Perception of Estonian vowel categories 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 report on the perception of Estonian vowel categories by L2 learners of Estonian whose L1 is Russian. The Estonian vowel system includes nine vowels whereas Russian has six. Five of the Estonian vowels have counterparts in Russian: /i/, /e/, /u/, /o/ and /a/, the new vowel categories for L2 speakers are /ü/, /ö/, /ä/, and partly /õ/.

For the perceptual experiments four-formant vowel stimuli were synthesized including nine Estonian prototype vowels and the intermediate steps between prototypes; the stimulus set covered 14 vowel category boundaries. The experiments involving native Estonian and non-native (Russian as L1) subjects showed that (1) Estonian vowels /i/, /e/, /u/ and /o/ assimilate well with their Russian counterparts; (2) Estonian /a/ and /ä/ assimilate with the allophones of Russian /a/; (3) Estonian /ü/, /ö/ and /õ/ assimilate partly with Russian /ɨ /; due to the close phonetic distance the L2 subjects' ability to discriminate these categories is poor.

Index Terms: categorical perception, vowel, category boundary, Estonian, Russian.

1. Introduction

Adult second language (L2) learners face difficulties in pronunciation and perception of L2 speech segments, as has been shown in numerous studies. It is mainly due to the "phonological deafness" towards phonetic contrasts of L2, which develops after the acquisition of the phonetic inventory of a child's first language (L1) during the first years of life [1]; it has been found that vowel categories are established during the first 6 months of life [2], [3] and consonant categories at the age of 10 to 12 month [4]. Phonological deafness remains rather stable during later life-span and negatively affects the perception of foreign phonetic contrasts as well as the acquisition of correct L2 pronunciation [5]. The role of perception in the acquisition of phonological categories in L2 is addressed in two models – Perceptual Assimilation Model (PAM) [6] and Speech Learning Model (SLM) [7]. According to them 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 they will assimilate with their L1 counterparts; (2) for L2 sounds that are dissimilar to the closest L1 sounds, a new category will be created. Based on Best's and Flege's models, Escudero and Boersma [8] have discussed different perceptual assimilation patterns (single-, two- and multiple-category assimilation) which may occur in the case of different L2&L1 combinations.

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 close counterparts in Russian (5 vowels), and second, a three-to-one assimilation which results in the origination of three new categories.

In our paper we study the perception of Estonian vowel categories by native and non-native (Russian) speakers. The perception experiments with synthetic stimuli are designed to discover the category boundaries of Estonian vowels in the F1&F2 space.

2. Estonian versus Russian vowel system

The Estonian vowel system includes nine vowels /i ü u e ö õ o ä a/ which are characterized by the articulatory features as shown in Table 1 [9].

Table 1. *Articulatory features of Estonian vowels.*

	Front		Back	
	Unrounded	Rounded	Unrounded	Rounded
High	/i/ [i]	/ü/ [y]	/õ/ [ɯ]	/u/ [u]
Mid	/e/ [e]	/ö/ [ø]	[ə] [ɤ]	/o/ [o]
Low	/ä/ [æ]		/a/ [ɑ]	

The back vowel /õ/ can be realized depending on the speaker as a mid back vowel [ɤ], a high back vowel [ɯ] or a mid central vowel [ə]. The quality differences of Estonian stressed vowels in different quantity degrees do not exceed 1 Bark in the F1-F2 vowel space, but vowels in unstressed syllables tend to be reduced [10], [11].

The Russian vowel inventory includes six vowels /i ɨ u ɛ o a/, an ambiguous status has the vowel /ɨ/ which is often classified as a positional allophone of /i/ occurring only in the position after a non-palatalized consonant. According to tongue height, Russian vowels are divided as high – /i ɨ u/, mid – /ɛ o/, and low – /a/; in the front-back dimension they are classified as front – /i e/, central – /a/, and back – /u o/. The classification of /ɨ/ is problematic – typically it is pronounced as a central vowel [ɨ] e.g., *мы* [tɨ] 'you', but in the context of velar stops and post-alveolar fricatives it is pronounced as the back vowel [ɯ] e.g., *шишка* [ʃɯʃkʌ] 'cone' [12]. However, articulatorily /ɨ/ is closer to /u/ rather than to /i/ and therefore it is classified as a front-shifted back vowel [13]. In addition, due to the large contextual variability of the central vowel /a/, it is realized as [æ], e.g., *сядь* [s'æt'] 'sit down' in-between of palatalized consonants, as [a], e.g., *дать* [dat'] 'to give' before a palatalized consonant, or as [ɑ] when it is followed by [ɬ], e.g., *палка* [pɑɬkʌ] 'stick' [12]. Since the articulatory features of /a/ are more similar to back vowels than to front ones, it is classified often as a back vowel [14].

10.21437/Interspeech.2010-540

The quality of Russian vowels in unstressed position is reduced when compared to those in stressed position, especially in the case of the vowel /a/ [15].

By comparing the vowels of Estonian and Russian, one can see 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 can not be considered as counterparts. Our earlier study [16] has revealed several deviations in the production of Estonian vowels by Russian subjects. An analysis of the formant frequencies of vowels in L1 and L2 speech has shown that in L2 speech:

(1) the vowel /ä/ is located a little bit higher, vowels /ö/ and /ü/ are more centralized and the vowel /õ/ has moved towards the front vowels when compared to those of L1,

(2) the quality of vowels /õ/ and /ü/ is most vague – they are located too close to each other and the large standard deviation of F2 shows great variability in production,

(3) the location of vowels /i, e, a, o, u/ corresponds well to the location of native Estonian vowels.

These results suggest that L2 subjects have difficulties in the acquisition of Estonian vowel contrasts which are new for Russian. The L2 perception data on Estonian vowel category discrimination may allow us to explain the deviations in L2 production.

3. Hypotheses

The hypotheses that we have proposed for the study are the following.

The perceived phonetic distance between Estonian and Russian counterparts /i/, /e/, /u/ and /o/ is small and therefore single-category one-to-one assimilation is very likely to occur:

Estonian		Russian
/i/	→	/i/
/e/	→	/ɛ/
/u/	→	/u/
/o/	→	/o/

The low-vowel area in Estonian is divided between two vowels: the front vowel /æ/ and back vowel /ɑ/ whereas there is only one low vowel /a/ in Russian. It suggests that a two-to-one assimilation should take place. However, as Russian /a/ has positional allophones [æ] and [ɑ], the phonetic contrast of the corresponding Estonian vowels is not new for Russian subjects and their discrimination is expected to be straightforward. According to SLM, the mapping of L1 and L2 vowels takes place on the allophonic level, not on the abstract phonemic level [7]. Thus, instead of a two-to-one assimilation the one-to-one mapping on the allophonic level is highly probable:

Estonian		Russian
/ɑ/	→	[ɑ]
/æ/	→	[æ]

The closest counterpart for Estonian vowels /y/, /ø/ and /ɤ/ is Russian /ɨ/, which suggests the three-to-one assimilation pattern:

Estonian		Russian
/y/	↘	
/ø/	→	/ɨ/
/ɤ/	↗	

The perceived phonetic distance between mapping vowels is rather small resulting in a poor discrimination of these vowels by Russian subjects and consequently, in unfavorable conditions for the creation of new perceptual categories.

In the case of the one-to-one assimilation of L1 and L2 vowels the categorization functions of L2 subjects are expected to be similar with those of L1 subjects; in the case of the many-to-one assimilation patterns certain differences in the categorization curves of L1 and L2 groups are expected, and should reveal how well the L2 subjects have acquired the new vowel contrasts.

4. Methods and material

4.1. Stimuli set and procedure

The stimuli set is comprised of 14 vowel continua between 9 Estonian vowel prototypes. Each continuum includes 16-18 synthetic vowels created by interpolating stepwise between prototypical values of the first four formants for each vowel pair and calculating evenly spaced steps through the F1/F2/F3/F4 continua (Figure 1).

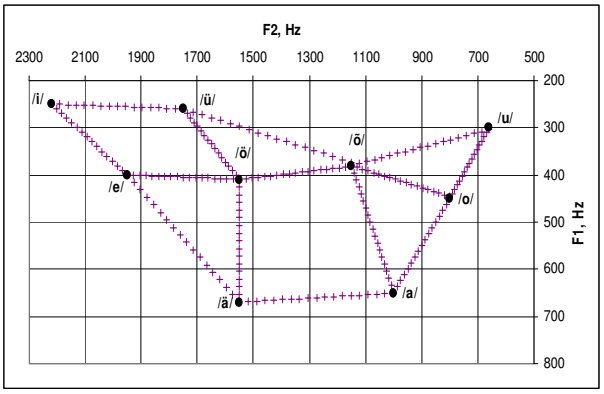

Figure 1. *Estonian prototype vowels (•) and stimuli steps (+) in F1&F2 space.*

The stimuli were synthesized with KlattWorks [17], an implementation of a Klatt-type formant synthesizer. All stimuli were synthesized with a constant F0 of 120 Hz and with a duration of 160 ms. The formant values of the prototype vowels (Table 2) are approximations of vowel formants measured in isolated vowels pronounced by six male native speakers of Estonian [18].

Table 2. *Formant frequencies (in Hz) of Estonian vowel prototypes.*

	/i/	/e/	/ä/	/ü/	/ö/	/õ/	/u/	/o/	/a/
F1	250	400	670	260	410	380	300	450	650
F2	2220	1950	1550	1750	1550	1150	660	800	1000
F3	3000	2580	2400	2200	2400	2160	2250	2460	2250
F4	3430	3350	3400	3220	3220	3220	3220	3300	3220

All testing was conducted in a sound-isolated room and stimuli were presented to subjects via high-quality headphones. The test was administered with Praat's [19] multiple forced-choice test facility; each vowel was repeated three times in random order with no replay option. 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 the subjects. The duration of the test turned out to be 20-25 minutes.

4.2. Subjects

Six (3 male, 3 female) native and ten (5 male, 5 female) non-native speakers of Estonian with Russian as L1, were involved in the study. L1 subjects (age 21-54, mean 28.5) 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 with no reported hearing problems.

The L2 subjects (age 21-33, mean 25.7) were born in monolingual Russian speaking families living in the north-east of Estonia or in the capital area; they were educated in the Russian-language at the 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 as they all have Estonian friends or colleagues, but at home they communicate in Russian (except for one subject). In the self-assessment of their proficiency of Estonian two subjects rated themselves as "very good", five subjects as "good", and three subjects as "average". According to perceptual assessment a moderate to slight foreign accent was noticed in their L2 speech. None of the subjects reported any language impairment.

4.3. Data analysis

The output of the perception experiment is a categorization function for each vowel continuum that plots the score of one of the two alternative responses across the continuum. The individual categorization curves were obtained by interpolating the response data of each subject with a *probit* function and the locations of individual vowel category boundaries (expressed in stimulus number) were calculated as the 50% cross-over point of the fitted curve. In addition, the slope of the categorization curve was calculated. The average categorization functions for L1 and L2 groups have been found by averaging the individual categorization functions in both groups resulting in a less steep slope of the average curve compared to the individual curves.

The location of the category boundary and the slopes of categorization curves were subjected to quantitative analysis using one-way ANOVAs with language background as an independent variable.

To evaluate the reliability of raters, the intraclass correlation measure (ICC) describing the consistency of responses within a group, and the pair-wise Pearson's correlation coefficient were calculated.

5. Results

The vowel category boundary locations and the slopes of the averaged categorization functions of 14 vowel continua (Table 3) show minor differences in category boundaries and slightly larger differences in slopes. The independent sample t-test of the location of the category boundaries between L1 and L2 groups showed no significant differences in the category boundaries of the averaged curves ($t = 0.4182$, $df = 21.695$, $p = 0.68$), also the one-way ANOVA revealed no group effect for the category boundary [$F(1, 26) = 0.1749$, $p = 0.68$].

However, the difference in slopes of averaged categorization curves in two groups turned out to be significant ($t = -3.5135$, $df = 25.99$, $p < 0.005$), the analysis of variance confirmed a significant group effect for slopes [$F(1, 26) = 12.345$, $p < 0.005$].

A more detailed analysis of subject-specific categorization functions in the two groups revealed that the difference in boundary location is significant only in the case of the **/ö/-/ä/** vowel pair ($t = -2.565$, $df = 13.93$, $p < 0.05$; $F(1,14) = 5.192$, $p < 0.05$); the differences in the slopes of the categorization curves is significant in the case of five vowel pairs: **/i/-/e/** ($t = -2.405$, $df = 9.65$, $p < 0.05$; $F(1,14) = 6.162$, $p < 0.05$), **/ö/-/ü/** ($t = -3.007$, $df = 5.462$, $p < 0.005$; $F(1,14) = 14.551$, $p < 0.005$), **/õ/-/ü/** ($t = -4.402$, $df = 5.972$, $p < 0.005$; $F(1,14) = 28.871$, $p < 0.001$), **/õ/-/ö/** ($t = -4.634$, $df = 5.921$, $p < 0.005$; $F(1,14) = 32.227$, $p < 0.001$), and **/õ/-/u/** ($t = -4.35$, $df = 6.729$, $p < 0.005$; $F(1,14) = 25.673$, $p < 0.001$).

Table 3. *Vowel category boundary locations (expressed in stimulus number) and slopes of the averaged categorization curves in 14 vowel pairs of L1 and L2 groups.*

	L1		L2	
Vowel pair	Category boundary	Slope	Category boundary	Slope
/i/-/ü/	8.59	-0.261	8.88	-0.283
/a/-/ä/	10.18	-0.198	9.43	-0.207
/u/-/o/	9.36	-0.175	10.07	-0.167
/e/-/ö/	8.04	-0.209	7.50	-0.193
/õ/-/o/	9.38	-0.227	9.65	-0.173
/õ/-/ü/ ***	9.35	-0.213	8.10	-0.141
/ö/-/ü/ **	8.68	-0.186	7.82	-0.111
/ö/-/ä/ *	9.90	-0.366	10.70	-0.290
/õ/-/a/	9.81	-0.240	9.77	-0.143
/a/-/o/	8.50	-0.338	8.33	-0.237
/i/-/e/ *	8.33	-0.269	8.65	-0.132
/õ/-/u/ ***	9.21	-0.302	9.01	-0.156
/õ/-/ö/ ***	8.64	-0.257	7.13	-0.087
/e/-/ä/	10.55	-0.333	11.24	-0.126

*'***' p<0.001 '**' p< 0.01 '*' p<0.05*

As expected, all L1 subjects showed consistent categorical perception in all vowel continua: the pair-wise Pearson's correlation test and the intraclass correlation measure showed high inter-subject correlation ($r = 0.81$ to 1.0, $p < 0.001$), ICC = 0.85 to 0.97).

The consistency of L2 responses was much poorer, ($r = 0.70$ to 1.0, $p < 0.01$), ICC = 0.68 to 0.96), however, the endpoints of continua were correctly identified in all vowel pairs by all L2 subjects.

6. Discussion

When interpreting the results we suppose that the location of category boundary and the slope of the categorization function in each vowel continuum are the most informative features exposing the differences in L1 and L2 perception and thus revealing how well L2 subjects have acquired the vowel categories of L1. In general, the results of the perception test tend to support our hypothesis. The Estonian vowels /i/, /e/, /u/ and /o/ show a quite good one-to-one match with their Russian counterparts and also the assimilation of Estonian vowels /a/ and /ä/ with the allophones of Russian /a/ ([ɑ] and [æ], respectively) seems to hold true since there were no large differences between the L1 and L2 groups in the category

boundary locations nor in the slopes of the categorization curves in most corresponding vowel continua. However, some differences turned out to be significant: (1) the slope of the categorization function in the /i/-/e/ continuum is less steep within the L2 group ($p < 0.05$) revealing a fuzzy category boundary between /i/ and /e/ (Figure 2, upper left); (2) the category boundary between /ö/ and /ä/ in the L2 group is shifted slightly lower in the close-open dimension ($p < 0.05$) (Figure 2, upper right); (3) the less steep categorization curve of the L2 group in the continuum /õ/-/u/ ($p < 0.001$) (Figure 2, left mid) can be interpreted as a fuzzy boundary between the vowel categories /õ/ and /u/, which is likely affected by the vaguely established category of vowel /õ/.

As expected, L1 and L2 subjects exhibited the largest differences ($p < 0.001$) in categorization in the case of vowel continua involving combinations of Estonian vowels /õ/, /ö/ and /ü/, i.e., in /õ/-/ö/, /õ/-/ü/ and /ö/-/ü/. The more gentle slopes of the categorization curves (Figure 2, mid right, bottom left, bottom right) suggest that for Russian subjects the perceptual contrasts between these vowels are difficult to acquire since it involves the splitting of the single native category /ɨ/ into three new categories.

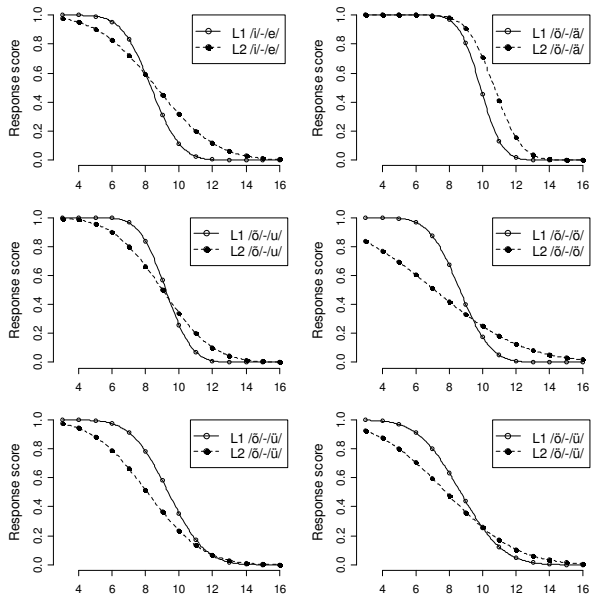

Figure 2. *Averaged vowel categorization functions for vowel continua showing significant differences between L1 and L2 subject groups: /i/-/e/, /ö/-/ä/, /õ/-/u/, /õ/-/ö/, /õ/-/ü/ and /ö/-/ü/.*

7. Conclusions

This study has explored the differences in perception of Estonian vowel categories by native and non-native subjects with Russian as their L1. Two characteristics – the location of category boundary and the slope of the categorization function – in 14 vowel continua have been calculated and exploited in the comparison of perception results in two subject groups. It was found that the category boundaries of Estonian vowels which have Russian counterparts on the phonemic or allophonic levels were perceived almost similarly by the L1 and L2 groups; in the case of new vowel categories (i.e., vowels /ü/, /ö/ and /õ/) the L2 subjects' perception deviated significantly from the perception of L1 demonstrating a much poorer discrimination ability of new vowel contrasts and resulting in much gentler slopes of the categorization functions.

8. 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 Programme for Estonian Language Technology. All subjects who participated in the perception tests are gratefully acknowledged.

9. References

[1] Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., Näätänen, R. "Development of language-specific phoneme representations in the infant brain", Nature Neuroscience, 1, 351-353, 1998.

[2] Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B. "Linguistic experience alters phonetic perception in infants by six months of age", Science, 255/5044, 606-608, 1992.

[3] Polka, L., Werker, J.F. "Developmental changes in perception of non-native vowel contrasts", Journal of Experimental Psychology: Human Perception and Performance, 20, 421-435, 1994.

[4] Werker, J.F., Tees, R.C. "Cross-language speech perception: evidence for perceptual reorganization during the first year of life", Infant Behavior and Development, 7, 49-63, 1984.

[5] 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.

[6] 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.

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

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

[9] Eek, A. "Eesti keele foneetika I", Tallinn: TTÜ Kirjastus, 2008.

[10] 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.

[11] Lippus, P. "Variation in Vowel Quality as a Feature of Estonian Quantity" 2010 (in print).

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

[13] Любимова Н.А. „Обучение русскому произношению", Москва: Русский язык,1977.

[14] Бондарко Л.В. „Звуковой строй современного русского языка", Москва: Просвещение, 1977.

[15] Бондарко Л.В. Фонетика современного русского языка", Санкт-Петербург, 1998.

[16] Meister, L., Meister, E. "Acoustic correlates of Russian accent in Estonian", in Kokkinakis,G., Fakotakis,N., Dermatas,E., Potapova,R. [Eds], SPECOM 2005: 10th International Conference Speech and Computer, University of Patras, 437-440, 2005.

[17] McMurray, B. "KlattWorks: A [somewhat] new systematic approach to formant-based speech synthesis for empirical research", in preparation.

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

[19] Boersma, P., Weenink, D. "Praat: doing phonetics by computer", Version 5.1.15. [Computer program] Online:http://www.praat.org, accessed on 31.08.2009.