

# Word segmentation in monolingual and bilingual infant learners of English and French

Linda Polka & Megha Sundara

School of Communication Sciences & Disorders, McGill University, Montreal

E-mail: [Linda.polka@mcgill.ca](mailto:Linda.polka@mcgill.ca); [msunda@po-box.mcgill.ca](mailto:msunda@po-box.mcgill.ca)

## ABSTRACT

Word segmentation skills emerge during infancy, but it is unclear to what extent this ability is shaped by experience listening to a specific language or language type. This issue was explored by comparing segmentation of bi-syllabic words in monolingual and bilingual 7.5-month-old learners of French and English. In a native-language condition, monolingual infants segmented bi-syllabic words with the predominant stress pattern of their native language. Monolingual French infants also segmented in a different dialect of French, whereas both monolingual groups failed in a cross-language test, i.e. English infants failed to segment in French and vice versa. These findings support the hypothesis that word segmentation is shaped by infant sensitivity to the rhythmic structure of their native language. Our finding that bilingual infants segment bi-syllabic words in two native languages at the same age as their monolingual peers shows that dual language exposure does not delay the emergence of this skill.

## Introduction

A fundamental skill in language processing is the ability to extract and retain words from fluent continuous speech. This is not a trivial task. Word boundaries are rarely marked by pauses in connected speech and the particular boundary cues and distributional patterns that support word segmentation are language-specific. Hence, the perceiver's task is to discover the information that is most efficient for segmenting words in their native language or languages. There are many questions to answer to understand the development of this skill.

We know that rhythmic structure contributes to word segmentation processes in adults [1]. Given the salience of prosody in early infancy, Cutler [1] and others [2] have claimed that rhythmic structure is accessed and utilized to segment words from a very early age. Research on word segmentation in infants provides more direct support for this view. The first studies of infant word segmentation pointed to a perceptual bias based on native-language prosody. Jusczyk and colleagues [2, 3] observed that by 7.5 months of age English-learning infants segment one-syllable words and two-syllable words that have a trochaic stress pattern (stress on the first syllable), the predominant stress pattern for English disyllables. English

words may also have iambic stress (stress on the last syllable), but this pattern occurs less frequently. At 7.5 months, English babies do not segment bi-syllabic iambic words, but they can do so by approximately 9 months of age. Jusczyk et al [2] provided further evidence that English-learning 7.5-month-olds use their familiarity with native language rhythm to begin to solve the segmentation problem. They found that at 7.5 months, English babies demonstrate a trochaic bias in their word segmentation even when this results in a mis-segmentation.

These findings led Jusczyk et al to hypothesize that infants approach the word segmentation problem by exploiting native-language rhythmic cues. One prediction following from this view is that infants should be able to apply their native-language rhythmic strategy to successfully segment words in a language with a similar rhythmic structure but not to a language with different rhythmic structure. Several findings uphold this prediction. Houston et al [4] has shown that at 9 months English babies can segment bi-syllabic trochaic words in Dutch, and Dutch babies can segment bi-syllabic trochaic words in English. Data reported by Höhle et al [5] have shown similar results with respect to German and English. Research on infant word segmentation in a rhythmically different language is quite limited. Newman & Jusczyk [reported in 6] found that English-learning babies could not segment Chinese words from a Chinese passage, although data on Chinese-learning infants was not available.

Presently, there is no data on segmentation in babies learning non-stress timed languages. Thus, findings to date may be interpreted as evidence that the word segmentation skills are supported by general cognitive or speech processing abilities that emerge around 8 months of age. Therefore, the hypothesis that word segmentation in infants is shaped by experience listening to a specific language-type is not firmly supported. To address this basic issue, we report a series of experiments that compare segmentation of bi-syllabic words in infants who are learning rhythmically different languages, including infants from English-speaking, French-speaking, and bilingual (French/English) families.

## Experiment 1: Native language segmentation

By 7.5 months of age, infants learning a stress-timed language are able to segment bi-syllabic words that have

the predominant stress pattern of their native language. Experiment 1 was designed to assess whether infants learning a syllable-timed language in which bi-syllabic words have a fixed iambic stress pattern display comparable word segmentation skills. To do so, we tested both French-learning and English-learning infants on segmentation of bi-syllabic words in their native language.

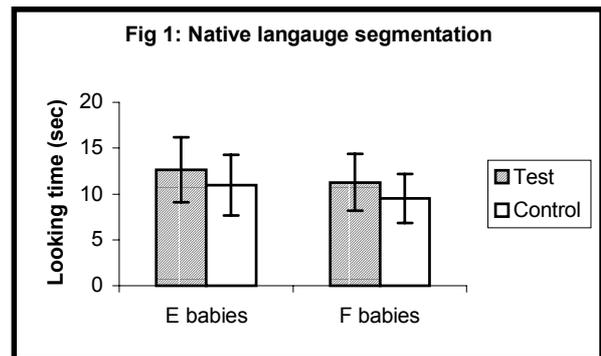
**Procedure:** We implemented the HPP to assess word segmentation as described in Jusczyk and Aslin [3]. This procedure has 2 stages, a familiarization stage followed by a test stage. In the familiarization stage the infant hears a word repeated when they look at a flashing light; two different words are presented (in random order) during this stage until the infant has accumulated 30 seconds of listening time to each word. In the test stage, the infant hears 4 passages. Two passages are test passages in which one of the familiarized words occurs repeatedly throughout the passage and two passages are control passages in which a novel word occurs repeatedly throughout. In the test stage, we record how long the infant chooses to listen (by looking) to each passage. We then compare the average listening time to the control and test passages. If the infant detected the occurrence of the familiarized word in the test passage, we expect them to listen longer to the test than the control passages.

**Stimuli:** Two sets of word lists and passages were developed, one in French and one in English. A female bilingual speaker of Canadian English and Canadian French produced both stimulus sets. She had acquired English and French from birth and used both languages regularly throughout her life; she had no detectable accent in either language. In each language she recorded 4 bi-syllabic words and 4 passages. The English words were ‘hamlet’, ‘kingdom’, ‘doctor’ and ‘candle’; the French words were ‘beret’, ‘surprise’, ‘devis’ and ‘guitar’. Each passage featured one of the bi-syllabic words. Each passage consisted of 6 sentences with the target word occurring once per sentence; twice at beginning; twice in middle, and twice in sentence-final position. Acoustic analysis of the target words in the passages showed that the first syllable of each English word was longer, louder, and higher in pitch compared to the second syllable. Acoustic analysis of the French words in the passages revealed that the second syllable of each French word was longer than the first syllable. The duration difference observed for the French words was larger compared to the duration difference observed for the English words. There were no differences in the pitch or amplitude across the two syllables in the French words.

Two test conditions were prepared using the English stimuli. In condition A, the words ‘doctor’ and ‘candle’ were presented during familiarization stage; in condition B the words ‘kingdom’ and ‘hamlet’ were presented during the familiarization stage. In both conditions the 4 English passages were presented during the test stage; recall that each test passage featured one of the 4 English words. Similar conditions were developed in French. In condition

A, the words ‘guitare’ and ‘devis’ were presented in familiarization; in condition B the words ‘beret’ and ‘surprise’ during familiarization. In both conditions, the 4 French test passages were presented in the test stage; recall that each test passage featured one of the 4 French words.

**Subjects:** Sixteen Canadian English-learning infants and sixteen Canadian French-learning infants were tested. Infants in both language groups were between 7mo and 8 days and 8mo and 15 days of age and were full-term.



**Results and Conclusions:** As shown in Figure 1, both English and French babies listened significantly longer to the test passages than the control passages. In the English group, 10 of the 16 infants showed the expected pattern; in the French group 15 out of 16 showed the expected pattern. These findings show that infants who are learning either English or French are able to segment bi-syllabic words with the characteristic or predominant word stress pattern of their native language. Thus, although the rhythmic properties of words differ in English and French, it appears that word segmentation skills emerge on a similar schedule in babies learning either French or English.

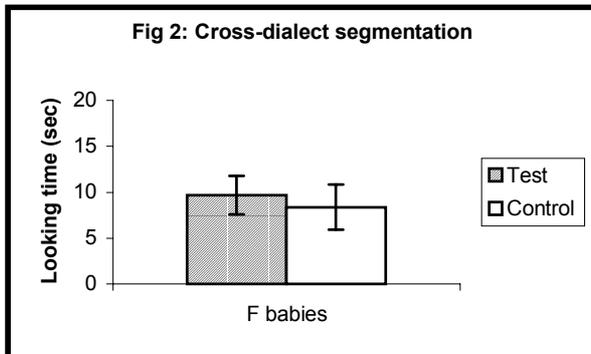
## Experiment 2: Cross-dialect segmentation

Experiment 2 was conducted to replicate and extend the word segmentation findings with French infants. To do so, we tested Canadian French-learning 8-month-olds using European-French stimuli to determine whether Canadian French-learning infants are able to segment in another dialect of French.

**Procedure and Stimuli:** The procedure was identical to Experiment 1. For this experiment, a native French speaker from Paris recorded the same French words and passages used in Experiment 1. This European French speaker had a faster speaking rate than the Canadian French speaker recorded in experiment 1. As in Experiment 1, acoustic analysis of the words in the passages showed that the second syllable of each French word was longer than the first syllable. These duration differences were smaller for the European French words compared to the Canadian French words, but still larger than the duration differences observed in the English words. Neither amplitude nor pitch differed across syllables in the words spoken by this

European French speaker.

**Subjects:** Sixteen French-learning infants between 7mo 14 days and 8mo 15 days of age were tested. All infants were full-term.



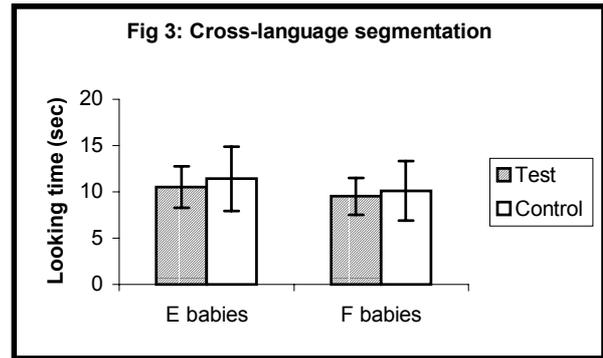
**Results and Conclusions:** As shown in Figure 2, French babies tested on European French speech listened significantly longer to the test passages than the control passages. Twelve out of 16 infants tested showed this pattern. Thus, infants who are acquiring Canadian French can segment bi-syllabic words in a European dialect of French. The present findings show that 8-month-olds learning a syllable-timed language are able to transfer their segmentation abilities to a rhythmically similar dialect, just as babies learning a stressed-time language are able to transfer their segmentation skills to a different stress-timed language.

### Experiment 3: Cross-language segmentation

Findings from experiment 1 and 2 are consistent with an account of segmentation abilities based either on general cognitive skill or specific language experience. We designed experiment 3 to disambiguate between these two hypotheses. In experiment 3, we investigated whether or not English- and French-learning 8-month-olds can segment in a rhythmically different language. To do so, we used a cross-language design, testing English-learning infants on the French stimuli and French-learning infants on English stimuli. A general cognitive-skill based account would predict that infants would be able to transfer skills across languages from different rhythm groups. Alternatively, a specific-language experience based account would predict that infants would fail to transfer skills across languages from different rhythm groups.

**Procedure and Stimuli:** Identical to Experiment 1.

**Subjects:** Sixteen Canadian English-learning infants and sixteen Canadian French-learning infants were tested. Infants in both language groups were between 7mo and 5 days and 8 mo and 15days of age and were full-term



**Results and Conclusions:** As shown in Figure 3, English infants did not listen significantly longer to the test passages when listening to French; roughly half listen longer to the test passages and half listen longer to control passages. Likewise, French infants did not listen significantly longer to the test passages when listening to English; 3 listened longer to the test passages; 3 had almost identical listening times for test and control passages; and 5 listened longer to the control passages. Previous research has shown that infants can transfer their segmentation to a rhythmically similar language or dialect. These findings show that infants cannot transfer their segmentation skills to a rhythmically different language. This finding confirms that infant word segmentation skills are shaped by native-language experience and hence require more than general cognitive or speech processing abilities.

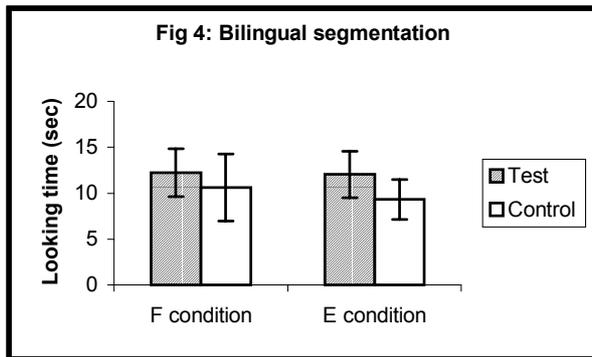
### Experiment 4: Segmentation in bilingual infants

With respect to language experience, it is also interesting to consider how word segmentation might develop in infants who are acquiring two rhythmically different languages simultaneously. Is it more demanding to learn to segment in two languages? Do infants being raised bilingually receive enough input in each of their native languages to support development of segmentation skills on the same schedule as their monolingual peers? Recent findings by Bosch and Sebastian-Galles [7] indicate that bilingual infants can discriminate their two native languages. However, to our knowledge, there have been no studies on word segmentation in infants who are being raised bilingually. Experiment 4 was designed to see if babies being raised bilingually in French and English segment bi-syllabic words in both languages at the same age as their monolingual peers.

**Procedure and stimuli:** The stimuli and procedure are identical to experiment 1, except that in this experiment each infant was tested on both French and English stimuli. Infants were tested in each language on a different day one week apart with order counterbalanced across infants.

**Subjects:** We tested 9 bilingual infants who are being raised with consistent, in most cases daily, exposure to English and French from their primary caregivers. We estimated the rate of exposure to each language; infants

were included only if they had 30 to 70 percent of their total language exposure in each of the two languages.



**Results and Conclusions:** As shown in Figure 4, bilingual infants listened significantly longer to the test passages than the control passages in both English and French. So far 6 of the 9 babies tested show this pattern. These results suggest that 8-month-olds with experience listening to both English and French from birth can segment bi-syllabic words in both languages. If upheld in a larger sample, this finding attests to the sophistication of infant speech processing skills.

### General Discussion

Overall, our findings indicate that at 8 months of age both French-learning and English-learning infants are able to segment bi-syllabic words in their native language. Infants are able to transfer this skill to a rhythmically similar dialect but not to a rhythmically different language, consistent with Jusczyk's hypothesis that word segmentation is strongly influenced by rhythmic structure of the infant's native language. Preliminary results for infants being raised bilingually in French and English, indicate that these infants are able to segment bi-syllabic words in both of their native languages at the same age as their monolingual peers. These findings further illuminate the remarkable capacity of infants for language learning, showing that infants can deal with the demanding task of segmenting words in two rhythmically different languages.

The findings presented here indicate that a trochaic bias is not the default strategy for all babies. Rather, this bias appears to be the language-specific implementation of a general rhythmic-based strategy that is used by babies learning stress-timed languages. The next stage of our research will explore how infants learning French begin to solve the segmentation problem. In particular, we are interested in finding out whether French-learning babies rely initially on an iambic template or adopt a syllable-based strategy. As well, we will be exploring how bilingual infants solve the segmentation problem to find out if they build two segmentation strategies that are comparable to their monolingual peers or adopt a unique approach that is optimal for dual language input.

### REFERENCES

- [1] A. Cutler, "Segmentation problems, rhythmic solutions," *Lingua*, vol. 92, pp. 81-104, 1994.
- [2] P. W. Jusczyk, D Houston and M Newsome, "The beginnings of word segmentation in English learning infants," *Cognitive Psychology*, vol. 39, pp. 159-207, 1999.
- [3] P. W. Jusczyk and R. N. Aslin, "Infants' detection of the sound patterns of words in fluent speech," *Cognitive Psychology*, vol. 29, pp. 1-23, 1995.
- [4] D. Houston, P. W. Jusczyk, C. Kuijpers, R. Coolen, A. Cutler, "Both Dutch- and English- learning 9-month-olds segment Dutch words from fluent speech," *Psychological Bulletin and Review*, vol. 7, pp. 504-509, 2000.
- [5] J. Weissenborn, B. Höhle, S. Bartels, B. Herold, M. Hofmann, "Development of prosodic competence in German infants," In 13<sup>th</sup> Biennial International Conference on Infant Studies, 2002.
- [6] P. W. Jusczyk, "A reply to Littman and Denenberg" *Science*, vol. 280, pp. 1176-1177, 1998.
- [7] L. Bosch and N. Sebastien-Galles, "Evidence of early language discrimination abilities in infants from bilingual environments" *Infancy*, vol. 2, pp. 29-49, 2001.