Production and perception of Estonian vowels by native and non-native speakers

Lya Meister, Einar Meister

Laboratory of Phonetics and Speech Technology
Institute of Cybernetics at Tallinn University of Technology, Estonia
lya@phon.ioc.ee, einar@ioc.ee

Abstract

The aim of the paper is to study the production of Estonian vowel categories by L2 speakers of Estonian with a Russian-language background and to compare the results with perception data from the same subjects. Ten native Estonian subjects and ten L2 speakers participated in both the reading of an Estonian text corpus and the perception experiment.

It was found that mostly the production and perception results show similar patterns and thus lend support to the common standpoint that L2 perception predicts the accuracy of L2 production. However, evidence was found that despite the correct perceptual identification of L2 vowels, in L2 production the native categorical vowel representation outweighs the newer L2 category pattern.

Index Terms: L2 speech production, categorical perception, category boundary, Estonian, Russian

1. Introduction

When acquiring the vowel system of a second language, adult L2 learners have to overcome their phonological deafness towards the phonetic contrasts in L2 that negatively affects the perception and production of L2 vowels [1]. Depending on the differences between the vowel systems of L1 and the target language, foreign language learners will have several difficulties as predicted by the Speech Learning Model (SLM) [2] and the Perceptual Assimilation Model (PAM-L2) [3], as reported in a number of papers. Both SLM and PAM-L2 state that the ability to perceive and distinguish L2 sounds depends on the phonetic distance between similar categories in L1 and L2: (1) L2 sounds which are acoustically and perceptually close to those of L1 are hard to discriminate and will assimilate with their L1 counterparts; (2) for L2 sounds that are dissimilar to the closest L1 sounds, a new category will be created. In [4] different perceptual assimilation patterns (single-, two- and multiple-category assimilation) have been discussed which may occur in the case of different L2&L1 combinations. In the case of Russian as L1 and Estonian as the target language, the L2 speakers have to accommodate their native six-vowel system to the nine-vowel system of Estonian. Five Estonian vowels (/i/, /e/, /u/, /o/ and /a/) have close counterparts in Russian; of the remaining four vowels /ü/, /ö/ and /ä/ are the new vowel categories for the L2 speakers. An ambiguous case is the Russian vowel /ɨ/ which can be considered a counterpart to the Estonian /õ/, however, it is a rather deceptive match since the quality of these two vowels is very different. In the acquisition of the Estonian vowel system by Russian subjects two processes are likely to take place: first, a one-to-one assimilation of those Estonian vowels which have phonemic or allophonic counterparts in Russian. Secondly, a three-to-one assimilation which ideally should result in a splitting of the acoustic space occupied by one native category into three subspaces and ending up with the origination of three new L2 categories.

SLM [2] hypothesizes that "the production of a sound eventually corresponds to the properties represented in its phonetic category representation". In line with this is the generally accepted view that L2 perception development is a prerequisite to L2 production. However, several studies report contradicting evidence (for an overview see [5]).

In our paper we study the production of Estonian vowels by native and non-native speakers with a Russian-language background and compare the production data with the perception results of the Estonian vowel categories reported partly in [6].

2. Vowel systems in Estonian and Russian

The Estonian vowel inventory includes nine vowels /i ü u e ö õ o ä a/ whereas Russian has six /i ɨ u ɛ o a/; the binary articulatory features of vowels in both languages are given in Table 1.

Table 1. *Articulatory features of Estonian and Russian vowels.*

	Front		Back		
	Unrounded	Rounded	Unrounded	Rounded	
Estonian	/i/ [i]	/ü/ [y]	[ɯ]	/u/ [u]	High
	/e/ [e]	/ö/ [ø]	/õ/[1] [ɘ] [ɤ]	/o/ [o]	Mid
	/ä/ [æ]		/a/ [ɑ]		Low
Russian	/и/ [i]		/ы/[2] [ɨ] [ɯ]	/y/ [u]	High
	/э/[e] [ɛ]			/o/ [ɔ]	Mid
			/a/[3] [æ] [ɑ]		Low

[1] The Estonian back vowel /õ/ can be realized as a mid back vowel [ɤ], a high back vowel [ɯ] or a mid central vowel [ɘ].
[2] The Russian high vowel /ы/ is typically pronounced as a central vowel [ɨ] or in the context of velar stops and post-alveolar fricatives as the back vowel [ɯ] [7].
[3] The Russian low vowel /a/ exhibits large contextual variability – in between palatalized consonants it is realized as [æ], before a palatalized consonant as [a], or as [ɑ] when it is followed by [ɨ] [7].

The comparison of the vowel systems of Estonian and Russian shows that there are counterparts in both languages – /i u e o a/, and then the vowels that are missing in Russian – /ä, ü, ö/. A special case is the vowel /õ/ – since the quality of the Estonian /õ/ and Russian /ɨ/ is very different, they cannot be considered counterparts. The acquisition of Estonian vowel contrasts which are new for L2 subjects is problematic and therefore may have consequences in L2 production. As an illustration, Figure 1 represents the locations of Estonian [8] and Russian [9] male vowels in the acoustic space of F1 & F2 measured from hyperarticulated isolated pronunciation.

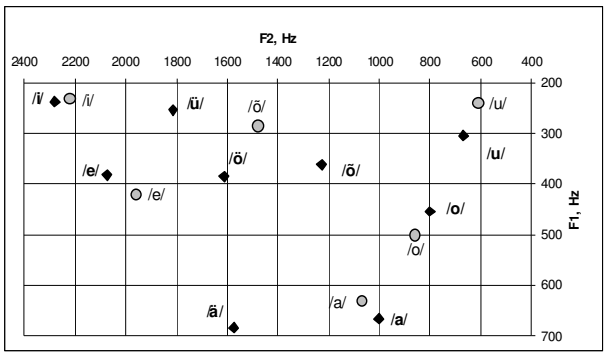

Figure 1. *Estonian (◆) and Russian (●) hyperarticulated vowels in F1 & F2 acoustic space.*

3. Production of Estonian vowels

3.1. Speech material

A subset of the Estonian Foreign Accent Corpus was used in the acoustic analysis. The corpus has been collected in a sound-proof room using high quality recording facilities (ATM33a condenser mic, sampling frequency 44.1 kHz, 16 bit). Currently, approximately 140 L2 subjects of Estonian with different language backgrounds (about 20) have been recorded while reading an Estonian text corpus. As a reference, 20 native Estonian subjects have been recorded. The text corpus includes 140 sentences that include the main phonological oppositions of Estonian, two short passages, and four prompts to elicit spontaneous speech (self-introduction and three pictures). The subset used in the current study includes 27 read sentences involving triplets of segmentally identical disyllabic words in the quantities Q1, Q2 and Q3, representing the structures CVCV, CVVCV and CVV:CV. The disyllabic words were selected to ensure that all Estonian vowels occur in primary-stressed syllables of all three CVCV-structures, e.g., Q1 *koli* /koli/ 'scrap', nom.sg.; Q2 *kooli* /kooli/ 'school', sg.gen.; Q3 *kooli* /koo:li/ 'school', sg.part.; etc. Thus, three instances of all Estonian vowels in similar phonetic/prosodic context are allocated to acoustic analysis.

3.2. Subjects

Ten (5 male, 5 female) native and ten (5 male, 5 female) non-native speakers of Estonian (with Russian as their L1) were involved in the study. Both groups participated in the perception experiments and were later recorded for the Estonian Foreign Accent Corpus. L1 subjects (age 21-54, median 26.5) came from monolingual Estonian-speaking families living in the capital area. All subjects have or are currently acquiring a university degree, and they represent standard Estonian pronunciation.

The L2 subjects (age 21-33, median 24.5) were born in monolingual Russian speaking families living in the north-east of Estonia or in the capital area; they were educated in Russian at basic and high schools levels and have or are currently acquiring a university degree in Estonia. Most of the L2 subjects started to learn Estonian in school at the age of 6-13, one subject at the age of 16 and one at the age of 20. All L2 subjects use Estonian almost every day at university or at their place of work, but at home they communicate in Russian (except for one subject). In the self-assessment of their Estonian proficiency (on a scale with options "basic", "average", "good", and "very good") two subjects rated themselves as "very good", five subjects as "good", and three subjects as "average". L1 subjects rated the degree of accentedness of L2 subjects in a continuous clip of speech (ca. 20 sec from each subject) using a 6-point scale (1 (no accent) to 6 (very strong accent)). For eight L2 subjects the ratings ranged from 2 to 4 (very weak to moderate), for two subjects the ratings 5 to 6 (strong to very strong) were given. None of the subjects reported any language impairment.

3.3. Formant frequencies of L1 and L2 vowels

Formant frequencies F1 and F2, and the duration of vowels were measured by using an adapted Praat [10] script. To facilitate automatic processing the signals were segmented beforehand on the word and phone levels. The formant frequencies were measured in the middle of the vowel, when necessary the formant values were verified by manual measurements in the spectrogram. Individual formant values in three CVCV structures were pooled together, separately for male and female speakers in both subject groups; the results are presented in Tables 2 and 3 and in Figures 2 and 3, for male and female subjects, correspondingly. Pooling formant data in vowels over different contexts is justified, as in Estonian the quality differences of the stressed vowels in three quantity degrees do not exceed 1 Bark in read speech [11]. However, the quality of L2 vowels in different quantity degrees may vary differently from L1 vowels; this certainly deserves further analysis.

Although the formant values of male and female vowels in the two groups reveal some natural subject-specific variability, both Figure 2 and 3 saliently point out the differences between the two groups in the production of vowels /ü/, /ö/ and /õ/, i.e., in the vowels representing the new categories for L2 subjects.

Table 2. *Formant frequencies of Estonian vowels in male L1 and L2 subjects ('***' p < 0.001, '**' p < 0.01, '*' p < 0.05).*

	L1				L2				L1-L2
Vowel	F1	F2	σ_{F1}	σ_{F2}	F1	F2	σ_{F1}	σ_{F2}	p
i	297	2113	30.6	139.7	304	2108	35.7	148.8	-
ü	304	1654	24.6	142.1	321	1804	28.4	156.3	$*_{F2}$
u	339	693	32.6	84.5	316	748	24.8	102.4	$*_{F1}$
e	443	1815	40.3	133.8	466	1791	42.2	128.5	-
ö	416	1533	20.1	77.1	442	1570	34.0	112.9	$*_{F1}$
õ	424	1289	38.8	78.0	375	1644	29.1	125.2	$***_{F1F2}$
o	444	818	39.1	98.3	417	805	31.0	140.8	$*_{F1}$
ä	595	1490	58.0	90.1	571	1614	55.1	126.0	$**_{F2}$
a	588	1109	55.0	78.4	579	1171	44.2	91.8	-

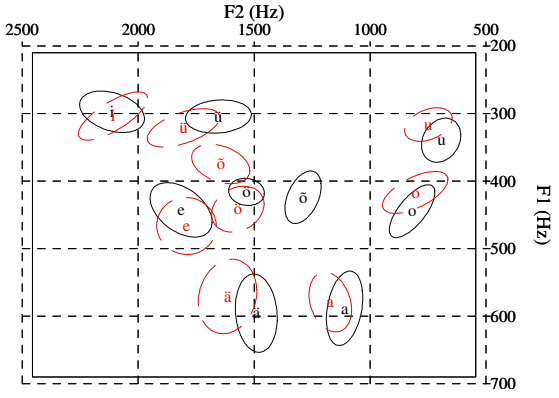

Figure 2. *Estonian vowels of male speakers in L1 (black solid line) and L2 (red dashed line).*

Table 3. *Formant frequencies of Estonian vowels in female L1 and L2 subjects.*

	L1				L2				L1-L2
Vowel	F1	F2	σ_{F1}	σ_{F2}	F1	F2	σ_{F1}	σ_{F2}	p
i	376	2653	42.2	157.3	336	2573	47.9	226.9	$*_{F1}$
ü	370	2093	36.1	132.9	351	1849	52.2	204.4	$***_{F2}$
u	381	779	43.3	86.5	370	762	44.6	153.9	-
e	555	2280	53.8	242.1	524	2265	65.6	236.2	-
ö	480	1834	35.0	98.6	512	1616	56.0	171.6	$*_{F1}$ $***_{F2}$
õ	491	1524	44.2	119.7	438	2068	57.9	193.4	$**_{F1}$ $***_{F2}$
o	501	969	43.1	124.3	499	860	45.2	103.8	$*_{F2}$
ä	812	1781	64.2	124.4	744	1748	58.2	151.9	$**_{F1}$
a	768	1291	63.4	129.3	757	1382	81.1	182.2	-

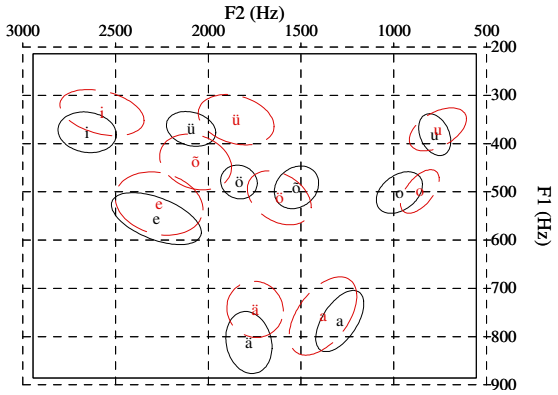

Figure 3. *Estonian vowels of female speakers in L1 (black solid line) and L2 (red dashed line).*

In order to make the formant values of male and female subjects comparable, the vowel normalization method proposed in [12] was applied. The normalized formant values of vowels in the L1 and L2 groups (male and female data pooled together) are presented in Table 4 and plotted on the F1&F2 normalized acoustic space of Figure 4. After normalization the general pattern of vowel location (Figure 4) remains pretty much the same as seen before in Figures 2 and 3, but it provides more robust data for bringing out the differences in L1 and L2 vowel production.

Table 4. *Normalized formant frequencies of Estonian vowels in the L1 and L2 groups.*

	L1				L2				L1-L2
Vowel	F1N	F2N	σF1	σF2	F1N	F2N	σF1	σF2	p
i	-1.09	1.61	0.23	0.15	-1.12	1.46	0.27	0.24	$**_{F2}$
ü	-1.08	0.64	0.20	0.27	-1.0	0.46	0.22	0.35	$*_{F2}$
u	-0.87	-1.55	0.22	0.16	-0.93	-1.57	0.30	0.26	-
e	0.17	0.96	0.28	0.28	0.25	0.87	0.29	0.29	-
ö	-0.20	0.29	0.23	0.19	0.11	0.02	0.18	0.26	$***_{F1F2}$
õ	-0.12	-0.25	0.27	0.15	-0.43	0.54	0.27	0.30	$***_{F1F2}$
o	0.004	-1.25	0.30	0.20	-0.02	-1.41	0.27	0.19	$**_{F2}$
ä	1.68	0.19	0.30	0.16	1.53	0.19	0.22	0.22	$*_{F1}$
a	1.50	-0.65	0.27	0.17	1.61	-0.57	0.21	0.25	-

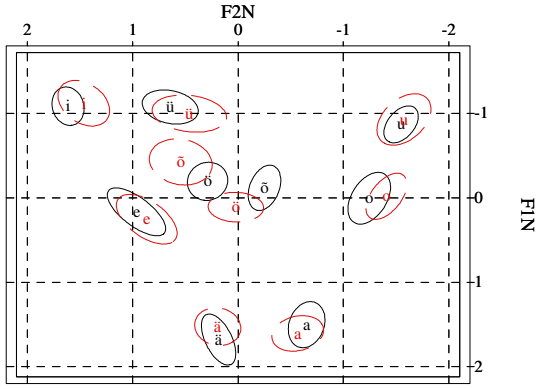

Figure 4. *Estonian vowels of L1 (black solid line) and L2 (red dashed line) groups in normalized F1&F2 space.*

The main findings are:

(1) differences between the two subject groups in the case of vowels /u/, /e/ and /a/ are minor and statistically insignificant;

(2) vowels /i/, /ü/ and /o/ in the L2 group deviate from those of L1 by being shifted backwards along the front-back dimension (in case of /i/ and /o/ $p < 0.01$; for /ü/ $p < 0.05$);

(3) vowel /ä/ in the L2 group is located higher than in the L1 group ($p < 0.05$);

(4) vowels /ö/ and /õ/ in the L2 group deviate from L1 vowels very significantly ($p < 0.001$) in both open-close and front-back dimensions – L2 /ö/ is more back and more open, and L2 /õ/ is more frontal and more closed; they have even switched their positions in the front-back direction.

4. Perception of Estonian vowels by L2 subjects

As the perception data has been previously reported in [6], here we give only a short overview of the experiment design and present the results in a slightly modified form.

4.1. Stimuli set and procedure

The stimulus set comprised 14 synthesized vowel continua between 9 Estonian vowel prototypes based on the formant values of native male vowels. Each continuum involved 16-18 synthetic vowels spaced equally between the prototypes. All stimuli were synthesized with a constant F0 of 100 Hz and with a duration of 160 ms. In the test, listeners had to decide on vowel category in a binary identification task by clicking on one of the two boxes labeled with the vowels of a vowel pair. In total 729 stimuli (16-18 stimuli x 3 repetitions x 14 vowel continua) were presented to each subject.

The categorization curves were obtained by interpolating the response data with a *probit* function and the locations of vowel category boundaries (expressed in stimulus number) were calculated as the 50% cross-over point of the fitted curve; the width of the category boundary was calculated as the difference between stimulus numbers corresponding to the score values of 0.8 and 0.2 from the curve.

4.2. Perception results

The locations and widths of vowel category boundaries in 14 vowel continua are reported in Table 5. As expected, L1 and L2 subjects exhibited the largest differences in the perception of vowel continua involving combinations of the Estonian vowels /õ/, /ö/, /ü/ and /ä/. The wider width and differences in the location of the category boundaries in /õ/-/ö/, /õ/-/ü/ and

/ö/-/ü/ suggest that for L2 subjects the perceptual contrasts between these vowels are difficult to acquire since they involve the splitting of the single native category /ɨ/ into three new categories. The fuzzy L2 category boundaries within the /õ/-/ö/-/ü/ subspace are the likely explanation for the deviations in the L2 production of these vowels. Likewise, the vague L2 pattern of /ä/ manifests itself with the much larger widths of /a/-/ä/ and /e/-/ä/ boundaries when compared to those of L1.

Table 5. *Median category boundaries of L1 and L2 groups.*

Vowel pair	L1 median boundary		L2 median boundary		L1-L2 difference[1]	
	Location	Width	Location	Width	p_L	p_W
/i/-/ü/	7.5	0.98	8.5	1.24	*	
/i/-/e/	8.2	1.15	8.1	1.77		
/e/-/ö/	7.5	1.57	7.6	1.61		
/ö/-/ä/	9.8	1.15	10.9	1.07	.	
/õ/-/u/	8.8	1.15	9.3	2.07		*
/u/-/o/	9.3	1.97	10.1	1.77		
/õ/-/o/	10.2	1.36	9.8	1.74		
/a/-/o/	8.5	0.98	8.0	0.98		
/a/-/ä/	10.5	1.07	10.8	2.05	.	**
/õ/-/a/	9.8	0.98	9.5	1.19		
/e/-/ä/	10.6	1.15	11.6	1.71		*
/ö/-/ü/	9.1	1.98	8.0	3.52	.	**
/õ/-/ü/	8.5	1.48	8.0	2.61		**
/õ/-/ö/	8.7	1.15	7.2	2.91	*	**

[1] p_L – significance of boundary location, p_W – significance of boundary width.

5. Discussion

We hold the position that L2 deviations in the location and width of L1 category boundaries reveal how well L2 subjects have acquired the L2 vowel categories, and this, in turn, results in deviating production of L2 vowels. In fact, the perception experiment was not about discovering the best perceptual vowel prototypes that could be directly evidenced in speech production; instead it explored the L1 vs. L2 differences in vowel category boundaries. However, we think that the results of the two experiments can be successfully compared and that they both point out the differences in L1 and L2 categories. In general, our production data is in line with the results of the perception test and thus tends to support the SLM hypothesis that the production of a sound corresponds to the properties of its phonetic category representation. The Estonian vowels /i/, /e/, /u/ and /o/ as well as /a/ and /ä/ showed a quite good one-to-one match with their Russian counterparts on the phonemic or allophonic levels in both production and perception; the deviating L2 patterns in the new vowel categories /ü/, /ö/ and /õ/ occurred likewise in production and perception. Yet, there is an interesting difference in the production and perception of vowels /ö/ and /õ/ – in the perception test all L2 subjects correctly distinguished the stimuli at the ends of the /ö/-/õ/ continuum representing the prototypes of the front vowel /ö/ and the back vowel /õ/ and demonstrated a systematic effect of categorical perception along the continuum. However, as the acoustic analysis shows, once in production these vowels were reversed in the front-back dimension. The location of Estonian /õ/ in L2 subjects corresponds approximately to the location of their native vowel /ɨ/ in the front-back dimension (cf. Figures 1 and 4) and in production the native vowel categorical representation outplays the new (more vaguely established) L2 pattern.

6. Conclusions

The current study aimed at exploring the differences in production and perception of Estonian vowel categories by native and non-native subjects with Russian as their L1. The production data represented the formant frequencies measured in all Estonian vowels in stressed syllables of three CVCV-structures; the perception data was obtained from the vowel category identification experiment where the location and width of category boundaries in 14 vowel continua were found. Mainly, the production and perception data are in line with each other and tend to support the SLM hypothesis of correspondence between the production of a sound and its abstract categorical representation. However, evidence of phonological asymmetry was found. Despite the correct perceptual identification of the front vowel /ö/ and the back vowel /õ/, in L2 production these vowels were reversed in the front-back dimension.

7. Acknowledgements

This research was partially supported by the target-financed theme No. 0322709s06 of the Estonian Ministry of Education and Research and by the National Program for Estonian Language Technology. All subjects who participated in the study are gratefully acknowledged.

8. References

[1] Dupoux, E., Peperkamp, S., "Fossil markers of language development: phonological 'deafnesses' in adult speech processing", in B. Laks, J. & Durand [Eds], Phonetics, Phonology, and Cognition. Oxford University Press, 168-190, 2002.

[2] Flege, J.E., "Second Language Speech Learning: Theory, Findings, and Problems", in W. Strange [Ed], Speech Perception and Linguistic Experience: Issues in Gross-Language Research, Timonium: York Press, 233-275, 1995.

[3] Best, C.T., Tyler, M.D., "Nonnative and second language speech perception: Commonalities and complementarities", in M.J. Munro & O.-S. Bohn [Eds], Second language speech learning: The role of language experience in speech perception and production, Amsterdam: John Benjamins, 13-34, 2007.

[4] Escudero, P., Boersma, P., "The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish", in B. Skarabela, S. Fish, A.H.-J. Do [Eds], Proceedings of the 26th Annual Boston University Conference on Language Development, pp. 208-219, 2002.

[5] Escudero, P., "Linguistic Perception and Second Language Acquisition. Explaining the attainment of optimal phonological categorization", Doctoral dissertation, Utrecht University, LOT Dissertation Series 113, 2005.

[6] Meister, L., Meister, E., "Perception of Estonian vowel categories by native and non-native speakers", in Proceedings of INTERSPEECH'2010, 1870-1873.

[7] Бондарко Л. В.; Вербицкая Л. А.; Гордина М. В., "Основы общей фонетики", Санкт-Петербург, 2000.

[8] Liiv, G., Remmel, M., "On acoustic distinctions in the Estonian vowel system", Soviet Fenno-Ugric Studies, 1, 7–23, 1970.

[9] Fant, G., "Description and Analysis of Contemporary Standard Russian. Acoustic theory and speech production with calculations based on X-ray studies of Russian Articulations", The Hague: Mouton, 1970.

[10] Boersma, P., Weenink, D., "Praat: doing phonetics by computer", (Version 5.2) [Computer program] Online: http://www.praat.org, accessed on 01.11.2010.

[11] Eek, A., Meister, E., "Quality of Standard Estonian vowels in stressed and unstressed syllables of the feet in three distinctive quantity degrees", Linguistica Uralica, 34/3, 226-233, 1998.

[12] Lobanov, B.M. 1971. "Classification of Russian vowels spoken by different listeners", Journal of the Acoustical Society of America 49:606-608.