t/ that begins words in English, such as *tap* (i.e., [tʰ]) is not found in other positions in English words, such as the /t/s in *stop* or *hat*. A listener who is sensitive to the distribution of these allophones could use this information in deciding whether a word boundary has occurred or not.

Of course, for allophonic cues to be useful in word segmentation by infants, several conditions must hold. First, there must be some indication that allophones are orderly manifestations of phonemic contrasts in the language, rather than simply some random acoustic variation produced by speakers. Second, the distribution of these allophones in fluent speech must correlate with word boundaries. Third, infants must be able to discriminate one allophone from another. Fourth, infants must be sensitive to the systematic distribution of these distinct allophones within native language words. Fifth, they must use these allophonic cues as markers of word boundaries during on-line speech processing. One might argue that a further condition should be added to this list—namely, that infants recognize that particular allophones can also be perceived as variants of the same phoneme. However, this last condition is not necessary for allophonic cues to be useful as markers of word boundaries. What is critical is that infants recognize that talkers are engaging in a systematic variation in producing these elements and, critically, that this variation is tied to their positioning within words.

Regarding the first of the five conditions mentioned above, linguists have long noted the existence of allophones as systematic variants of phonemes and have incorporated these into their descriptions of the sound structures of language (e.g., Ladefoged, 1975). Similarly, there are indications in the previous literature of a correlation between allophonic variants and English word boundaries (Bolinger & Gerstman, 1957; Church, 1987a; Hockett, 1958; Lehiste, 1960; Umeda & Coker, 1974). With respect to the third condition, the results of an investigation by Hohne and Jusczyk (1994) indicate that English-

learning 2-month-olds can discriminate the kinds of allophonic differences that could cue the location of word boundaries. Specifically, the infants discriminated the kinds of allophonic variants of /t/ and /r/ that distinguish pairs of items such as *nitrate* and *night rate*. For *nitrate*, the first *t* is aspirated, released, and retroflexed, whereas the *r* is devoiced, suggesting that it is part of a cluster. By comparison, the first *t* in *night rate* is unaspirated and unreleased, suggesting that it is syllable final, whereas the following *r* is voiced, suggesting that it is syllable initial. The infants were able to discriminate these allophonic pairs even when they were surrounded by phonetic contexts that were acoustically identical. Hence, 2-month-olds have the capacity to discriminate allophonic distinctions that could signal the presence or the absence of word boundaries in English.

The fourth condition, mentioned above, for using allophonic cues in word segmentation is that infants be sensitive to how allophones are distributed within words. The evidence on this point is suggestive but indirect. By 9 months, English-learning infants have been shown to be sensitive to the frequency with which certain phonetic sequences occur within syllables in their language (see, e.g., Jusczyk, Luce, & Charles-Luce, 1994). Moreover, by 7.5 months, English-learners have some demonstrated capacity for segmenting wordlike units from fluent speech on the basis of the location of stressed syllables (Newsome & Jusczyk, 1995). Attention to the phonetic information that occurs at the beginnings and endings of such units could provide learners with information about the way allophones are typically distributed within such units. Knowledge of the contexts in which particular allophones frequently appear could then be used as potential cues to the position of likely word boundaries.

The primary goal of the present investigation is to provide more direct evidence about this fourth condition—namely, when do English-learning infants display sensitivity to the way that allophones are typically distributed within words. A demonstration of this sensitivity would be consistent with, although it would not prove, the possibility that they also meet the fifth condition (i.e., that they actually use allophonic cues to segment words from fluent speech).

EXPERIMENT 1

There are a number of reasons why an infant might be able to distinguish a pair of allophones on the kind of discrimination task used by Hohne and Jusczyk (1994) but might fail to use this information in locating word boundaries in fluent speech. First, although infants might discriminate the relevant allophones, they may not have associated them with any particular phonetic contexts. Indeed, in order to learn how these allophones are distributed in words, it would be useful for the child to receive single-word utterances at least occasionally. There are indications that at least some proportion of the input directed to the learner consists of single-word utterances

(Woodward & Aslin, 1990). Moreover, as noted above, 7.5-month-olds do show some capacity for segmenting some words from fluent speech (Jusczyk & Aslin, 1995; Newsome & Jusczyk, 1995). Second, even though an infant might have the capacity to perceive a phonetic contrast when discriminating a pair of isolated utterances, the infant might not be able to make full use of these capacities under more complicated circumstances (Jusczyk, 1997; Stager & Werker, 1997), such as when these items are embedded in fluent speech.

To explore the possible use of allophonic cues in word segmentation by infants, we tested 9-month-olds because, as previous studies have demonstrated, they already have some ability to segment words. We used the same pair of items (*nitrate* and *night rate*) as that in the Hohne and Jusczyk (1994) study. This pair has often been mentioned in discussions of how allophonic cues can signal word boundaries (Hockett, 1958; Lehiste, 1960; Nakatani & Dukes, 1977). In addition, to compare word segmentation based on allophonic cues to more general segmentation processes, we included a second pair of items, *hamlet* and *doctor*, which differed in other ways. Previous research by Newsome and Jusczyk (1995) indicated that infants are capable of segmenting these items from fluent speech contexts.

We employed the modified version of the Headturn Preference Procedure that Jusczyk and Aslin (1995) used to examine infants' detection of words in fluent speech. Because the critical focus of the present investigation is on allophonic cues, each infant was familiarized with a pair of items. One of these items was either *nitrate* or *night rate*; the other was either *hamlet* or *doctor*. During the test phase, the infants heard both a *nitrate* passage and a *night rate* passage, as well as a *hamlet* and a *doctor* passage. As in previous studies (e.g., Jusczyk & Aslin, 1995), word segmentation abilities were indexed by listening preferences for the passages containing the items that the infants had heard during familiarization. Thus, if the infants are attentive to the distribution of allophonic cues within words, familiarization with *nitrate* should lead to significantly longer listening times to the *nitrate* passage than to the *night rate* passage (and vice versa for infants familiarized with *night rate*).

Method

Participants. The participants were 24 American infants (13 males, 11 females) from monolingual English-speaking homes. The infants were approximately 9 months old, with a mean age of 38 weeks, 6 days (range, 36 weeks, 6 days to 40 weeks, 6 days). To obtain the 24 participants for the study, it was necessary to test 34. Some of the infants were excluded, for the following reasons: crying (5), unresponsiveness to the flashing lights (4), and looking times averaging less than 3 sec (1).

Stimuli. A female talker, who was a native speaker of American English from western New York, recorded four different six-sentence passages (see Table 1). She was encouraged to read the passages in a lively voice, as if she were reading them to a small child. The recordings were made in a sound-attenuated room with a Shure microphone. The critical passages were digitized on a VAXStation Model 3176 computer at a sampling rate of 10 kHz via

Table 1
Test Passages for Experiments 1 and 4

1. Nitrates are not something that everyone needs. My teacher told us all about nitrates. Farmers use nitrates to help grow their crops. Nitrates are used to preserve food. This storeroom has many different kinds of nitrates. There were boxes of nitrates on all of the shelves.
2. Your hamlet lies just over the hill. Far away from here near the sea is an old hamlet. People from the hamlet like to fish. Another hamlet is in the country. People from that hamlet really like to farm. They grow so much that theirs is a very big hamlet.
3. Night Rates can help us to save some money. Businesses try to use night rates to send their packages. Even the airlines have cheaper night rates. The man wrote out the night rates on the blackboard. Night rates at this hotel are expensive. Many people look for the best night rates.
4. The doctor saw you the other day. He's much younger than the old doctor. I think your doctor is very nice. He showed another doctor your picture. That doctor thought you grew a lot. He was a big doctor himself.

a 12-bit analog-to-digital converter. The average duration of the passages was 18.02 sec (ranging from 17.92 sec for the *doctor* passage to 20.6 sec for the *nitrates* passage). Once it was determined that the passages were acceptable for use in the experiment, the talker was asked to record versions of the isolated words to be used during the familiarization phase of the experiment. For each target word, the talker was asked to repeat the item with some variation 15 times in a row, in a lively voice, and as if naming the object for an infant. These lists were then digitized on the computer in the same way as the sentences. The average duration of the lists was 23.98 sec (ranging from 23.75 sec for the *night rates* list to 24.35 sec for the *doctor* list). Digitized versions of the passages and the lists were transferred to a PDP-11/73 computer for playback during the experiment.

Detailed acoustic measurement of critical properties of the *nitrates* and *night rates* stimuli were carried out using a CSL workstation (Kay Elemetrics). Measurements of the isolated tokens of *nitrates* and *night rates* revealed a number of interesting differences. The duration of the [a'] vowel was significantly longer [$t(28) = 4.24, p < .0005$] in *night rates* (193 msec) than in *nitrates* (150 msec). There was also evidence of significantly longer burst [$t(28) = 18.56, p < .0001$] and closure [$t(28) = 23.19, p < .0001$] durations for the initial *t* in *nitrates* (162 and 74 msec, respectively) than for *night rates* (7 and 8 msec, respectively). These last differences are consistent with the claims that the initial *t* in *nitrates* is released and aspirated, whereas that of *night rates* is unreleased and unaspirated. Acoustic measurements of the *r* in these tokens indicated that friction was present in every single token of *nitrates* and in none of the tokens of *night rates*. Finally, the overall duration of voicing of *r* was significantly longer [$t(28) = 10.76, p < .0001$] for *night rates* (98 msec) than for *nitrates* (47 msec). Hence, the isolated tokens of *nitrates* and *night rates* did contain distinctive allophonic cues.

Similar measurements were made for the tokens of *nitrates* and *night rates* that occurred in the passages. In contrast to the isolated tokens, the duration of the [a'] vowel did not differ significantly [$t(28) = 1.61, p < .15$] for *night rates* (138 msec) and *nitrates* (130 msec). However, as for the isolated words, there was evidence for the words in the passages of significantly longer burst [$t(28) = 14.01, p < .0001$] and closure [$t(28) = 24.02, p < .0001$] durations for the initial *t* in *nitrates* (38 and 51 msec, respectively) than for *night rates* (8 and 5 msec, respectively). With respect to *r*, there was once again evidence of friction in every single token of *nitrates*, but there was no friction for any token of *night rates*. However, the overall duration of voicing of *r* was not significantly longer [$t(28) = 1.04, p < .30$] for *night rates* (51 msec) than for *nitrates* (44 msec). Thus, the critical distinguishing properties that were available to in-

fants in both the isolated and the passage contexts were the burst and closure durations of *t* and the presence or absence of friction in *r*.

Finally, it is worth noting that the average overall durations of the target words in the passages (*nitrates*, 577 msec; *night rates*, 594 msec) were considerably shorter than the isolated tokens of the same words (*nitrates*, 1,252 msec; *night rates*, 1,146 msec). Thus, the infants could not identify the familiarized target in the passages by performing a simple acoustic match from the isolated tokens presented during the familiarization period.

Design. Half of the infants heard the words *night rates* and *doctor* during the familiarization phase, and the other half heard the words *nitrates* and *hamlet*. During the test phase, all the infants heard four blocks of the same four passages. Each block contained a different random ordering of the passages corresponding to *night rates*, *nitrates*, *doctor*, and *hamlet*.

Apparatus. A PDP-11/73 controlled the presentation of the stimuli and recorded the observers' coding of the infants' headturn responses. 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.

Procedure. The experiment was conducted in a three-sided test booth constructed out of 4 × 6 ft pegboard panels on three sides and open at the back. An observer looked through one of the existing pegboard holes in the front panel to monitor the infant's headturns. Except for a small section for viewing the infant, the remainder of the pegboard panels were backed with white cardboard, to guard against the possibility that the infant might respond to movements behind the panel. The test booth had a red light and a loudspeaker mounted at eye level on each of the side panels 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 used 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 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 headturns, and terminated a trial when the infant looked away for more than 2 sec. Information about the direction and duration of headturns and the total trial was stored in a data file on the computer. 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 rhyming and control lists were calculated by the computer after the completion of each session.

A version of the headturn preference procedure was used (for an extensive discussion concerning the reliability of this procedure, see Jusczyk, 1998; Kemler Nelson et al., 1995; Polka, Jusczyk, & Rvachew, 1995). 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 was begun by blinking the green light on the center panel until the infant had oriented in that direction. Then, the center 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 sec (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 sec, 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 sample. The flashing red light remained on for the entire duration of the trial.

Each experimental session began with a familiarization phase in which infants heard repetitions of two of the target items on alternating trials until they accumulated 30 sec of listening time to each one. If the infants achieved the familiarization criterion for one item, but not for the other, the trials continued to alternate until the criterion was achieved for both. The location of the loudspeaker from which the words were emitted was varied from trial to trial, with a different random order being used for each infant.

The test phase began immediately after the familiarization criterion was attained. The stimuli for the test phase consisted of the four six-sentence passages. The order of each of the sentences within a passage was fixed, and each trial always began with the first of the six sentences in the passage. The test trials were blocked in groups of four, so that each passage occurred once per block. The order of the passages within a block was randomized. Each infant was tested on four blocks, for a total of 16 test trials.

An observer hidden behind the center panel looked through the peephole and recorded the direction and duration of the infant's headturns, using a response box. The observer was not informed as to which items served as familiarization words for a given infant. The loudness levels for the samples were set by a second assistant, who was not involved in the observations, at 72 ± 2 dB (C) SPL. Both the observer and the infant's caregiver wore foam earplugs and listened to masking music over tight-fitting closed headphones (SONY MDR-V600). The masker consisted of loud instrumental music, which had been recorded with few silent periods. Caregivers and observers reported that, with this masker, they were unaware of either the location or the nature of the stimulus on the trial. Reliability checks between the live observer and the observers of the videotapes of each session are high, with correlations ranging from .92 to .96 (Kemler Nelson et al., 1995).

Results and Discussion

Mean listening times to the four different passages were calculated for each infant across the four blocks of trials. These data were submitted to a repeated measures analysis of variance (ANOVA) of a 2 (experience: familiar vs. unfamiliar) \times 2 (word type: allophonic vs. other) design. The analysis revealed a significant main effect for experience [$F(1,23) = 4.63, p < .05$], indicating that the listening times were significantly longer for the familiar items. The main effect of word type [$F(1,23) < 1.00$] was not significant. The interaction between these two factors was marginally significant [$F(1,23) = 4.05, p = .056$]. The average listening times across all subjects are displayed in Figure 1. A series of planned comparisons, using contrast tests based on the ANOVA, was carried out to explore possible differences for the familiar and unfamiliar items in the *night rates/nitrates* and *doctor/hamlet* pairs. The difference between the familiar and the unfamiliar items was significant for the *doctor/hamlet* pair [$F(1,23) = 11.27, p < .003$], but not for the *nitrates/night rates* pair [$F(1,23) < 1.00$]. The latter result suggests that 9-month-olds did not use allophonic cues to match the item from the familiarization phase to the correct *night rates* or *nitrates* passage during the test phase.

The present results provide some further confirmation that infants can segment bisyllabic words from fluent speech contexts, thus replicating the finding reported by Newsome and Jusczyk (1995). At the same time, there was no indication that the 9-month-olds were able to take advantage of allophonic cues in recognizing *nitrates* or

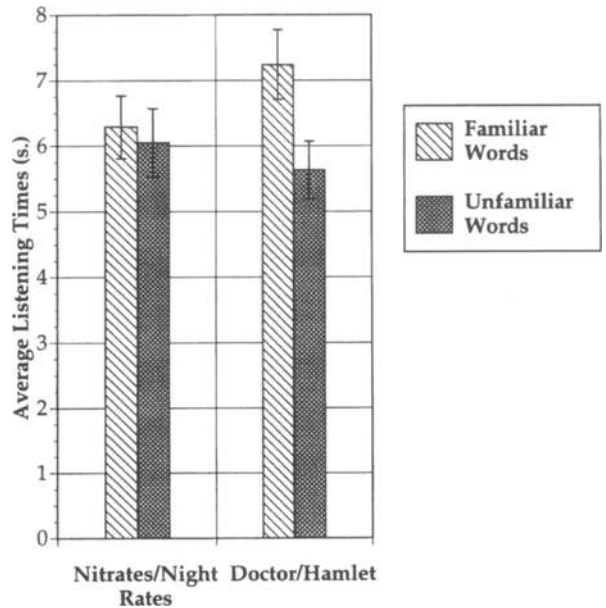


Figure 1. Nine-month-olds' average listening times (and standard error bars) for the passages with the familiar and the unfamiliar target words in Experiment 1.

night rates in fluent speech contexts, even though prior research suggests that these sorts of differences are discriminable for 2-month-olds (Hohne & Jusczyk, 1994). As was indicated earlier, there are a number of possible explanations for the infants' failure to correctly segment *nitrates* and *night rates* from fluent speech. One possibility is that 9-month-olds have not yet identified the contexts in which the critical allophones typically appear. A second possibility is that the greater processing demands associated with fluent speech perception limit the extent to which infants can fully utilize their discriminative capacities (Jusczyk, 1997; Stager & Werker, 1997). If the latter explanation is true, it is possible that infants might show some ability to use the allophonic cues, if the processing demands were simplified.

EXPERIMENT 2

For infants to detect the occurrence of the familiarized items in the test passages, they must be able to encode and retain an accurate representation of these items. *Night rates* and *nitrates* are very similar to each other with respect to their phonetic properties. Hence, correctly matching these items requires a rather detailed representation of the familiarized item. There is evidence from previous studies using this paradigm that infants do retain a detailed representation of the familiarized words. For example, Jusczyk and Aslin (1995) found that 7.5-month-olds who were familiarized with an item like *tup* did not later generalize to an item in a fluent speech passage that differed from it only by a single phonetic feature—namely, the word *cup*. More recently, Tincoff and Jusczyk (1996) extended these findings by showing that

infants familiarized with *cut* did not subsequently generalize to *cup*. However, one difference between these prior studies and our first experiment is that, in the latter, the infants had to encode and retain a pair of bisyllabic items from the familiarization phase. Thus, the memory demands were greater in Experiment 1 than in the previous studies. The additional syllables in our stimulus materials could have affected the amount of phonetic detail that the infants were able to encode.

One way to make the stimulus materials in the present investigation more comparable with those used in previous studies is to use monosyllabic items during the familiarization phase. For example, assuming that infants can use allophonic cues in word segmentation, then if they are familiarized with the word *night*, they should find a match to it in the *night rates* passage, but not in the *nitrates* passage (where the syllabification suggests that the first *t* is actually part of a cluster in the onset of the second syllable). To further ensure that the memory demands during the familiarization period were reduced, we made a modification to the other pair of items by replacing *hamlet* with *dock*. Hence, in the present experiment, all the infants heard *night* and *dock* during the familiarization period and were tested on *night rates*, *nitrates*, *dock*, and *doctor* passages. Our prediction was that infants familiarized with these new items would have significantly longer listening times for the *night rates* and *dock* passages than for the *nitrates* and *doctor* passages, respectively.

Method

Participants. The participants were 24 American infants (12 males, 12 females) from monolingual English-speaking homes. The infants were approximately 9 months old, with a mean age of 39 weeks, 3 days (range, 36 weeks, 3 days to 43 weeks, 0 days). To obtain the 24 participants for the study, it was necessary to test 29. Some of the infants were excluded, for the following reasons: crying (2), parental interference (2), and looking times averaging less than 3 sec (1).

Stimuli. The female talker from the previous experiment recorded one new six-sentence passage for the word *dock* (see Table 2). The duration of this new passage, 17.51 sec, was comparable with those of the other three passages. In addition, she recorded 15 versions each of the isolated words *night* and *dock*. Once again, she was instructed to speak as if she were addressing a small child and to repeat the items with some variation in prosody. The durations of the familiarization lists were 20.17 sec (*night*) and 18.61 sec (*dock*). Acoustic measurements of the burst and closure durations of *t* in the isolated tokens of *night* (1 and 6 msec, respectively) indicated that these did not differ significantly [$t(19) = 0.75, p > .40$, and $t(19) = 0.57, p > .50$, respectively] from those of *night* in the *night rates* passages (2 and 14 msec, respectively). However, as was expected, the burst and closure durations of *t* in the isolated tokens of *night* did differ significantly [$t(19) = 22.47, p <$

.0001, and $t(19) = 5.38, p < .0001$] from those in the *nitrates* passages (38 and 51 msec, respectively).

Design. With the exception of the fact that all the infants heard the same two words during familiarization, the design was the same as that in the preceding experiment. Thus, in this experiment, all the infants were familiarized with isolated tokens of *night* and *dock* and were tested on passages containing *nitrates*, *night rates*, *dock*, and *doctor*.

Apparatus and Procedure. The apparatus and the procedure were the same as those in Experiment 1.

Results and Discussion

Once again, mean listening times to the four different passages were calculated for each infant across the four blocks of trials. These data were submitted to a repeated measures ANOVA of a 2 (experience: familiar vs. unfamiliar) \times 2 (word type: allophonic vs. other) design. The analysis revealed a significant main effect for experience [$F(1,23) = 9.18, p < .001$], indicating that the listening times were significantly longer for the familiar items. Neither the main effect of word type [$F(1,23) = 2.03, p > .15$] nor the interaction between experience and word type was significant [$F(1,23) = 2.33, p > .10$]. The average listening times across all subjects are displayed in Figure 2. A series of planned comparisons, using contrast tests based on the ANOVA, was carried out to explore possible differences for the familiar and unfamiliar items in the *night rates/nitrates* and *dock/doctor* pairs. The difference between the familiar and the unfamiliar items was significant for the *dock/doctor* pair [$F(1,23) = 10.52, p < .004$], but not for the *nitrates/night rates* pair [$F(1,23) < 1.18, p > .25$]. The latter result indicates, once again, that 9-month-olds did not use allophonic cues to match *night* from the familiarization phase to the *night rates* passage during the test phase. In fact, there was little evidence that *night* was recognized in either the *night rates* or the *nitrates* passages. Post hoc contrast tests indicated that listening times to the *nitrates* and *night rates* passages differed significantly from the familiar *dock* passage [$F(1,23) = 11.78, p < .005$], but not from the unfamiliar *doctor* passage [$F(1,23) < 1.00$].

Reducing the memory load during the familiarization phase did not produce any measurable gain in 9-month-olds' ability to use allophonic cues in segmenting words from fluent speech. Therefore, it seems likely that infants at this age are not sensitive to how these allophonic cues are distributed within words. By comparison, the infants in the present study had no difficulty in matching *dock* to its occurrence in the *dock*, as opposed to the *doctor* passage. The latter finding replicates one reported by Newsome and Jusczyk (1995), who found that 7.5-month-olds familiarized with *can* did not generalize to *candle*. Newsome and Jusczyk interpreted this finding as being a result of the role that distributional cues play in word segmentation—namely, the infants were familiarized with *can*, but during the test phase, the distributional cues indicated that the *can* in *candle* always co-occurred with the same following syllable, suggesting that the two formed a single unit. Nine-month-olds' attention to distributional

Table 2
Test Passage for *Dock* in Experiment 2

The dock by the lake has sailboats. Motor boats stay at the old dock. That dock gets very busy. Fishing from the dock looks like fun. People at your dock like to swim, too. A family brought their boat to the new dock.

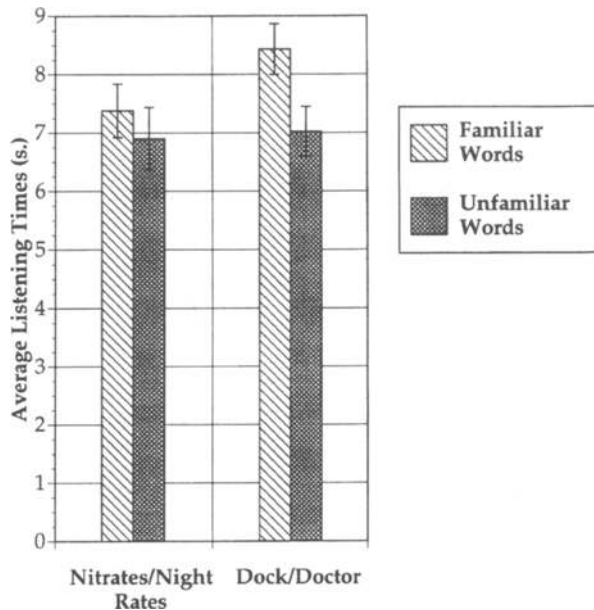


Figure 2. Nine-month-olds' average listening times (and standard error bars) for the passages with the familiar and the unfamiliar target words in Experiment 2.

cues in the present experiment could also help to explain why they did not detect a better match to *night* in the *night rates* passage than in the *nitrates* passage. Specifically, *rates* consistently co-occurred with *night* in the passages. Hence, *night rates* might have been perceived as a single unit in much the same way that *candle* was. The following experiment was designed to explore this possibility.

EXPERIMENT 3

In the previous experiment, the infants who were familiarized with *night* were just as likely to listen to a *nitrates* passage as they were to listen to a *night rates* passage. The fact that the allophonic properties of the final [t] in *night* matched those of *night rates*, but not those of *nitrates*, did not significantly affect their listening times to the passages. However, suppose that the distributional properties of the passages favored the extract of *night* in one passage, but not in the other? Would 9-month-olds still treat a *nitrates* passage as a potential match after familiarization with *night*?

There is evidence that distributional cues are helpful for word segmentation in artificial language experiments with adults (Dahan & Brent, in press; Saffran, Newport, & Aslin, 1996). Moreover, 8-month-olds exposed to 2 min of a continuous stream of four trisyllabic nonsense words appeared to draw on distributional cues (i.e., the frequency with which one syllable followed another) to recognize novel from unfamiliar "words" (Saffran, Aslin, & Newport, 1996). Still, it would be useful to know whether these findings with artificial languages also transfer to settings in which listeners are dealing with native lan-

guage input. As was noted above, Newsome and Jusczyk (1995) have argued that it was English-learners' sensitivity to distributional cues that led them not to detect a familiarized target such as *can* in *candle* passages. However, if the infants truly are using distributional cues in segmenting fluent speech, they should also be able to use these cues to detect the real occurrence of familiarized items, such as *night*, in passages.

To investigate this possibility, we replaced the *night rates* passage with a new one containing the word *night*. In the new passage, *night* was always followed by a different noun (e.g., *night caps*, *night games*). If English-learning 9-month-olds can use distributional cues in segmenting fluent speech, then after familiarization to *night*, they should listen longer to the new *night* passage than to the *nitrates* passage.

Method

Participants. The participants were 24 American infants (13 males, 11 females) from monolingual English-speaking homes. The infants were approximately 9 months old, with a mean age of 40 weeks, 2 days (range, 37 weeks, 6 days to 43 weeks, 0 days). To obtain the 24 participants for the study, it was necessary to test 31. Some of the infants were excluded, for the following reasons: crying (2), failure to look at the flashing lights (2), and looking times averaging less than 3 sec (3).

Stimuli. The female talker from the previous experiment recorded one new six-sentence passage (*night + X*) for the word *night* (see Table 3). This new passage replaced the *night rates* passage used in the previous two experiments. The duration of this new passage, 18.12 sec, was comparable with the other three passages (i.e., *nitrates*, *dock*, and *doctor*). Once again, the talker was instructed to speak as if she were addressing a small child. Acoustic measurements of *night* in the new passages indicated differences in both burst [$t(28) = 16.44, p < .0001$] and closure [$t(28) = 25.80, p < .0001$] durations for the *t*, as compared with the initial *t* of *nitrates*. In particular, the burst and closure durations of *t* in *nitrates* (162 and 74 msec, respectively) were significantly longer than those for the *t* of *night + X* (1 and 1 msec, respectively). Hence, for the words in the passages, the same burst and closure cues differentiating the *t* in *night* from the initial *t* in *nitrates* were available as in the previous two experiments. Critically, comparisons of the burst and closure durations of *t* in the isolated tokens (1 and 6 msec, respectively) indicated that neither of these differed significantly [$t(19) = 0.67, p > .50$, and $t(19) = 0.28, p > .75$] from those of the *night + X* words in the passages.

Design. As in Experiment 2, all the infants were familiarized with isolated tokens of *night* and *dock*; then they were tested on the passages containing *night + X*, *nitrates*, *dock*, and *doctor*.

Apparatus and Procedure. The apparatus and procedure were the same as those in the previous experiments.

Results and Discussion

Once again, mean listening times to the four different passages were calculated for each infant across the four

Table 3
Test Passage for *Night + X* in Experiment 3

Many interesting things happen during night time. Night schools teach people about computers. Teams play night games at baseball fields. Night lights help the players to see. Some of us even put special clothes on, like night caps. My friend wears a different night gown every day.

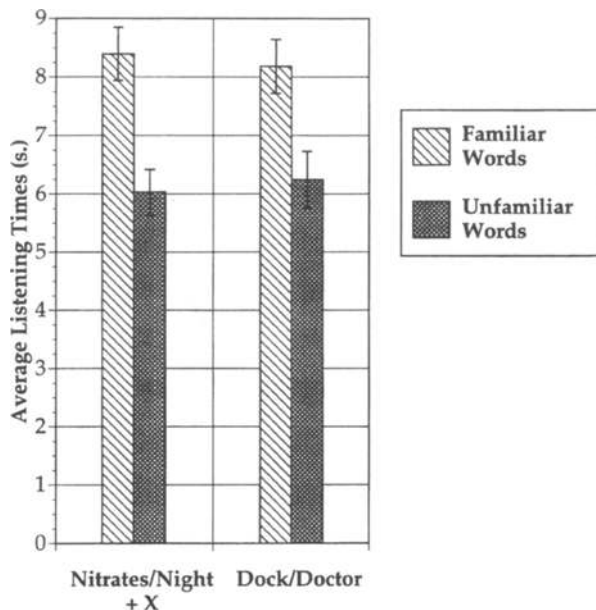


Figure 3. Nine-month-olds' average listening times (and standard error bars) for passages with the familiar and unfamiliar target words in Experiment 3.

blocks of trials. These data were submitted to a repeated measures ANOVA of a 2 (experience: familiar vs. unfamiliar) \times 2 (word type: allophonic vs. other) design. The analysis revealed a significant main effect for experience [$F(1,23) = 39.12, p < .0001$], indicating that the listening times were significantly longer for the familiar items. Neither the main effect of word type [$F(1,23) > 1.00$] nor the interaction between experience and word type [$F(1,23) > 1.00$] was significant. The average listening times across all the subjects are displayed in Figure 3. A series of planned comparisons, using contrast tests based on the ANOVA, was carried out to explore possible differences for the familiar and the unfamiliar items in the *night + X/nitrates* and *dock/doctor* pairs. The differences between the familiar and the unfamiliar items were significant for both the *night + X/nitrates* [$F(1,23) = 17.29, p < .0005$] and the *dock/doctor* [$F(1,23) = 17.34, p < .0005$] pairs. The finding of a significant listening preference for the *dock* passage over the *doctor* passage replicates that of the previous experiment. By comparison, the significant listening preference observed for the *night + X* passage over the *nitrates* passage contrasts with the pattern of findings in the two previous experiments. This latter finding suggests that 9-month-olds were able to use distributional cues to detect the occurrence of *night* in the *night + X* passage.

Use of distributional cues to word boundaries requires that the listener hear the word in the context of different words. In this way, the transitional probability of the same syllable preceding or following the target word is reduced. Hence, one might expect that the effectiveness of distri-

butional cues for a word target will increase, the more often infants hear a word in varied contexts. To investigate this possibility, we conducted an analysis of performance across the four blocks of trials. A repeated measures ANOVA of a 4 (blocks) \times 2 (experience: familiar vs. unfamiliar) \times 2 (word type: allophonic vs. other) design yielded significant main effects for blocks [$F(1,693) = 14.77, p < .0001$] and experience [$F(1,23) = 30.52, p < .0001$], along with a marginal three-way interaction [$F(3,69) = 2.40, p < .08$]. Analyses of performance on the individual blocks indicated that significantly longer listening times for the *night + X* passage than for the *nitrates* passage did not emerge until Blocks 3 and 4. Hence, there is some indication that infants were better able to detect the target word in the passages as they gained increased experience with the distributional contexts.

Together with the results of the previous experiments, the present findings suggest that, although 9-month-olds do not take advantage of allophonic cues, they can use information about distributional contexts in detecting word boundaries in fluent speech. Thus, the present results provide further support for the view that language learners use distributional cues in segmenting words from fluent speech (Brent, 1997; Brent & Cartwright, 1996; Dahan & Brent, in press; Newsome & Jusczyk, 1995; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996; Suomi, 1993). Indeed, one explanation of why infants in Experiment 2, familiarized with *night*, did not show a preference for the *night rates* passage is that the distributional cues favored treating *night rates* as a single word rather than as a sequence of two words (the first of which matches the familiar *night*).

The apparent failure of 9-month-olds in Experiments 1 and 2 to use allophonic cues to detect the occurrence of familiar words in passages raises a number of questions. First, is it possible that 9-month-olds are able to use allophonic cues in some instances, but that the *night rates/nitrates* pair is not one of these? Although we cannot definitely rule out this possibility, we do have some reason to believe that 9-month-olds do not use allophonic cues for other pairs of items. In another experiment, in addition to testing twenty-four 9-month-olds on *night rates* and *nitrates*, we also tested them on *gray ties* and *great eyes*. In neither instance did the infants display a significant listening preference for the passage that matched the items heard during familiarization [$t(23) = 0.06$, for *night rates/nitrates* and $t(23) = 1.30, p > .20$, for *gray ties/great eyes*]. Of course, these pairs do not exhaust the range of possible allophonic cues that 9-month-olds could potentially use in word segmentation. Still another possibility to consider is that not only are 9-month-olds unable to use allophonic cues to word boundaries, but also that language learners at any age do not use these types of cues. Although there is evidence that infants can use stress-based cues (Echols et al., 1997; Morgan & Saffran, 1995; Newsome & Jusczyk, 1995) and distributional cues (Saffran, Aslin, & Newport, 1996) in word seg-

mentation, there is no indication yet that they use allophonic cues. For this reason, we thought that it would be useful to investigate whether older infants show any sensitivity to how these allophonic cues are distributed within words.

EXPERIMENT 4

Previous work in our laboratory suggests that some interesting changes occur in English-learners' word segmentation abilities between 7.5 and 10.5 months of age. For instance, Newsome and Jusczyk (1995) observed that 7.5-month-olds are able to segment words with strong/weak stress patterns from fluent speech, but not words with weak/strong patterns. Yet, Myers et al. (1996) found that 10.5-month-olds were just as apt to detect interruptions of words with weak/strong stress patterns as they were to detect ones with strong/weak patterns. Similarly, Houston, Newsome, and Jusczyk (1995) showed that 10.5-month-olds can detect weak/strong words in fluent speech contexts. Thus, it appears that, between 7.5 and 10.5 months, English-learners develop some ability to segment weak/strong words. Moreover, in order to segment weak/strong words from fluent speech, 10.5-month-olds cannot rely solely on a stress-based strategy, such as MSS, where onsets of words are identified with the occurrence of strong syllables. Rather, they must draw on other potential sources of information about word boundaries, possibly including allophonic cues. Consequently, it seemed worthwhile to examine whether 10.5-month-olds show any sensitivity to allophonic cues in detecting occurrences of *night rates* and *nitrates* in fluent speech.

Method

Participants. The participants were 24 American infants (11 males, 13 females) from monolingual English-speaking homes. The infants were approximately 10.5 months old, with a mean age of 45 weeks, 1 day (range, 42 weeks, 0 days to 48 weeks, 6 days). To obtain the 24 participants for the study, it was necessary to test 32. Some of the infants were excluded, for the following reasons: crying (2), sleeping (1), and looking times averaging less than 3 sec (5).

Stimuli. The stimulus set from Experiment 1 was used.

Design, Apparatus, and Procedure. The design, apparatus, and procedure were identical to those in Experiment 1. Thus, infants were familiarized with *night rates* and *doctor* or with *nitrates* and *hamlet* and were tested on passages containing *night rates*, *nitrates*, *doctor*, and *hamlet*.

Results and Discussion

Once again, mean listening times to the four different passages were calculated for each infant across the four blocks of trials. These data were submitted to a repeated measures ANOVA of a 2 (experience: familiar vs. unfamiliar) \times 2 (word type: allophonic vs. other) design. The analysis revealed a significant main effect for experience [$F(1,23) = 25.48, p < .0001$], indicating that the listening times were significantly longer for the familiar items. The main effect of word type was also significant [$F(1,23) = 7.98, p < .01$], a result of longer listening times overall

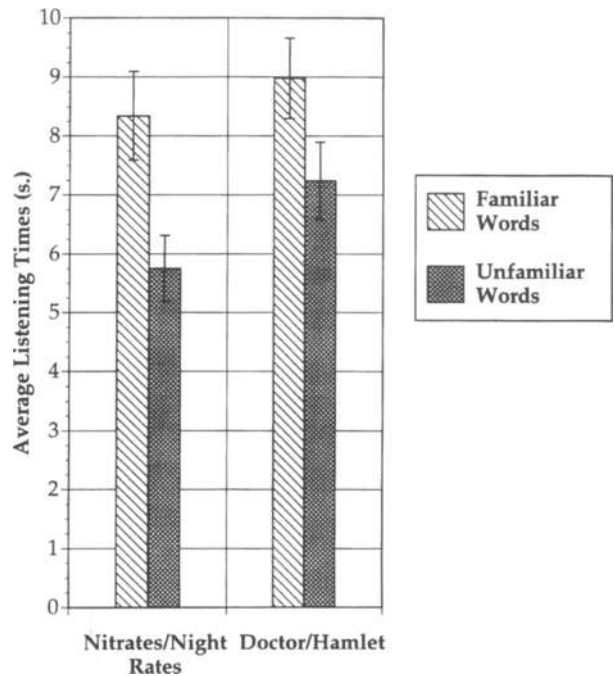


Figure 4. Nine-month-olds' average listening times (and standard error bars) for the passages with the familiar and the unfamiliar target words in Experiment 4.

to the *doctor/hamlet* pair than to the *night rates/nitrates* pair. However, the interaction between experience and word type was not significant [$F(1,23) = 1.67, p > .20$]. The average listening times across all subjects are displayed in Figure 4. A series of planned comparisons, using contrast tests based on the ANOVA, was carried out to explore possible differences for the familiar and the unfamiliar items in the *night rates/nitrates* and *doctor/hamlet* pairs. The differences between the familiar and the unfamiliar items were significant for both the *night rates/nitrates* [$F(1,23) = 31.30, p < .0001$] and the *doctor/hamlet* [$F(1,23) = 14.18, p < .001$] pairs. The finding of a significant listening preference for the familiar item in the *doctor/hamlet* pair replicates that of Experiment 1. By comparison, the significant listening preference that the 10.5-month-olds displayed for the familiar item of the *night rates/nitrates* pair contrasts with the performance of the 9-month-olds in Experiment 1. This latter finding suggests that 10.5-month-olds were sensitive to the distribution of allophonic cues in these words in the fluent speech contexts.

To explore further the apparent developmental trend in sensitivity to how these allophonic cues are distributed within words, we submitted the mean listening times in Experiments 1 and 4 for the familiar and unfamiliar items in the *night rates/nitrates* pair to mixed ANOVA of a 2 (experience: familiar vs. unfamiliar) \times 2 (age: 9 vs. 10.5 months) design. As was expected, there was a significant interaction between experience and age [$F(1,46) =$

11.74, $p < .01$], which was attributable to the fact that the difference between the familiar and the unfamiliar items was significant for the 10.5-month-olds, but not for the 9-month-olds. Hence, sensitivity to how these allophonic cues are distributed within words appears to develop between 9 and 10.5 months of age.

GENERAL DISCUSSION

As was noted earlier, a number of conditions must be met to justify the claim that infants use allophonic cues to segment words in fluent speech. We identified five such conditions: (1) Allophones are not random acoustic variants, but orderly manifestations of phonemic contrasts; (2) the distribution of some of these allophones correlates with word boundaries; (3) infants are capable of discriminating these kinds of allophonic differences; (4) infants are sensitive to how such allophones are distributed within words; and (5) they use these allophonic cues in segmenting words during on-line speech processing. Previous research had provided empirical support for the first three of these conditions. The focus of the present investigation was to determine whether there was empirical support for the fourth condition. Our findings indicate that although 10.5-month-olds display sensitivity to the distribution of allophonic cues within words in fluent speech contexts, 9-month-olds do not.

Do the present findings allow us to conclude further that 10.5-month-olds are actually using allophonic cues in segmenting words? Certainly, the performance of these older infants is consistent with this claim. However, further empirical evidence is required to definitively establish the validity of the claim. For example, in future studies, it will be necessary to demonstrate that the 10.5-month-olds perform better in segmenting the same words from fluent speech when these allophonic cues are present than when they are not present. One way to investigate this issue is to expose infants to passages first and see whether they make use of the word boundary cue in segmenting *rates* from *night rates*, rather than from *nitrates*. Also, will infants respond to *traits* after hearing a passage with *nitrates*, but not after hearing a passage with *night rates*? More generally, probing the kinds of isolated words that infants respond to after their initial exposure to fluent speech passages could help determine whether the sequences that infants extract are affected by the presence of allophonic cues to word boundaries.

Although they did not display sensitivity to how allophones were distributed in *nitrates* and *night rates*, 9-month-olds did consistently show evidence of detecting some types of familiarized words in fluent speech contexts. In particular, they showed listening preferences for the familiar items in the *doctor/hamlet* and *dock/doctor* pairs. Moreover, they demonstrated an ability to use distributional cues to word boundaries in Experiment 3.

The use of allophonic cues to locate boundaries depends on some prior knowledge of which allophones are likely to occur in particular contexts. Infants could gain

this knowledge by noting the distribution of allophones in single word utterances or in the word-like units that they begin to segment from fluent speech at around 7.5 months of age (Jusczyk & Aslin, 1995; Newsome & Jusczyk, 1995; Saffran, Aslin, & Newport, 1996). There certainly is some empirical support for the notion that infants at this age are sensitive to the distributional frequencies of elements within words (Jusczyk et al., 1994; Saffran, Aslin, & Newport, 1996). In fact, a recent investigation has shown that English-learning 9-month-olds are sensitive to the way in which phonotactic sequences line up with likely word boundaries (Mattys, Jusczyk, Luce, & Morgan, 1999). Specifically, Mattys et al. explored how infants responded to CC sequences that were more likely to occur within words or between words in English. They found that when such sequences occurred at locations in which stress-based cues indicated a potential word boundary, infants favored the between-word CC sequences. However, they favored the within-word CC sequences whenever stress-based cues suggested the absence of a word boundary. One implication of the present results is that, between 9 and 10.5 months, infants begin to track the kind of dependencies that hold between certain allophones and likely word boundaries.

Why might English-learners be slower to use allophonic cues to word boundaries than they are to use other types of cues? One possible factor is that these kinds of allophonic differences may be difficult to detect in fluent speech. Although 2-month-olds can discriminate isolated versions of *night rate* and *nirate* (Hohne & Jusczyk, 1994), the added demands associated with processing a stream of continuous speech may prevent 9-month-olds from fully utilizing these capacities. Another possible factor is that the infant learner may require experience with a sufficient number of instances of words in order to learn the mapping between allophones and the contexts in which they appear. An unsegmented utterance provides the learner with two opportunities to map allophones onto contexts, namely—the utterance-initial and the utterance-final positions. By comparison, once the learner can break the input into smaller chunks, using stress-based (Jusczyk, 1997) or distributional cues (Brent & Cartwright, 1996), more contexts are potentially available for learning the mapping of allophones to contexts. In essence this tendency is a *divide and conquer* strategy (Jusczyk, in press), whereby access to a greater number of smaller chunks facilitates the process of extracting other potential cues to word boundaries.

The present findings are consistent with others (Echols et al., 1997; Jusczyk & Aslin, 1995; Morgan & Saffran, 1995; Newsome & Jusczyk, 1995; Saffran, Aslin, & Newport, 1996) in demonstrating that word segmentation capacities are developing in English learners between 7.5 and 10.5 months of age. However, these previous investigations focused on infants' use of either stress-based (Echols et al., 1997; Newsome & Jusczyk, 1995) or distributional (Morgan & Saffran, 1995; Newsome & Jusczyk, 1995; Saffran, Aslin, & Newport, 1996)

cues in word segmentation. The present study provides an indication that 10.5-month-olds can also use allophonic cues as a basis for distinguishing words occurring in fluent speech contexts. At the same time, English-learning infants' capacities for using allophonic cues appear to develop after their ability to use stress-based (7.5 months in Newsome & Jusczyk's study) and distributional cues (8 months in Saffran, Aslin, & Newport's, investigation).

The foregoing discussion suggests that it is unlikely that allophonic cues are the primary means by which English-learners begin to segment words from fluent speech. Nevertheless, allophonic cues provide an additional source of information that listeners can draw on in word segmentation. Of the various kinds of cues to word segmentation that have been suggested (e.g., stress-based, distributional, allophonic, and phonotactic cues), none is completely reliable by itself. For example, an English-listener who relied solely on a strategy of identifying word onsets with the occurrence of stressed syllables would continually mis-segment words beginning with weak syllables. Hence, it seems likely that listeners draw on multiple sources for information about word boundaries in fluent speech and then go with the weight of the evidence (Jusczyk, 1997). Indeed, several computer simulation models of word segmentation have employed a multiple-cue approach with some degree of success (Brent & Cartwright, 1996; Christiansen, Allen, & Seidenberg, 1998). What is not yet clear is the extent to which language learners actually draw on multiple cues in segmenting words from fluent speech. To determine this, future investigations will need to explore how infants' word segmentation skills are affected when different kinds of cues conflict. In this way, it may be possible to identify whether infants tend to rely on one type of word boundary cue more than on another.

REFERENCES

- BATES, E., MARCHMAN, V., THAL, D., FENSON, L., DALE, P., REZNICK, J. S., REILLY, J., & HARTUNG, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, **21**, 85-124.
- BERNSTEIN, J., & FRANCO, H. (1996). Speech recognition by computer. In N. J. Lass (Ed.), *Principles of experimental phonetics* (pp. 408-434). St. Louis: Mosby.
- BOLINGER, D. L., & GERSTMAN, L. J. (1957). Disjuncture as a cue to constraints. *Word*, **13**, 246-255.
- BRENT, M. R. (1997). Toward a unified model of lexical acquisition and lexical access. *Journal of Psycholinguistic Research*, **26**, 363-375.
- BRENT, M. R., & CARTWRIGHT, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, **61**, 93-125.
- CHRISTIANSEN, M. H., ALLEN, J., & SEIDENBERG, M. S. (1998). Learning to segment speech using multiple cues. A connectionist model. *Language & Cognitive Processes*, **13**, 221-268.
- CHURCH, K. [W.] (1987a). Phonological parsing and lexical retrieval. *Cognition*, **25**, 53-69.
- CHURCH, K. W. (1987b). *Phonological parsing in speech recognition*. Dordrecht: Kluwer.
- COLE, R. A., & JAKIMIK, J. (1978). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* Hillsdale, NJ: Erlbaum.
- COLE, R. [A.], & JAKIMIK, J. (1980). How are syllables used to recognize words? *Journal of the Acoustical Society of America*, **67**, 965-970.
- CUTLER, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, **20**, 55-60.
- CUTLER, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 105-121). Cambridge, MA: MIT Press.
- CUTLER, A., & BUTTERFIELD, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory & Language*, **31**, 218-236.
- CUTLER, A., & CARTER, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & 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 & Performance*, **14**, 113-121.
- DAHAN, D., & BRENT, M. R. (in press). On the discovery of novel word-like units from utterances: An artificial language study with implications for native-language acquisition. *Journal of Experimental Psychology: General*.
- ECHOLS, C. H., CROWHURST, M. J., & CHILDERS, J. B. (1997). Perception of rhythmic units in speech by infants and adults. *Journal of Memory & Language*, **36**, 202-225.
- HOCKETT, C. F. (1958). *A course in modern linguistics*. New York: Macmillan.
- HOHNE, E. A., & JUSZYK, P. W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception & Psychophysics*, **56**, 613-623.
- HOUSTON, D., NEWSOME, M., & JUSZYK, P. W. (1995, November). *Infants' strategies of speech segmentation. Clues from weak/strong words*. Paper presented at the 20th Annual Boston University Conference on Language Acquisition, Boston.
- JUSZYK, P. W. (1996, October). *Investigations of the word segmentation abilities of infants*. Paper presented at the 4th International Conference on Spoken Language Processing, Philadelphia.
- JUSZYK, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- JUSZYK, P. W. (1998). Using the headturn preference procedure to study language acquisition. In C. Rovee-Collier, L. P. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 188-204). Stamford, CT: Ablex.
- JUSZYK, P. W. (in press). Dividing and conquering the linguistic input. In M. C. Gruber, K. Olson, & T. Wysocki (Eds.), *CLS 34. Vol. II The panels*. Chicago: University of Chicago Press.
- JUSZYK, P. W., & ASLIN, R. N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, **29**, 1-23.
- JUSZYK, P. W., CUTLER, A., & REDANZ, N. (1993). Preference for the predominant stress patterns of English words. *Child Development*, **64**, 675-687.
- JUSZYK, P. W., LUCE, P. A., & CHARLES-LUCE, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory & Language*, **33**, 630-645.
- KEMLER NELSON, D. G., JUSZYK, P. W., MANDEL, D. R., MYERS, J., TURK, A., & GERKEN, L. A. (1995). The Headturn Preference Procedure for testing auditory perception. *Infant Behavior & Development*, **18**, 111-116.
- KLATT, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, **7**, 279-312.
- KLATT, D. H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169-226). Cambridge, MA: MIT Press.
- LADEFOGED, P. (1975). *A course in phonetics*. New York: Harcourt Brace Jovanovich.
- LEHISTE, I. (1960). *An acoustic-phonetic study of internal open juncture*. New York: Karger.
- LIBERMAN, A. M., & STUDDERT-KENNEDY, M. G. (1978). Phonetic perception. In R. Held, H. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology: Perception* (Vol. 8, pp. 143-178). Berlin: Springer-Verlag.
- MARCUS, S. M. (1984). Recognizing speech: On the mapping from

- sound to word. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 151-163). Hillsdale, NJ: Erlbaum.
- MATTYS, S. L., JUSCZYK, P. W., LUCE, P. A., & MORGAN, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, **38**, 465-494.
- MILLS, C. B. (1980). Effects of the match between listener expectancies and coarticulatory cues on the perception of speech. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 528-535.
- MORGAN, J. L., & SAFFRAN, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, **66**, 911-936.
- MYERS, J., JUSCZYK, P. W., KEMLER NELSON, D. G., CHARLES-LUCE, J., WOODWARD, A., & HIRSH-PASEK, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language*, **23**, 1-30.
- NAKATANI, L., & DUKES, K. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America*, **62**, 714-719.
- NEWSOME, M., & JUSCZYK, P. W. (1995). Do infants use stress as a cue for segmenting fluent speech? In D. MacLaughlin & S. McEwen (Eds.), *Proceedings of the 19th Annual Boston University Conference on Language Development* (Vol. 2, pp. 415-426). Somerville, MA: Cascadilla Press.
- POLKA, L., JUSCZYK, P. W., & RVACHEW, S. (1995). Methods for studying speech perception in infants and children. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 49-89). Timonium, MD: York Press.
- REDDY, R. (1976). Speech recognition by machine: A review. *Proceedings of the IEEE*, **64**, 501-531.
- SAFFRAN, J. R., ASLIN, R. N., & NEWPORT, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926-1928.
- SAFFRAN, J. R., NEWPORT, E. L., & ASLIN, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory & Language*, **35**, 606-621.
- STAGER, C. L., & WERKER, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, **388**, 381-382.
- SUOMI, K. (1993). An outline of a developmental model of adult phonological organization and behavior. *Journal of Phonetics*, **21**, 29-60.
- TINCOFF, R., & JUSCZYK, P. W. (1996, July). *Are word-final sounds perceptually salient for infants?* Paper presented at the LabPhon V, Northwestern University.
- UMEDA, N., & COKER, C. H. (1974). Allophonic variation in American English. *Journal of Phonetics*, **2**, 1-5.
- WAIBEL, A. (1986). Suprasegmentals in very large vocabulary word recognition speech perceptions. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines* (Vol. 1, pp. 159-186). New York: Academic Press.
- WOODWARD, J. Z., & ASLIN, R. N. (1990, April). *Segmentation cues in maternal speech to infants*. Paper presented at the 7th biennial meeting of the International Conference on Infant Studies, Montreal.

(Manuscript received August 20, 1997;
revision accepted for publication September 20, 1998.)