5. Acoustic features of quantity in early recordings of Estonian[1]

Pärtel Lippus

5.1. Introduction

Starting with the early work of Ilse Lehiste[2] and Georg Liiv[3], for the past 60 years the three-way quantity system of Estonian has been the most extensively studied prosodic feature of the language. On a general level the issue that continues to attract researchers' interest is whether three-way oppositions are possible in human languages at all[4]. There has been a discussion among Estonian linguists of whether the three-way quantity system of Estonian should be regarded as a feature of phonemes, syllables, or feet[5]. Numerous experiments[6] have indicated that quantity is a property of the disyllabic foot. In Estonian, lexical stress is fixed on the first syllable and disyllabic feet can occur in short, long or overlong quantity degree (below, respectively Q1, Q2 and Q3). Phonologically, it is the stressed vowel (cf. [vilu] 'chilly, sg. nom.' – [vi:lu] 'slice, sg. gen.' – [vi::lu] 'slice, sg. part.'), syllable-medial consonants (e.g. [kalʲi] 'kvass, sg. nom.' – [kalʲli] 'hug, sg. nom.' – [kalʲ:li] 'precious, sg. gen.'), or a combination of a stressed vowel and the following consonant (e.g. [sate] 'fallout, sg. nom.' – [sa:tte] 'get, 2nd pers. pl.' – [sa:t:te] 'broadcast, sg. gen.') that carries quantity – unstressed syllables do not have length opposition[7].

---

1 The research underlying this chapter was partly funded from grant 7904 of the Estonian Science Foundation.

2 Ilse Lehiste, Segmental and syllabic quantity in Estonian, in: American Studies in Uralic Linguistics, Bloomington, vol. 1 (1960), pp. 21–82.

3 Georg Liiv, Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid [Duration and melodic types of vowels in the three quantity degrees of Estonian], in: Keel ja Kirjandus, no. 7 (1961), pp. 412–424, and no. 8 (1961), pp. 480–490.

4 Cf. Bert Remijsen / Leoma Gilley, Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect), in: Journal of Phonetics, vol. 36 (2008), pp. 318–344.

5 Cf. Arvo Eek / Einar Meister, Foneetilisi katseid ja arutlusi kvantiteedi alalt (I). Häälikukestusi muutvad kontekstid ja välde [Phonetics experiments and reflections regarding quantity (I). Contexts and quantity degrees that change the duration of sounds], in: Keel ja Kirjandus, no. 11 (2003), pp. 815–837 and no. 12 (2003), pp. 904–918.

6 See overview in Eek / Meister, Foneetilisi katseid (see previous footnote).

7 Tiit-Rein Viitso, Phonology, morphology and word formation, in: Estonian language, ed. by Mati Erelt, Tallinn: Esto 2003, pp. 9–92.

Due to a certain amount of foot isochrony[8], the duration of the second syllable (S2) compensates for the variation of the first syllable (S1) such that the unstressed syllable is the longest in Q1 and the shortest in Q3. There is no quantity opposition in monosyllabic words and, although the variation of length in unstressed syllables is not phonological, it is vital for perceiving quantity – perception tests have shown that quantity opposition is not perceived if the second syllable of a disyllabic word is not presented[9]. The duration of syllable onset consonants is mainly dependent on the local speech rate, while quantity can be described as the ratio of the duration of the rhyme of S1 and S2[10], or by comparing the duration of the nucleus of the stressed syllable with the weighted sum of segment durations within the foot[11].

In addition to the temporal structure of the disyllabic foot (the primary feature of quantity), variation of the pitch contour has also been studied extensively. As a rule, pitch in Q1 and Q2 is rising-falling with the peak at the end of S1, while in Q3 the peak occurs early in S1[12], or as relatively flat in

---

8 Ilse Lehiste, Prosodic change in progress: from quantity language to accent language, in: Development in prosodic systems, ed. by Paula Fikkert / Haike Jacobs, Berlin / New York: Mouton de Gruyter 2003, pp. 47–66; Francis Nolan / Eva Liina Asu, The pairwise variability index and coexisting rhythms in language, in: Phonetica, vol. 66 (2009), pp. 64–77.

9 Eek / Meister, Foneetilisi katseid (see footnote 5).

10 Eva Liina Asu / Pärtel Lippus / Pire Teras / Tuuli Tuisk, The realization of Estonian quantity characteristics in spontaneous speech, in: Nordic prosody. Proceedings of the 10th conference, Helsinki 2008, ed. by Martti Vainio / Reijo Aulanko / Olli Aaltonen, Frankfurt: Peter Lang 2009, pp. 49–56; Eek / Meister, Foneetilisi katseid (see footnote 4); Arvo Eek / Einar Meister, Foneetilisi katseid ja arutlusi kvantiteedi alalt (II). Takt, silp ja välde [Phonetics experiments and reflections regarding quantity (II). Foot, syllable, quantity degree], in: Keel ja Kirjandus, no. 4 (2004), pp. 251–271, and no. 5 (2004), pp. 336–357; Lehiste, Segmental and syllabic quantity (see footnote 2); Ilse Lehiste, Search for phonetic correlates in Estonian prosody, in: Estonian prosody: papers from a symposium, ed. by Ilse Lehiste / Jaan Ross, Tallinn: Institute of Estonian Language 1997, pp. 11–35; Lehiste, Prosodic change (see footnote 8); Liiv, Eesti keele kolme vältusastme (see footnote 3); Pärtel Lippus, Variation in vowel quality as a feature of Estonian quantity, in: Speech prosody 2010, 100877: 1–4 [Proceedings of the 2010 speech prosody conference (Chicago, USA, 11–14 May 2010) – techn. ed.], http://www.speechprosody2010.illinois.edu/papers/100877.pdf, accessed 26 May 2010.

11 For instance, see: Hartmund Traunmüller / Diana Krull, The effect of local speaking rate on the perception of quantity in Estonian, in: Phonetica, vol. 60 (2003), pp. 187–207; Diana Krull / Hartmund Traunmüller / Pier Marco Bertinetto, Local speaking rate and perceived quantity: an experiment with Italian listeners, in: Lund University, Centre for Languages & Literature, Dept. of Linguistics & Phonetics Working Papers, vol. 52 (2006), pp. 81–84.

12 Lehiste, Segmental and syllabic quantity (see footnote 2); Lehiste, Search for phonetic correlates (see footnote 10); Lehiste, Prosodic change (see footnote 8); Liiv, *Eesti keele kolme vältusastme* (see footnote 3); Mart Remmel, The phonetic scope of Estonian: some specifications, Preprint KKI-5, Tallinn: Academy of Sciences of the Estonian SSR, Institute of Language and Literature 1975.

S1 and falling at the syllable boundary in the case of Q1 and Q2, and falling from the beginning of S1 in the case of Q3[13].

It is thought that the pitch cue is of vital importance for making the distinction between Q2 and Q3[14]. Some researchers, however, disagree with this view, pointing out that there are many cases where there is a voiceless consonant in the stressed syllable coda, due to which the possible locations of an early or a late peak are unvoiced (e.g. [kat:ta])[15]. Nevertheless, recent perception studies[16] show that although discrimination between Q2 and Q3 may be affected in the case of conflicting temporal and pitch cues, it cannot be ruled out simply because the pitch cue is absent. The weight of the pitch cue appears to vary in accordance with the dialectal background of the speaker – Estonian speakers from North and West Estonia rely more on the pitch cue than those from South and East Estonia[17].

The development of the Estonian quantity system from a typical Finnic short-long opposition to the modern three-way system is thought to have resulted from a number of sound changes including apocopations and syncopations that took place between the 13th and 16th century[18]. The

---

13 Asu / Lippus / Teras / Tuisk, The realization (see footnote 10).

14 Arvo Eek, Estonian quantity: notes on the perception of duration, in: Estonian Papers in Phonetics 1979, ed. by Arvo Eek, Tallinn: Academy of Sciences of the Estonian SSR, Institute of Language and Literature 1980, pp. 5–29; Ilse Lehiste, Experiments with synthetic speech concerning quantity in Estonian, in: Congressus tertius internationalis Fenno-Ugristarum Tallinae habitus, 17–23 August 1970, pars I: Acta Linguistica, ed. by Valmen Hallap, Tallinn: Valgus 1975, pp. 254–269; Lehiste, Search for phonetic correlates (see footnote 10); Lehiste, Prosodic change (see footnote 8); Ilse Lehiste / Douglas G. Danforth, Foneettisten vihjeiden hierarkia viron kvantiteetin havaitsemisessa [The hierarchy of phonetic cues in the perception of Estonian quantity], in: Virittäjä, no. 4 (1977), pp. 404–411.

15 Traunmüller / Krull, The effect (see footnote 11).

16 Pärtel Lippus / Karl Pajusalu, Regional variation in the perception of Estonian quantity, in: Nordic prosody. Proceedings of the 10th conference, Helsinki 2008, ed. by Martti Vainio / Reijo Aulanko / Olli Aaltonen, Frankfurt: Peter Lang 2009, pp. 151–157; Pärtel Lippus / Karl Pajusalu / Jüri Allik, The tonal component in perception of the Estonian quantity, in: Proceedings of the 16th international congress of phonetic sciences in Saarbrücken, Germany, 6–10 August 2007, ID: 1029, http://www.icphs2007.de, accessed 30 September 2007; Pärtel Lippus / Karl Pajusalu / Jüri Allik, The tonal component of Estonian quantity in native and non-native perception, in: Journal of Phonetics, vol. 37 (2009), pp. 388–396; Pärtel Lippus / Karl Pajusalu / Jüri Allik, The role of pitch cue in the perception of the Estonian long quantity, in: Prosodic categories: production, perception and comprehension, ed. by Sónia Frota / Gorka Elordieta / Pilar Prieto, Studies in natural language and linguistic theory, Dordrecht / Heidelberg / London / New York: Springer 2011, pp. 231–242.

17 Lippus / Pajusalu, Regional variation (see footnote 16).

18 Arnold Kask, Eesti keele ajalooline grammatika [Historical grammar of Estonian], Tartu: Tartu Riiklik Ülikool 1972.

over-length of Q3 arose due to compensatory lengthening when trisyllabic words lost one of their unstressed syllables and their stressed syllables acquired the properties of what was previously a disyllabic sequence. Lehiste[19] claims that Estonian is still undergoing a prosodic change from a quantity language to an accent language: its short/long opposition is manifested by durational means only, whereas for the long/overlong opposition, pitch accent is brought into play as an additional marker. The recordings from the Berlin archives give us a unique opportunity to see whether there have been developments in Estonian word prosody during the past century. This chapter analyses the segmental duration and pitch patterns of disyllabic words used by informants in the Berlin recordings and compares the results to those of present-day speakers.

5.2. Materials and methods

From the Estonian material in the Berlin archives, the present analysis focuses on the recordings of eight informants who were asked to read a fairy-tale or a story from the Bible. The reason for choosing these particular recordings and disregarding others lies in the fact that, from the different types of material recorded, they represented the most natural speech pattern. The Berlin material also includes recordings of word-lists which could be considered as "data for linguistic purposes", many of which took the form of triplets consisting of the three principal case forms of nouns whose nominative was a monosyllabic foot and whose genitive and partitive represented a disyllabic opposition of Q2 and Q3. Most of the triplets, however, do not include a word in Q1 and cannot thus be considered a minimal triplet showing three-way quantity opposition. Even more serious, however, is the difficulty that consists in the high signal-to-noise level of phonograph recordings, which makes unstressed syllables of words uttered in isolation nearly undetectable. Detection of unstressed syllables is considerably easier in a recording consisting of fluent speech, in which such syllables are bounded by stressed ones.

The Berlin data selected for analysis in this chapter features eight male speakers (below referred to as "the Berlin group"): one from Saaremaa, three from Võrumaa and four from Tartumaa. The texts that the speakers from Saaremaa and Tartumaa were asked to read were probably written in standard (North) Estonian whereas the text read by the Võrumaa speakers was in South Estonian. The speaker from Saaremaa had a strong Saaremaa accent, while the speakers from Tartumaa exhibited no significant prosodic features that could be linked to a specific dialect.

---

19 Lehiste, Prosodic change (see footnote 8).

To compare the Berlin data with contemporary Estonian, contemporary recordings of informants of a similar dialectal background were chosen from the material in the Phonetic Corpus of Estonian Spontaneous Speech. As a side project, the informants whose spontaneous conversations had been recorded for the corpus were also asked to read the story *Põhjatuul ja päike* [The north wind and the sun] in Standard Estonian. The recordings were made in 2006–2009. In total there are 90 recordings from 45 speakers, since every subject was asked to read the story twice. For the study reported here, eight male speakers with a dialectal background similar to that of the informants in the Berlin group were selected (below: "the contemporary group"): three from Saaremaa, three from Võrumaa and two from Tartu.

A total of 1110 disyllabic words were extracted from the data. The pitch accent and the quantity characterising the words were assessed on an auditory basis by the author of this chapter. In the Berlin data the number of disyllabic words was 653. Of these, 551 had an H*+L pitch accent, 31 an L*+H pitch accent and 71 were deaccented (i.e., there was no marked pitch movement, the word was usually unstressed and preceded by a stressed word that carried a pitch accent). In the contemporary data there were 457 words: 288 H*+L, 59 L*+H, and 110 deaccented. Only words with an H*+L accent are analysed below: the samples of the other types were insufficient for a valid statistical analysis. The words with a long vowel in a closed S1 were not included in the analysis, because the number of tokens was insufficient. For the same reason, the words with an open S2 have been included in the group of words whose S2 is closed. For a detailed classification of the data selected for analysis, see Table 5.1.

Table 5.1. Number of analysed tokens.

		Number of speakers	Q1	Q2	Q2	Q3	Q3
			CVCV(C)	CV:CV(C)	CVCCV(C)	CV::CV(C)	CVC:CV(C)
Berlin	Saaremaa	1	20	5	8	8	17
	Võru	3	70	16	56	21	56
	Tartu	4	35	23	48	23	43
Contemp.	Saaremaa	3	16	6	29	9	29
	Võru	3	11	8	24	10	19
	Tartu	2	5	8	21	8	18

The selected data were analyzed in Praat[20]. Phoneme boundaries were tagged manually, after which segment durations and pitch contours were extracted with a script created by the author. Finally, five pitch points were selected: the beginning of the S1 nucleus, the end of S1, the beginning and end of the S2 nucleus, and the overall peak. The pitch in the selected points was checked manually. To facilitate the comparison of speakers, pitch values were converted from the logarithmic Hertz scale to a linear semitone (st) scale whose zero point lies at 50 Hz.

5.3. Results

5.3.1 Temporal characteristics

First of all, the analysis showed that the speech rate in the Berlin group was approximately 20% slower than in the contemporary group. This is demonstrated by comparing the mean duration of the words, which is 450 ms for the Berlin group and 370 ms for the contemporary group. Naturally, there is a certain degree of inaccuracy in this calculation due to the difference of distribution of the quantity degree in the Berlin and the contemporary group. Yet, the ANOVA test shows that the variation between the two groups is significant – $F(1, 668) = 114.00$, $p < 0.001$, while the difference between the dialects is not – $F(2, 667) = 2.80$, $p = 0.06$.

For both groups, the extracted words were divided into four groups according to their position within the phrase: phrase-initial, phrase-internal, phrase-final, and single-word phrases. The single-word phrases were left out of the analysis because they are highly focused and widely different in terms of word duration. The duration of the words appearing in different positions within the phrase (initial, internal, final) was compared to check whether word position had an effect on duration. In the Berlin group, the analysis did not reveal any positional effects – $F(2, 446) = 2.63$, $p = 0.076$. In the contemporary group, however, a significant phrase-final lengthening was demonstrated – $F(2, 218) = 51.99$, $p < 0.001$. The mean duration values in the contemporary group were 340 ms for words in phrase-initial and phrase-internal positions and 450 ms for words in phrase-final positions.

Syllable onset consonants are known to vary mainly in accordance with the speech rate and less with quantity. It was surprising to find a small variation in the duration of onset consonants (C1) between the quantity degrees. The difference was especially marked in the contemporary group: in Q1, Q2

---

20 Paul Boersma / David Weenink, Praat: doing phonetics by computer (Version 5.1.34), http://www.praat.org, accessed 31 May 2010.

Table 5.2. Mean segment durations and standard deviations for the dialects in the Berlin and the contemporary group.

		Berlin						Contemporary					
		Saaremaa mean	s.d.	Tartu mean	s.d.	Võru mean	s.d.	Saaremaa mean	s.d.	Tartu mean	s.d.	Võru mean	s.d.
S1 onset	Q1	73	17	64	18	82	22	75	19	71	6	81	43
	Q2 V	51	8	78	28	78	19	49	12	58	16	67	16
	Q2 C	62	12	69	24	75	18	55	25	62	22	62	18
	Q3 V	62	22	71	21	74	12	72	12	66	10	60	16
	Q3 C	74	18	66	17	70	20	53	24	56	17	51	20
S1 rhyme	Q1	90	21	86	26	94	26	66	17	57	12	67	25
	Q2 V	170	15	161	39	166	25	116	10	114	14	107	13
	Q2 C	161	27	153	33	165	33	129	15	119	21	126	21
		(100 + 62)		(96 + 57)		(101 + 63)		(77 + 51)		(67 + 53)		(75 + 51)	
	Q3 V	199	40	218	50	211	36	124	12	132	28	135	18
	Q3 C	225	33	226	50	246	43	184	38	188	44	176	47
		(110 + 116)		(95 + 130)		(113 + 133)		(80 + 103)		(74 + 114)		(66 + 111)	
S2 onset	Q1	58	14	51	16	65	13	47	14	49	6	50	14
	Q2 V	58	11	52	11	68	19	42	11	47	10	34	2
	Q2 C	56	11	51	15	59	13	40	15	44	14	42	14
	Q3 V	81	24	54	10	71	17	62	13	57	17	62	16
	Q3 C	58	13	63	16	63	14	58	18	60	17	57	13
S2 nucleus	Q1	119	34	109	29	131	30	103	27	85	27	84	31
	Q2 V	101	12	79	18	96	32	68	8	61	16	52	7
	Q2 C	99	26	75	18	97	33	73	20	61	20	65	24
	Q3 V	74	23	69	15	84	14	55	13	62	13	55	15
	Q3 C	63	18	70	22	81	22	58	18	55	22	46	14
S2 coda	Q1	98	29	70	19	64	14	100	19	98	13	100	48
	Q2 V	78	-	72	21	87	26	54	26	60	22	61	12
	Q2 C	88	9	75	17	81	23	57	22	70	16	75	33
	Q3 V	79	11	78	19	86	21	54	6	55	7	65	10
	Q3 C	84	15	80	18	68	15	64	19	64	23	72	34

and Q3 words, C1 was 77 ms, 59 ms and 57 ms respectively. The difference is smaller in the Berlin group – 75 ms, 73 ms and 70 ms respectively. Variation was more limited in the onset consonant of the unstressed syllable (C2). In the Berlin group, the mean C2 was 60 ms while in the contemporary group was is 50 ms. The duration of all segments is presented in Table 5.2.

V1 is phonologically short in Q1 words as well as in those Q2 and Q3 words which start with a closed syllable. V1 is long in Q2 and Q3 words which have an open S1. In the Berlin group, the mean duration of a short, long and overlong V1 is 98 ms, 164 ms and 213 ms respectively. Thus, long vowels in the Berlin group tend to be 66 ms longer than short ones (Q2/Q1 ratio is 1.67) and the duration of overlong vowels exceeds that of long ones by 49 ms (Q3/Q1 ratio is 2.17). In the contemporary group the mean duration of short, long, and overlong first vowels is 72 ms, 112 ms and 131 ms respectively. This means that a long first vowel in that group is on average by 40 ms longer than a short one (Q2/Q1 ratio is 1.55), while the duration of an overlong V1 exceeds that of a long one by only 19 ms (Q3/Q1 ratio is 1.82).

In the words with a closed S1, the coda consonant of the syllable is phonologically short in Q2 words and long in Q3 words. The duration of S1 offset consonants in Q2 and Q3 words should be comparable to that of short and long vowels in Q1 and Q2 words, respectively. In the Berlin group the mean duration of short consonants in that position in Q2 words is 57 ms and of corresponding long consonants in Q3 words 130 ms, while in the contemporary group the corresponding values are 49 ms and 88 ms, respectively.

The rhyme of open S1 and closed S1 in Q2 words in both groups is almost equal in length: the respective figures for the Berlin group are 164 ms and 159 ms and for the contemporary group, 112 ms and 123 ms. The difference between open and closed syllables is not significant – $F(1, 250) = 0.168$, $p = 0.682$. In Q3 the S1 rhyme is significantly longer in words with a short V1 and a closed S1 than in words with an overlong V1 and an open S1 – ANOVA shows a significant difference $F(1, 259) = 34.16$, $p < 0.001$. The corresponding figures for the Berlin group are 213 ms in the case of open S1 vs. 236 ms in the case of closed S1, and for the contemporary group, 131 ms vs. 161 ms, respectively.

Although the S2 vowel does not carry any phonological length, foot isochrony requires it to compensate the length of S1, such that V2 is half-long in Q1 words, short in Q2 words and extra-short in Q3 words. The mean duration of V2 in Q1, Q2 and Q3 words in the Berlin group is 123 ms, 88 ms and 75 ms respectively. The corresponding figures for the contemporary group are 87 ms, 62 ms and 49 ms. A significant effect ($F(1, 668) = 58.38$,

$p < 0.001$) on the duration of V2 is also exerted by the type of S2: V2 duration is almost 20% less when S2 is closed. Regardless of this and due to the limited sample of this study, for the purposes of analysis, words with a closed S2 were grouped with those that have an open S2. Furthermore, the existence or absence of a final consonant in S2 affects the preceding vowels similarly in all quantity degrees.

The S2 final consonant is phonologically short (as are syllable-initial ones) and is not affected by the quantity degree. In the Berlin group, the mean C3 duration is 79 ms and there is no significant variation due to the quantity of the word ($F(2, 446) = 0.81$, $p = 0.372$) or the position of the word in the phrase ($F(2, 446) = 0.20$, $p = 0.819$). In the contemporary group, the effect of word position is significant ($F(2, 218) = 19.36$, $p < 0.001$): the mean duration of C3 in phrase-initial, phrase-internal and phrase-final position is, respectively, 75 ms, 60 ms and 100 ms. Quantity does not affect the duration of C3 in the contemporary group either ($F(2, 218) = 1.21$, $p = 0.276$).

The ratio of S1 and S2 rhyme duration (S1/S2 ratio, see Figure 5.1) is the most commonly used characteristic for comparing quantity degrees of disyllabic feet. In both the Berlin and the contemporary group the mean S1/S2 ratio for Q1 and Q2 was 0.77 and 1.98 respectively. The difference between the degrees is significant ($F(1, 407) = 566.22$, $p < 0.001$), while that between the Berlin and the contemporary group is not ($F(1, 407) = 0.005$, $p = 0.947$). No significant difference was found between words with an open S1 and those with a closed one ($F(1, 250) = 0.441$, $p = 0.507$). The S1/S2 ratio for Q3 in words with an open S1 is different from that of words with a closed S1 ($F(1, 259) = 30.33$, $p < 0.001$) – the respective figures are 2.76 and 3.47.

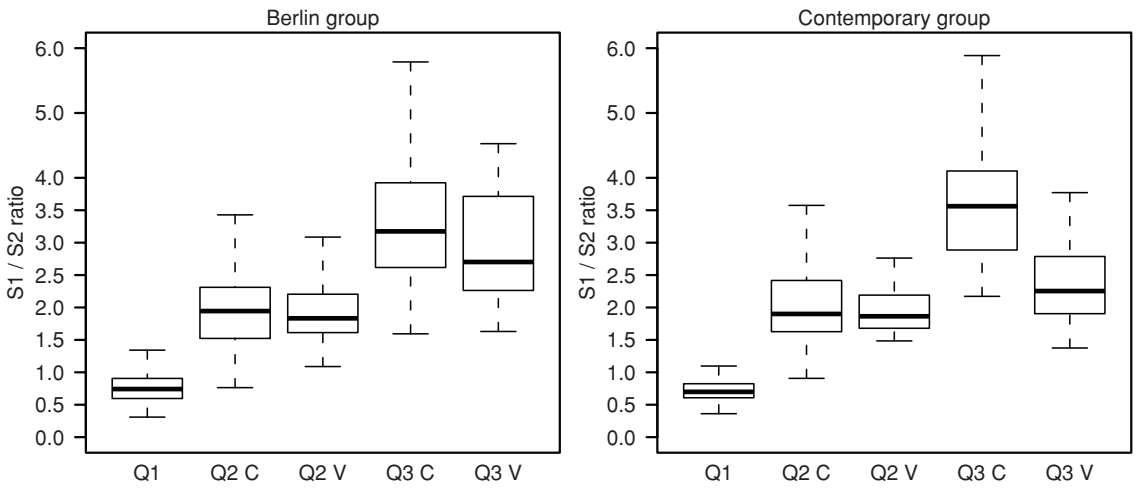

Figure 5.1. S1/S2 ratio in the Berlin (left) and the contemporary (right) group. The letter C is added to the quantity reference of groups of words with a closed S1, while V is used to denote an open S1. The fat line corresponds to the median, the box to the first and third quartiles, and the whiskers to the 95% confidence intervals.

5.3.2 Tonal characteristics

A comparison of the mean pitch of the two groups showed that in the Berlin group the mean pitch was 25 st or 214 Hz, while in the contemporary group it was only 16 st or 128 Hz (standard deviation in both groups was 4 st). This suggests that 100 years ago Estonian was spoken or at least read with a considerably higher pitch than today.

The pitch in the Berlin group is rising-falling, with the F0 in all quantity degrees peaking in the second half of S1 or later. In the contemporary group the speakers seemed to produce two patterns: in some words the F0 peak is in the very beginning of V1 and in others, in the middle of the word. At first sight it seemed that the peak alignment does not depend on the quantity degree of the word. To see the regularities, the pitch patterns in the contemporary group were divided into two sets: words where the peak was in the first half of S1 and those where the peak occurs later. When we look at the data in Table 5.3, we can see that each dialect group has its own particular pattern. In the Saaremaa dialect, the peak in most Q1 and Q2 words occurs later in the word while the peak in most Q3 words is in the first half of the first syllable (the chi-square test shows a significant difference between the quantity degrees: $\chi^2(2, N = 89) = 15.95, p < 0.001$). In the Tartu dialect, the peak usually appears in the beginning of the word and there are no significant differences between quantity degrees ($\chi^2(2, N = 60) = 3.63, p = 0.163$). In the Võru dialect, words with Q1 and Q2 showed a random distribution of early and late peaks (the peaks in Q1 and Q2 words do not exhibit a regular pattern: $\chi^2(1, N = 43) = 0.07, p = 0.795$), whereas in Q3 the peak tended to appear in the beginning of the word (the distribution difference between peaks in Q2 and Q3 words was significant: $\chi^2(1, N = 29) = 4.33, p < 0.05$).

Table 5.3. The contemporary group: number of words with an early (first half of S1) or late pitch peak

Location of peak	Saaremaa			Tartu			Võru		
	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
Early	6	10	28	2	20	21	5	16	22
Late	10	25	10	3	9	5	6	16	7

In the Berlin group the pitch is always rising-falling. On the average, the peak is the latest in the Saaremaa dialect and the earliest in the Tartu dialect (respectively, 110 ms and 77 ms from the beginning of V1), with the Võru dialect falling in the middle (95 ms from the beginning of V1). The difference between all dialects is significant ($F(2, 446) = 11.14, p < 0.001$), whereas that between the quantity degrees it is not ($F(2, 446) = 3.41, p = 0.066$).

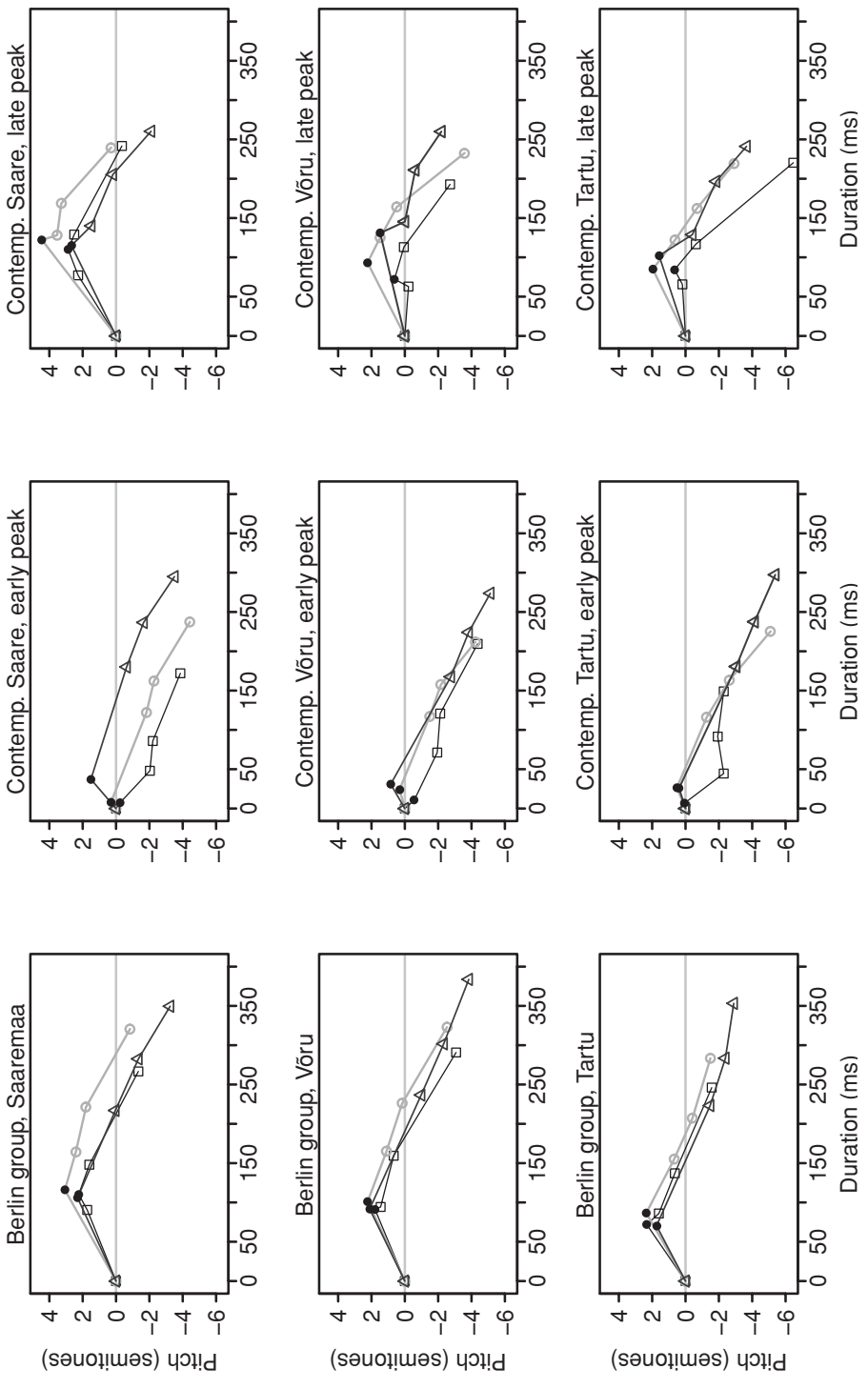

Figure 5.2. Pitch curves for the Berlin and the contemporary group. The zero point on both axes corresponds to the beginning of V1 (squares Q1, circles Q2, and triangles Q3). Peaks are marked with black filled circles. The left column represents data from the Berlin group, the middle column data from the contemporary group (words whose F0 peaks at the beginning of V1) and the right column data from the contemporary group (words with the F0 peak in the beginning and with the F0 peak in the middle of the word).

In the contemporary Saaremaa dialect the pitch is usually rising-falling, with the peak at approximately 118 ms from the beginning of V1. There is no difference between quantity degrees ($F(2, 218) = 0.37$, $p = 0.693$). In the Tartu and Võru dialect the peak is at 95 ms from the beginning of V1 and there is no significant difference between the two dialects ($F(1, 159) = 0.50$, $p = 0.484$). The difference between the contemporary Saaremaa dialect on one hand and the Tartu and Võru dialects on the other is significant ($F(2, 218) = 3.39$, $p < 0.05$).

Data from the contemporary group showed that when the pitch in a word was falling from the beginning of V1, there was no significant difference between the dialect groups ($F(2, 127) = 0.69$, $p = 0.507$). At the same time, these cases demonstrated a significant difference between Q1 and Q3 and between Q2 and Q3 ($F(2, 127) = 6.88$, $p < 0.005$) – on the average the peak occurs 8 ms after the beginning of V1 in Q1 and Q2, while the corresponding value in the case of Q3 was 26 ms.

In the Berlin group, the pitch rises about 2 st from 25 st at the beginning of V1 to achieve a peak of 27 st. There is no dialect difference in pitch rise ($F(2, 446) = 1.91$, $p = 0.149$). In the contemporary group, pitch rises by about 3.7 st (from 16.3 st to 20 st) in the Saaremaa dialect, while the corresponding increase in the Tartu and Võru dialect is 1.7 st (from 17.5 st to 19.2 st). The difference between Saaremaa dialect on one hand and the Tartu and Võru dialect on the other is significant ($F(2, 218) = 15.67$, $p < 0.001$). The difference between the Berlin group and the Tartu and Võru dialects of the contemporary group is also significant ($F(1, 493) = 4.20$, $p < 0.05$).

In both the Berlin and the contemporary group, the pitch fall from the peak to the end of the foot seems to vary in accordance with foot quantity ($F(2, 667) = 10.84$, $p < 0.001$) and the position of the word in the phrase ($F(2, 667) = 47.87$, $p < 0.001$). In phrase-initial and phrase-internal positions, the pitch in Q1, Q2 and Q3 words falls by 3.6, 4.3 and 4.8 st respectively. In the phrase-final position, pitch fall amounts to 7 st in all quantities and there is no significant difference between the quantity degrees ($F(2, 165) = 1.29$, $p = 0.278$).

5.4. Discussion

It is hard to say whether the reading style or the prosody itself has changed during the past hundred years. What is clear is that the speakers in the hundred-year-old recordings spoke more slowly and with a higher pitch than contemporary speakers do. Of course some of the differences can to a certain extent be ascribed to developments in technology, namely the recording devices. We can assume that for the speakers in the Berlin group the situation

of being recorded with a phonograph device was a first-time experience, causing anxiety and stress, whereas to be recorded today in various media and by means of a variety of devices is very much part of our everyday lives. On the other hand, according to the author's impression, the reading style in the Berlin data reminds one of the narrating style used in early broadcasts from the 1930s–1960s by trained broadcast reporters. Unfortunately the prosody of broadcast speech has not been studied in Estonian and we do not know what distinguishes the prosody of everyday usage from that of a media broadcast.

The temporal characteristics of word prosody are relatively similar in the two groups. Due to the difference in speech rate between the Berlin and the contemporary group it is difficult to compare the absolute values of the segments. It can be affirmed, however, that both groups share a series of prosodic patterns. For instance, in both groups the syllable onsets are short consonants and their duration is comparable to the duration of short vowels, while long vowels in Q2 words are about 1.5 times longer than short ones. The groups diverge in that in the Berlin group the long vowels in Q3 are more than twice as long as short vowels, whereas in the contemporary group they are only about 1.8 times longer, which means that the difference between long vowels in Q2 words and their counterparts in Q3 words is much smaller in the contemporary group. As for the S1/S2 ratio, which in most cases was surprisingly similar in the two groups, decreases considerably in the contemporary group in the case of Q3 words with a long open S1. This means that the distinction between Q2 and Q3 in the contemporary group is less marked and additional characteristics (such as pitch) may be needed to establish the difference.

In the Berlin group the pitch contour is linked to the absolute time scale: in all quantity degrees, the pitch rises by 2 st during approximately a hundred milliseconds starting from the beginning of the word, and then starts to fall. Of course, as the duration of the stressed syllable varies between the quantity degrees, this also affects the location of the F0 peak within the stressed syllable boundaries. Thus, when we examine the F0 peak location in the stressed syllable, it matches the descriptions discussed in the introductory section of this chapter. In the contemporary group, the situation is different. There are two types of pitch patterns: the pitch either starts as a falling one from the beginning of the word or it may rise and fall similarly to the pattern in the Berlin data. Neither of these patterns is linked to a particular quantity degree, yet there is a tendency to use the rising-falling pattern in Q1 and Q2 words and a falling pattern in Q3 words.

Apparently, the differences in quantity realisation between the dialects are rather small. In the Berlin group the only difference is that the peak occurs the latest in Saaremaa dialect and the earliest in the Tartu dialect, with the Võru dialect falling in the middle. In the contemporary group, the late peak

in the Saaremaa dialect also stands out clearly, although the rising-falling pattern now appears to be regularly used in Q1 and Q2 words, while Q3 words tend to be frequently produced with the falling pattern. In the Võru dialect the two patterns are equally used with Q1 and Q2 words, whereas the falling pattern is more frequent in Q3 words. Only in the Tartu dialect is the falling pattern preferred with all quantity degrees, making the rising-falling pattern the less frequently used one. These results support the conclusions drawn on the basis of a series of perception tests conducted by Lippus and Pajusalu[21], which showed that Estonian listeners from North and West Estonia preferred to judge quantity by reference to the pitch cue while speakers from East and South Estonia based their decisions mainly on temporal cues.

5.5. Conclusion

The syllable duration ratio introduced by Ilse Lehiste in the early 1960s for describing the Estonian three-way quantity system has turned out to be the most stable characteristic of the Estonian quantity system. The similarity of the syllable ratios derived from the Berlin data collected in the beginning of the 20th century and from the contemporary data is simply surprising. At the same time the reading style has changed – in the contemporary data, the speech rate is considerably faster and the pitch lower than in the Berlin data.

Since the first experimental studies of Estonian quantity in the early 1960s, pitch contour has been claimed to be a descriptive feature which, in terms of discriminating between Q3 and Q2, was second in importance only to the syllable duration ratio mentioned above. Yet, in the Berlin data, the contrastive power of the pitch contour was less marked, because pitch was linked to the absolute time scale. The pitch in the Berlin data analysed was always rising-falling, peaking in the middle of the word. In the contemporary data, on the other hand, the differences between the temporal structure of Q2 and Q3 are smaller, and different pitch patterns are often used to emphasise the opposition. The two different pitch patterns are used with all quantity degrees, although the contour with the peak in the beginning of the word is more frequently used with Q3 words.

---

21 Lippus / Pajusalu, Regional variation (see footnote 16).