NARROWING THE DISTANCE TO LANGUAGE: ONE STEP AT A TIME

PETER W. JUSCZYK
Departments of Psychology and Cognitive Science, Johns Hopkins University

Infants’ earliest attempts at word segmentation appear to be guided by a single source of information (e.g., English-learners initially rely on the predominant stress pattern of words). This initial strategy successfully identifies many potential words in the input, but mis-segments others. However, simply breaking the input into smaller chunks helps learners to identify other possible cues to the location of word boundaries in utterances. Because no one source of information is completely reliable, listeners must eventually rely on multiple cues to segment words. The development of such skills is not critical for developing a native language vocabulary, but also for acquiring the grammatical organization of utterances. Tracking familiar sound patterns, such as function words and grammatical morphemes, may help in learning about syntactic organization. One factor that facilitates learning about the distribution of such elements is sensitivity to boundaries of prosodic phrases. Access to such linguistically-relevant chunks also helps in tracking the distribution of words in the input.

Educational Objectives: The reader will be introduced to basic findings concerning infant speech perception capacities including (a) the nature of the perceptual capacities that infants possess; (b) changes in speech perception capacities that come about with increased exposure to a particular native language; (c) how infants’ growing understanding of native language sound organization is used in developing word segmentation strategies; and (d) how developing word segmentation abilities may facilitate infants’ discovery of the grammatical organization of utterances.

KEY WORDS: Phonotactics, Context-sensitive allophones, Transitional probabilities, Co-occurrence patterns, Word segmentation, Function words

The early perceptual capacities of infants for discriminating and categorizing speech sounds are well-documented (Aslin, Jusczyk & Pisoni, 1997). Before the age of 6 months, infants demonstrate extraordinary abilities for discriminating subtle differences among speech sounds. They discriminate many different kinds of phonetic contrasts that are present in their own native language.

Address correspondence to Dr. P.W. Jusczyk, Department of Psychology, Ames Hall, Johns Hopkins University, Baltimore, MD 21218-2686. Phone: (410) 516-3836; Fax: (410) 516-4478; E-mail: <Jusczyk@JHU.EDU>.
(Eimas, 1974, 1975; Eimas & Miller, 1980; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Levitt, Jusczyk, Murray, & Carden, 1988; Swoboda, Morse, & Leavitt, 1976; Trehub, 1973). Moreover, they also display some capacity to discriminate contrasts that do not occur in their native language (Streeter, 1976; Trehub, 1976). However, the speech-processing abilities of young infants extend well beyond what is required for picking up distinctions among speech sounds. Specifically, they show some abilities to categorize speech sounds in linguistically relevant ways. For example, infants 6 months of age and younger have some ability to generalize from one talker’s utterances to those of another talker (Jusczyk, Pisoni, & Mullennix, 1992b; Kuhl, 1979, 1983) and to compensate for changes in speaking rates (Eimas & Miller, 1980; Miller & Eimas, 1983). Similarly, soon after birth infants demonstrate some abilities to distinguish utterances in their native language from those of another language with a different rhythmic organization (Bosch & Sebastián-Gallés, 1997; Christophe & Morton, 1998; Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988; Nazzi, Bertoncini, & Mehler, 1998). These early speech-processing capacities of infants enable them to begin to categorize the linguistic input and to track patterns that regularly occur in it. Thus, these capacities provide the foundation for learning about the organization and structure of the native language.

Toward the middle of the first year, the speech-processing capacities of infants begin to evolve from general ones that could apply equally to any linguistic input to ones that are more closely adapted to the organization and structure of sound patterns in the native language. One indication of the increased specialization of these capacities for dealing with the native language is the fact that sensitivity to certain non-native speech contrasts has been shown to decline during the second half year (Best, Lafleur, & McRoberts, 1995; Werker & Desjardins, 1995; Werker & Lalonde, 1988; Werker & Tees, 1984). At the same time, infants also appear to be learning about many features of native language sound organization. For instance, between 6 and 9 months, English-learners begin to display sensitivity to the predominant stress pattern (i.e., strong/weak) of English words (Jusczyk, Cutler, & Redanz, 1993a). Thus, although 6-month-olds are as likely to listen to English words with the less frequently-occurring weak/strong stress pattern (e.g., default) as to ones with strong/weak patterns (e.g., falter), 9-month-olds display a significant listening preference for the latter. Furthermore, during this same period, infants are also becoming more closely attuned to the distribution of both segmental elements in native language utterances. For example, Dutch and English words have very similar prosodic characteristics; words in these languages differ primarily in terms of the phonetic segments and the ordering of these segments (i.e., their phonotactics). At 6-months of age, an American infant listens approximately equally to lists of unfamiliar Dutch versus unfamiliar English words. However, by 9 months, both English- and Dutch-learning infants listen signifi-
cantly longer to the words from their own native language (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993b). Hence, the infants have learned something about the kinds of sounds and sound sequences that are likely to appear in words in their native language. Moreover, infants appear to be developing detailed knowledge of the frequency with which certain phonotactic patterns are likely to occur in words. Thus, English-learning 9-month-olds listen longer to CVC’s with frequently-occurring phonotactic sequences than to ones with less frequently-occurring sequences (Jusczyk, Luce, & Charles Luce, 1994).

THE BEGINNINGS OF WORD SEGMENTATION

Infants’ developing knowledge of the segmental and suprasegmental organization of native language utterances appears to play a critical role in how they begin to segment words from fluent speech. To build a lexicon, it is crucial that learners have some ability to extract words from fluent speech. Relatively few of the words addressed to infants occur as isolated words (van de Weijer, 1998), even when mothers are explicitly asked to teach their infants new words (Aslin, Woodward, LaMendola, & Bever, 1996; Woodward & Aslin, 1990). There is mounting evidence that word segmentation abilities develop considerably between 7.5–10.5-months of age. Jusczyk and Aslin (1995) familiarized infants with different tokens of two words (e.g., “cup” and “dog,” or “feet” and “bike”), and then tested whether they would listen significantly longer to passages that contained these words than to ones that did not. One of the familiarized target words appeared in every sentence of a particular test passage; the other target word appeared in another passage. In addition, there were two other similarly structured passages without either of the familiarized target words. When 6-month-olds were tested, they were equally likely to listen to the passages without the targets as they were to listen to the ones with the targets. However, 7.5-month-olds listened significantly longer to the passages with the target words, suggesting that they were able to detect the targets in the fluent speech contexts.

To determine whether infants at this age would still succeed in segmenting the target words if their initial exposure to these words came in fluent speech contexts, Jusczyk and Aslin conducted an additional experiment. This time, the 7.5-month-olds were familiarized with two of the passages (e.g., the “feet” passage and the “bike” passage), and then were tested on repetitions of four different isolated words (two of which were the same as the targets in the passages). Once again, the infants listened significantly longer to the words that matched those heard during the familiarization phase. Consequently, Jusczyk and Aslin concluded that 7.5-month-olds do have some rudimentary ability to segment the sound patterns of words from fluent speech.

A number of questions can be raised about these findings. For example, are infants actually segmenting words? Clearly, if by words, we mean that the in-
fants are extracting sound patterns with a particular meaning attached, the answer is no. Rather, what infants appear to be extracting is the sound pattern of a potential word in the lexicon. But can we be sure that even the latter claim is true? Might infants only be attending to some salient aspect of the sound patterns that they are familiarized with, such as the vowels? In fact, 7.5-month-olds familiarized with items such as “zeet” and “gike” (Jusczyk & Aslin, 1995) or with “feek” and “bipe” (Tincoff & Jusczyk, 1996) did not show any tendency to match these patterns to passages containing the words feet and bike. In both of these cases, the familiarization items differed minimally in their phonetic properties from the words that appeared in the test passages. Therefore, it appears that English-learning 7.5-month-olds are encoding detailed information about the phonetic properties of the sound patterns of the familiarized words.

Another question that arises in conjunction with these findings is the extent to which infants show any long-term retention of the sound patterns that they are extracting from fluent speech. If these kinds of investigations are tapping the processes involved in building a lexicon, then one might expect that infants might retain information about the patterns that they are familiarized with. Houston, Jusczyk, and Tager (1998) explored this hypothesis by familiarizing 7.5-month-olds with a pair of words, and then waiting 24 hours to test the infants on passages with or without the familiarized targets. Infants tested under these conditions performed about as well as infants who heard the passages without any delay. In another investigation (Jusczyk & Hohne, 1997), 8-month-olds heard audio versions of three stories ten times during a two-week period. Then, after a two-week delay, they were brought into the laboratory and tested on lists of words that had either appeared in the stories or were foils that were matched to in their phonetic characteristics and overall frequency to the story words. The infants who had heard the stories listened longer to the story words than to the foils, whereas a control group of infants who had not heard the stories showed no preference for either type. Taken together, the findings of these two studies suggest that infants are storing information about the sound patterns of potential lexical items, even before they attach meanings to these items.

Still another issue that arises with respect to these early word segmentation abilities concerns the extent to which infants can generalize from words produced by one talker to those of another. There are some indications that 7.5-month-olds’ abilities are limited in this domain. Houston et al. (1998) familiarized infants at this age with isolated repetitions of two target words produced by one talker, and then tested them on passages produced by a different talker. As long as the two talkers were of the same gender, the infants successfully detected the target words in the passages. However, when the talkers were of different genders, the 7.5-month-olds did not listen significantly longer to the passages containing the familiarized targets. By comparison, an
older group of infants, 10.5-month-olds, did generalize across talkers from different genders. Hence, these findings suggest that the word segmentation abilities of 7.5-month-olds are not fully formed, but rather that they continue to develop over the course of several months.

**HOW DO INFANTS SEGMENT WORDS FROM FLUENT SPEECH?**

Because languages differ in their sound organization, the most effective cues for segmenting words in one language will almost certainly differ from those that are optimal in another language. There are a number of different potential sources of information that could be used as indicators of word boundaries in fluent speech. These include: (a) metrical stress cues; (b) phonotactic cues; (c) context-sensitive allophones; and (d) statistical and/or distributional properties. Proponents of metrical stress cues have noted that if the words in the language have a predominant stress pattern, listeners could use this information in a first-pass segmentation strategy. For instance, Cutler and her colleagues have suggested such a Metrical Segmentation Strategy (MSS) for English, whereby listeners might assume that strong syllables mark the onsets of new words in fluent speech (Cutler, 1990; Cutler & Butterfield, 1992; Cutler & Norris, 1988). Other investigators have pointed to the potential utility of phonotactic cues in word segmentation (Brent & Cartwright, 1996; Cairns, Shillcock, Chater, & Levy, 1997). Thus, some types of phonotactic sequences (e.g., [vm], [db], [fh]) never occur within the same English syllable, and rarely occur within words. So when listeners encounter these in fluent speech, they are likely to belong to different words. Similarly, some allophonic variants of particular phonemes are found in certain positions of words and not others. Listeners who are sensitive to the contexts in which these allophones are likely to occur could use this information in inferring the locus of potential word boundaries (Church, 1987; Hockett, 1955; Lehiste, 1960). Finally, listeners might also be able to take advantage of their knowledge of certain co-occurrence relations among syllables to determine whether these are likely or unlikely to occur together within the same word (Saffran, Newport, & Aslin, 1996). For instance, if the transitional probability of one syllable following another is low, the two likely belong to different words. Conversely, if the transitional probabilities are relatively high, the two syllables likely come from the same word.

None of the potential cues to word boundaries just mentioned is completely foolproof for English. Each of these can be shown to fail in one case or another. For example, a complete reliance on MSS would lead an English listener to miss the onset of new words whenever they begin with weak (i.e., unstressed) syllables. Similarly, a listener who postulates a word boundary in a phonotactic sequence such as [vm] will incorrectly divide “pavement” into
two words. Clearly, then, listeners must draw on some combination of these potential cues in segmenting words from fluent speech. Recent investigations suggest that infants begin learning about these various word segmentation cues between 7–11 months of age.

Jusczyk, Houston, and Newsome (in press) have conducted an extensive investigation of English-learners’ use of metrical stress cues in word segmentation. In particular, they examined the possibility that infants might identify the onsets of new words in fluent speech with the occurrence of strong syllables. To test this hypothesis, they familiarized infants with words with either strong/weak (e.g., “kingdom” and “hamlet”) or weak/strong (e.g., “surprise” and “guitar”) stress patterns. The MSS predicts that infants should have an easier time with the former than the latter. Moreover, it predicts that weak/strong words are likely to be missegmented at the strong syllable boundary. These predictions were confirmed when Jusczyk et al. (in press) tested 7.5-month-olds on such items. Infants familiarized with strong/weak words correctly detected these words in fluent speech passages, but they did not detect familiarized weak/strong words when these appeared in such passages. Moreover, 7.5-month-olds familiarized with “prize” and “tar” responded as though they heard these items in passages with the words “surprise” and “guitar.” Therefore, 7.5-month-olds do appear to be identifying word onsets with the occurrence of strong syllables in fluent speech. However, Jusczyk et al. (in press) also found that by 10.5-months, the infants were able to segment weak/strong words from fluent speech. The latter finding suggests that the older infants are relying on more than just metrical stress cues to segment words.

As mentioned above, attention to co-occurrence relations among syllables is another potential source of information about word boundaries. Saffran, Newport, and Aslin (1996; Aslin, Saffran, & Newport, 1998) have shown that 8-month-olds are capable of using this kind of information to infer word boundaries in fluent speech. They familiarized infants for two minutes with a continuous string of connected CV syllables and found that the infants subsequently responded differentially to syllables that consistently co-occurred with each other in the familiarization sequence versus groupings of syllables that only co-occurred with each other relatively infrequently. Similarly, in their investigation, Jusczyk et al. (in press) also found that when a weak/strong word in a passage was consistently followed by the same weak syllable (e.g., “guitar is”), 7.5-month-olds responded as though they detected a word with a strong/weak pattern (i.e., “taris”). Thus, there is evidence that infants do attend to co-occurrence relations in extracting words from fluent speech.

Given that infants are sensitive to such co-occurrence patterns in fluent speech, might they simply respond to any kind of consistently occurring sequence, regardless of whether it comes from the same word or from parts of different words? Suppose the word “oats” is always preceded in a passage by words that end in [b] (e.g., “web,” “crib,” “scrub,” etc.) Will infants segment
the word “boats” from such passages? Mattys and Jusczyk (in preparation) found that English-learning 7.5-month-olds did not incorrectly segment “boats” from such contexts, even though they could correctly segment “boats” from ones in which it really did appear as a lexical item. What appears to inhibit infants from incorrectly segmenting “boats” in a context such as “scrub oats” is the presence of allophonic information that indicates that the word final [b] in “scrub” is not the onset of a new word. Indeed, evidence from other investigations suggests that between 7.5–10.5 months, infants are developing sensitivity to allophonic (Jusczyk, Hohne, & Bauman, in press) and to phonotactic cues (Mattys, Jusczyk, Luce, & Morgan, in press; Morgan, 1998).

In summary, although infants may initially segment words by relying on one source of information, such as English-learners use of metrical stress, developing effective and efficient word segmentation strategies depends on multiple sources of information. Nevertheless, even though reliance on metrical stress cues may sometime lead English-learners to mis-segment words, it may also provide them with increased opportunities to learn about other potential cues to word boundaries. Specifically, using metrical stress cues to divide the input into smaller chunks may allow English-learners to observe that certain kinds of allophones and phonotactic sequences are likely to occur at the beginnings of these chunks, but others are not. In this way, learners may gain knowledge of the distributional properties of allophones and phonotactic sequences.

**BEYOND NOUNS: SEGMENTING OTHER TYPES OF WORDS**

Although an ability to extract nouns from fluent speech is certainly useful, learners also need to be able to extract other kinds of words as well if they are to become fluent users of a particular language. Of course, if all parts of speech had exactly the same sound structure, then an ability to extract nouns would naturally be indicative of an ability to segment other types of words. However, in English, it has been noted that although nouns typically have stress on their initial syllable, verbs often have stress on non-initial syllables (Cassidy & Kelly, 1991; Kelly, 1988, 1992). For example, when the English word “discount” occurs as a noun, the accent falls on the initial syllable; when the same word occurs as a verb, the accent occurs on the second syllable.

Because English-learners initially have an easier time in segmenting words that begin with strong syllables, this stress difference between nouns and verbs could lead to better detection of nouns than verbs in fluent speech. To explore this possibility, Nazzi, Jusczyk, and Bhagirath (in preparation) examined English-learners’ abilities to segment verbs with weak/strong stress patterns, such as “incites” and “discounts,” from fluent speech. They began by testing 10.5-month-olds because of the previous findings indicating that En-
English-learners do not appear to segment nouns with weak/strong stress patterns until this age (Jusczyk et al., submitted). They were surprised to find that 10.5-month-olds did not segment the weak/strong verbs from fluent speech, although 13.5-month-olds did. To determine why English-learning 10.5-month-olds are able to segment weak/strong nouns, but not weak/strong verbs from fluent speech, Nazzi et al. compared their materials with those used in the Jusczyk et al. study. They found that the target words occurred in almost exactly the same contexts in both types of passages. However, the nouns were more likely to be pitch accented in the sentences than were the verbs—an observation that has been made by many investigators who have studied pitch accent (Shattuck-Hufnagel, personal communication). This difference in pitch prominence of nouns and verbs could be one of a number of factors that accounts for why early vocabulary acquisition in English is skewed towards nouns rather than verbs.

Learning to detect the occurrence of nouns and verbs in fluent speech is important in building a lexicon, but to learn about their syntactic properties, infants need to know more about how such words are distributed within utterances. One possibility is that learners discover how such words are positioned with respect to units that are marked prosodically in the input. There is evidence that even 2-month-olds show some sensitivity to the prosodic marking of clauses (Mandel, Jusczyk, & Kemler Nelson, 1994; Mandel, Kemler Nelson, & Jusczyk, 1996), and English-learning 9-month-olds display sensitivity to prosodic phrases (Gerken, Jusczyk, & Mandel, 1994; Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward, & Piwoz, 1992a). Thus, infants may be able to track how the nouns and verbs they detect in fluent speech line up with prosodic phrase boundaries. For example, is a particular word found more typically towards the beginning or end of a prosodic phrase?

Another possibility is that nouns and verbs are marked in relation to other words that are more likely to have a fixed position within such phrasal units, such as function words and other grammatical morphemes. For the latter possibility to be feasible, infants would need some ability to detect function words in fluent speech. Yet, the omissions of these types of words from children’s early utterances has often been interpreted as an indication that they are hard for learners to perceive (Echols & Newport, 1992; Gleitman, Gleitman, Landau, & Wanner, 1988). However, more recent studies have suggested that even when such function words are absent in speech, children still display some sensitivity to the occurrence of these words in speech directed toward them (Gerken, 1994a, 1994b; Gerken, Landau, & Remez, 1990). In a study using an AEP measure, Shafer, Gerken, Shucard, & Shucard (1992) found that 11-month-olds reacted to the substitutions of nonsense syllables for function words in sentences. More recently, these findings have been replicated and extended using a behavioral measure (the Headturn Preference Procedure). Shady, Jusczyk, and Gerken (1998) found that 10.5-month-olds listened sig-
nificantly longer to passages with real function words than to ones where nonsense words were substituted for the function words. This pattern of responding held, even when the nonsense words had very similar phonetic properties to real English function words. By comparison, in another experiment in which nonsense words were substituted for content words (instead of function words), no significant differences in listening times to these two types of passages occurred. These findings suggest that infants in the earlier experiment were not simply responding to the occurrence of nonsense words in some of the passages, but rather they had developed some expectations about the function words that should occur in English utterances.

Shady and Jusczyk (in preparation) examined whether 10.5-month-olds might also show some ability to segment particular function words from fluent speech. In one experiment, they found that infants who had been familiarized with either “this” or “that” listened significantly longer to passages containing these targets. An even more impressive result was obtained in a second experiment that explored whether infants familiarized with “a” or “the” would detect these items in fluent speech. When “a” and “the” occur in fluent speech, they tend to be greatly reduced and phonetically altered compared to when they are produced in citation form. Nevertheless, the 10.5-month-olds did listen significantly longer to the passage that contained the appropriate target word.

What these findings demonstrate is that, by 10.5-months, English-learners are beginning to track certain function words in utterances. Evidence from other investigations suggests that infants may not discover the actual positioning of these words in utterances until between 14- and 16-months (Shady et al., 1998). Nevertheless, the very fact that infants are sensitive to the occurrence of these words allows them to begin to learn about how these words are distributed within prosodic phrases. With respect to “a” and “the,” one observation that they may make is that when these words occur near a prosodic phrase boundary, it will almost certainly always be at the beginning rather than at the end of such a phrase. Moreover, attending to the fact that “a” and “the” occur immediately before some known adjectives and nouns, but not before verbs, could prove useful in identifying the likely syntactic category of some unknown words in an utterance. In this way, the ability to segment and identify certain function words can contribute to the infant’s discovery of the syntactic organization of the native language.

SUMMARY AND CONCLUSIONS

During the second half of the first year, infants acquire much information about the structure and organization of sound patterns in their native language. Their knowledge of this information appears to play a pivotal role in developing abilities to segment words from fluent speech. Word segmentation abilities develop considerably between 7–11 months of age. At the same time, in-
fants also appear to be developing an increased sensitivity to the marking of linguistically relevant units, such as prosodic phrases, in utterances. Together the ability to segment words plus the ability to divide utterances into coherent chunks may allow the infant to learn about the distribution properties of certain elements that recur frequently in speech. Consistent positioning of function words and other such elements within such phrasal units may help in delimiting their possible syntactic roles, as might the frequent co-occurrence of certain elements within such units. In this manner, infants may use information in the speech signal, one step at a time, to move closer to an understanding of the syntactic organization of native language utterances.

REFERENCES


tion morphemes. Paper presented at the Boston University Conference on Language Development, Boston, MA.


CONTINUING EDUCATION

Narrowing the Distance to Language: One Step at a Time

QUESTIONS

1. The speech perception capacities of infants in the first 6 months of life
   a. Are limited to detecting only gross distinctions between sounds.
b. Require considerable native language input to begin functioning.
c. Are effective in distinguishing phonetic contrasts in the native language, but not in foreign languages.
d. Allow infants to discriminate phonetic contrasts that could occur in any of the world’s languages.
e. Are good enough for segmenting words from fluent speech.

2. Which of the following is not an indication that speech perception capacities become more specialized in the course of development?
a. Six-month-olds can generalize from speech by one talker to that of another talker.
b. Nine-month-old English learning infants are more likely to listen to unfamiliar English words than to unfamiliar Dutch words.
d. English-learning 7.5-month-olds segment some types of words from fluent speech.
e. English-learning 9-month-olds are more likely to listen to words with strong/weak rather than weak/strong stress patterns.

3. Which of the following is not true of word segmentation abilities?
a. Infants who are familiarized with words on one day will listen longer to passages containing these words on the next day.
b. English-learning 7.5-month-olds are able to segment words with strong/weak stress patterns from fluent speech.
c. English-learning 10.5-month-olds are able to segment words with weak/strong stress patterns from fluent speech.
d. English-learning 7.5-month-olds can generalize from isolated words produced by a male talker to ones in passages produced by a female talker.
e. An infant who is familiarized with “cut” will not respond to a passage containing the word “cup.”

4. All of the following have been suggested as possible cues to word boundaries in fluent speech except
a. Phonotactic cues
b. Context-sensitive allophones
c. Syntactic categories
d. Metrical stress cues
e. Statistical and/or distributional cues

5. Which of the following is not true of function words?
a. Children often omit these words in their early productions.
b. Infants begin segmenting these words in speech earlier than they do nouns.
c. Eleven-month-olds notice when nonsense words are substituted for real function words in sentences.

d. An ability to detect function words could help learners narrow down the syntactic categories of other unknown words in speech.

e. Learners are not sensitive to the correct positioning of function words until between 14 and 16 months of age.