

# The Rhythm-based Prosodic Bootstrapping Hypothesis of Early Language Acquisition: Does It Work for Learning for All Languages?

REIKO MAZUKA

*RIKEN Brain Science Institute and Duke University*

**Abstract:** Research in infant speech perception has demonstrated that young infants are sensitive to prosodic properties of language that are relevant to linguistic rhythm. In the Rhythm-based Prosodic Bootstrapping Hypothesis, it is proposed that infants' early sensitivity to the rhythmic properties of a language will enable them to adopt a metrical speech segmentation strategy appropriate for their language. Review of the literature shows that while infants are capable of discriminating languages from different rhythm classes, this does not necessarily lead them to adopt metrical segmentation strategies appropriate to each and every particular class. It is argued that the rhythm of a language may be salient for infants in all languages, but how this sensitivity is linked to other aspects of language acquisition needs to be re-considered.

**Key words:** prosody, bootstrapping, acquisition, Japanese address for correspondence

## 1. Introduction

In recent years, research in infant speech perception has shown that infants are sensitive to many of the aspects of speech that surround them (cf. Saffran, Werker and Werner 2006). On the segmental side, cumulative research has shown that infants are born with a capability for discriminating among many (but not all) of the contrasts that exist in human languages. During the second half of their first year, this language-general discrimination ability becomes increasingly attuned to the contrasts in their own language, such that their sensitivity to foreign contrasts declines while their ability to discriminate among contrasts in their first language becomes sharper (Werker and Yeung 2005 for review). At the same time, numerous studies have shown that infants are highly sensitive to the suprasegmental components of speech sounds as well. In fact, infants' sensitivity to the prosodic aspects of speech has been claimed to play a critical role in language acquisition, i.e., it allows infants to bootstrap into other aspects of language. This view is sometimes called the "prosodic bootstrapping hypothesis." Although a number of variations of this hypothesis have been put forth, the most concrete version (e.g. Nazzi and Ramus 2003) proposes that the rhythmic organization of a language provides the child learner with a means of segmenting the speech stream into linguistically

significant units. Yet, the research on this hypothesis has exclusively been based on data either from languages with stressed-timed rhythm (i.e. English, Dutch and German), or from those with syllable-timed rhythm (French, Italian and Spanish). As discussed below, Japanese is said to have a third type of rhythm: mora-timed rhythm. In this paper, we will review the research on that version of the prosodic bootstrapping hypothesis that focuses on the rhythmic class of a language as a key feature, and evaluate whether the hypothesis can account for acquiring languages in all three classes of rhythm, that is, in stress-timed, syllable-timed and mora-timed languages. In order for the hypothesis to be valid, it should apply to infants learning any language, including Japanese. In section 2 below, we will first discuss the notion of linguistic rhythm and the acoustic cues that have been argued to correlate with the linguistic rhythm. In section 3, we will introduce the hypothesis of prosodic bootstrapping in early language acquisition, and discuss one particular version of the hypothesis proposed by Nazzi and Ramus (2003), which uses linguistic rhythm as the key feature. We will then review empirical evidence for this hypothesis from stress-timed, syllable-timed and mora-timed languages in section 4. In section 5, we will argue that although infants may have an innate sensitivity to the rhythmic properties of language, the specific proposal at the heart of the Rhythm-based Prosodic Bootstrapping Hypothesis may be inadequate to account for first language acquisition in all three rhythmic classes.

## 2. The Rhythmic Classes of Language

The languages of the world have been described as falling into different rhythmic classes. The first division was made between the stress-timed rhythm of languages such as English, German, and Dutch on the one hand, versus the syllable-timed rhythm of languages such as French, Italian, and Spanish on the other (Pike 1945). In a language with stress-timed rhythm, speakers are said to produce utterances such that stressed syllables come at approximately equal intervals, while in a syllable-timed language, each syllable occurs at a relatively constant rate (Abercrombie 1967). Yet a third type of rhythm is proposed for languages such as Japanese, which is called the “mora-timed” rhythm (Ladefoged 1975; Port, Dalby and O’Dell 1987). In a mora-timed language, each mora is uttered at a regular rate.

- (1) Foot (Ft)  
 |  
 Syllable ( $\sigma$ )  
 |  
 Mora ( $\mu$ )

The mora is a sub-syllabic unit that determines the weight of a syllable, as shown in the prosodic hierarchy in (1). The rhythm of a syllable-timed language is determined by syllables, while the rhythm of a stress-timed language is determined by a foot that alternates stressed and unstressed syllables. Thus, the rhythmic units for the three rhythmic classes are hierarchically related.

The traditional definition of a language’s rhythm is based on an intuitive judg-

ment that the rhythmic unit controls the timing, i.e. that the rhythmic units in each rhythm class are uttered at a regular rate. However, empirical studies that measured actual utterances generally have not been supportive of the isochronies for the proposed rhythmic units (Dauer 1983; Roach 1982; Waner and Arai 2001). For example, Dauer (1983) measured inter-stress intervals in English, Thai, Spanish, Italian and Greek. The study found that the inter-stress intervals of English, a stress-timed language, are not any more regular than those of Spanish, a syllable-timed language.

Lack of empirical support for the isochronies of rhythmic units, however, does not necessarily mean that rhythmic classes do not exist, and proposals have been made to identify other acoustic correlates of rhythmic classes. Dauer (1983) argued that syllable structures are more complex and varied in stress-timed languages than syllable-timed languages, while vowel reduction rarely occurs in syllable-timed languages. Ramus, Nespore and Mehler (1999) measured such variables as the proportion of vocalic intervals (%V), the standard deviation of vocalic intervals ( $\Delta V$ ), and the standard deviation of consonant intervals ( $\Delta C$ ) for sample utterances from English, Dutch, Polish, French, Spanish, Italian, Catalan, and Japanese. When the eight languages were plotted on a graph, with %V on the x-axis and  $\Delta C$  on the y-axis, the stress-timed and syllable-timed languages each formed distinct clusters, separate from the mora-timed Japanese.

In a related study, Grabe and Low (2002) used variability of duration in a pairwise comparison of successive measurements (Pairwise Variability Index, PVI) to compare a wide range of languages. They measured vowel duration and computed nPVI (PVI normalized for each speaker) and rPVI (raw PVI) for intervocalic duration in 18 languages. The results showed that, although prototypical stress-timed languages (English, German and Dutch) and prototypical syllable-timed languages (French and Spanish) fell into distinct clusters, there were languages that did not fit in either of these clusters clearly. Japanese, interestingly, was shown to fit with syllable-timed languages.

These studies suggest that the acoustic correlates of rhythmic classes may not lie in the isochronies of the rhythmic units themselves. Whether or not there is a single acoustic cue that is capable of distinguishing the rhythmic classes of world's languages is still under debate.

### **3. The Rhythm-based Prosodic Bootstrapping Hypothesis**

Research in infant speech perception has shown that infants are sensitive to the prosodic properties of language from very early in life, even before birth. Studies have shown that fetuses can perceive ambient sounds, and that they actually learn some of the prosodic properties of their mothers' speech. For example, Decasper et al. (1994) have shown that, immediately after birth, infants can discern nursery rhymes read aloud by their mothers between the 33 and 37<sup>th</sup> weeks of pregnancy from those not read. At birth, infants also show a preference for their own mothers' speech over that of other women (Spence and Freeman 1996), for their (mothers') native language over foreign languages (Moon, Cooper and Fifer 1993), and

for infant-directed speech over adult-directed speech (Cooper and Aslin 1990).

Among the various prosodic properties of language, its rhythmic characteristics have been shown to be extremely salient to infants. For example, Nazzi, Bertoncini and Mehler (1998) have found that French neonates are capable of discriminating between languages that differ in their rhythmic classes, i.e. between stress-timed languages such as English and Dutch and syllable-timed languages such as Italian and Spanish. At the same time, infants were unable to discriminate between two languages (English and Dutch) that were in the same rhythmic class. The infants in Nazzi et al.'s study had no prior exposure to the languages used in the study, either prenatally or postnatally. Thus, their results indicate that infants' sensitivity to the prosodic properties differentiating stress-timed rhythm from syllable-timed rhythm must be present without prior experience with particular languages.

From this and other evidence, it has been proposed that infants are predisposed to attend to the rhythmic properties of language, and that this predilection gives infants an avenue into other aspects of language learning. One version of the proposal, articulated most clearly in Nazzi and Ramus (2003), states that "infants' sensitivity to rhythm at the utterance/supra-segmental level will allow them to specify the type of rhythm of their native language, and develop the procedure appropriate to its segmentation" (p. 236). The proposal is based on research in adult speech perception, where it has been proposed that adults' speech segmentation is influenced by the metrical system of their native language, and that different metrical segmentation strategies are associated with the three rhythm classes (cf. Cutler et al. 1986; Cutler and Norris 1988; Otake et al. 1993). In this paper, we will refer to this version of the hypothesis as the "Rhythm-based Prosodic Bootstrapping Hypothesis."

#### **4. Evidence for the Rhythm-based Prosodic Bootstrapping Hypothesis**

As Nazzi and Ramus argue, three things must be true in order for the Prosodic Bootstrapping Hypothesis to work: 1) acoustic cues must be present in the input for infants to exploit, 2) infants have to be sensitive to those cues, and 3) infants must use such cues in the process of learning their language. The first point was addressed above. That is, the acoustic cues for rhythmic class may not be found in the straightforward isochronies for rhythmic units. But there may be other cues that correlate with rhythmic class that would provide sufficient information to distinguish among the rhythmic classes.

##### **4.1. Discrimination of languages from different rhythm classes**

The second question is whether infants are sufficiently sensitive to the relevant acoustic cues necessary for prosodic bootstrapping. Two types of evidence are relevant here: infants' sensitivity to cues that distinguish languages from different rhythmic classes, and infants' sensitivity to the prosodic units that define the rhythm of their language. As discussed above, Nazzi et al. (1998) demonstrated that French neonates are capable of discriminating between stress-timed languages (English and Dutch) and syllable-timed languages (Spanish and Italian)

even when they had no prior exposure to any of the languages tested. The fact that French neonates can discriminate between languages from these two rhythm classes suggests that there must be reliable cues that permit a contrast to be made between syllable-timed and stress-timed languages. The languages tested in Nazzi et al. (1998) are “prototypical” stress-timed and syllable-timed languages. As Grabe and Low (2002) point out, however, there are many languages that do not fall neatly into either of these categories. We do not know if infants would be equally apt at discriminating between languages when the languages in question are less prototypical.

Data on mora-timed rhythm is limited, partly because Japanese is the only language that has so far been clearly identified as a mora-timed language. In Nazzi et al. (1998), French neonates were found to distinguish Japanese from English, even when the stimuli were low-pass filtered, indicating that infants were able to discriminate the two languages by prosodic cues alone. This suggests that infants are capable of distinguishing mora-timed rhythm from stress-timed rhythm at birth. Christophe and Morton (1998) showed that English-learning infants at 2 and 3 months of age had difficulty discriminating between French and Japanese,<sup>1</sup> suggesting that discrimination between the mora-timed Japanese and a syllable-timed language may be more difficult. But by 5 months of age, English-learning infants are able to distinguish Japanese from Italian (another syllable-timed language) (Nazzi, Jusczyk and Johnson 2000). These studies show that the prosodic cues that allow one to discriminate Japanese from stress- and syllable-timed languages are available, and furthermore, are accessible to infants without specific experience with Japanese. But it seems that distinguishing a mora-timed language (i.e., Japanese) from a syllable-timed language is more difficult than distinguishing it from a stress-timed language. This is not surprising if one considers that mora-timed rhythm is sometimes argued to be a sub-type of syllable-timed rhythm, and that it is more similar to syllable-timed rhythm than stress-timed rhythm in acoustic measures such as rPVI (Grabe and Low 2002).

Discrimination between languages within the same rhythmic class, by contrast, appears to require some experience with the specific language. Nazzi et al. (2000) showed that even at 5 months of age, infants are not capable of discriminating between languages from the same rhythmic class if both are foreign to them. They are capable of discriminating between them only when one of them is their native language.

For the rhythm-based prosodic bootstrapping hypothesis to work, the first step an infant must take is to identify the rhythm class of the language to be learned. So far, existing data are consistent with the hypothesis, since infants seem to be capable of distinguishing languages from the three rhythm classes without specific

---

<sup>1</sup> Nazzi and Ramus (2003) argue that the lack of discrimination could be due to the methodology Christophe and Morton used in their study. In contrast to the High Amplitude Sucking method previous studies have employed, Christophe and Morton used the habituation/dishabituation method due to the age of the infants in their study.

experience with a particular language.

#### 4.2. Access to rhythm bearing units

The second type of evidence we need to consider is infants' sensitivity to the prosodic units that bear the rhythm in each rhythm class. As discussed above, the foot, the syllable, and the mora are the rhythm-bearing units for stress-timed, syllable-timed, and mora-timed rhythms, respectively. In Nazzi and Ramus' proposal, identifying the rhythm class of their language would allow infants to adopt a metrical segmentation strategy that is appropriate for the rhythm class of their language. In order for this to happen, infants need to be able to access the metrical unit for their language.

Of the three rhythm-bearing units, existing data suggest most strongly that syllables are accessible to infants at birth. Bertoncini et al. (1995) reported that French neonates were able to detect a change in number of syllables from two to three. Van Ooijen et al. (1997) showed that French neonates are capable of discriminating weak-strong syllable sequences from monosyllables, providing additional evidence that infants are sensitive to syllables at birth.

Access to the stress-alternating foot may be slightly more difficult. Van Ooijen et al. (1997) found that neonates were unable to discriminate weak-strong bisyllabic sequences from strong-strong ones. Since French is a weight-insensitive language, one cannot rule out the possibility that the neonates in their study were influenced by prenatal exposure to French. But Jusczyk and Thompson (1978) demonstrated that by 2 months of age at the latest, American infants are able to distinguish strong-weak from weak-strong words, suggesting that the foot is also a unit accessible to infants learning a stress-timed language, long before they are ready to segment words.

By contrast, data available to date point to a different pattern for the mora. Bertoncini et al. (1995) also tested French neonates for their ability to discriminate 2-mora words from 3-mora words, while keeping the number of syllables constant. Their results showed that French neonates did not perform this discrimination, suggesting that the mora may not be as easily accessible as the syllable. No published data on Japanese infants are available yet. But data reported at conferences do not point to early sensitivity to moraic units among Japanese infants. For example, Mazuka and Hayashi (2006) reported that 8 to 10-month-old Japanese infants show preference for a list of three-mora, Heavy-Light bisyllabic words such as /toNna/, over two-mora, Light-Light bisyllabic counterparts such as /tona/. But infants at 4–6 months did not show such a preference. The results show that by 8–10 months of age, they were capable of discriminating between 2- and 3-mora words. But it suggests that younger infants may not be capable of such a distinction. However, since Mazuka and Hayashi were testing infants' preferences, we cannot rule out the possibility that the 4 to 6 month-old infants were capable of discriminating between the lists but had no preference. Mugitani, Kobayashi and Amano (2005) reported that 12-month-old Japanese infants are capable of telling two-mora words from three-mora words when they differ minimally in the dura-

tion of an obstruent closure.

In sum, the evidence does not suggest that the mora is a unit that is easily accessible to infants. It appears as though it may not be until well into the second half of the first year that Japanese infants have access to moraic units. If true, this will present a serious challenge to the Prosodic Bootstrapping Hypothesis. We will return to this point later.

### 4.3 Consequences for rhythm-based prosodic bootstrapping

The third requirement for the Rhythm-based Bootstrapping Hypothesis is that infants actually use the relevant acoustic cues. In the present case, the question is whether or not infants in different rhythm classes use the metrical segmentation strategy appropriate for their language. For this question, evidence comes overwhelmingly from stress-timed languages. English-learning infants are found to use stressed syllables in segmenting continuous speech stimuli. Using Jusczyk and Aslin's (1995) familiarization preference paradigm, Jusczyk, Houston and Newsome (1999) presented 7.5-month-old American infants with strong-weak bisyllabic words embedded within four stories for familiarization. When in the testing phase the infants were presented with a list of familiar words (i.e. the words presented in the familiarization stories) and a list of unfamiliar words, they showed a preference for the familiar words. Six-month-old infants did not show such preference. The results indicate that 7.5-month-old English-learning infants were capable of segmenting trochaic (SW) words from the stories they heard during the familiarization phase. Infants at the same age, however, did not show any sign of segmentation for iambic (WS) words. Houston et al. (2000) found that both English and Dutch 9-month-old infants were able to segment strong-weak Dutch words from Dutch stories, showing that the importance of strong syllables is not limited to English, and that such a preference can be found even in non-L1 stimuli as long as it shares the same rhythmic property.

The evidence for a foot-based segmentation strategy has been found in statistical learning tasks as well. By 8 months of age, infants have been found to be sensitive to the transitional probability of adjacent syllables and can use it to segment "words" from continuous speech when it is the only cue available (Saffran, Aslin and Newport 1996; Aslin, Saffran and Newport 1998). When stress cues were pitted against statistical cues, however, 8-month-old infants relied on stress cues, ignoring the statistical cues (Johnson and Jusczyk 2001). Thiessen and Saffran (2003) showed that at 7 months, infants are able to use statistical cues to segment words when the familiarization stimuli were made of either trochaic or iambic words. But by 9 months, English-learning infants showed a preference for trochaic sequences, irrespective of whether the familiarization stimuli were trochaic or iambic (but see also Thiessen and Saffran 2004).

The influence of stress-timed rhythm is not limited to word segmentation, and a wide range of evidence has been reported from many aspects of the acquisition of stress-timed languages. For example, Jusczyk, Cutler and Redanz (1993) found that 9-month-old English-learning infants show a preference for lists of nonsense

words with a strong-weak stress pattern compared to words with a weak-strong pattern. But they found no preference at 6 months. In early production, English-learning children tend to produce utterances that fit the trochaic template. While children are still in the telegraphic stage, they tend to produce utterances that fit a strong-weak alternation pattern (Allen and Hawkins 1979, 1980; Echols and Newport 1992; Gerken, Landau and Ramez 1990; see also Demuth 1996). This pattern was found in an elicited production task as well (Gerken 1994).

Evidence from syllable-timed languages is not so clear. In a recent paper, Nazzi et al. (2006) reported that French infants showed a very different pattern of word segmentation from English- or Dutch-learning infants. Using the same experimental paradigm as Jusczyk and Aslin (1995), they found that 8-month-old French infants showed no evidence of word or syllable segmentation. By 12 months, French infants seemed to be able to segment final or initial syllables from bisyllabic words, but only at 16 months of age did they show signs of segmenting bisyllabic words. Nazzi et al. (2006) argued that this is evidence showing that French infants are developing a syllable-based segmentation strategy, consistent with the Rhythm-based Prosodic Bootstrapping Hypothesis.

By contrast, Polka, Sundara and Blue (2002) reported that French-Canadian infants in Montreal showed an iambic segmentation pattern at 7.5 months of age. This is the same age at which English-learning infants have been reported to show a trochaic segmentation pattern (Jusczyk, Houston and Newsome 1999), and significantly younger than Nazzi et al.'s Parisian French infants, who showed segmentation of bisyllabic words only at 16 months of age. Although it is difficult to evaluate the details of this study since we only have access to the conference abstract, the difference between the two French studies may be attributable to different dialects of French (Nazzi et al. 2006).

To date, no empirical data are available as to whether Japanese infants use a "mora-based" segmentation strategy. But it would not be possible to use such a strategy unless the mora is accessible. Thus, if Japanese infants' access to the mora is limited early on, a mora-based segmentation strategy would not be an option for them. With older children, there have been reports suggesting that the mora may not play as dominant a role in their phonology as in adult phonology. For example, there is a language game called 'shiritori' (fit the tail) in Japanese, where one takes the last "sound" of a word the previous player produced, and finds a word that starts with that sound. When adults and older children play this game, the last mora is taken as the last "sound", e.g. /bataH/ (butter) may be followed by /ame/ (candy). A four-year-old child in Kubozono's (1993) study, by contrast, took the first mora of the last syllable /ta/ as the last sound, e.g. /takoyaki/ (octopus dumplings). This shows that Japanese children, unlike adults, are not treating a non-syllabic mora as an independent unit of sound. These and other data have led Kubozono and Honma (2002) to argue that syllables (in addition to morae) play a prominent role in Japanese phonology.



### 5. Does the Rhythm-based Prosodic Bootstrapping Hypothesis Work?

So far, we have reviewed the evidence for rhythm-based prosodic bootstrapping. Cumulative data seem to suggest that infants are sensitive to the distinction between rhythm classes. The fact that infants are able to distinguish among these classes without any exposure to languages from different rhythm classes suggests that this ability may be an innate capacity of human infants. However, once infants recognize to which rhythm class their language belongs, it is not clear whether they can readily access the prosodic units by which their language's rhythm is defined, let alone use these units to segment speech. Infants learning stress-timed languages seem to be able to access the stress pattern of words (the foot) early on, and use it for word segmentation, while infants learning syllable-timed languages can access syllables at birth but do not seem to use them for word segmentation until later. We have no data from infants learning a mora-timed language. But it is unlikely that Japanese infants can access the mora early on, or use it for segmentation.

The unevenness of the findings across the three rhythm classes presents serious challenges to the Rhythm-based Prosodic Bootstrapping Hypothesis. The basic idea of the bootstrapping hypothesis is that one type of information, for instance the prosodic cues of a language's rhythm, accessible to infants early in development, enables infants to learn something within another domain which would otherwise be difficult to learn; thus the term "bootstrapping." If it is only helpful for infants learning stress-timed languages, it will not work as a bootstrapping mechanism.

The timing at which French infants were said to begin to show a "syllable-based" segmentation strategy is also a problem (Nazzi et al. 2006). As discussed above, French infants should be able to discern that their language is syllable-timed, and they should also be able to access syllables at birth. Then, they should also be able to deduce that the optimal segmentation strategy for them is syllable-based. Yet, it is not until sometime between 8 and 12 months of age that French infants come to be able to use the syllable-based segmentation strategy. By 12 months, infants already have a sizable recognition vocabulary, and most of them begin to produce recognizable words. If acquisition of initial recognition vocabulary precedes their mastery of the syllable-based segmentation strategy, it means that the infants were able to learn their initial vocabulary by using some other strategy than the metrical segmentation strategy. If they already have some skill that allows them to segment "words" without the help of a syllable-based segmentation strategy, then they do not require bootstrapping.

The situation for Japanese would be even more troublesome. If the mora is indeed not readily accessible by young Japanese infants, they would be required to do the reverse. That is, Japanese infants cannot use a mora-based segmentation strategy until they have "learned" to pay attention to the mora. Instead, they would have to use some other strategy to segment speech, and use that information to learn that the mora is something they need to pay attention to in their language. Of course, we cannot draw any conclusions without directly testing Japanese infants. But this eventuality would be particularly damaging for the Rhythm-based

### Prosodic Bootstrapping Hypothesis.

Faced with these problems, one may argue that we should abandon the basic idea of rhythm-based prosodic bootstrapping. However, the problem is not inherent in the idea of bootstrapping per se, but rather in the specific consequences the hypothesis has been argued to have. The empirical data on infants' ability to discriminate between languages on the basis of their rhythm seem to be robust. As discussed above, numerous studies have shown that infants are highly sensitive to the prosodic aspects of human speech, and that they seem to be attuned to pay attention to the rhythmic aspects of their own language. It would not be surprising if infants were biologically endowed to take advantage of such sensitivities in learning language. The problem is that recognition of a rhythm class does not seem to lead infants automatically to utilize the metrical segmentation strategy based on that rhythm. The precise mechanism by which infants in stress-timed languages are led to demonstrate a foot-based segmentation strategy is not clear, nor is the reason why French infants do not use a syllable-based segmentation strategy. But it is very possible, and even likely, that there are other roles that the prosodic cues of their language could play, for instance in facilitating the acquisition of syntactic structure (e.g. Mazuka 1996, 1998; Gervain et al. 2007).

## 6. Conclusion

In this paper, we have examined whether the Rhythm-based Prosodic Bootstrapping Hypothesis can account for data from all three rhythm classes. We argued that although evidence for infants' sensitivity to the cues of rhythm classes may be robust, data do not seem to support the claim that infants in each rhythm class use a metrical segmentation strategy appropriate to the rhythm class of their language.

Review of the literature also highlights the unevenness of research across the three rhythm classes. Considering the importance that linguistic rhythm plays in the research of early language development, it is not surprising that a large amount of research has been conducted into the acquisition of rhythm by infants. Yet, virtually no data is available on infant learning of one of those three rhythm classes: the mora-timed rhythm class. Research on how Japanese infants learn the mora-timed rhythm of their language can significantly advance our understanding of how linguistic rhythm contributes to the early stages of language acquisition.

## References

- Abercrombie, David (1967) *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Allen, George D. and Sara Hawkins (1979) Trochaic rhythm in children's speech. In: Harry Hollien and Patricia A. Hollien (eds.) *Current issues in phonetic sciences*, 927–933. Amsterdam: John Benjamins.
- Allen, George D. and Sara Hawkins (1980) Phonological rhythm: Definition and development. In: Grace H. Yeni-Komshian, James F. Kavanagh and Charles A. Ferguson (eds.) *Child phonology, vol. 1: Production*, 227–256. New York: Academic Press.
- Aslin, Richard N., Jenny R. Saffran and Elissa L. Newport (1998) Computation of condi-

- tional probability statistics by 8-month-old infants. *Psychological Science* 9: 321–324.
- Bertoncini, Josiane, Caroline Floccia, Thierry Nazzi and Jacques Mehler (1995) Morae and syllables: Rhythmic basis of speech representations in neonates. *Language and Speech* 38: 311–329.
- Christophe, Anne and John Morton (1998) Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science* 1: 215–219.
- Cooper, Robin P. and Richard N. Aslin (1990) Preference for infant-directed speech in the first month after birth. *Child Development* 61: 1584–1595.
- Cutler, Anne, Jacques Mehler, Dennis G. Norris and Juan Segui (1986) The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25: 385–400.
- Cutler, Anne and Dennis Norris (1988) The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113–121.
- Dauer, Rebecca M. (1983) Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51–62.
- DeCasper, Anthony J., Jean-Pierre Lecanuet, Marie-Claire Busnel, Carolyn Granier-Deferre and Roselyne Maugeais (1994) Fetal reactions to recurrent maternal speech. *Infant Behavior and Development* 17: 159–164.
- Demuth, Katherine (1996) Prosodic structure of early words. In James Morgan and Katherine Demuth (eds.) *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 171–184. Mahwah, NJ: Laurence Erlbaum Associates.
- Echols, Catharine H. and Elissa L. Newport (1992) The role of stress and position in determining first words. *Language Acquisition* 2: 189–220.
- Gerken, LouAnn, Barbara Landau and Robert E. Remez (1990) Function morphemes in young children's speech perceptions and production. *Developmental Psychology* 26: 204–216.
- Gerken, LouAnn (1994) Young children's representation of prosodic phonology. *Journal of Memory and Language* 33: 19–38.
- Gervain, Judit, Marina Nespou, Reiko Mazuka, Ryota Horie and Jacques Mehler (2007) Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. Ms.
- Grabe, Esther and Ee Ling Low (2002) Durational variability in speech and the rhythm class hypothesis. In: Carlos Gussenhoven and Natasha Warner (eds.) *Laboratory phonology 7*, 515–546. Berlin: Mouton de Gruyter.
- Houston, Derek M., Peter W. Jusczyk, Cecile Kuijpers, Riet Coolen and Anne Cutler (2000) Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review* 7(3): 504–509.
- Johnson, Elizabeth K. and Peter W. Jusczyk (2001) Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language* 44: 548–567.
- Jusczyk, Peter W. and Elizabeth J. Thompson (1978) Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. *Perception & Psychophysics* 23: 105–109.
- Jusczyk, Peter W., Anne Cutler and Nancy J. Redanz (1993) Preference for the predominant stress patterns of English words. *Child Development* 64: 675–687.
- Jusczyk, Peter W. and Richard N. Aslin (1995) Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 29: 1–23.

- Jusczyk, Peter W., Derek M. Houston and Mary Newsome (1999) The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39: 159–207.
- Kubozono, Haruo (1993) Kodomo no shirotori to mora no kakutoku [Children's "shirotori game" and acquisition of mora]. In: Shosuke Haraguchi (ed.) *Nihongo no moora to onsetukoozoo ni kansuru soogoteiki kenkyuu (2)* [A comprehensive study of the structure of mora and syllable in Japanese]: Research Report of Kakenhi 04207104, 130–137.
- Kubozono, Haruo and Takeshi Honma (2002) *Moora to onsetsu* [Mora and syllable]. Tokyo: Kenkyusha.
- Ladefoged, Peter (1975) *A course in phonetics*. New York: Harcourt Brace Jovanovich.
- Mazuka, Reiko (1996) Prosodic contributions to early setting of a grammatical parameter. In: James Morgan and Katherine Demuth (eds.) *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mazuka, Reiko (1998) *The development of language processing strategies: A cross-linguistic study between Japanese and English*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mazuka, Reiko and Akiko Hayashi (2006) Japanese infants prefer to listen to phonological forms of typical child-directed vocabulary. Paper presented at XVth Biennial International Conference on Infant Studies. Kyoto, Japan.
- Moon, Christine, Robin P. Cooper and William P. Fifer (1993) 2-day-olds prefer their native language. *Infant Behavior and Development* 16: 495–500.
- Mugitani, Ryoko, Tessei Kobayashi and Shigeaki Amano (2005) Nihongo gakushuu nyuuji no sokuon chikaku tokusei [Japanese learning infants' perception of geminate obstruents]. Poster presented at Acoustical Society of Japan.
- Nazzi, Thierry, Josiane Bertoncini and Jacques Mehler (1998) Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24: 756–766.
- Nazzi, Thierry, Peter W. Jusczyk and Elizabeth K. Johnson (2000) Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language* 43: 1–19.
- Nazzi, Thierry and Franck Ramus (2003) Perception and acquisition of linguistic rhythm by infants. *Speech Communication* 41: 233–243.
- Nazzi, Thierry, Iakimova, Galina Iakimova, Bertoncini, Josiane, Frédonie, Sévrine and Alcantara, Carmela, (2006) Early segmentation of fluent speech by infants acquiring French: Emerging evidence for cross-linguistic differences. *Journal of Memory and Language* 54: 283–299.
- Otake, Takashi, Giyoo, Hatano, Anne Cutler and Jacques Mehler (1993) Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language* 32: 258–278.
- Pike, Kenneth L. (1945) *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Polka, Linda, Megha Sundara and Stephanie Blue (2002) The role of language experience in word segmentation: A comparison of English, French, and bilingual infants. *Journal of the Acoustical Society of America* 111: 2455–2455.
- Port, Robert, Jonathan Dalby and Michael O'Dell (1987) Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81: 1574–1585.
- Ramus, Franck, Marina Nespor and Jacques Mehler (1999) Correlates of linguistic rhythm in the speech signal. *Cognition* 72: 1–28.
- Roach, Peter (1982) On the distinction between "stress-timed" and "syllable-timed" languages. In: David Crystal (ed.) *Linguistic controversies*, 73–79. London: Edward Arnold.
- Saffran, Jenny R., Richard N. Aslin and Elissa L. Newport (1996) Statistical learning by 8-

- month-old infants. *Science* 274: 1926–1928.
- Saffran, Jenny R., Janet F. Werker and Lynne A. Werner (2006) The infant's auditory world: Hearing, speech and the beginnings of language. In: Deanna Kuhn and Robert S. Siegler (eds.) *Handbook of child psychology, Volume 2: Cognition, perception, and language*, 58–108. New York: Wiley.
- Spence, Melanie J. and Mark S. Freeman (1996) Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development* 19: 199–212.
- Thiessen, Erik D. and Jenny R. Saffran (2003) When cues collide: Use of stress and statistical cues to word boundaries by 7–9-month-old infants. *Developmental Psychology* 39: 706–716.
- Thiessen, Erik D. and Jenny R. Saffran (2004) Infants' acquisition of stress-based word segmentation strategies. Alejna Brugos, Linnea Micciulla and Christine E. Smith (eds.) *BUCLD 28 Proceedings*, 608–619. Somerville, MA: Cascadia Press.
- Van Ooijen, B., Josiane Bertoncini, A. Sansavini and Jacques Mehler (1997) Do weak syllables count for newborns? *Journal of the Acoustical Society of America* 102: 3735–3741.
- Warner, Natasha and Takayuki Arai (2001) Japanese mora-timing: A review. *Phonetica* 58: 1–25.
- Werker, Janet F. and Henny H. Yeung (2005) Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences* 9: 520–527.

[Received 13 June 2007; Accepted 20 June 2007]

## [要 旨]

### リズム特性に基づく韻律ブートストラッピング仮説: 強勢拍リズム以外の言語での可能性

馬塚れい子

理研脳科学総合研究センター／デューク大学

近年の乳児音声知覚研究により、乳児は言語のリズムを規定する韻律情報に敏感であることがわかってきた。この特徴に注目して、乳児は韻律情報から自分の獲得しようとする言語リズムを特定し、そのリズムに適応した音声分節化方略を採用することが出来るという仮説が提唱されている。これは、リズム特性ベースの韻律ブートストラッピング仮説と呼ばれる。強勢拍リズム、音節リズム、モーラリズムの言語における乳児の言語間弁別、連続音声の分節実験の結果を比較すると、乳児はリズム特性による言語間弁別はできるが、どの言語においても、そのリズム特性に適した分節化方略を用いているわけではない。乳児はリズム特性には敏感であるが、それを音声分節化方略とリンクすることは修正が必要である。