Word-prosodic features in Estonian conversational speech: Some preliminary results

Diana Krull
Stockholm University, Institute of Linguistics
S-106 91 Stockholm, Sweden

ABSTRACT

Estonian has three distinctive degrees of quantity: short, long and overlong. This paper reports an investigation on the temporal and tonal correlates of quantity in the natural conversation of one Estonian speaker. The results show statistically significant differences between quantities only for the temporal correlates. The tonal correlates had a considerable overlapping between quantities.

INTRODUCTION

Over the years, a large amount of investigation has been addressed to the Estonian quantity system. A number of theories about the nature of the three distinctive quantities and their acoustic correlates have been put forward. (For an overview, see Lehiste 1970; a new theory is presented by Eek and Help 1987). The phonetic material on which the theories could be based, consisted of so called "laboratory speech", that is, words or sentences prepared by the investigator and read by the speaker. More recent work has shown that more naturally produced speech can differ considerably from such laboratory productions (see e.g. Lindgren, Krull and Engstrand (1987). In particular, language dependent differences in the stability of temporal cues to quantity have been found (Engstrand 1992). The temporal cues for Swedish seem to be less stable than in Finnish. The probable reason is that in Swedish, quantity is not signalled by temporal cues alone, there is also a clear difference in vowel quality.

It can be hypothesized that if an acoustic parameter is used as a primary or only cue to phonological distinctions in a language, then the freedom of the speakers of the language will be restricted. That is, differences in the acoustic parameter should remain robust in any speaking style. Therefore, an investigation of natural Estonian conversation is important for two reasons: It can help to test the hypothesis, and, at the same time, throw some more light on the question relative importance of different cues to quantity in Estonian.

Estonian has three phonologically distinct degrees of quantity, short (Q1), long (2) and overlong (Q3). They are signalled by the duration ratio between the first (main stressed) and the second syllable of a word. The typical ratio for Q1 is 2:3, for Q2 3:2 and for Q3 2:1 (Lehiste 1960). To distinguish Q3 from Q2, listeners use an additional, tonal cue (Lehiste 1970): falling F0 for Q3 and flat or slightly rising for Q2. Earlier results with words read in isolation and in a carrier phrase (Krull 1993) showed that the duration relation between syllables remained stable even when the syllables involved were shortened. In most cases, also the falling F0 contour remained stable.

Other tonal cues described in the literature but not studied in Krull op. cit., are F0 stepdown from the end of V1 to the beginning of V2 (Lehiste 1970b), and an earlier location in time of an F0 maximum within V1 for Q3 (Eek 1990).

The aim of the present study is to assess the stability of these cues in natural conversational Estonian speech, to begin with, of one speaker.

METHOD

The subject was a male phonetician, native speaker of standard Estonian, resident in Estonia. Seated in an anechoic chamber together with the author and prompted by a few short questions, he related episodes from his childhood, schooldays and travels. The talk - over an hour and a half of a lively near-monologue - was recorded digitally. Lexical non-compound Q2 and Q3 words of the form (C)V1CV2(CV3...) were located and sampled at a rate of 10 kHz/s. V1 was a long or overlong vowel/diphthong, C a short consonant. This form was chosen because the exact duration of single short consonants in syllable initial position is of no consequence for the quantity degree and therefore V1 and V2 can be used to represent syllables (Lehiste 1960). The three-way quantity contrast is connected to the two initial syllables of a word, of which the first carries the main stress.

For the analysis, the Kay CSL 3400 system was used. First, the duration of the vowels was measured. The beginning and end of the vowel were defined as the onset and and offset of a clear formant pattern. In the case of stops, the burst release was defined as the beginning of the vowel. The aspirative phase sometimes appearing at the end of vowels, especially in word final position, was not included. Thereafter, the duration ratio V1/V2 was calculated.

Next, F0 was measured at the beginning and end of V1 and V2. After stops, the beginning was now defined as the voice onset instead of the burst. If there was an F0 maximum within V1, its frequency and temporal location were marked. Utterance final words were excluded. The remaining material consisted of 157 words, 48 of Q1, 45 of Q2 and 64 of Q3. Words where V1 was an (over)long vowel were measured separately from words where it was a diphthong. Moreover, disyllabic words - which were in a clear majority - were measured separately from words of three and four syllables.

Mann Whitney U-tests showed no significant change in syllable duration due to the lengthening of the word as was reported by Krull (1993) for laboratory speech, nor any significant difference between long vowels and diphthongs. Therefore, these groups were analyzed together.

RESULTS

Duration ratios and F0 data are shown in Table 1. The data were, in general, not normally distributed, therefore median values are given instead of means, and the Mann-Whitney U-test was used. The test showed statistically significant differences between Q1-Q2, and between Q2-Q3, both for the durations of V1 and V2 and for the V1/V2 ratios. For all cases $p<.001$. The ratios are shown in the so called "box and whiskers" plot in Figure 1.

The corresponding differences in F0-change, on the other hand, were not statistically significant, although there was a tendency for F0 to fall more during V1 in Q3 than in Q2 words. There was also a tendency for F0 to stay unchanged between the end of V1 and the beginning of V2 in Q3 words, and to fall in connection with Q2, but even here the difference was small and the variation considerable.

An F0 maximum within V1 was found in 27% of the Q2 words and in 59% of the Q3 words. The median F0 rise was about 8.9% of the vowel-initial value for Q2 and 7.5% for Q3. The location of the maximum on the time axis was earlier in Q3 than in Q2 words: for Q2 median location of the maximum lay at a point 55% of the entire duration of V1, for Q3 the corresponding value was 44%. However, the difference in location was not statistically significant.

Finally, the material was checked for possible correlations between the different acoustic correlates to quantity. No such correlation - positive or negative - was found.

Table 1. *Median, mininimum and maximum of the duration of V1, V2, and of the V1/V2 ratio. Percent F0 change between the beginning and end of V1, and between the end of V1 and the beginning of V2.*

		Duration			ΔF0 (%)	
	N	V1 (ms)	V2 (ms)	V1/V2	V1	V1-V2
Q1 med	48	82	109	.7	0.3	2.4
min		56	69	.3	-7.0	-35.3
max		115	194	1.3	44.4	28.9
Q2 med	45	134	78	1.8	.8	-6.1
min		74	37	1.0	-40.1	-42.2
max		244	149	3.2	52.5	21.9
Q3 med	64	180	56	3.2	-4.5	-2.5
min		48	23	1.0	-33.8	-28.3
max		323	132	8.1	62.5	32.2

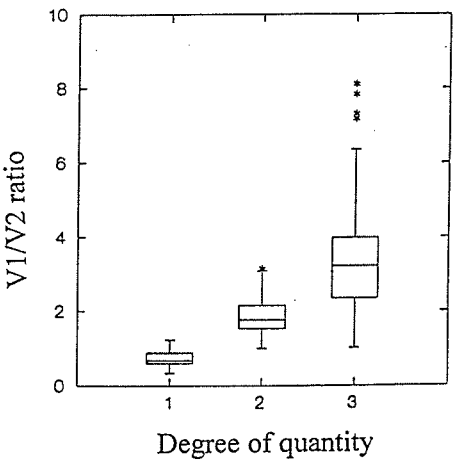

Figure 1. *Box-and-whiskers plot of the V1/V2 ratio in Q1, Q2 and Q3. The boxes represent the spread of 50% of the sample, 25% (quartile) above and 25% below the median. The whiskers (above and below the boxes) are drawn to the nearest value not beyond a standard span of the quartiles. The points beyond that are represented by asterisks and circles (For details see Velleman and Hoaglin 1981, p. 65f).*

DISCUSSION

The hypothesis presented in Introduction, if true, would predict that the temporal correlates to Q1 and Q2, and at least one - temporal or tonal - correlate to the distinction between Q2 and Q3 would remain stable. That is, there should be some acoustic differences between quantities found in "laboratory speech" that remain stable even in a more spontaneous speaking style.

The results of this investigation showed that - for this one speaker - the most stable acoustic difference between quantities was temporal. Not only the duration ratios, but even the absolute durations of V1 and V2 were significantly different between quantity degrees. Although there was always a certain overlap, it involved relatively few items. A comparison of these results with temporal data from CV(V(V))CV words read in a carrier phrase by the same speaker (Eek 1975) showed, moreover, that the temporal differences between quantities were enhanced rather than weakened in a more natural speech style.

The tonal cues, on the other hand, were present only as a statistically non- significant tendency. The stability of the temporal correlates to quantity in Estonian therefore suggests that they may be the most crucial cues for signalling quantity. However, further investigation with more speakers is necessary.

REFERENCES

Eek, A. (1975). "Observation on the duration of some word structures II". *Estonian Papers in Phonetics (EPP)* 1975, Academy of Sciences of the Estonisn S.S.R., Tallinn, 7-51.

Eek, A. (1990). "Units of temporal organization and word accents in Estonian". *Linguistica Uralica XXVI*, 251-264.

Eek, A. and Help, T. (1987). "The interrelationship between phonological and phonetic sound changes: A great rhythm shift of old Estonian." *Proceedings of the 11th ICPhS,* Tallinn 1987, Vol. 6 pp. 218-233.

Engstrand (1992). "Durational correlates of quantity in Swedish and Finnish: Data from natural speech". *Papers from the Sixth Swedish Phonetic Conference, Gothenburg, May 20-22, 1192*, 47-50.

Krull, D. (1993). "Temporal and tinal correlates to quantity in Estonian", *PERILUS XV*, Institute of Linguistics, Stockholm University.

Lehiste, I. (1960). "Segmental and syllabic quanity in Estonian", *American Studies in Uralic Linguistics.* Indiana University, Bloomington, Vol. 1, 21-82.

Lehiste, I. (1970a). *Suprasegmentals*. MIT Press, Cambridge, Mass.

Lehiste, I. (1970b). "Experiments with synthetic speech concerning quantity in Estonian", *Congressus Tertius Internationalis Fenno-Ugristarum*, Pars I, Acta Linguistica Tallinn.

Lindgren, R., Krull, D. och Engstrand, O. (1987). "Akustiska studier av fonetisk variation i svenskan", i P. Linell, V. Adelsvärd, T. Nilsson och P.A. Pettersson (eds.), *Svenskans beskrivniong **16**. Förhandlingar för att dryfta frågor rörande svenskans beskrivning.* Vol **2**, 326-338.

Velleman, P. and Hoaglin, D. (1981). *Applications, Basics, and Computing of Exploratory Data Analysis*. D. Duxbury Press, Boston, Mass.