https://doi.org/10.3176/lu.1998.3.11

ARVO EEK, EINAR MEISTER (Tallinn)

QUALITY OF STANDARD ESTONIAN VOWELS IN STRESSED AND UNSTRESSED SYLLABLES OF THE FEET IN THREE DISTINCTIVE QUANTITY DEGREES*

Introduction

In recent works we have reached a conclusion that, in Estonian phonology, it would be reasonable to define segmental phonemes as phonologically relevant representatives of the corresponding sound types. We proposed that there are no long (and also overlong) segmental phonemes in Estonian. All vowel and consonant phonemes can occur in a short and a long duration degree. A long degree is represented by sequences of two identical segmental phonemes (V, VV and C, CC). The three-way contrast is projected to foot level. Estonian has three phonological foot patterns called quantity degrees (Eek, Meister 1997; 1998).

There is a strong argument in favour of this analysis: long monophthongs behave like diphthongs, and geminates as consonant clusters in word prosody. This treatment is supported by a widespread opinion that quality differences between vowels of different length are not so big that short and long vowels should be considered as different phonemes (cf. e.g. English where short and long vowels are qualitatively different enough to be considered as different phonemes; cf. Wiik 1965 : 57—60). However, G. Liiv's (1962) data show a regular change in the stressed-syllable vowel quality accompanying with an increase of the foot quantity degree, e.g. /i/, /ä/, /u/ and /a/ gradually move to a more peripheral position in the acoustic space of vowels; /e/ moves higher and forwards; /ü/, /ö/ and /õ/ tend to move to higher and backward positions, and /o/ shifts backwards. Unfortunately G. Liiv's data do not allow to estimate the magnitude of these quantity dependent shifts in vowel quality and to test their perceptual relevance.

We presume that in this respect the situation is similar to Finnish where short vowels of the stressed syllable overlap to a great extent with corresponding long vowels in the acoustic space of F1 vs. F2 (see Wiik 1965 : Table 1), and as the psychoacoustic differences between short vs. long vowels are below 1 Bark, they should belong to one and the same category of vowel quality (Iivonen 1987 : Figure 19).

The second question we address is: which type — Finnic or Germanic — char-

* The support from the Estonian Science Foundation has made the present research possible.

acterises quality reduction of Standard Estonian vowels in the unstressed second syllable of Q1, Q2 and Q3 feet? Kalevi Wiik (1965 : 131—134) has mentioned that the formant positions of Finnish unstressed vowels are roughly similar to those of stressed ones, whereas the shortest unstressed vowels, following a heavy syllable, tend to centralise slightly more than the half-long unstressed vowels following a light syllable (cf. also Iivonen 1988 : Figure 5). Moreover, J. Local and R. Ogden (1998 : 21) have added that values of F1 and F2 for the Finnish short and half-long unstressed /a/ are almost identical irrespective of the nature of the first syllable. We suppose that the described Finnic-type of vowel quality reduction is also valid at least for the Estonian half-long unstressed vowels in Q1 feet but this weak reduction type is most likely not applicable to the extra short unstressed vowels in Q3 feet. Presumably Q3 feet are best characterised by the Germanic-type of reduction (cf. e.g. a strong centralised quality reduction in English, Wiik 1965 : 133), being rather dependent on the length and tenseness of the first peaked syllable, controlled probably by two relatively autonomous mechanisms, than directly related to the concrete duration of an unstressed-syllable vowel.

Methods and speakers

We have measured 227 CV(V)CV words where intervocalic C's are non-palatalised/palatalised alveolar consonants (with some exceptions). These words are taken from the text corpus of the Estonian Phonetic Database (18 filler sentences from Block S1: vowels in the feet of differnt quantity degrees, and 6 sentences from Passages). Speech recording was directed to disk with a sampling rate of 20 kHz, and was controlled by the EUROPEC software. Recordings were made in a sound-treated chamber using the Sony ECM44 microphone. Durations and frequencies of F1—F4 of vowels are measured on wideband spectrograms using Kay Elemetrics CSL 4300 system (formant frequencies are measured at the quasi-stationary phase defined visually on a wideband spectrogram). Measurements are based on the material of 3 male speakers whose speech represents correct Estonian pronunciation (AE — b. 1937, from Western dialect area; AA — b. 1955, from Western dialect area; MK — b. 1975, from Tallinn).

Results and discussion

Data about stressed-syllable vowels are presented in Table 1 and Figure 1.

The stressed-syllable high/low vowels in a Q1 and Q3 foot, as the shortest/longest vowels, have also qualitatively occupied the marginal positions in the production space of F1 vs. F2: in a Q3 foot (if compared with Q1) /i/ and /ü/ have moved higher and forwards; /u/ has moved higher and backwards; /ä/ has become lower and /a/ has moved lower and backwards. The stressed vowel in Q3 feet is the most peripheral and the stressed vowel in Q1 feet is most centralised (Figure 1). The stressed vowels /i ü u ä a/ in Q2 feet are located in the proximity to the corresponding stressed vowels in Q3 feet.

Quantity induced qualitative differences of the stressed mid vowels are as follows: the long stressed /e/ is higher and more anterior than that of the short /e/; the long stressed /o/ is higher and more back-positioned than the short /o/; the long stressed /õ/ is more back than the short one; the stressed /ö/ in a Q3 foot is the lowest and backmost. All these quality differences of the stressed vowels are relatively small and do not exceed 1 Bark. The only noteworthy exception: the short stressed /õ/ has moved considerably forwards. Shifting a short /õ/ into the /ö/ space is obviously due to the unbalanced phonemic context of the measured words because in the majority of them /õ/ was followed by a palatalised alveolar consonant which shifts the vowel forward.

Table 1

Average durations (ms) and formant frequencies (Hz) of the stressed-syllable vowels of Q1, Q2 and Q3 feet (3 male speakers)

		Q1 foot					Q2 foot					Q3 foot				
		D	F1	F2	F3	N	D	F1	F2	F3	N	D	F1	F2	F3	N
/i/	x̄	69	251	2079	2696	15	141	236	2270	2906	6	202	224	2271	3024	6
	sd	17	27	180	212		13	13	160	264		42	10	192	204	
	min	43	203	1863	2318		122	227	1979	2436		154	203	2005	2690	
	max	118	329	2385	3020		162	253	2411	3248		255	228	2436	3197	
/e/	x̄	82	493	1729	2492	11	147	454	1827	2434	6	223	435	2009	2546	6
	sd	17	79	220	177		20	76	252	182		40	44	209	177	
	min	60	363	1409	2233		117	318	1421	2157		185	380	1675	2284	
	max	117	609	2131	2791		173	532	2106	2614		274	507	2208	2791	
/ä/	x̄	94	655	1551	2446	9	182	683	1569	2445	9	251	676	1548	2453	6
	sd	22	65	160	190		25	60	168	201		72	75	182	154	
	min	70	583	1345	1954		148	590	1345	2030		182	558	1319	2284	
	max	137	786	1776	2588		223	761	1751	2639		365	761	1725	2664	
/ü/	x̄	85	288	1624	2220	12	167	253	1704	2136	6	213	240	1703	2128	6
	sd	27	89	132	147		25	16	167	194		32	25	121	192	
	min	46	228	1370	1909		144	228	1395	1852		167	203	1590	1878	
	max	138	558	1903	2436		208	279	1903	2309		254	272	1878	2309	
/ö/	x̄	97	477	1552	2322	6	141	446	1561	2306	10	200	491	1525	2283	8
	sd	27	78	159	205		30	69	100	133		39	54	110	169	
	min	76	380	1345	1954		85	355	1421	1972		146	431	1370	2030	
	max	122	583	1827	2512		176	558	1751	2411		266	558	1649	2487	
/u/	x̄	82	267	655	2447	11	148	253	605		7	193	244	596		6
	sd	15	24	64	207		23	29	44			36	19	53		
	min	63	228	558	2208		118	228	532			166	228	558		
	max	114	304	761	2569		187	304	681			265	273	685		
/o/	x̄	82	486	884	2335	11	167	402	807	2531	6	229	444	831	2525	8
	sd	25	32	77	165		38	52	37	56		63	53	50	73	
	min	50	454	786	1903		134	329	761	2461		153	380	761	2411	
	max	130	558	1040	2538		236	456	862	2588		329	558	913	2614	
/õ/	x̄	83	424	1395	2452	8	147	411	1120	2558	6	182	455	1184	2436	3
	sd	19	71	210	226		21	81	38	162		2	1	29	217	
	min	56	304	1142	1979		125	279	1065	2363		180	454	1167	2233	
	max	110	507	1649	2690		177	507	1167	2791		184	456	1218	2664	
/a/	x̄	90	654	1142	2428	14	157	631	1066	2497	6	245	671	1071	2460	6
	sd	21	76	80	93		35	58	53	106		65	53	63	84	
	min	61	499	1040	2309		119	558	1015	2385		188	609	989	2360	
	max	139	736	1363	2563		219	710	1142	2639		341	727	1142	2538	

In the two-formant production space the unrounded /i/ has occupied a part of the perception space of the rounded /ü/ and the unrounded /e/ is located in the perception space of the rounded /ö/. When F3 and F4 are removed from the spectrum, /i/ is perceived as /ü/ (Eek, Meister 1994 : 64; cf. Schwartz, Beautemps, Abry, Escudier 1993) and /e/ as /ö/. Figure 1 shows that /i/ — /ü/ and /e/ — /ö/ differ essentially on the basis of F3. The distinction of /i/ — /ü/ as well as /e/ — /ö/ increases namely owing to the fact that F3 is located near the relatively strong F2 of rounded front vowels which amplifies the cumulative effect of F2 and F3, whereas in the case of unrounded front vowels the proximity of the strong F4 to the dominant F3 can direct the listener's attention to the upper part of the spectrum. Therefore, the perceptual parameter F2′ seems to describe well the perceptual phenomena governing this contrast.

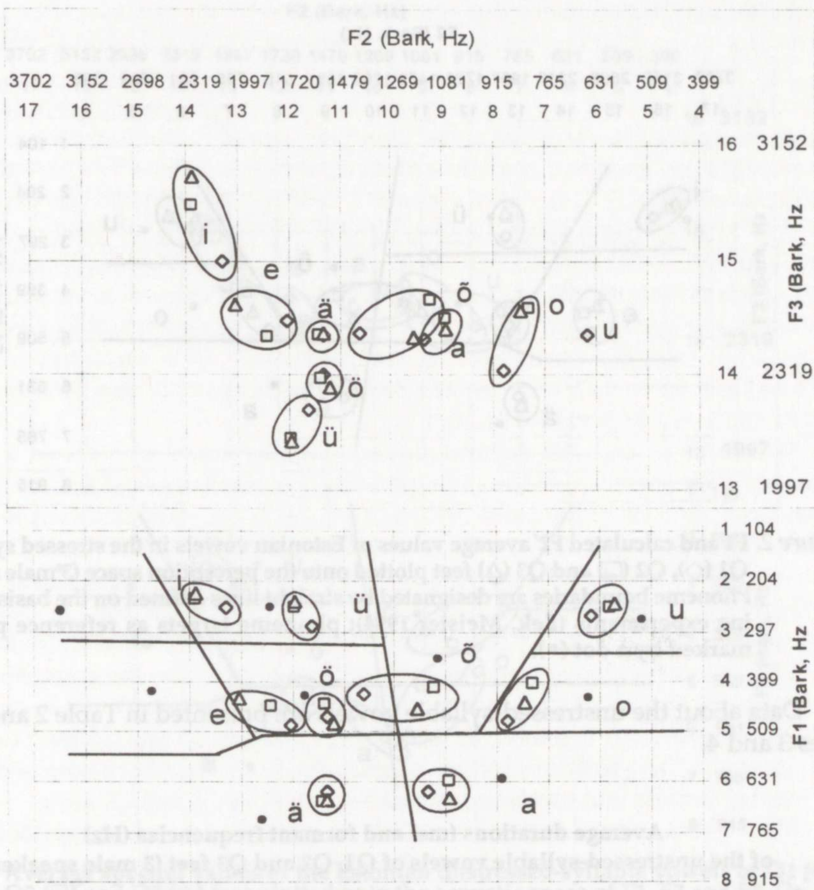

Figure 1. **Average spectral values of the Estonian stressed-syllable vowels of Q1 (◇), Q2 (□) and Q3 (△) feet plotted onto the acoustic space of F1, F2 and F3. On the lower part of the Figure phoneme boundaries are designated by straight lines defined on the basis of matching experiments (Eek, Meister 1994 : Fig. 6); phoneme targets as reference points in the two-formant perception space are marked by a dot (•).**

In order to estimate the quality differences of stressed vowels in the perception space we calculated the "effective" second formant (F2′) values on the basis of the first four formants. For the calculation of the F2′ we used the formula from Bladon, Fant 1978 : 3. The positions of stressed-syllable vowels in the perception space are presented in Figure 2.

In perception experiments with two phonetically trained listeners, all the measured words, separated from the sentence context, were identified according to the read text. Stressed-syllable long vowels, separated from Q2 and Q3 words, were identified as the corresponding intended vowel types. The same is also valid for the stressed-syllable vowels separated from Q1 feet. Exceptions: short isolated /õ/ was perceived 37% of the cases as /ö/ (before palatalised alveolars); short /ü/ 16% as /ö/ (in the word *ühe*) and short /e/ 18% as /ö/. On the basis of average values, short and long stressed-syllable vowels are situated in the perception space of the corresponding phonemes, whereas the phonemes are best characterised by a long vowel in the Q3 foot and worst by a short vowel in the Q1 foot. However, these differences do not exceed 1 Bark. Long and short vowels differ little in quality. Therefore, defining them as different phonemes based only on quality is not justified.

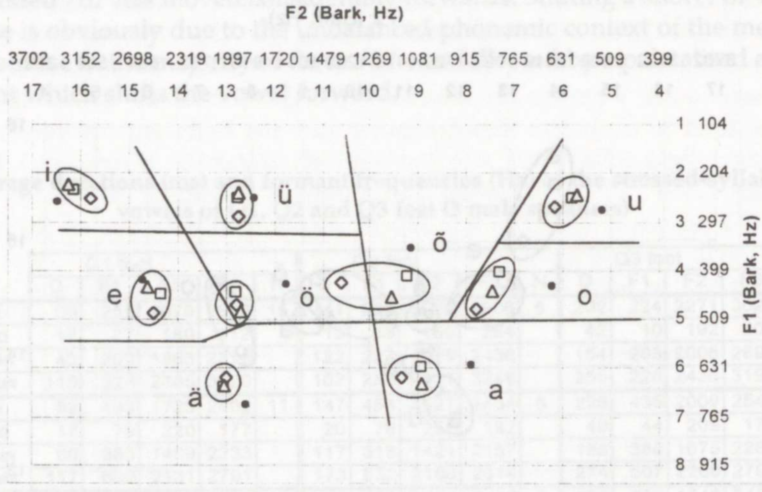

Figure 2. **F1 and calculated F2′ average values of Estonian vowels in the stressed syllables of Q1 (◇), Q2 (□) and Q3 (△) feet plotted onto the perception space (3 male speakers). Phoneme boundaries are designated by straight lines defined on the basis of matching experiments (Eek, Meister 1994); phoneme targets as reference points are marked by a dot (•).**

Data about the unstressed-syllable vowels are presented in Table 2 and on Figures 3 and 4.

Table 2

Average durations (ms) and formant frequencies (Hz) of the unstressed-syllable vowels of Q1, Q2 and Q3 feet (3 male speakers)

		Q1 foot					Q2 foot					Q3 foot				
		D	F1	F2	F3	N	D	F1	F2	F3	N	D	F1	F2	F3	N
/i/	$\bar{x}$	121	252	2145	2730	40	84	263	2075	2675	37	67	289	1961	2530	35
	sd	34	29	186	192		20	42	189	203		20	54	190	223	
	min	55	203	1776	2309		55	203	1751	2284		38	203	1649	1878	
	max	221	329	2411	3020		143	380	2385	2944		126	431	2309	2893	
/e/	$\bar{x}$	115	565	1653	2460	21	81	526	1569	2386	7	83	549	1567	2369	12
	sd	35	84	167	135		12	84	178	148		22	91	135	181	
	min	63	380	1269	2208		64	409	1370	2106		48	431	1319	1954	
	max	199	710	1928	2614		92	634	1776	2563		134	710	1725	2563	
/u/	$\bar{x}$	106	340	822	2415	11	76	416	982	2309	8	73	383	1084	2421	17
	sd	31	109	219	122		31	40	197	424		23	84	163	22	
	min	70	227	532	2258		51	355	685	1852		43	272	609	2411	
	max	169	499	1142	2588		146	482	1181	2690		138	558	1272	2461	
/o/	$\bar{x}$	70	482	901	2373	2	90	520	1095	2416	6	77	634	1142	2512	1
	sd	6	0	18	305		58	74	109	227		0	0	0	0	
	min	66	482	888	2157		40	456	939	2157		77	634	1142	2512	
	max	74	482	913	2588		198	636	1192	2614		77	634	1142	2512	
/a/	$\bar{x}$	116	668	1176	2450	10	93	652	1231	2487	8	85	579	1359	2440	12
	sd	35	100	83	125		14	84	86	95		26	95	93	156	
	min	72	482	1065	2233		73	499	1167	2309		56	431	1218	2030	
	max	163	761	1345	2614		115	761	1421	2588		150	710	1548	2588	

In an unstressed syllable, contrary to the stressed one, the "half-long" vowels of Q1 feet are the least reduced and the "extra short" vowels of Q3 feet the most centralised leaving the unstressed vowels of Q2 feet in an intermediate position (cf. Figures 3 and 4). Especially the unstressed back vowels of a Q3 foot are extremely centralised.

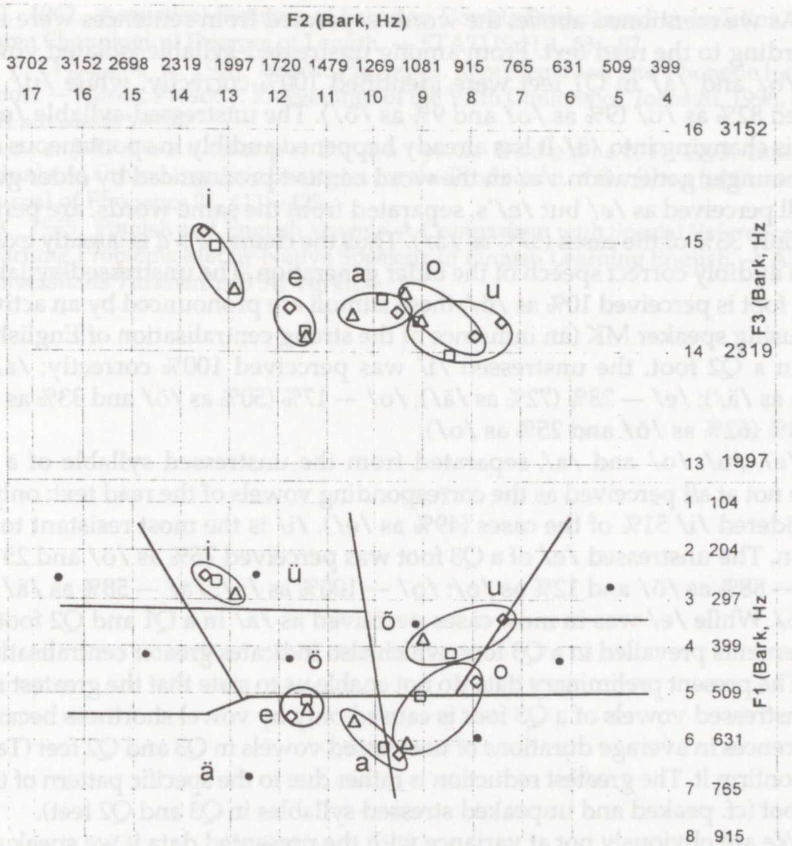

Figure 3. **Average spectral values of the Estonian unstressed-syllable vowels of Q1 (◇), Q2 (□) and Q3 (△) feet plotted onto the acoustic space of F1, F2 and F3. On the lower part of the Figure phoneme boundaries are designated by straight lines defined on the basis of matching experiments (Eek, Meister 1994); phoneme targets as reference points in the two-formant perception space are marked by a dot (•).**

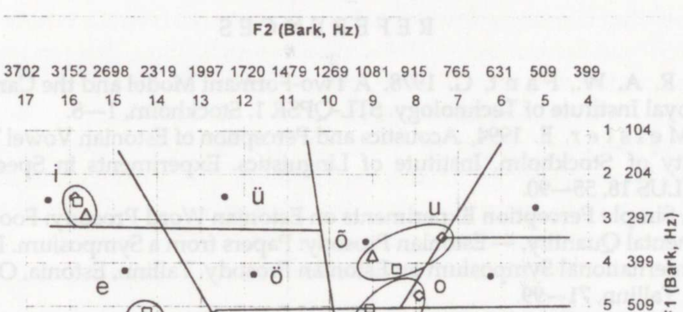

Figure 4. **F1 and calculated F2′ average values of Estonian vowels in the unstressed syllables of Q1 (◇), Q2 (□) and Q3 (△) feet plotted onto the perception space (3 male speakers). Phoneme boundaries are designated by straight lines defined on the basis of matching experiments (Eek, Meister 1994); phoneme targets are marked by a dot (•).**

As we mentioned above, the words separated from sentences were identified according to the read text. From among unstressed-syllable isolated vowels only /i/, /o/ and /a/ in Q1 feet were identified 100% correctly, while /u/ was perceived 82% as /u/ (9% as /o/ and 9% as /õ/). The unstressed-syllable /e/ of a Q1 foot is changing into /ä/. It has already happened audibly in spontaneous speech of the younger generation. /e/ in the word context pronounced by older generation is still perceived as /e/ but /e/'s, separated from the same words, are perceived as /e/ only 33% of the cases (57% as /ä/). Thus the change *e* > *ä* is latently existing also in an audibly correct speech of the older generation. The unstressed-syllable /e/ of a Q1 foot is perceived 10% as /ö/; these stimuli are pronounced by an actively English using speaker MK (an influence of the strong centralisation of English?).

In a Q2 foot, the unstressed /i/ was perceived 100% correctly; /a/ — 88% (12% as /ä/); /e/ — 28% (72% as /ä/); /o/ — 17% (50% as /õ/ and 33% as /a/); /u/ — 13% (62% as /õ/ and 25% as /o/).

/e/, /u/, /o/ and /a/, separated from the unstressed syllable of a Q3 foot, were not at all perceived as the corresponding vowels of the read text; only /i/ was considered /i/ 51% of the cases (49% as /e/). /i/ is the most resistant to centralisation. The unstressed /e/ of a Q3 foot was perceived 75% as /ö/ and 25% as /ä/; /u/ — 88% as /õ/ and 12% as /o/; /o/ — 100% as /a/; /a/ — 58% as /ä/ and 42% as /ö/. While /e/ was in most cases perceived as /ä/ in a Q1 and Q2 foot, the /ö/ judgements prevailed in a Q3 foot, which also indicates greater centralisation of Q3.

The present preliminary data do not enable us to state that the greatest reduction of unstressed vowels of a Q3 foot is caused only by vowel shortness because small differences in average durations of unstressed vowels in Q3 and Q2 feet (Table 2) do not confirm it. The greatest reduction is rather due to the specific pattern of the whole Q3 foot (cf. peaked and unpeaked stressed syllables in Q3 and Q2 feet).

We are obviously not at variance with the presented data if we speak about the so-called Finnic-type of reduction in the case of unstressed vowels in the Q1 foot and about the so-called Germanic-type of reduction in the Q3 foot (especially among back vowels).

REFERENCES

Bladon, R. A. W., Fant, G. 1978, A Two-Formant Model and the Cardinal Vowels. — Royal Institute of Technology. STL-QPSR 1, Stockholm, 1—8.

Eek, A., Meister, E. 1994, Acoustics and Perception of Estonian Vowel Types. — University of Stockholm, Institute of Linguistics. Experiments in Speech Processes. PERILUS 18, 55—90.

—— 1997, Simple Perception Experiments on Estonian Word Prosody: Foot Structure vs. Segmental Quantity. — Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody, Tallinn, Estonia, October 29—30, 1996, Tallinn, 71—99.

—— 1998, Estonian Speech in the Babel Multilanguage Database: Phonetic-Phonological Problems Revealed in the Text Corpus. — Proceedings of the Workshop on Speech Database Development for Central and Eastern European Languages. The First International Conference on Language Resources and Evaluation, Granada.

Iivonen, A. 1987, The Critical Band in the Explanation of the Number of Possible Vowels and Psychoacoustical Vowel Distances, Helsinki (Helsingin yliopiston fonetiikan laitoksen monisteita 12).

—— 1988, Vokaaliposition määräytyminen formanttikartassa. — Fonetiikan päivät, Espoo 1988. Papers from the 15th Meeting of Finnish Phoneticians, Otaniemi (Helsinki University of Technology, Faculty of Electrical Engineering, Acoustics Laboratory, Publication 31), 39—63.

L i i v, G. 1962, Acoustical Features of Estonian Vowels Pronounced in Isolation and in Three Phonological Degrees of Length. — ETATÜS 11 1, 63—97.

L o c a l, J., O g d e n, R. 1998, Nordic Prosodies. Representation and Phonetic Interpretation. — Nordic Prosody. Proceedings of the VIIth Conference. Joensuu, 1996, Frankfurt am Main, 9—24.

S c h w a r t z, J.-L, B e a u t e m p s, D., A b r y, C., E s c u d i e r, P. 1993, Inter-Individual and Cross-Linguistic Strategies for the Production of the [i] vs. [y] Contrast. — Journal of Phonetics 21, 411—425.

W i i k, K. 1965, Finnish and English Vowels. A Comparison with Special Reference to the Learning Problems Met by Native Speakers of Finnish Learning English. — Annales Universitatis Turkuensis B 94, Turku.