

FOOT ISOCHRONY IN ESTONIAN

Diana Krull

Department of Linguistics, Stockholm University, Stockholm, Sweden

ABSTRACT

The duration of Estonian disyllabic feet varies relatively little with the three degrees of quantity. This is due to the duration distribution between the syllables: the first syllable increases with the degree of quantity and the second decreases. The amount of isochrony varies with speaking style: in conversational speech, foot duration is more dependent on the degree of quantity than in list reading. Moreover, in list reading the difference in foot duration between short and long quantity is greater than between long and overlong. The present paper shows that in conversational speech this is true only for words where the bearer of the quantity contrast is a vowel and the F0 contour can constitute an additional cue to overlong quantity. In words where the bearer of the quantity contrast is a consonant, the difference in foot duration is largest between long and overlong quantity. The rhythmic pattern in oral production of poetry is similar to that of conversational speech.

1. INTRODUCTION

Estonian has three phonologically distinct quantities: short (Q1), long (Q2) and overlong (Q3). The domain of quantity is a disyllabic unit. The duration of the first syllable – which carries the stress – increases with rising degree of quantity while the duration of second syllable decreases. As a result, the total duration of the unit changes relatively little with degree of quantity. What varies is the duration ratio between the two syllables: typical ratios are 2:3 for Q1, 3:2 for Q2 and 1:2 for Q3. Investigation of word initial disyllabic units in Estonian has shown that the difference in duration ratio with the degree of quantity is very robust; it is preserved even in conversational speech where the absolute duration of the syllables undergoes considerable changes [3], [5].

The tendency to foot isochrony in Estonian has been commented on by Eek, among others: "The problems concerning duration ratios (...) are related to isochrony. In the Estonian language (...) mono- and disyllabic feet, irrespective of the accent¹ type and segmental duration, differ durationally less than intra-feet segments." [2]. However, Estonian rhythmic feet can contain either two or three syllables, and the tendency to isochrony does not include trisyllabic feet – they are usually longer than their disyllabic counterparts. Eek comments on this as follows: "The whole third syllable is inert in defining accents. It seems as if we were dealing, on the one hand, with stress isochrony which comprises the whole stress foot, and on the other hand, with isochrony characterizing the accents." Eek is of the opinion that this "genuine Lappo-Finnic foot isochrony" works even today and cites Wiik: "Foot was originally defined as the sequence from the beginning of the first vowel to the end of the second syllable vowel"[13].

The question arises: is isochrony a goal to be achieved and the syllabic duration ratios within the foot a means to reach this goal? Or are the syllabic ratios important in

themselves and isochrony a spin off effect?

The present investigation includes disyllabic units that characterize the different degrees of quantity: disyllabic words and the two initial syllables of trisyllabic words. The goal is to assess the duration relations between the disyllabic units of different degrees of quantity in conversational speech and compare them with corresponding relations in orally produced poetry and in list reading. Are the duration relationships between disyllabic units as stable as the intra-syllabic duration ratios or do they vary with speaking style?

2. METHOD

The material consists of conversational speech produced by five native speakers of Standard Estonian, one male and four females. The recordings were made at the Phonetics Laboratory, Stockholm University. The experimenter's role was to ask questions and keep the conversations going. Two of the subjects were present simultaneously, of the remaining three, each maintained a separate conversation with the experimenter. The recordings consist of 60-90 minutes of conversation. Part of the material has been used in earlier publications [3], [5], [6].

In word initial position, disyllabic units of seven different (C1)V1C2V2 combinations are possible. These can be divided into three types: (1) V1 lengthens with quantity degree, (2) C2 lengthens, and (3) both V1 and C2 lengthen. Only types (1) and (2) were chosen for this investigation; for simplicity, they will henceforth be referred to as Vcv and vCv type words. They can be schematized and exemplified as follows.

	Q1	Q2	Q3
Word type			
Vcv(cv...)	/sata/	/saata/	/saa:ta/
	'hundred'	'send' 2 nd p. sg. imper.	'to get',
vCv(cv...)	/lake/	/lakke/	/lak:ke/
	'bare'	'thin gruel' nom. sg.	'ceiling' ill. sg.

Table 1. Examples of word types included in this paper.

Vcv type words in this study include both long vowels and diphthongs, in vCv type words consonant clusters are excluded. To ascertain that the disyllabic unit ends with an open syllable, the third syllable always begins with a short consonant.

The selected words were digitized at 10 kHz. The duration measurements were performed with Kay Elemetrics CSL 3400, beginning with the onset of V1 and ending with the offset of V2. The mean durations of the measured units were calculated for each word type and length and the results for the speakers were pooled.

The material used for comparison consists of duration

data from papers published by other investigators: (1) lists of words in sentence context [12], [1]; (2) words occurring in oral production of poetry (free verse, trochaic and dactylic verse, folk songs) [11]. Only words of the same form as the ones shown in Table 1 are included.

3. RESULTS

Figure 1 shows duration ratios of disyllabic words from the conversational speech of five subjects. The ratios have been calculated between V1CV2 units, that is, the whole word minus the initial consonant. Dark bars show the Q1/Q2 ratio, and light bars the Q2/Q3 ratio. Vcv type words are shown to the left, and vCv type words to the right. Figure 2 shows the corresponding data for the two initial syllables of trisyllabic words. It can be seen that in both cases there is a difference between the two word types: in Vcv type words the Q2/Q3 ratio is higher than Q1/Q2, that is, the difference in duration is larger between Q1 and Q2 than between Q2 and Q3. In vCv type words, on the contrary, the difference between Q2 and Q3 is larger than between Q1 and Q2. The same is true for the initial two syllables of trisyllabic words (Figure 2). The corresponding numerical data are shown in Table 2.

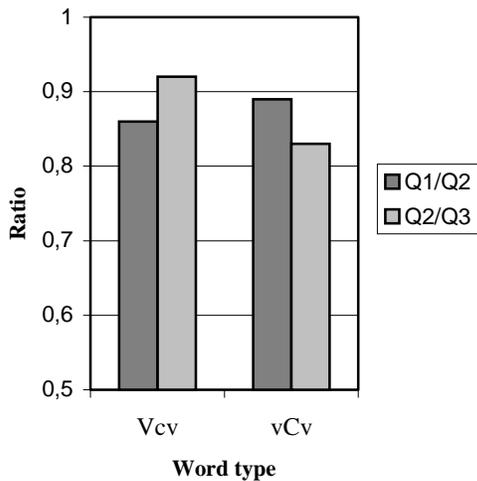


Figure 1. Duration ratios between disyllabic words of different degrees of quantity. Conversational speech.

Comparing the data from conversational speech with list reading shows that there is more isochrony in the list data (Table 3). In particular, Q2 and Q3 words have very nearly the same duration, Q2 may even be longer than Q3. In both Vcv and vCv type words, the difference between Q1 and Q2 is larger (Figure 3). A comparison of conversational speech with the reading of poetry reveals a greater temporal similarity between these two styles than either of them has to list reading. Firstly, there is no great difference in the amount of isochrony between conversational speech and poetry reading (Table 2, Table 3). Secondly, and more importantly, both styles demonstrate the same kind of difference between Vcv and vCv type words (see Figure 1 and Figure 4).

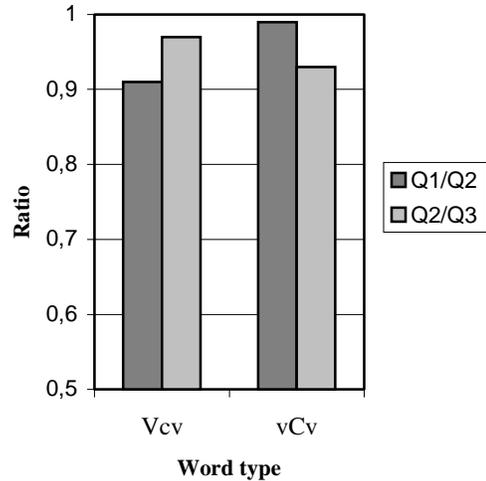


Figure 2. Duration ratios of the two initial syllables of trisyllabic words between different degrees of quantity. Conversational speech.

Number of speakers	Word length (syllables)	Word type	Total number of items	Q1/Q2	Q2/Q3
5	2	Vcv	248	0.86	0.92
	3		210	0.91	0.99
3	2	vCv	96	0.89	0.83
	3		129	0.97	0.93
1 (same as in Table 3, list 2)	2	Vcv	81	0.85	0.90
	3		46	0.89	0.93
	2	vCv	60	0.87	0.81
	3		50	0.98	0.95

Table 2. Duration ratios between disyllabic units of different degrees of quantity in conversational speech. The units are disyllabic words and the two initial syllables of trisyllabic words.

Speaking style	Number of speakers	Word length (syll)	Word type	Q1/Q2	Q2/Q3
List 1 [12]			Vcv	0.85	0.99
List 2 [1]	1	2	Vcv	0.84	1.04
		2	vCv	0.92	1.02
Reading of poetry [11]	5	2	Vcv	0.84	0.95
		3	Vcv	0.81	0.93
		2	vCv	0.88	0.82
		3	vCv	0.86	0.88

Table 3. Duration ratios between V1CV2 units of different degrees of quantity in list reading and in oral production of poetry.

The material available also makes it possible to compare the effect of speaking style on the duration relations in the speech of one particular subject: the speaker in Table 3, list 2

is identical with one of the subjects in conversational speech (Table 2, the four last rows). The difference between the duration patterns of Vcv and vCv type words is quite salient in his conversational speech but is not found in list reading. There is also less isochrony in his conversational speech.

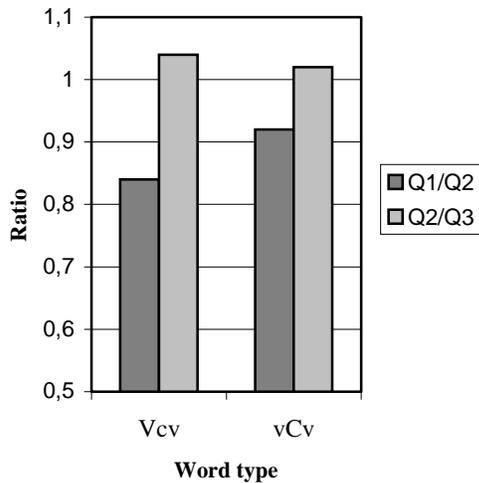


Figure 3. Duration ratios between disyllabic words of different degrees of quantity. List reading.

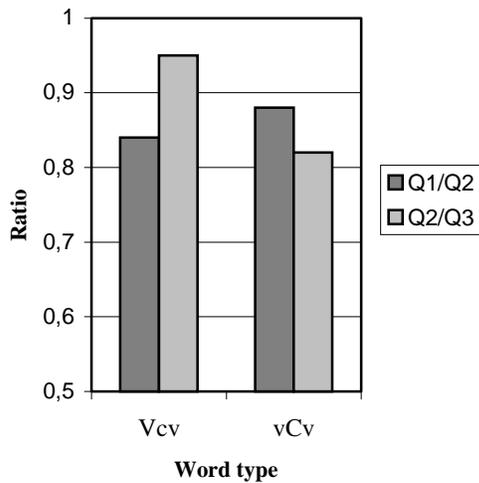


Figure 4. Duration ratios between disyllabic words of different degrees of quantity. Oral production of poetry.

There are differences in the amount of isochrony also between speakers within the same speaking style. For example, Lehiste reports duration data from six readers of trochaic verse, where the data from one speaker gives a ratio of 0.85 between Q2 and Q3 feet, while another has 0.99; the mean value is 0.94 [10]. In an earlier study of trochaic verse with two speakers, Lehiste reports relatively large differences between different degrees of quantity [8]: the duration ratios between trochaic feet calculated on her data (initial consonant

included) are 0.69 for Q1/Q2 and 0.87 for Q2/Q3, which is larger than in conversational speech. There is no separate information for different word types in this study.

As could be expected, there are also differences between speakers in conversational speech: for example, one speaker's Q2/Q3 ratio in Vcv words is 1.00, another's 0.86. In vCv words, there is less variation: the corresponding ratios are 0.87 and 0.80.

Speaking style	Word length	Word type	Q1	Q2	Q3
List 2 [1]	2	Vcv	309	368	354
	2	vCv	300	327	322
Poetry [11]	2	Vcv	346	414	438
	3		302	372	402
	2	vCv	(346)	394	480
	3		(302)	351	400
Conver- sation	2	Vcv	235	273	304
	3		242	273	292
	2	vCv	(235)	271	333
	3		(242)	247	259

Table 4. Mean V1CV2 durations (ms) in different speaking styles. The data in List 2 and Conversation are from the same speaker.

Table 4 shows some examples of the durations of the disyllabic units in different speaking styles. (The relatively large difference between the durations of Vcv and vCv type words in list reading is mainly due to a difference in the duration of the initial consonant which is not included here). The main difference between the speaking styles lies in the mean durations of the units in conversational speech on the one hand and list reading and poetry on the other: the durations are much shorter in conversational speech. The difference is least pronounced in Q3 units. Another difference is that the disyllabic words are markedly longer than the initial disyllabic units of a trisyllabic words in both list reading and in the reading of poetry. In conversational speech the difference is small and may sometimes be non-existent.

4. DISCUSSION

The results of this investigation show that the duration relations between disyllabic feet are not as stable as the intra-syllabic duration ratios. They vary with speaking style and may even vary within the same style. Isochrony appears strongest in list reading, at least when the list consists of words of the same length.

Of particular interest is the difference found between Vcv and vCv type words: in Vcv words the duration of a Q2 foot is very close to that of a Q3 foot; in vCv words, on the other hand, the difference in duration between Q2 and Q3 feet is larger than the difference between Q1 and Q2. This phenomenon appears in conversational speech and, to a lesser degree, in the reading of poetry, and has – as far as is known to the present author – not been reported before. It can be explained by the fact that in Vcv type words the quantity distinction is carried by the vowel, and Q3 can therefore be distinguished from Q2 by both temporal and tonal cues: Q3

words have a falling F0 contour in V1, while in Q2 words F0 is flat or rising. In vCv words, on the other hand, the quantity distinction is carried by the consonant, and V1 is always short. Therefore the Q2-Q3 distinction has to depend mainly on temporal relations. Eek has also noted a certain "peakiness" in the first syllable in connection with Q3 [2]. Such "peakiness" is not possible in connection with, for example, stop consonants.

This does not explain the fact that the difference between Vcv and vCv words only occurs in conversational speech and in oral production of poetry, but is absent in list reading. A possible reason could be that reading lists of disyllabic words can give rise to a certain steady rhythm, even though the words appear in sentence initial position (as in [1]) or in a carrier phrase. Another reason may be that reading lists of words or sentences is not meant to carry a semantic message in the same way as poetry or conversation.

In Introduction the question was posed whether it is foot isochrony or the intersyllabic duration ratio that is of primary importance in Estonian word prosody. Different opinions have been presented that have relevance to this question. Eek, for example, sees the "basic duration" of disyllabic structures as something to be preserved [1]; he notes that vowel duration increases less between Q2 and Q3 than does the duration of the intervocalic consonant in order to preserve this basic duration. More recently, Eek has shown that listeners can perceive the Q2-Q3 distinction even when part of V1 – and thus the characteristic syllabic ratio – is removed [2]. The large difference (43% on the average) in the duration of the initial consonant between Vcv and vCv type words in Eek's study of disyllabic words from 1974 [1] can also serve as an argument for the importance of isochrony. The initial C has no importance for the perception of quantity, therefore the sole function of this difference in its duration seems to be to equalize the durations of the Vcv and vCv type units. The material is so large – a total of 1033 words – that the phenomenon can hardly be accidental.

The most radical view against the importance of syllabic ratios is put forward by Hint [4], who sees the first – stressed – syllable as the sole definer of the degree of quantity, and the duration of V2 something that is dependent on V1 and not relevant in itself.

Against that point of view stands the fact that syllabic ratios are the most stable characteristic that have hitherto been shown to exist in connection with the degrees of quantity. If their main function is to uphold the isochrony of disyllabic units then they are not very effective: the present study has shown that foot isochrony can be weak in some speaking conditions. Moreover, it does not seem probable that the syllabic ratios would be as stable as they are if they did not have some other function, such as serving as a cue to perception. That this is in fact so, has been shown by Lehiste with synthetic stimuli. However, although necessary, the syllabic ratios alone do not provide sufficient cues for the recognition of Q3 [7]. Krull's perception test with isolated words extracted from conversational speech confirms this conclusion [6].

The difference in the relative foot durations between Vcv and vCv type words discussed above can be taken as another

indication that isochrony may not be of primary importance in Estonian word prosody. The relative size of the Q1-Q2 and Q2-Q3 differences is not relevant for isochrony, but a help to enhance the difference between Q2 and Q3 when the F0 cue is missing. There seems to be a constant interplay between different acoustic correlates to quantity. Or as Lehiste expresses it [11]: "Aspects of the system involve not only the presence or absence of a feature, but the relationship of the features to each other and the whole."

NOTES

1. Eek uses the term "accent" instead of the traditional "quantity", and A1, A2, A3 instead of Q1, Q2, Q3.

REFERENCES

- [1] Eek, A. 1974. Observations on the durations of some word structures: Part I. *Estonian Papers in Phonetics EPP 1974*, Tallinn.
- [2] Eek, A. and Meister, E. 1997. Simple perception experiments on Estonian Prosody: Foot structure vs. segmental quantity. In Lehiste, I. and Ross, J. (eds.), *Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody. Tallinn, Estonia, October 29-30, 1996*, Tallinn, 71-99.
- [3] Engstrand, O. and Krull, D. 1994. Durational correlates to quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. In Engstrand, O. and Kohler, K. (eds.), *Festschrift on the Occasion of Björn Lindblom's 60th Birthday. Phonetica* 51, No. 1-3, 80-91.
- [4] Hint, M. 1998. Why syllabic quantity? Why not the foot? In Klaus, V. (ed.), *Proceedings of the Finnic Phonetics Symposium, August 11-14, 1998, Pärnu, Estonia, Linguistica Uralica XXXIV:3*.
- [5] Krull, D. 1997. Prepausal lengthening in Estonian: Evidence from conversational speech. In Lehiste, I. and Ross, J. (eds.), *Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody. Tallinn, Estonia, October 29-30, 1996*, Tallinn, 136-148.
- [6] Krull, D. 1998. Perception of Estonian word prosody. A study of words extracted from conversational speech. In Klaus, V. (ed.), *Proceedings of the Finnic Phonetics Symposium, August 11-14, 1998, Pärnu, Estonia, Linguistica Uralica XXXIV:3*, 162-166.
- [7] Lehiste, I. 1960. Segmental and syllabic quantity in Estonian. *American Studies in Uralic Linguistics*, Vol. I, Bloomington, Indiana University, 21-82.
- [8] Lehiste, I. 1990. Phonetic investigation of metrical structure in orally produced poetry. *Journal of Phonetics* 18, 123-133.
- [9] Lehiste, I. 1995. *Keel ja Kirjandus* 8, 525-533, and 9, 596-603.
- [10] Lehiste, I. 1997. The structure of trisyllabic words. In Lehiste, I. and Ross, J. (eds.), *Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody. Tallinn, Estonia, October 29-30, 1996*, Tallinn, 149-164.
- [11] Lehiste, I. 1997. Search for phonetic correlates in Estonian prosody. In Lehiste, I. and Ross, J. (eds.), *Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody. Tallinn, Estonia, October 29-30, 1996*, 11-33.
- [12] Liiv, G. 1962. On the acoustic composition of Estonian vowels of three degrees of length. *Eesti NSV teaduste Akadeemia toimetised, XI, Ühiskonnateaduste seeria* 3, 271-290
- [13] Wiik, K. 1995. The Baltic Sea prosodic area revisited. In Suhonen, S. (ed.) *The Fenno-Baltic Cultural Area. Castrenianumin toimitteita* 49, Helsinki, 75-92.