

INVESTIGATIONS OF THE WORD SEGMENTATION ABILITIES OF INFANTS

Peter W. Jusczyk

Department of Psychology and Center for Cognitive Science
State University of New York at Buffalo
Buffalo, NY 14260-4110

ABSTRACT

Evidence that I presented at the last meeting in Yokohama indicated that English-learning infants first show some capacity for segmenting words from fluent speech at about 7.5 months of age. Further studies that we have conducted suggest that English-learning infants initially rely on a prosodically based strategy which may cause them to mis-segment words beginning with weak syllables. However, by 10.5 months, English-learners appear to draw on other potential sources of information to locate word boundaries, even for words beginning with weak syllables. Our studies suggest that infants do retain information about the sound patterns of words for up to 24 hours, and that they generalize across different talker's pronunciations of the same words.

1. INTRODUCTION

A critical ability for any infant engaged in learning language is to segment the speech stream correctly into words. Although the past 25 years of research have revealed a great deal about the speech perception capacities of young infants, relatively little is known about how and when they begin to segment words from speech. This picture has begun to change recently as the development of new test procedures has made it possible to investigate how infants respond to utterances longer than just a few syllables. Jusczyk and Aslin [1] used one such procedure to investigate whether infants who were familiarized with a pair of isolated monosyllabic words, such as "cup" and "dog", subsequently recognized these words when they occurred in fluent speech contexts. Their results indicated that 7.5-month-olds, but not 6-month-olds, did recognize the familiarized words in the fluent speech contexts. Moreover, another group of 7.5-month-olds who were familiarized with passages containing two target words later displayed recognition of isolated versions of the target words. Hence, even when infants' first encountered the target words in a fluent speech context, they were apparently able to segment them well enough to connect them with isolated versions of the same words.

Although Jusczyk and Aslin's study demonstrated that English-learning 7.5-month-olds have some capacity for detecting and recognizing words in fluent speech, it did not identify what cues the infants were drawing on for information about segmenting the utterances into words. At the last meeting, I described some results from an investigation that Mary Newsome, Derek Houston and I conducted to determine whether English learners rely on prosodic cues to identify word boundaries in fluent speech. Our decision to explore prosodic cues to word boundaries was based on several factors. First,

Cutler and her colleagues [2-4] have pointed out that a high proportion of content words in English conversational speech begin with strong syllables. Hence, a strategy that assumes each strong syllable marks the onset of a new word might provide a reasonable first guess about the location of word boundaries in English utterances. Indeed, data from studies with English-speaking adults suggest that they may follow what Cutler and her colleagues have termed the Metrical Segmentation Strategy (MSS) [3, 5, 6]. Moreover, studies with English-learners indicate that by 9-months of age, infants show listening preferences for words with the predominant stress pattern, i.e., ones that begin with strong syllables [7]. This last finding suggests that infants at this age are at least sensitive to the kind of prosodic information that could potentially serve as a marker of word onsets in fluent speech.

In the investigation that I reported, Newsome, Houston, and I examined the ability of 7.5-month-olds to detect bisyllabic words such as "doctor", "candle", "hamlet", and "kingdom" in fluent speech contexts. The use of these items allowed us to test not only whether infants detect these longer items in fluent speech, but also to investigate whether they respond to partial matches involving only the strong syllables of these items. For example, after hearing "candle" and "doctor", would they respond to "can" and "dock"? Our findings showed that (1) 7 1/2-month-olds detect the occurrence of such bisyllabic items in fluent speech contexts, and (2) infants at this age require more than a match of the strong syllables for these items (i.e., familiarization with "doctor" and "candle" does not lead to a tendency to listen longer to samples with "can" and "dock"). Thus, these findings indicate that the infants were doing more than simply responding to the strong syllables of the familiar items. Rather, they suggest that the infants were responding to the sound patterns of the whole words in the fluent speech contexts.

However, we also observed that there were some limits on 7.5-month-olds' abilities to detect bisyllables in fluent speech contexts. In particular, infants at this age who were familiarized with words beginning with a weak syllable (such as "beret", "surprise", "guitar", and "device") did not give evidence of recognizing these words in fluent speech contexts. We suggested that their apparent failure might be attributable to a tendency to assume that strong syllables mark the onset of new words. If so, then an infant hearing a passage containing a word such as "guitar" might interpose a word boundary before the syllable "tar". The result would be that such infants would be more likely to perceive a match between an isolated utterance of "tar" and a passage containing "guitar" than they would to match an isolated utterance of "guitar" to the same passage. In fact, when we did familiarize a group of 7.5-month-olds with just the strong syllables of our weak/strong bisyllables, we

found that they did give evidence of matching the isolated strong syllable (e.g., "vice") to the passage containing the comparable weak/strong word (e.g., "device").

The interpretation that Newsome, Houston, and I offered for these studies with bisyllabic words is that English-learning infants' first attempts at segmenting words from fluent speech may involve the use of a strong/weak template. Although such a template would provide infants with a start in segmenting words that begin with strong syllables, it would result in mis-segmentations of words beginning with weak syllables. Moreover, it might even lead infants to misparse the input so as to detect apparent strong/weak words even in places where there are not any such words. This could occur if a strong syllable were consistently followed by a weak syllable from a following word (e.g., if infants heard "guitar" consistently followed by "is", they might detect "tariz" as a word). Since the last conference, we have conducted exactly such a study with 7.5-month-olds. As predicted by our hypothesis, infants exposed to passages in which a weak/strong word such as "surprise" is consistently followed by the same weak syllable "in" will subsequently falsely recognize the pseudoword "prizin". Hence, although 7.5-month-old English-learners are able to segment the speech stream into word-sized units, their skills at doing so fall short of what might be considered to be true word segmentation.

2. WHEN DO ENGLISH-LEARNERS BEGIN TO SEGMENT WORDS WITH WEAK INITIAL SYLLABLES?

Eventually, the infant learning English must be able to identify words with weak syllable onsets. For this to happen, the infant needs to move beyond any segmentation strategy that associates all word onsets with the occurrence of strong syllables. Other sources of information that might provide clues to the onsets of words include phonotactic and allophonic cues [8]. Thus, one way in which English-learners could improve their word recognition performance is by drawing on multiple sources of information to detect word boundaries.

With respect to phonotactic cues, there is evidence that, during the second half of the first year, infants are developing sensitivity to these features of the native language. For example, with respect to phonotactic cues, it has been shown that Dutch and American 9-month-olds show clear listening preferences for words with sound patterns that are consistent with the phonotactic constraints of their native language. Thus, the American 9-month-olds listened longer to words with English phonotactic patterns than to ones with Dutch phonotactic patterns, whereas Dutch 9-month-olds preferred the Dutch patterns to the English ones [9]. Because the native language words used in this study were ones chosen to be unfamiliar to infants, these results indicate that infants have learned some thing about the kinds of phonotactic sequences that are apt to appear in native language words. Of course, while sensitivity to these properties is a prerequisite to being able to use this information in word segmentation, it does not necessarily follow that infants actually draw on their knowledge of phonotactics to find word boundaries.

In the case of allophonic cues to word boundaries, there is not only evidence that infants are sensitive to these sources of information, but there is also some indication that it does figure in 10.5-month-olds' segmentation of fluent speech. In a study with 2-month-olds, Hohne and Jusczyk [10] found that infants at this age were capable of discriminating the kind of allophonic distinctions that mark the presence or absence of word boundaries in pairs of items such as "nitrate" and "night rate". In a follow-up study with older infants, Jusczyk, Hohne, & Bauman [11] found that even though 9-month-olds might have the capacity to perceive the relevant allophonic differences that cue a distinction between "nitrates" and "night rates", these infants failed to distinguish between these items in fluent speech contexts. That is, 9-month-olds who were familiarized with "nitrates" were just as apt to respond to a passage containing "night rates" as they were to one which contained "nitrates". However, by 10.5-months of age, there was evidence that infants did use the allophonic cues in segmenting the fluent speech passages. In particular, the infants listened longer to the passage containing the word that correctly matched the one which they had been familiarized with. When familiarized with "nitrates" they listened significantly longer to the passage containing "nitrates" than they did to the one containing "night rates".

Given that it appears that older infants are beginning to use nonprosodic sources of information to segment words in fluent speech, Houston and I decided to examine whether 10.5-month-olds might show some capacity to correctly detect words with initial weak syllables. In our first experiment, we familiarized infants with a pair of words (such as "guitar" and "device"), and then measured their listening times to four different six-sentence passages. Two of the passages included repetitions of the familiarized words ("guitar" or "device"); the other two, contained repetitions of two novel words (i.e., "surprise" or "beret"). In contrast to their counterparts at 7.5-months of age, the 10.5-month-olds did listen significantly longer to the passages which included the familiarized words. Hence, 10.5-month-olds do show some ability to segment words beginning with weak syllables from fluent speech.

To gain a better indication of 10.5-month-old English-learners' capacities for segmenting words in fluent speech, we conducted two additional experiments. First, we sought to determine whether these older infants were actually matching the entire familiarized weak/strong word or just the strong syllable of this word to the item that occurred in the passage. To explore this, we familiarized the infants with repetitions of a pair of strong syllables (such as "prize" and "ray") and then presented them with the four different six-sentence passages containing the words "surprise", "beret", "device", and "guitar". Note that two of the passages contained words with strong syllables that matched the syllables that infants had been familiarized with, whereas the other two passages did not. If the 10.5-month-olds in our first experiment had been simply matching the strong syllables of weak/strong words, then the infants in this new experiment should have shown the same pattern of results. Namely, they should have had longer listening times for the passages which included weak/strong words with the same strong syllables as heard during familiarization. In fact, the infants displayed no significant listening preferences for these passages. This result suggests that the infants in the first experiment were indeed segmenting the whole weak/strong

word and not just the strong syllable from the fluent speech passages.

Next, we examined whether, like the 7.5-month-olds, the 10.5-month-olds would mis-segment words from fluent speech when the same weak/strong word was always followed by a weak syllable from a different word. We familiarized the infants with a pair of pseudo words (such as "prizin" and "tariz", or "viceto" and "rayon"), and then presented them with four different six-sentence passages. Two of these contained repetitions of the word sequences "guitar is" or "surprise in", whereas the other two contained repetitions of "device to" or "beret on". Unlike their younger counterparts, the 10.5-month-olds were not fooled by our passages. That is, they did not show any tendency to match the familiarized pseudowords to the passages with the comparable two word sequences.

Hence, by 10.5 months, English-learners appear to use more than just the location of strong syllables to identify the onsets of words in fluent speech. Their behavior suggests that they have moved beyond an approximation to word segmentation based on a strong/weak template to something that approaches the kind of word segmentation abilities that are found in mature speaker/hearers of English. Thus, by 10.5 months, English-learning infants appear to be drawing on multiple cues to word boundaries in fluent speech (see Morgan and Saffran [12] for a similar suggestion). Thus, one possible account of how English-learners develop word segmentation skills is that an initial, prosodically-based, strategy allows them to break fluent speech into smaller chunks. It may be easier for the infant to detect other kinds of potential cues to word boundaries (such as allophonic and phonotactic properties) within these chunks. For example, with these smaller units, infants may be in a better position to note which types of sounds and sound sequences are most apt to occur at both edges of these smaller sized units. Noting the distributional properties of these sounds and sound sequences would be a prerequisite for using this information in segmenting fluent speech.

3. EXPLORING THE GENERALIZABILITY AND DURABILITY OF INFANTS EARLY REPRESENTATIONS OF WORDS

It is impressive that infants who are exposed to repetitions of a pair of words, and then immediately presented with fluent speech passages, are more apt to attend to those containing the familiarized words. However, to what extent does the behavior that infants display in our test situation generalize to their behavior in the real world? For example, can we assume that an infant who hears a particular word repeated on some occasion will actually store some representation of its sound pattern that will allow him or her to recognize the word in fluent speech on some later occasion? Furthermore, is the infant able to generalize from hearing a word repeated by one talker to subsequent productions by other talkers of the same word in fluent speech contexts? Derek Houston, Jennifer Tager, and I have begun to investigate these questions.

First of all, we have examined whether 7.5-month-olds who are familiarized with repetitions of a particular word from one talker will respond to the presence of the same word in fluent

speech produced by another talker. In our first experiment, we used speech produced by two different female talkers. Both talkers recorded the repetitions of the words to be used during familiarization and the full set of test passages. Half of the infants were familiarized with the isolated tokens produced by talker DM and tested on the passages produced by MS. The other half were familiarized with words produced by MS and tested on the passages produced by DM. The words and passages that we recorded were the same as those used by Jusczyk and Aslin [1]. That is, infants were familiarized with either "cup" and "dog" or with "feet" and "bike", and tested on four passages containing each of these words. Despite the fact that there was a change in talker's voice between the familiarization and test phases of the experiment, our results exactly replicated those of Jusczyk and Aslin. Thus, when infants were familiarized with a particular pair of words, they listened significantly longer to the passages containing these words than they did to the passages containing the novel words. Thus, 7.5-month-olds do generalize from tokens produced by one talker to those produced by another talker of the same gender. We are currently testing whether they will generalize across different genders of talkers from a male voice during familiarization to a female during the test phase, and vice versa. Preliminary evidence indicates that 7.5-month-olds are also able to generalize in this situation.

We have also begun to examine how durable infants' representations are of the words heard during familiarization. Our results indicate that when 7.5-month-olds are exposed to repetitions of a particular word for 30 seconds on one day, they appear to store some long-term representation of the sound patterns of the word that is sufficient to detect its occurrence in fluent speech contexts the next day. Thus, infants who are familiarized with "cup" and "dog" on one day, listen longer to the passages containing "cup" and "dog" than they do to ones containing "feet" and "bike". To this point, we have only used utterances produced by the same talker on both days. However, the next step of our investigation is to examine whether infants generalize from tokens produced by one talker on the first day to those produced by a different talker on the next day.

In any case, the findings from these most recent investigations suggest that 7.5-month-old English-learners are able to generalize from tokens produced by one talker to those produced in fluent speech by a different talker, and that they are doing some long-term encoding of the sound patterns of repeated words. Both of these abilities are important prerequisites for building up lexical entries of words in the native language. Hence, it should not be surprising that infants' abilities to extract words from fluent speech contexts are developing at a time when they are showing the earliest signs of word comprehension [13, 14].

4. SUMMARY AND CONCLUSIONS

I have reviewed evidence from a series of studies examining the ability of English learning infants to recognize words in fluent speech contexts. Our results indicate that English-learners' first steps toward segmenting words from fluent speech appear to be based on prosodic cues. Specifically, they appear to identify the onsets of words with the occurrence of strong syllables. This initial segmentation of fluent speech is helpful in locating words with strong/weak stress patterns, but

does lead to mis-segmentations of words with weak/strong stress patterns. By 10.5 months, there is evidence that English-learners are able to draw on other sources of information about word boundaries, with the result that they are successful in detecting the occurrence of familiar words with weak/strong stress patterns in fluent speech. Finally, evidence from our most recent investigations suggests that infants' representations of the sound patterns of repeated words is both durable and generalizable across tokens produced by different talkers.

5. ACKNOWLEDGMENTS

The studies reported here were supported by a research grant from NIH (HD#15795) to the author. I am grateful to Ann Marie Jusczyk for comments that she made on an earlier version of this paper.

6. REFERENCES

1. Jusczyk, P. W. and R. N. Aslin, *Infants' detection of sound patterns of words in fluent speech*. *Cognitive Psychology*, 1995. **29**: p. 1-23.
2. Cutler, A. and D. M. Carter, *The predominance of strong initial syllables in the English vocabulary*. *Computer Speech and Language*, 1987. **2**: p. 133-142.
3. Cutler, A. and D. G. Norris, *The role of strong syllables in segmentation for lexical access*. *Journal of Experimental Psychology: Human Perception and Performance*, 1988. **14**: p. 113-121.
4. Cutler, A., *Exploiting prosodic probabilities in speech segmentation*, in *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, G. T. M. Altmann, Editor. 1990, MIT Press: Cambridge. p. 105-121.
5. Cutler, A. and S. Butterfield, *Rhythmic cues to speech segmentation: Evidence from juncture misperception*. *Journal of Memory and Language*, 1992. **31**: p. 218-236.
6. Cutler, A., *Segmentation problems, rhythmic solutions*. *Lingua*, 1994. **92**: p. 81-104.
7. Jusczyk, P. W., A. Cutler, and N. Redanz, *Preference for the predominant stress patterns of English words*. *Child Development*, 1993. **64**: p. 675-687.
8. Church, K. W., *Phonological parsing in speech recognition*. 1987, Dordrecht: Kluwer Academic Publishers.
9. Jusczyk, P. W., *et al.*, *Infants' sensitivity to the sound patterns of native language words*. *Journal of Memory and Language*, 1993. **32**: p. 402-420.
10. Hohne, E. A. and P. W. Jusczyk, *Two-month-old infants' sensitivity to allophonic differences*. *Perception & Psychophysics*, 1994. **56**: p. 613-623.
11. Jusczyk, P. W., E. A. Hohne, and A. L. Bauman. *Infants' sensitivity to word juncture cues*. in *36th Annual Meeting of the Psychonomic Society*. 1995. Los Angeles, CA.
12. Morgan, J. L. and J. R. Saffran, *Emerging integration of sequential and suprasegmental information in preverbal speech segmentation*. *Child Development*, 1995. **66**: p. 911-936.
13. Benedict, H., *Early lexical development: Comprehension and production*. *Journal of Child Language*, 1979. **6**: p. 183-201.
14. Huttenlocher, J., *The origins of language comprehension*, in *Theories in cognitive psychology*, R. L. Solso, Editor. 1974, Wiley: New York. p. 331-368.