How Infants Adapt Speech-Processing Capacities to Native-Language Structure

Peter W. Jusczyk[1]

Departments of Psychology and Cognitive Science, Johns Hopkins University, Baltimore, Maryland

Abstract

As infants learn the sound organization of their native language, they use this developing knowledge to make their first attempts to extract the underlying structure of utterances. Although these first attempts fail to capture the full complexity of features that adults use in perceiving and producing utterances, they provide learners with the opportunity to discover additional cues to the underlying structure of the language. Three examples of this developmental pattern are considered: learning the rhythmic organization of the native language, segmenting words from fluent speech, and identifying the correct units of grammatical organization.

Keywords

infant speech perception; word segmentation; prosodic bootstrapping

Infants' excellent abilities to discriminate speech sounds provide them with the foundation for learning about different native-language sound categories. That these initial abilities for discriminating speech sounds are general, as opposed to specialized for perceiving a particular native language, is evident from infants' discrimination of speech contrasts that do not occur in their native language. Nevertheless, within their first year of life, infants' discriminative capacities become more refined and adapted to processing the particular sound organization of their native language (see Jusczyk, 1997, for a review of these early findings). The pattern evident in the development of speech discrimination abilities (i.e., general capacities to categorize elements of the input, followed by the adaptation of these capacities to process the sound organization of a particular language more efficiently) is one repeated at different points during language acquisition. Three additional examples of this developmental pattern are discussed here.

LEARNING RHYTHMIC PROPERTIES OF ONE'S LANGUAGE

Many infants grow up hearing more than one language spoken in their environment. This situation could complicate language acquisition because unless infants keep utterances from different languages separate, they may draw the wrong generalizations about the structure of these languages. What information might infants use to distinguish utterances in one language from those of another? One possibility is that infants are attuned to the rhythmic properties of language and use this information in discriminating utterances from different languages (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998). This hypothesis was developed after Mehler et al. (1988) reported that even newborns have some ability to discriminate utterances in one language (e.g., French) from those in another language (e.g., Russian). Of course, several different speech properties distinguish French from Russian (e.g., differences in the inventories of phonetic elements, the sequences of segments that are permissible, and prosodic properties such as rhythm, pitch contours, and intonation patterns). In another experiment, Mehler et al. played their speech samples through a special filter that cut out any sound infor-

Copyright © 2002 American Psychological Society

Downloaded from cdp.sagepub.com at PENNSYLVANIA STATE UNIV on May 10, 2016

mation above 400 Hz. This removed information about the phonetic segments but left the prosody intact. The infants still discriminated the Russian and French utterances, an indication that the prosodic properties of languages are sufficient for infants to discriminate them.

A subsequent study by Nazzi et al. (1998) suggested that the critical information lies in the rhythmic properties of languages. Three general rhythmic classes have been described for languages: stress-timed, syllable-timed, and mora-timed. Languages such as English and German, which alternate strong syllables (ones containing full vowels) and weak syllables (ones with reduced vowels), are considered to be stress-timed, whereas languages such as French and Spanish, in which each syllable receives equal stress, are said to be syllable-timed. Languages such as Japanese and Tamil are considered to be mora-timed. The mora is a rhythmic unit that can be either syllabic or subsyllabic. In English, a mora roughly corresponds to a syllable consisting of a consonant followed by a short vowel (e.g., "the"), although it may also consist of a consonant alone (e.g., "them" has two moras) or a lengthening of a vowel (e.g., "thee" also has two moras). Nazzi et al. found that French newborns discriminated two languages from different rhythmic classes (e.g., English from Japanese), but not two languages from the same rhythmic class (e.g., English from Dutch). Moreover, during the first few months of life, experience with a particular language does not help in distinguishing it from other languages in the same rhythmic class (e.g., English-learning 2-month-olds do not discriminate Dutch from English).

Therefore, infants begin with a general ability to discriminate languages in one rhythmic class from those in another. However, they do not yet demonstrate an ability to discriminate their native language from another in the same rhythmic class. By 5 months, infants have sufficient familiarity with the rhythmic properties of their native language to distinguish it from other languages in the same rhythmic class (Nazzi, Jusczyk, & Johnson, 2000). For example, American 5-month-olds not only distinguish their native language from Dutch, but also even distinguish their dialect of English (American) from another dialect of English (British). So, although rhythmic properties provide the original basis for infants' discrimination of different languages, by 5 months, they seem to have learned something specific about the properties of native-language utterances. That this knowledge is specific to the native language is shown by their inability to discriminate two unfamiliar languages, German and Dutch, that both belong to the same rhythmic class as their native language.

LEARNING TO SEGMENT WORDS FROM SPEECH

Most utterances are produced without clear pauses in between words. Because most words that infants hear are not produced in isolation, but in the context of other words, a critical step in acquiring language is learning to segment words from speech. An initial study demonstrated that English-learning infants begin to segment words from fluent speech at around 7.5 months of age (Jusczyk & Aslin, 1995). Subsequent research has focused on the information that infants use to locate word boundaries. Early on, they appear to rely heavily on prosodic stress cues to segment words. The predominant stress pattern of content words in English is an initial stressed syllable followed by one or more unstressed syllables. Thus, a first-pass strategy for segmenting words in English is to posit that each stressed syllable marks the onset of a new word (Cutler & Norris, 1988). English-learning infants first display sensitivity to the predominant stress pattern of words between 6 and 9 months. Perhaps it is not surprising, then, that the words that 7.5-month-olds segment in fluent speech begin with a stressed syllable: They are able to segment words, such as "kingdom" and "hamlet," that begin with a stressed syllable, but not words, such as "beret" and "device," that begin with an unstressed syllable (Jusczyk, Houston, & Newsome, 1999). This pattern of behavior is consistent with the view that English-learning infants identify word onsets with stressed syllables. Of course, one disadvantage of this strategy is that it leads to the mis-segmentation of words that begin with unstressed syllables.

Thus, once again, infants begin with a first approximation of what they need to deal effectively with the sound structure of the native language. Segmenting speech at the onsets of stressed syllables certainly helps divide the speech stream into smaller chunks. Attention to how phonetic information patterns within these chunks may then help infants to learn about other potential cues to word boundaries. For example, listeners might use information about contexts in which particular allophones (acoustic variants of a phoneme) appear or knowledge of phonotactics (constraints on which orders of phonetic segments are permitted within words) as cues to likely word boundaries in fluent speech. Indeed, English learners can use phonotactic cues to word boundaries at age 9 months (Mattys & Jusczyk, 2001), whereas sensitivity to allophonic cues for word

Published by Blackwell Publishing Inc.
Downloaded from cdp.sagepub.com at PENNSYLVANIA STATE UNIV on May 10, 2016

segmentation develops by 10.5 months (Jusczyk, Hohne, & Bauman, 1999), the same age at which some ability to segment words beginning with unstressed syllables is demonstrated (Jusczyk, Houston, & Newsome, 1999). Thus, although their initial means of segmenting fluent speech are not fully satisfactory for correctly locating the boundaries of words, English learners' developing sensitivity to native-language sound organization leads to more effective segmentation over the course of a 3-month span.

IDENTIFYING UNITS OF GRAMMATICAL ORGANIZATION

Just as it is important in learning words to segment fluent speech correctly, so too is it crucial for learning the grammar of a language to identify the major constituents (i.e., clauses and phrases) of its organization. Any attempt to discover grammatical organization that begins with misparsed constituents (e.g., fragments of one phrase combined with fragments of an adjacent phrase) would be doomed to fail. It has been suggested that information in the speech signal could help learners to find the correct units of grammatical organization (Gleitman & Wanner, 1982). In order for this kind of bootstrapping from the speech signal to be viable, several conditions must hold. First, there must be information in speech that marks the important grammatical constituents. Second, learners must be capable of detecting such markers. Third, learners must use these markers in processing the linguistic input. In fact, all of these conditions appear to be met with respect to syntactic clauses. Thus, not only are cues to the location of clause boundaries (e.g., lengthening of pauses, falling pitch at the ends of clauses) present in speech addressed to infants, but also infants are sensitive to these cues (Hirsh-Pasek et al., 1987). Moreover, infants use the speech cues about clausal units to encode and remember the information contained in these units (Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000).

Although the case for infants' using speech cues to locate units corresponding to clauses is strong, the case for their using speech cues to locate syntactic phrases seems to be considerably weaker. In part, this is understandable because languages differ considerably in how they organize phrases within clauses. Languages that rely heavily on word order to convey syntactic information are likely to package words from the same phrase together, whereas this does not necessarily hold for languages that allow for freer word orders. Still, even within languages that rely on word order, syntactic phrases are not consistently marked in the input that the child receives, although when such information is available, English-learning 9-month-olds appear to detect it (Jusczyk, 1997), and use it in encoding and remembering information contained in such units (Soderstrom, Jusczyk, & Kemler Nelson, 2000). The problem is that the markers contained in speech relate to prosodic phrases, which do not correspond consistently to particular types of syntactic phrases. To understand this point, consider the following sentences:

(1) Molly / baked the bread.
(2) She baked / the bread.

In (1), the prosodic phrase boundary (indicated by "/") coincides with the syntactic boundary between the subject phrase and the predicate phrase. However, in (2), the prosodic phrase boundary occurs within the predicate phrase, between the verb and its object. English-learning 9-month-olds display listening preferences that accord with the prosodic organization of utterances (Gerken, Jusczyk, & Mandel, 1994). In other words, they prefer to have pauses inserted between prosodic phrases even when they don't correspond to the syntactic phrase boundary between subject and predicate phrases.

At first glance, the fact that prosodic boundaries do not mark the same syntactic boundaries in different sentences seems to rule out the possibility that speech information helps infants to discover syntactic phrases. However, dividing utterances into prosodic phrases could still aid in learning to identify syntactic phrases. The situation is not unlike the one noted for word segmentation. Infants' initial attempts to segment words correctly identified word boundaries for some types of words, but not for others. Parsing fluent speech into chunks beginning with a stressed syllable provides infants the opportunity to learn about other potential word segmentation cues. Might infants derive a similar benefit from grouping information into prosodic phrases?

One source of information within prosodic phrases that could facilitate the discovery of syntactic units is the occurrence of grammatical morphemes. (A morpheme is the smallest meaningful unit in a language; e.g., *walked* consists of the morpheme *walk* plus the morpheme *ed*, which indicates the past tense.) In English, certain function words typically occur only at certain locations inside phrasal units. For instance, "the" marks the beginning of a noun phrase, and is extremely unlikely to occur as the last word of a phrasal unit. Hence, grouping the speech input into prosodic phrases and noting regularities in how certain morphemes are distributed within such phrases may help delineate their syntactic

Copyright © 2002 American Psychological Society
Downloaded from cdp.sagepub.com at PENNSYLVANIA STATE UNIV on May 10, 2016

roles. Moreover, there is evidence that English-learning infants begin to track the occurrence of function words in utterances at age 10.5 months, if not earlier (Shafer, Shucard, Gerken, & Shucard, 1998). Hence, by referencing the positioning of particular function words to prosodic phrase boundaries, as well as noting particular content words that these function words co-occur with frequently within such phrases, infants could gain important information about the distributional properties of these words within utterances. In this way, an initial grouping of words into prosodic phrases may help infants to discover the syntactic organization of their native language.

CONCLUSIONS

Infants learn a great deal about different aspects of the sound organization of their native language in a relatively short span of time (i.e., their first year). In mastering the various necessary skills to become efficient in processing their native language, infants appear to follow a similar developmental pattern, which might be described as "divide and conquer." Their initial attempts yield some approximation of the speech-processing skill that they are trying to master, be it categorization of their language's sound units, discrimination of utterances from different languages, word segmentation, or extracting the clauses and phrases of utterances. Although only partially successful, these initial attempts allow infants to pick up the additional information that is required to align their speech-processing capacities to the particular sound organization of their native language. Still, many important questions remain about the developmental pattern described here. How do infants discover that their initial strategies are inadequate? How do they locate the additional sources of information required to improve their performance? How do infants correctly integrate these new sources of information with the ones used initially in processing some aspect of sound structure? Answers to these questions will further our understanding of language acquisition.

Recommended Reading

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*, 33–69.

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, *1*, 215–219.

Morgan, J.L., & Demuth, K. (1996). *Signal to syntax*. Mahwah, NJ: Erlbaum.

Peters, A. (1983). *The units of language acquisition*. Cambridge, England: Cambridge University Press.

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

Acknowledgments—This article is dedicated to the memory of Alvin M. Liberman, who was one of my great inspirations for entering this field of research. Preparation of this manuscript was aided by a Research Grant from the National Institute of Child Health and Human Development (No. 15795) and a Research Scientist Award from the National Institute of Mental Health (No. 01490). I am grateful to Ann Marie Jusczyk and Elizabeth Johnson for helpful comments on earlier versions of this manuscript.

Note

1. With great sadness, we regret to report that Peter W. Jusczyk passed away on August 23, 2001.

References

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

Gerken, L.A., Jusczyk, P.W., & Mandel, D.R. (1994). When prosody fails to cue syntactic structure: Nine-month-olds' sensitivity to phonological vs. syntactic phrases. *Cognition*, *51*, 237–265.

Gleitman, L., & Wanner, E. (1982). The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge, England: Cambridge University Press.

Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.

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

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

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

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

Mattys, S.L., & Jusczyk, P.W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, *27*, 644–655.

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

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

Nazzi, T., Jusczyk, P.W., & Johnson, E.K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*, 1–19.

Nazzi, T., Kemler Nelson, D.G., Jusczyk, P.W., & Jusczyk, A.M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, *1*, 123–147.

Shafer, V.L., Shucard, J.L., Gerken, L.A., & Shucard, D.W. (1998). "The" and the brain: An electrophysiological examination of infants' responses to familiar and unfamiliar function morphemes. *Journal of Speech, Language and Hearing Research*, *41*, 874–886.

Soderstrom, M., Jusczyk, P.W., & Kemler Nelson, D.G. (2000). Evidence for the use of phrasal packaging by English-learning 9-month-olds. In S.C. Howell, S.A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (Vol. 2, pp. 708–718). Somerville, MA: Cascadilla Press.

Downloaded from cdp.sagepub.com at PENNSYLVANIA STATE UNIV on May 10, 2016