

Effect of Pitch Movement Timing on Perceived Duration and Prominence in Estonian and English Listeners

Robert Allen Fox, Department of Speech and Hearing Science
Ilse Lehiste, Department of Linguistics
The Ohio State University, Columbus, OH 43210

ABSTRACT

This study is one in a series in which we explore the possible influence of a listener's linguistic background on his/her perception of suprasegmental cues in the speech signal. Two experiments investigated the effect of the timing of a pitch rise on the perception of a medial syllable's perceived duration and/or prominence by native speakers of Estonian or American English. Results indicated that English subjects tended to perceive tokens with a later pitch rise as longer and more prominent than those with an early pitch rise. The reverse was true for the Estonian subjects in terms of perception of prominence but for these subjects, the timing of the pitch rise had little effect upon perceived duration.

INTRODUCTION

In the past two years we have examined the perception of "prominence" in sequences of both speech (using the nonsense token [bab]) and non-speech (signal-correlated noise) tokens by native Estonian and American English listeners while independently manipulating individual token duration and amplitude (Lehiste & Fox, 1992). As is well known, Estonian is a quantity language in which durational differences can distinguish between words. Duration in English, on the other hand, is not independently contrastive although it serves as one of the phonetic characteristics of stressed syllables. Our results have indicated that Estonian listeners are more sensitive to token duration in making their "prominence" decisions to both speech and non-speech stimuli than are English listeners and support the contention that the linguistic background of listeners has a potentially significant effect on the perception of prominence—particularly in terms of the utilization of the cues of amplitude and duration.

In these experiments we did not introduce any pitch variations. However, it is the case that variations in pitch patterns can influence both the perception of vowel duration as well as prominence, independent of changes in either duration or amplitude. For example, Lehiste (1976) demonstrated that the presence of a pitch change in a speech token could result in longer perceived duration. More recently, Rump (1992) demonstrated that the timing of accent-leading pitch movements had a significant effect both on perceived prominence and perceived vowel length in Dutch subjects. For example, he found that syllables with very early pitch rises were perceived as relatively more prominent than were syllables with later pitch rises. However, for Dutch speakers the prominence judgments did not always parallel perceived durations. In particular, there was a slight tendency for syllables with late pitch rises to be perceived as longer than those with early pitch rises (in opposition to the prominence decisions). The situation was somewhat different for pitch falls in that syllables with a late pitch were perceived as both relatively more prominent and longer. This suggests that, at least for Dutch subjects, the relationship between perceived prominence and duration is complex. This basic conclusion is supported by Hermes (1991) who argued that the timing of pitch movements, per se, may not be the most important factor determining prominence in Dutch, but rather this timing may determine the kind of "pitch movement with which a syllable is accentuated."

The present study compares the effect of the timing of pitch movements and language background on the perception of both syllable duration (Experiment 1) and prominence (Experiment 2) by Estonian and English listeners.

METHOD

Stimuli

The stimuli consisted of sequences of five monosyllables ([bob]) in which the third monosyllable had a pitch rise. The pitch pattern was superimposed on a slow F0 declination across the entire sequence of monosyllables as shown in Figure 1. The pitch rise started at 105 Hz and rose (over 120 ms) to 140 Hz; it then declined to 131 Hz by the end of the third monosyllable. The pitch rise began either at the onset of the vowel (0 ms onset) or 60 ms following vowel onset—this variation is similar to Rump's (1992) pitch movement onset factor. The experiment also manipulated the vowel duration of the third monosyllable as well as the timing of the pitch rise. In particular, vowel duration could be either 400, 425, 450 or 475 ms. Please note that these 5-syllable sequences differed only in terms of the third syllable.

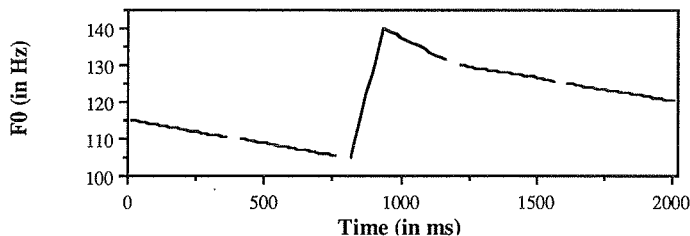


Figure 1. Schematic of the basic pitch contour used (onset differences not shown).

Procedure

In each experiment, subjects heard experimental trials which consisted an *anchor sequence* and a *comparison sequence* separated by a 500 ms interstimulus interval. There were two different anchors. In one anchor, the third syllable was 400 ms in duration with a pitch rise that began at vowel onset (0 ms). The third syllable of the second anchor was also 400 ms in duration, but had delayed pitch rise (beginning 60 ms after vowel onset). The comparison sequences represented either one of the anchors, or any of the other sequences (each of which had physically longer third syllables with a 0 or 60 ms pitch rise). The same set of comparison sequences were paired with each anchor sequence. In Experiment 1, subjects were required to indicate which sequence in a trial pair had the longer third syllable. In Experiment 2, subjects were required to indicate in which sequence the third syllable was "more prominent." Each possible pair of sequences was presented four times (twice with the anchor in first position, twice with the anchor in second position) in different random orders for the two experiments.

Subjects

There were two different subjects groups: twenty-three native speakers of English (living in Columbus OH) and thirty-three native speakers of Estonian (living in Tartu, Estonia). Subjects participated in both experiments (and always completed them in the order Experiment 1, Experiment 2).

RESULTS

As shown in Table 1, there was an overall tendency for both subject groups to perceive the second sequence as either longer or more prominent than the first sequence. This is very possibly due to a type of recency effect, making the relative "importance" of the second sequence (which was last heard) greater than that of the first sequence. However, since there was no significant difference between the two language groups in terms of this position effect (for either duration or prominence judgements), we will ignore positional variations in further descriptions of the data. Note, here and elsewhere, all significance tests were done using chi-square statistics.

Table 1. *Percentage of responses identifying either the first or second sequence as "longer" (in Experiment 1) or "more prominent" (in Experiment 2) across all experimental trials.*

	SEQUENCE POSITION	
	First	Second
Experiment 1		
<i>Estonian</i>	43.8	56.2
<i>English</i>	47.5	52.4
Experiment 2		
<i>Estonian</i>	42.8	57.7
<i>English</i>	44.9	55.1

One way to quickly determine the possible effect of the timing of the pitch rise upon perceived duration and prominence is to examine responses from those experimental trials which paired the different anchor sequences (a direct, "head-to-head" comparison). As shown in Table 2, the Estonian listeners judged the 0 ms anchor to be longer about as often as the 60 ms anchor. However, these same listeners judged the 0 ms anchor to be prominent 15% more often than the 60 ms anchor. The English subjects show a significantly different pattern for both duration and prominence judgements (at the .05 level). For the English responses, the 60 ms anchor is more often judged to be both longer and more prominent than the 0 ms anchor.

Table 2. *Percentage of responses identifying either the 0 ms or 60 ms pitch onset anchor tokens as "longer" (in Experiment 1) or "more prominent" (in Experiment 2) when the these two different anchors are compared directly ("head-to-head") in an experimental trial.*

	ANCHOR TOKENS	
	0 ms	60 ms
Experiment 1		
<i>Estonian</i>	51.6	48.4
<i>English</i>	41.3	58.7
Experiment 2		
<i>Estonian</i>	57.6	42.6
<i>English</i>	43.2	56.8

Shown in Table 3 is a breakdown of the responses when only one of the sequences in an experimental trial was an anchor. This table shows the percentage of time that the third syllable of the anchor sequence was identified as "longer" or "more prominent". Remember that in these experimental trials, the comparison sequence was always physically longer.

As expected, for both language groups, the number of times an anchor sequence was judged as being longer or more prominent decreased as the duration of the third syllable of the comparison sequence increased. In general, the Estonian subjects show little (non-significant) difference in duration judgments as a function of the timing of the pitch rise. The most critical comparison (highlighted in the table) is the mean response of 0 ms anchor vs. 60 ms comparison and 60 ms anchor vs. 0 ms comparison; that is, those trials in which the anchor token has a different pitch rise onset from the comparison sequence. For duration judgements, this comparison is almost identical for Estonian subjects (23.1% vs. 22.1%). This difference is only slightly larger (and only of borderline significance) in the prominence judgements (37.5% vs. 33.3%).

Again, however, the reverse is true of the English subjects, the 60 ms syllable is judged to be both longer and prominent more often than the 0 ms syllable. For example, in this critical comparison described above, English subjects show a significant difference both for the duration judgments (13.2% vs. 24.5%) and the prominence judgments (17.7% vs. 26.5%). It is interesting to note that for these data, English subjects seem to be more affected by the durational differences between the anchor syllable and the comparison syllable in making their prominence judgments than the Estonian subjects (i.e., the physically shorter anchor is identified more often as more

prominent by Estonian subjects than by English subjects), contrary to the pattern that might be expected given our results in Lehiste & Fox (1992). However, one should note that the stimuli in this previous experiment on perception of prominence had no F0 variations.

Table 3. *Percentage of responses identifying the anchor stimuli as "longer" or "more prominent" when compared with the longer comparison tokens. The critical comparison between anchor and comparison sequences that differ in pitch onset (are in bold italics).*

ANCHOR	COMPARISON TOKEN								Overall Mean
	0 ms Pitch Onset				60 ms Pitch Onset				
	425	450	475	Mean	425	450	475	Mean	
	Experiment 1 (Duration Comparison)								
<i>Estonian Data</i>									
0 ms	27.4	13.4	9.7	16.8	37.1	22.0	10.2	23.1	20.0
60 ms	32.3	19.4	14.5	22.1	35.5	14.0	13.4	21.0	21.6
Mean	29.9	16.4	12.1	19.5	36.3	18.0	11.8	22.1	
<i>English Data</i>									
0 ms	25.0	12.1	6.8	14.6	17.7	11.4	10.6	13.2	13.9
60 ms	42.4	22.0	9.1	24.5	28.0	11.4	7.6	15.7	20.1
Mean	33.7	17.1	8.0	19.6	22.9	11.4	9.1	14.5	
	Experiment 2 (Prominence Comparison)								
<i>Estonian Data</i>									
0 ms	34.1	33.3	31.7	33.0	35.0	41.4	36.0	37.5	35.3
60 ms	40.3	30.6	29.0	33.3	38.7	29.0	21.0	29.6	31.5
Mean	37.2	32.0	30.4	33.2	36.9	35.2	28.5	34.5	
<i>English Data</i>									
0 ms	24.2	27.3	18.9	23.5	22.0	15.2	15.9	17.7	20.6
60 ms	33.2	27.3	18.9	26.5	31.8	21.2	13.6	22.2	24.4
Mean	28.7	27.3	18.9	24.0	26.9	18.2	14.8	20.0	

Our data support the conclusion reached by Rump (1992) that differences in timing may produce differences in perceived prominence. However, there is a significant effect of language background. Unlike the Estonian listeners (or Rump's Dutch listeners), delaying the onset of the pitch rise tended to increase the perceived prominence of a syllable for English listeners. Our obtained increase in prominence when the pitch movement occurs later actually better matches the pattern obtained by Rump for a falling F0 contour. However, there are significant differences in the stimulus sets used in his study as opposed to ours (e.g., in his "early" timing condition the pitch rise occurs 120 ms before vowel onset rather than 0 ms in our study) and it is difficult to directly compare the two sets of prominence results.

ACKNOWLEDGMENTS

This work was partially supported by a grant from NIA/NIH to the first author (AG08353) and by a grant from NSF to the second author (BNS-8719891). The help of Dr. Jüri Allik of Tartu University in running the Estonian listening tests is gratefully acknowledged.

REFERENCES

- Hermes, D.J. (1991), "Prominence cased by rising and falling pitch movements with different positions in the syllable", In *IPO Annual Progress Report, Eindhoven*, Vol. 16, pp. 17-28.
- Lehiste, I. (1976), "Influence of fundamental frequency pattern on the perception of duration", *Journal of Phonetics*, Vol. 4, pp. 113-117.
- Lehiste, I. and Fox, R.A. (1992), "Perception of Prominence by Estonian and English Listeners", *Language and Speech*, Vol. 35, pp. 419-434.
- Rump, H.H. (1992), "Timing of pitch movements and perceived vowel duration", in *Proc. of the International Conference on Spoken Language Processing*, Vol. 2, pp. 1047-1050.