Final lengthening in infant directed speech may function as a cue to phrase constituents

Eeva Koponen and Francisco Lacerda

Department of Linguistics, Stockholm University

The aim of the current study was to investigate whether lengthening of word and phrase final vowels in infant directed speech (IDS) occurs to the same degree as lengthening of word and phrase final vowels in adult directed speech (ADS). The stimuli consisted of vowels embedded in a two-syllable nonsense word that were varied systematically with respect to phrase position and focal accent. At phrase level, the results showed that FL in accented words in IDS was significantly greater as compared to ADS. Earlier studies have shown that young children are sensitive to pauses as markers to phrase constituents. In line with these experiments, the perceptual importance of FL in IDS is discussed.

1. Introduction

Segment duration varies due to many linguistic and non-linguistic factors. The linguistic factors can be divided in factors at segment, syllable, word, and phrase level. Factors at segment level are e.g. degree of opening in the vowel, and manner of articulation of a following consonant. A factor at syllable level in Swedish is the complementary distribution of phonological length. At word level vowel duration varies as a function of stress, and number of syllables in the word. Vowel duration decreases as the number of following syllables increases (Lindblom & Rapp, 1973). At phrase level, segment duration varies as a function of focal accent, and number of words in the phrase. Vowel duration decreases as the number of following syllables carrying main stress increases (Lindblom & Rapp, 1973). Thus, segment duration in word and phrase final position is often longer than in all other positions in a word or a phrase.

FL occurs in several languages but it is still unclear whether FL occurs to the same extent in languages with quantity systems such as in Finnish as in Swedish (Oller, 1973). Further, FL does not seem to be restricted to speech. FL also occurs in music, bird song and insect chirps (Cooper, 1976; Lindblom, 1978) where a unit within music or animal vocalizations has longer duration when it, within a longer unit, occurs in a final position. Does FL as a wide spread phenomenon indicate that there is one common liable mechanism for it or does FL in different areas have independent explanations? Lindblom (1978), among others, discusses potential explanations: 1) FL is a learned, language specific, grammatically based phenomenon, 2) FL is a consequence of speech production constraints, and 3) FL is perceptually motivated because it signals constituent structure to the listener. Apart from FL in music demarcating "breath-groups" that are natural to produce, evidence for explanation 2) comes from speech production oriented models: "The power law model" (Lindblom, Lyberg, Holmgren, 1976), "The short-term memory model" (Lindblom et al., 1976), and "F_0-dependence model" (Lyberg, 1979). Also, the occurrence of FL in many languages suggest

that it is a consequence of universal speech production constraints, but since the degree of FL in world's languages may vary quantitatively, it can on the other hand be seen as a learned, language specific, but phonetically natural phenomenon (Lindblom, 1978). Further evidence for explanation 1) comes from an observation that FL is not present in speech of deaf speakers (Nickerson et al., 1974), nor in babbling as compared to phonetically comparable utterances produced by adult speakers (Oller & Smith, 1977). Thus omission of FL here suggests that FL is a learned behavior. What about perceptual motivation for FL? Let us draw our attention to FL in IDS, which – in line with speech in general (or music) – indicate that FL may segment the stream of syllables (or beats) into easily perceivable constituents.

Temporal measurements in Swedish IDS have shown longer vowel duration, and longer pauses between utterances as compared to Swedish ADS (Sundberg, 1998). In addition, pauses in IDS seem more often to coincide with syntactic boundaries. Evidence of young children being sensitive to pauses as markers to phrase constituents comes from a study by Hirsh-Pasek et al. (1987). The phrases in their study were manipulated by inserting a 1 sec long pause either at every phrase boundary or between words within phrases. Their results showed that the infants preferred to listen to the set of phrases with pause inserted at every phrase boundary. The children's segmentation task, according to the authors, was thus supported by the coincident pauses at phrase boundaries, a result in line with studies of adult speech production showing that FL occurs, not only before a pause, but also before syntactic boundaries (Klatt, 1975; Cooper, 1978). In addition, the perceptual relevance of FL is demonstrated by studies indicating that adult listeners expect FL and can recognize patterns that diverge from ones that usually occur in speech production (Lehiste, 1973; Klatt & Cooper 1975).

In the current study the occurrence of FL in IDS was assessed through measures of vowel duration in word initial and final, and in phrase initial, medial, and final positions in Swedish IDS and ADS. The aim of the current study was to investigate a) whether lengthening of word and phrase final vowels in IDS occurs to the same degree as lengthening of word and phrase final vowels in ADS, and b) whether the relationship between FL in IDS and ADS is affected by focal accent. The hypothesis was that a) the relationship between FL in IDS and ADS is linear to the lengthening of initial/medial vowels, i.e. FL in IDS occurs to the same degree as compared to ADS, and b) focal accent will have an increasing effect on FL.

2. Method

The stimuli were 144 occurrences of vowel [ʊ] (n72) and [a] (n72) in the word [kʊk:an]. The word's phrase position (italics), and focal accent (underlined) were systematically varied. The same set of phrases was read aloud in IDS (n36) and ADS style (n36) (see Table 1).

Table 1. The vowels [ʊ] and [a] embedded in the nonsense word [kʊk:an].

		[kʊk:an] phrase position		
		Initial	Medial	Final
Focal accent	Initial	*Kockan* är så söt och grann x 4	Titta söta *kockan* är här x 4	Titta här är den söta *kockan* x 4
	Medial	*Kockan* den hade man lekt med x 4	Se den lilla *kockan* som leker x 4	Se på den lilla rara *kockan* x 4
	Final	*Kockan* har så fina färger x 4	Titta på den fina *kockan* så rolig x 4	Titta på fina *kockan* x 4

All the utterances were produced by a female native speaker of Swedish, digitally recorded in an anechoic room (48 kHz, 16 bit) and analyzed in Swell Soundfile editor (AB Nyvalla DSP, 87-98). Syllable and vowel boundaries were identified using conventional acoustic landmarks. Each vowel label contained information about speech style, stress, segment's position in a word, focal accent, and word's position in a phrase.

3. Results

Reflecting the lower speech rate in IDS (4.53 syllables/s) as compared to ADS (5.57 syllables/s), the vowels in IDS were consistently longer than corresponding vowels in ADS. At word level, the results showed (Figure 1 Left) that the degree of lengthening (from ADS to IDS) of vowel [a] in the final syllable was similar (about 20 ms) to the lengthening of vowel [ʊ] in the initial syllable. There were no significant effects of speech style or of focal accent on vowel duration at word level (the variable "stress" was constant).
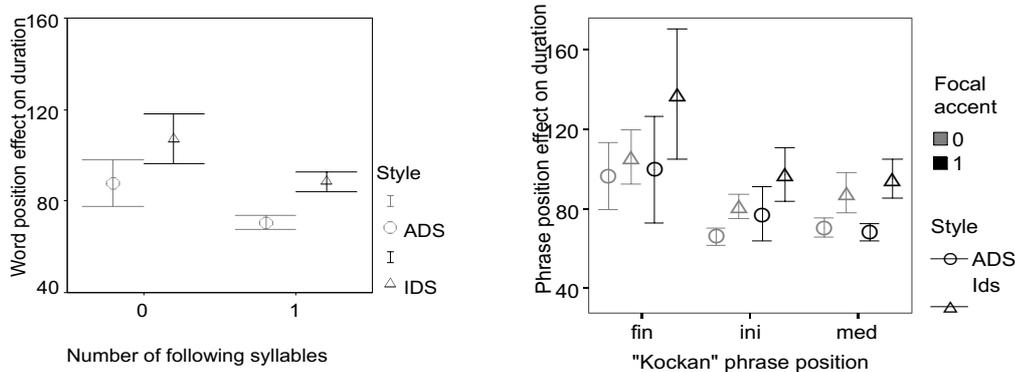

Figure 1. Left: Duration of vowel [a] (no. of following syllables 0), and vowel [ʊ] (no. of following syllables 1) (ms) in IDS and ADS. Right: Duration of vowels [ʊ] and [a] (ms) in phrase final, initial, and medial position, in accented (1) and non-accented (0) words in IDS and ADS. The bars indicate 95 % confidence intervals of the means.

At phrase level, the results showed (Figure 1 Right) among non-accented words that the degree of lengthening (from ADS to IDS) of vowels [a] and [ʊ] in phrase final position was similar (about 20 ms) to the lengthening of vowels [a] and [ʊ] in phrase initial and medial positions. Among accented words, the degree of lengthening (from ADS to IDS) of vowels [a] and [ʊ] in phrase final position was greater than lengthening of vowels [a] and [ʊ] in phrase initial and medial positions. In sum, in phrase final position there was a significant effect of speech style ($F(1,22) = 38.598$; $p < 0.001$) and a significant interaction between speech style and focal accent ($F(1,22) = 13.628$; $p < 0.001$).

4. Discussion

At word level, FL in IDS occurred to the same degree as FL in ADS. Lack of even greater FL effect in both speech styles may presumably be explained by the fact that all the words were stressed on their first syllable. Lengthening due to stress thus might level out some of

the FL effect. At phrase level, FL among non-accented words in IDS occurred to the same degree as in FL in ADS. However, among accented words, FL in IDS was significantly greater than FL in ADS. Hence, FL was amplified by focal accent.

To be able to recognize regularities in the ambient language, such as FL, allows the listener to make predictions about its overall syntactic organization. Such regularities may have a great perceptual importance for children: segmentation and potential understanding of input becomes easier and faster (Jusczyk et al., 1995). FL in Swedish IDS, especially in accentuated words in phrase final position, reinforces thus the hypothesis that FL may function as a perceptually important marker of constituent structure. In sum, just like punctuation in written texts help us in avoiding incorrect segmentation, it seems that potential prosodic cues, such as a pause, change in f_0, or FL as a marker of phrase boundary, may indicate to a listener how the utterance ought to be divided.

5. References

Cooper, W. E. (1978) Syntactic-to-Phonetic Coding. In *Language Production* (B. Butterwoth, editor). Academic Press, New York.

Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Wright Cassidy, K., Druss, B., and Kennedy, L. (1987) Clauses are perceptual units for young infants, *Cognition*, 26, 269-86.

Jusczyk, P., Hohne, E., Mandel, D. (1995) Picking Up Regularities in the Sound Structure of the Native Language. In *Speech Perception and Linguistic Experience Issues in Cross-Language Research* (W. Strange, editor), 91-119.

Klatt, D. (1975) Vowel lengthening is syntactically determined in connected discourse, *Journal of Phonetics*, 3, 129-140.

Klatt, D. & Cooper, W. E. (1975) Perception of Segment Duration in Sentence Contexts. In *Structure and Process in Speech Perception* (A. Cohen & S. Nooteboom, editors). Springer Verlag, Heidelberg.

Lehiste, I. (1973) Rhytmic units and syntactic units in production and perception, *Journal of the Acoustic Society of America*, 54, 1228-1234.

Lindblom, Rapp, K. (1973) Some Temporal Regularities of Spoken Swedish, *PILUS* 21, SU.

Lindblom, B., Lyberg, B., Holmgren, K. (1976) Durational Patterns of Swedish Phonology: Do they Reflect Short-Term Memory Processes?, unpublished ms., Stockholm University.

Lindblom, B. (1978) Final lengthening in Speech and Music. In *NordicProsody* (E. Gårding, G. Bruce and R. Bannert, editors), 85-102, Department of Linguistics, Lund University.

Lyberg, B. (1979) Final lengthening – partly a consequence of restrictions on the speed of fundamental frequency change?, *Journal of Phonetics* 7, 187-196.

Nickerson, R.S., Stevens, K.N., Boothroyd, A., and Rollins, A. (1974) Some Observations on Timing in the Speech of Deaf and Hearing Speakers. BBN Inc., Report 2905, Sept. 1974

Oller, D.K. (1973) The Effect of Position in Utterance on Speech Segment Duration in English, *Journal of the Acoustic Society of America,* 54, 1235-1247.

Oller, D.K. & Smith, B.L. (1977) Effect of Final-Syllable Position on Vowel Duration in infant Babbling, *Journal of the Acoustic Society of America,* 62 (4), 994-997.

Sundberg, U. (1998) Mother tongue – Phonetic Aspects of Infant-Directed Speech, Ph.D thesis, Stockholm University.