Lund University, Dept. of Linguistics
Working Papers 49 (2001), 90–93

Phonotactic and word accent cues for speech segmentation in Swedish one-year-olds

Eeva Koponen
Dept. of Linguistics, Stockholm University
eevak@ling.su.se

Abstract

Two experiments were designed to examine 12-month-old Swedish-learning infants' ability to use phonotactic and word accent cues to word boundaries. Twenty-five infants were tested using the so-called Visual Preference Procedure. The stimuli were CVC.CVC (experiment 1) and CVC.CVCCVC nonsense words (experiment 2). The word accent of the stimuli signalled either a single word-like unit or two separate units. Also, the cross-syllabic C.C clusters of the stimuli were either typically occurring within words or across word boundaries. The infants looking times at images of objects, presented along the nonsense words, were measured. Compared with results of similar experiments on English-learning infants, the results of the present study were not clear-cut: the infants did not specifically respond to word boundary cues. Instead, a preference for accent 2 stimuli and stimuli with typical word internal phonotactics was found.

1 Background and purpose

Unlike printed text, fluent speech does not provide systematic spaces between words. How do infants, who have not yet built up a vocabulary, manage to detect word units? A first step is presumably to learn language-specific probabilistic cues to word boundaries.

During the second half of the first year, infants learn how stress and segments are typically distributed within words (Mattys & Jusczyk, 2001). Specifically, cues to word boundaries are provided by prosody, such as word stress patterns, and by probabilistic phonotactics; i.e., distributional regularities in phoneme sequences. In addition, the phonetic realizations of phonemes, depending on their position in words and on the surrounding segments, are known to function as segmentation cues. When considered as *distributional frequencies*, all these cues can be thought of as probabilistic cues (Johnson & Jusczyk, in press).

Mattys et al. (1999) investigated 9-month-old English-learners' sensitivity to prosodic and phonotactic cues. The stimuli used, in their Headturn Preference experiments, were bisyllabic CVC.CVC nonsense words (henceforth non-words) with primary stress on either the first (strong-weak, S/W) or the second syllable (weak-strong, W/S). The cross-syllabic C.C clusters had a high probability of occurring either within words (e.g. *mo<u>ft</u>uth*) or at word junctures (e.g. *mo<u>fh</u>uth*). The infants' looking times at flashing lights, presented along the stimuli, were measured. It was found that, among non-words with the S/W stress pattern, infants preferred within-word clusters over between-word clusters; i.e., they may have interpreted the S/W stimuli as one unit. In contrast, among W/S stimuli, infants preferred between-word clusters over within-word clusters; i.e., they may have interpreted the W/S stimuli as two units, the second stressed syllable functioning as a word onset marker.

The purpose of the present study was to provide comparable data for Swedish-learning infants. Thus, Swedish one-year-olds' ability to use phonotactic and word accent cues was investigated.

2 Experiment 1

2.1 Method

A version of the Preferential Listening Procedure (Fernald, 1985), the Visual Preference Procedure, was used. The infant was seated on the parent's lap. A television monitor showed a sequence of images of objects. Each object was accompanied by a set of stimuli which segmentation cues were either in agreement or in conflict (see Table 1). A video camera recorded the infant watching the film. If the infant looked away from an object for consecutive 10 seconds, the experimenter changed the film to the next object, but if the infant kept looking, the object was shown for the entire duration of the audio file. The infants' exact looking times were measured off-line and used as a measure of attention to the respective audio files.

Nineteen infants (11 females, 8 males, mean age 372,5 days, age range 356 to 392 days) participated in experiment 1. The infants were drawn randomly from the National Swedish Address Register (SPAR). A total of 24 video recordings were made: 5 infants participated twice, and 14 participated once in experiment 1. One additional infant was tested, but the recording was interrupted because the infant was crying.

The stimuli were accent 2 (CV̀C.CV́C) and accent 1+1 non-words (CV́C.CV́C). In accordance with Swedish word prosody rules, the two-peaked fundamental frequency (f_0) of word accent 2 was expected to hold the CVC.CVC syllables together signaling a single word-like unit. In contrast, the two concatenated accent 1 CVC monosyllables were expected to signal two separate word-like units. The stimuli were recorded by a female native speaker of Swedish. To avoid coarticulatory effects, each accent 1 CVC syllable was recorded as a monosyllable (in a carrier phrase), and each syllable of the accent 2 non-words was recorded in a controlled coarticulatory context. Each of the first accent 1 CVC syllables was simply concatenated, without a pause, with each of the second accent 1 CVC syllables, as well as each of the first accent 2 CVC syllables was concatenated with each of the second accent 2 CVC syllables.

The cross-syllabic C.C clusters were chosen on the basis of statistics from Sundberg's (1998) Swedish Infant-Directed Speech (SWIDS) corpus. The clusters with a high probability of occurring within words (*mo<u>sk</u>ut, mo<u>st</u>ut, mo<u>nt</u>ut, mo<u>br</u>ut, mo<u>lk</u>ut, mo<u>tr</u>ut*) were expected to signal one word-like unit, and the C.C clusters with high probability of occurring between words (*mo<u>nd</u>ut, mo<u>ns</u>ut, mo<u>nh</u>ut, mo<u>nv</u>ut, mo<u>nk</u>ut, mo<u>nf</u>ut*) were expected to signal two units. The hypothesis was that the listening preference of the infants – irrespective of whether the cues in agreement or the cues in conflict should lead to longer listening – should be reversed with opposite phonotactic or prosodic cues to word boundaries.

Table 1. The segmentation cues of the stimuli were either in agreement, both indicating one unit or two units, or in conflict, one indicating one unit and the other two units.

	accent 2 (one unit)	**accent 1+1 (two units)**
within-word phonotactics (one unit)	agreement	conflict
between-word phonotactics (two units)	conflict	agreement

2.2 Results and discussion

The results of experiment 1 showed following (see Figure 1): among non-words with within-word phonotactics, accent 1+1 words were preferred over accent 2 words, among non-words with between-word phonotactics, accent 2 words were preferred over accent 1+1 words, among accent 2 non-words, words with within-word phonotactics were preferred over between-word phonotactics, and among accent 1+1 non-words, words with within-word phonotactics were preferred over between-word phonotactics. In sum, by altering the phonotactic cue, the infants' preference for word accent was reversed, but by altering the word accent, the infants' preference for phonotactics was not reversed. Instead, a preference for within-word phonotactics was found.

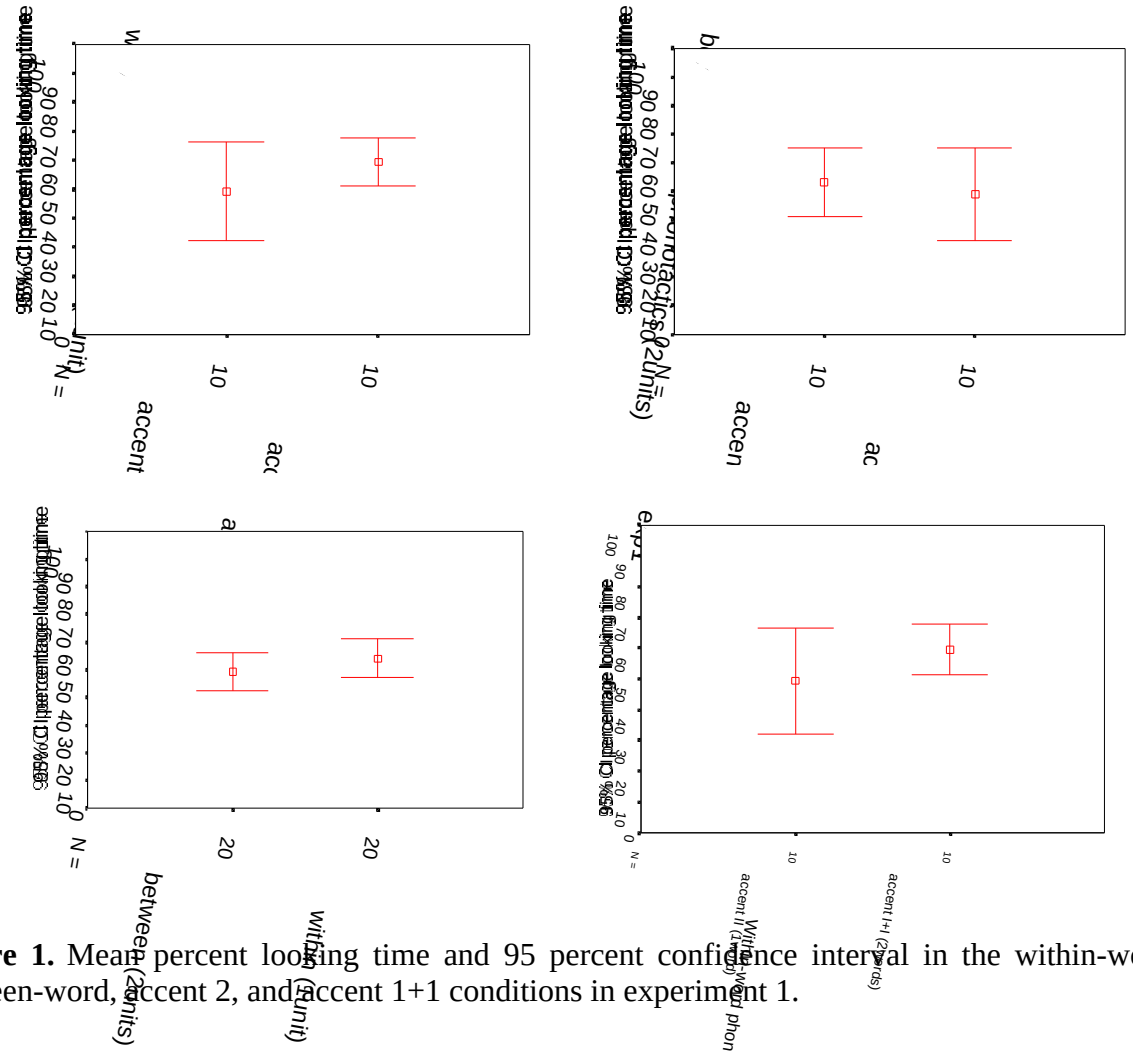

Figure 1. Mean percent looking time and 95 percent confidence interval in the within-word, between-word, accent 2, and accent 1+1 conditions in experiment 1.

A possible weakness of the stimuli in experiment 1 was that the word accent contrast was not very salient: the f_0 contours of accent 2 and accent 1+1 stimuli may have been too similar. Also the within-word stimuli chosen from the SWIDS corpus also had a relatively high probability of occurring between words, and vice versa. Finally, concatenation of the stimuli sometimes resulted in unnatural C.C boundaries, which might have influenced the infants' perception of one vs. two units. In consequence, experiment 2 was designed to avoid these possible flaws.

3 Experiment 2

3.1 Method

Sixteen infants (11 females, 5 males, mean age 379,5 days, age range 356 to 399 days) participated in experiment 2. A total of 16 video recordings were made: 10 infants had participated in experiment 1, and 6 participated for the first time. One additional infant was tested, but the recording was interrupted because the infant was crying.

The stimuli were accent 2 CV̀C.CVCCV́C non-words signaling one word-like unit and accent 1+2 CV́C.CV̀CCV́C non-words, which were expected to signal two separate word-like units more efficiently than the accent 1+1 stimuli used in experiment 1. Also, the phonotactic cues were expected to be more salient than those used in the previous experiment; the within-word C.C clusters used in this experiment (*bettattan*, *beskattan*, *bellattan*, *bengattan*, *bestattan*, *bennattan*) had a high probability of occurring within words, but a low probability of occurring between words; and the between-word C.C clusters (*berdattan*, *bendattan*, *bensattan*, *benhattan*,

benvattan, *benfattan*) had a high probability of occurring between words, but a low probability of occurring within words. Also, the stimuli were recorded as whole sequences, rather than being created using concatenation.

3.2 Results and discussion

The results of experiment 2 were similar to those in experiment 1, except that (see Figure 2) among non-words with within-word phonotactics, accent 2 words were preferred over accent 1+2 words. In sum, by altering the phonotactic cue, the infants' preference for word accent was not reversed. Instead, presumably due to the more distinct word accent contrast, a preference for accent 2 non-words was found.

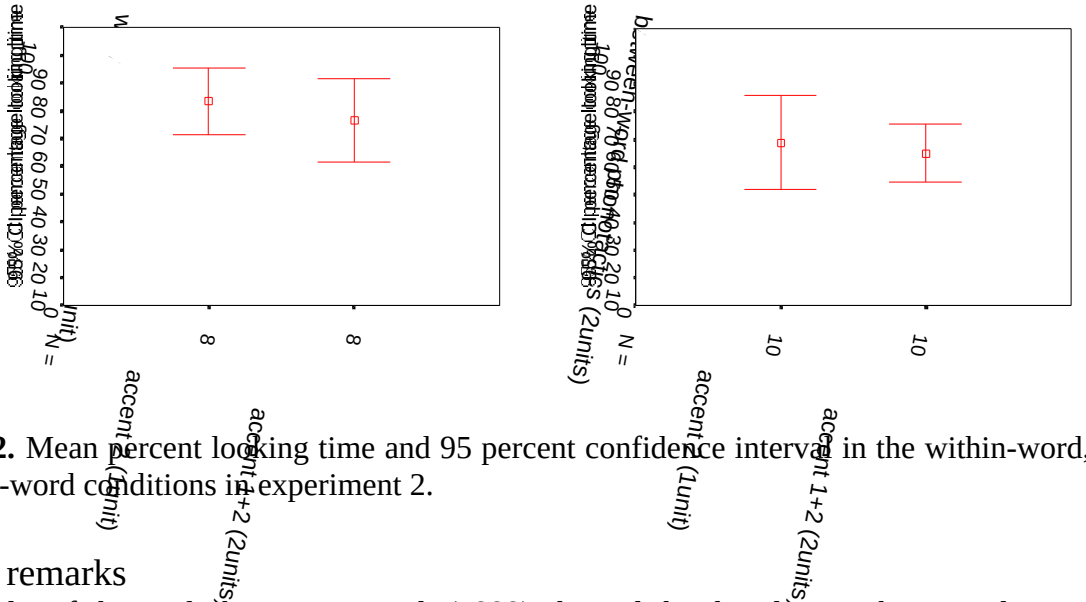

Figure 2. Mean percent looking time and 95 percent confidence interval in the within-word, and between-word conditions in experiment 2.

4 Final remarks

The results of the study by Mattys et al. (1999) showed that by altering the prosodic cue, the English-learning infants' preference for phonotactics was reversed. This was not confirmed in the present study on Swedish-learning infants.

Swedish children are massively exposed to accent 2 words. Accordingly, the preference for accent 2 non-words and within-word phonotactics may arise from infants storing information about frequently occurring prosodic and phonotactic word patterns in the ambient language.

Acknowledgements

This study was supported by Humanistiska fakultets särskilda forskningsresurser and Telefonaktiebolaget L M Ericsson.

References

Fernald, A. 1985. 'Four-month-old infants prefer to listen to motherese'. *Infant Behavior and Development* 8, 181-195.

Johnson, E.K., Jusczyk, P.W. in press. 'Word segmentation by 8-month olds: When speech cues count more than statistics'.

Mattys, S.L., Jusczyk, P.W. 2001. 'Phonotactic Cues for Segmentation of Fluent Speech by Infants'. *Cognition* 78, 91-121.

Mattys, S.L, Jusczyk, P.W., Luce, P.A., Morgan, J.L. 1999. 'Phonotactic and Prosodic Effects on Word Segmentation in Infants'. *Cognitive Psychology* 38.

Sundberg, U. 1998. Swedish Infant-Directed Speech (SWIDS) corpus, Department of Linguistics, Stockholm University.