

PRODUCTION OF ESTONIAN VOWELS BY JAPANESE SUBJECTS

Rena Nemoto, Einar Meister, Lya Meister

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

ABSTRACT

This study explores the production of Estonian vowel categories by second-language (L2) speakers of Estonian whose native language (L1) is Japanese. The Estonian vowel system includes nine vowels whereas Japanese has only five.

The results by six native Japanese and ten native Estonian subjects when reading Estonian target words in sentence show that Japanese subjects were successful in production of Estonian vowels which assimilate well to the corresponding Japanese vowels (/i/, /e/, /o/) and faced difficulties in discrimination of vowel contrasts involving single-category assimilation (Estonian /ü/, /ö/, /õ/, /u/ to Japanese /u/).

Keywords: Estonian, Japanese, L2, vocalic production.

1. INTRODUCTION

The acquisition of sounds of a second language is generally a difficult task for adult learners and typically results in a deviant production of L2 segments, as a great deal of research has shown. Theoretical models such as the Speech Learning Model (SLM) [9], the extension of the Perceptual Assimilation Model (PAM-L2) [1], and more recently, Second Language Linguistic Perception model (L2LP) [5] posit that the ability to perceive and distinguish L2 sounds depends on the phonetic distance between similar segments in L1 and L2. These models also make predictions of how L2 contrasts can be acquired in perception and production: (1) L2 sounds which are acoustically and perceptually close to those of L1 are difficult to discriminate and will assimilate with their L1 counterparts; (2) for those L2 sounds that are dissimilar to the closest L1 sounds, new categories will be created. In the case of different vowel inventories of L1 and L2 different perceptual assimilation patterns (single-, two- and multiple-category assimilation) may occur, as discussed in [6].

In this study we explore the acoustic characteristics of Estonian vowels produced by L2 speakers of Estonian whose L1 is Japanese. The Esto-

nian vowel system includes nine vowels whereas Japanese has only five; four L2 vowel categories are new for Japanese speakers. The combination of L1 and L2 vowel inventories can be rather challenging for Japanese subjects since several single-category assimilations (according to PAM) or perceptual equivalences (according to SLM) are likely to occur. Therefore, the formation of new vowel categories might be difficult. On the other hand, since the Japanese vowel space is rather sparse, there is enough room for new vowel categories [19].

Several studies have shown that L1 vowel inventory size effects the L2 perception, however, with contradictory results. Some studies suggest that if L1 vowel inventory is larger than the vowel inventory of the target language, then it may facilitate L2 vowel learning [11, 12]. Other studies have shown that subjects with smaller L1 vowel inventory than the target language outperform subjects with larger L1 vowel inventory [4, 7, 8]. There are a number of studies investigating the perception and production of native Japanese subjects learning vowel inventories larger than in Japanese, such as English [15, 20, 21] and French [13, 14].

This is the first study comparing Estonian and Japanese vowel systems and exploring the production of Estonian vowels by Japanese subjects. In the paper we try to answer the following research questions: (1) how Japanese speakers produce Estonian vowels in comparison to native Estonian speakers? and (2) how Japanese speakers produce Estonian vowels in comparison to their native vowels?

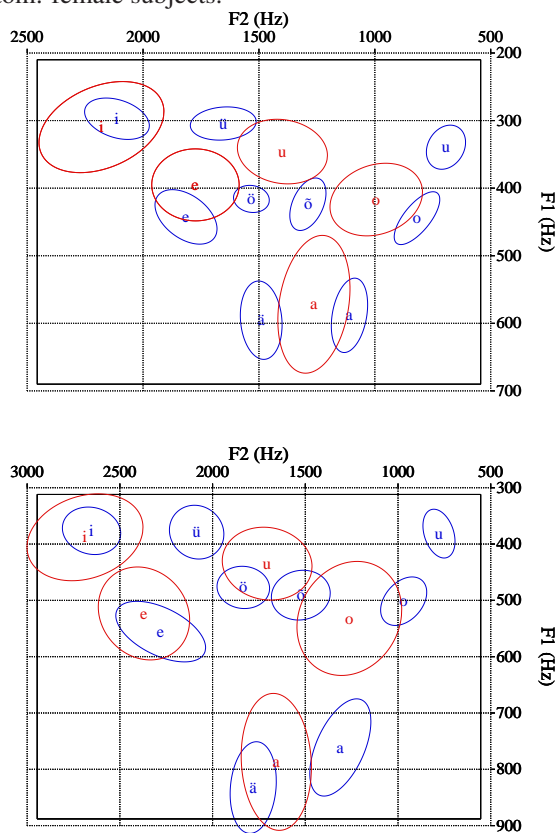
2. ESTONIAN VOWEL SYSTEM VERSUS JAPANESE VOWEL SYSTEM

The Estonian vowel system includes nine vowels /i, ü, e, ö, ä, a, õ, o, u/ transcribed here as [i, y, e, ø, æ, a, ʏ, o, u]. Estonian vowels are classified into three height levels: close /i, ü, u/, mid /e, ö, õ, o/ and open /ä, a/, and front /i, ü, e, ö, ä/ and back /a, õ, o, u/ vowels; the vowels /ü, u, ö, o/ are rounded. The unrounded back vowel /õ/ exhibits larger variations and can be realized as a mid back vowel [ʏ], a high back vowel [u] or a mid central vowel [ə].

Unlike Estonian, Japanese has a relatively sparse vowel system involving only the five vowels /i, e, a, o, u/ transcribed as [i, e, a, o, u]. The vowels vary in height: close /i, u/, mid /e, o/ and open /a/, and in front-back dimension: front /i, e/, central /a/, and back /o, u/, and there is only one rounded vowel: /o/.

Figure 1 compares the phonetic realizations of Estonian and Japanese vowels in read speech by native Estonian and Japanese speakers, respectively (see 3.1. for the description of subjects and corpus, and Table 1 for numeric values of formants).

Figure 1: Comparison of native Estonian (blue) and native Japanese (red) vowels (ellipses represent standard deviation). Top: male subjects; bottom: female subjects.



Based on the phonological descriptions and phonetic realizations of vowels in the two languages, and the predictions of SLM and PAM, we propose the following hypotheses:

1. The Estonian front vowel pairs /i, ü/ and /e, ö/ involve unrounded-rounded contrast. It might be relatively easy to learn the production of this new contrast which in turn might facilitate the formation of new perceptual patterns for /ü/ and /ö/, whereas for L2 /i/ and /e/ the native patterns will be used. However, it is more likely that the vowels /ü/ and /ö/

perceptually assimilate to Japanese /u/ and in this case the creation of new categories will be hindered.

2. The Japanese /a/ (central) is located between the Estonian /a/ (back) and /ä/ (front) and a single-category assimilation is likely to take place. The perceived phonetic distance between these L2 vowels is rather small resulting in a poor discrimination of these vowels and consequently, in unfavorable conditions for the formation of new categories.

3. Japanese /o/ is located rather close to the Estonian /o/, therefore the use of the Japanese native pattern should be accurate enough for L2 /o/.

4. Estonian /õ/ and /u/ are likely to assimilate to the Japanese /u/ (phonetically [u]). Japanese /u/ features less or no lip rounding and much higher F2 value (see Figure 1) than a typical Estonian /u/, and therefore it is perceptually close to Estonian /õ/ realized phonetically mainly as [ɤ], seldom as [u] or [ə]. We expect that the discrimination of these L2 vowels will be poor and the formation of new categories will be very difficult.

Both languages are quantity languages, i.e., they use the duration cue for manifesting phonological contrasts. Japanese has a binary quantity contrast resulting in qualitatively close short-long vowel pairs [10]. Estonian features a three-way quantity contrast referred to as short (Q1), long (Q2), and overlong (Q3) quantity degrees [16]. The quantity oppositions in vowels occur in the stressed syllable only; the quality of vowels in the three quantity degrees varies slightly – vowels in stressed syllables of Q1 feet are closer to the center and vowels of Q2 and Q3 feet are more peripheral [3, 17]. However, the main feature distinguishing Estonian quantity contrasts is duration rather than quality. Therefore, in this paper we will neglect the analysis of quality variations depending on the quantity degree.

3. METHOD

3.1. Speakers and corpus

The speech material of ten native Estonian speakers (five male, five female) and six native Japanese speakers (three male, three female) has been used. The L1 subjects (age 21-54, median 26.5) came from monolingual Estonian-speaking families living in the capital area. All subjects have acquired or were acquiring a university degree, and they represent standard Estonian pronunciation. The L2 subjects (age 32-45, median 36.5) were born in monolingual Japanese speaking families living in Japan and were educated in different high schools or universities in Japan. They started to learn Estonian after arrival to Estonia at the age of 19-43 (median

27.5). They have stayed in Estonian for several years and communicate in Estonian daily; in self-assessment they rated their proficiency in Estonian from “intermediate” (2 subjects) to “advanced” (4 subjects).

All subjects participated in the recordings of the Estonian Foreign Accent Corpus [18]. The corpus has been collected in a sound-proof room using high quality recording facilities. Subjects with different language backgrounds were recorded while reading an Estonian text corpus including 140 sentences involving the main phonological oppositions of Estonian, two short passages, and four prompts to elicit spontaneous speech. 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 structures, e.g., Q1 *koli /koli/* ‘scrap’, nom.sg.; Q2 *kooli /kooli/* ‘school’, sg.gen.; Q3 *kooli /koo:li/* ‘school’, sg.part.; etc. Hence, three formally identical repetitions of each Estonian vowel in each of the three structures ($9 \times 3 = 27$) were subjected to acoustic analysis.

In addition, a small Japanese corpus was recorded by the same Japanese subjects when reading the story “The North Wind and the Sun” in Japanese, and 20 phonetically rich Japanese sentences. This corpus was necessary as a reference for native Japanese vowels (see Figure 1) and for comparing native Japanese vowels with L2 Estonian vowels. All recordings were segmented in Praat [2] on word and segment levels.

3.2. Formant analysis

The F1 and F2 of the vowels were measured (with a Praat script) in the middle of vowels. For male and female subjects different settings for formant tracking were used, when necessary the formant values were verified by manual measurements in the spectrogram. For each speaker, mean formant values for each stressed vowel in each of the three word structures were pooled across the three formally identical repetitions for Estonian. As for Japanese, only vowels (not diphthongs) were measured.

4. RESULTS AND DISCUSSION

The mean F1 and F2 frequencies for the speaker groups are shown in Table 1 and the vowel charts in Figures 1, 2, and 3. An analysis of variance (ANOVA) was applied to compare the Estonian

vowels produced by the two groups.

Both male and female L2 subjects tended to produce L2 /i/ more front and close than L1 /i/; however, the L1-L2 difference was significant ($p < 0.05$) for the female speakers only. L2 /e/ and /o/ were produced very close to those of L1. For these three vowels L2 subjects successfully applied the corresponding native patterns. Concerning the L2 open vowels /a/ and /ä/, L2 males were not able to produce two different patterns, their L2 /a/ is close to Estonian /a/, and L2 /ä/ is too back (F2: $p < 0.001$) overlapping /a/. L2 female subjects discriminated two different patterns in their production (F1: $p < 0.05$, F2: $p < 0.01$), however, their /a/ is more front (F2: $p < 0.01$) and /ä/ more open (F1: $p < 0.05$) than the Estonian vowels.

Table 1: Mean formant frequencies of male (top) and female (bottom) subjects in native Estonian (L1-EE), L2 Estonian (L2-EE), and native Japanese (L1-JP) vowels. Significant differences between L1-EE and L2-EE vowel formants are marked as ‘****’ $p < 0.001$, ‘***’ $p < 0.01$, ‘*’ $p < 0.05$.

Male	L1-EE		L2-EE		L1-JP	
	F1	F2	F1	F2	F1	F2
i	297	2113	284	2240	310	2179
ü	304	1654	314	1480*	-	-
e	443	1815	451	1901	395	1774
ö	416	1533	363**	1363*	-	-
ä	595	1490	650	1237***	-	-
a	588	1109	582	1133	572	1263
õ	424	1289	402	1130***	-	-
o	444	818	411	872	417	994
u	339	693	359	1169***	346	1399

Female	L1-EE		L2-EE		L1-JP	
	F1	F2	F1	F2	F1	F2
i	376	2653	329*	2928*	388	2688
ü	370	2093	404	1826*	-	-
e	555	2280	556	2294	523	2370
ö	480	1834	419**	1768	-	-
ä	812	1781	902*	1797	-	-
a	768	1291	800	1520**	787	1656
õ	491	1524	417**	1476	-	-
o	501	969	536	1039	532	1263
u	381	779	413	1386***	435	1706

As expected, the most difficult thing for the L2 subjects was the contrasts among Estonian /ü/, /ö/, /õ/, and /u/ categories. L2 /ü/ was produced more back (F2: $p < 0.05$), L2 /ö/ more close (F1: $p < 0.01$), and L2 /u/ more front (F2: $p < 0.001$) than the corresponding L1 vowels; in addition, for L2 males /õ/ was more back (F2: $p < 0.05$). L2 /õ/ was produced differently by males and females: in males it is more back (F2: $p < 0.001$) and in females more close (F1: $p < 0.01$) when compared to L1 production. These L2 vowels manifested several overlaps, however, as

Figure 2: Vowels of male speakers. Top: L2 (red) and L1 Estonian (blue) vowels. Bottom: Estonian L2 (red) and L1 Japanese (black) vowels (ellipses represent standard deviation).

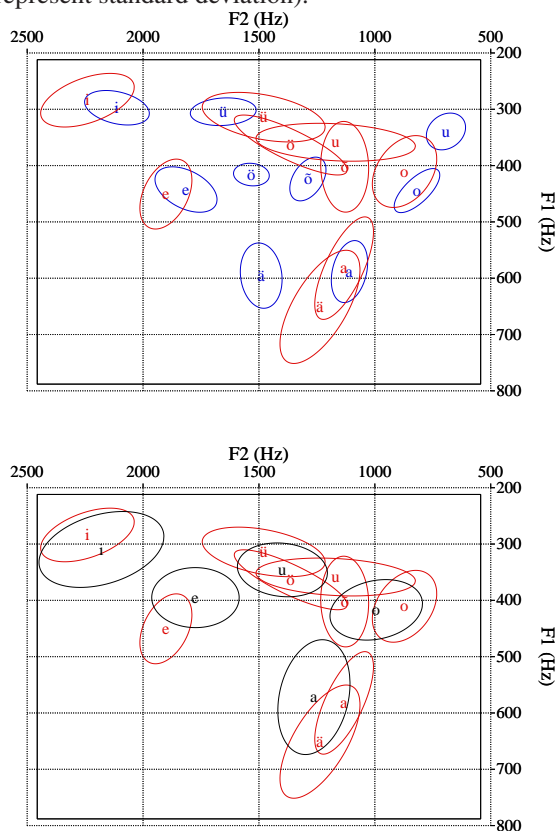
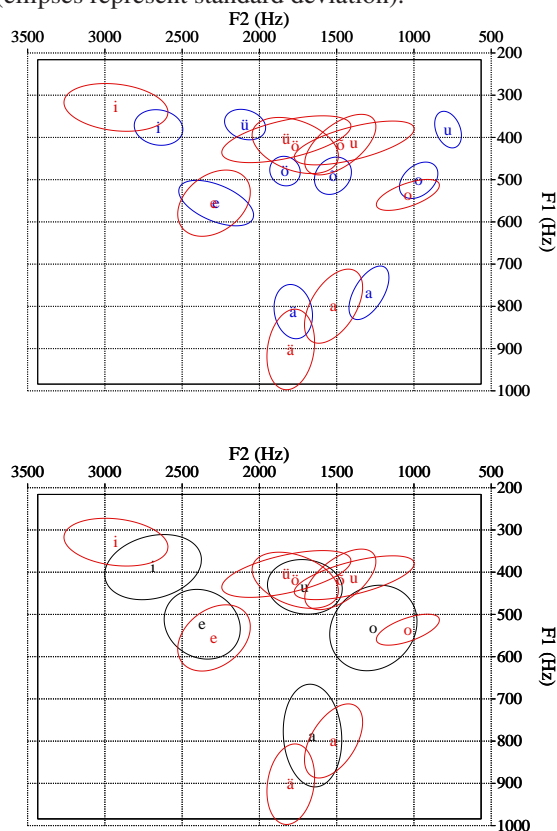


Figure 3: Vowels of female speakers. Top: L2 (red) and L1 Estonian (blue) vowels. Bottom: Estonian L2 (red) and L1 Japanese (black) vowels (ellipses represent standard deviation).



TukeyHSD test revealed, L2 males discriminated /ü/ and /õ/ vowels along F1 ($p=0.01$) and F2 ($p<0.05$), and L2 females /ü/ and /u/ along F2 ($p<0.05$).

Figures 2 and 3 (bottom) compare the location of L2 vowels and Japanese native vowels. In line with theoretical models (SLM, PAM, and L2LP), the figures explain the difficulties in forming new vowel categories for L2 vowels.

The hypotheses of this study were mainly confirmed by the results. An alternative scenario (see hypothesis 1) that it might be easy to acquire the unrounded-rounded contrast salient in Estonian front vowel pairs /i, ü/ and /e, õ/ was suppressed by the high degree of acoustic similarity between /ü, õ/ and Japanese /u/.

5. CONCLUSIONS AND PERSPECTIVES

The results of the study showed that Japanese subjects were successful in producing the Estonian vowels /a, e, i, o/ since these assimilate to the corresponding Japanese vowels. The acquisition of L2 /a-ä/ contrast leads to the splitting of the single na-

tive /a/-category which, however, was demonstrated by the L2 female subjects only. The most difficult L2 vowels are /ü/, /õ/, /õ/, and /u/ which to a great extent overlap in L2 production due to single-category assimilation to Japanese /u/. The high degree of acoustic similarity between these L2 contrasting vowels and a single L1 category causes unfavorable conditions for the creation of new categories and results in poor discrimination in L2 production.

Subsequent studies on perceptual discrimination of Estonian vowels may shed more light on the acquisition of Estonian vowel contrasts by Japanese speakers and explain the results of the current study.

6. ACKNOWLEDGEMENTS

This research was supported by the European Regional Development Fund through the Estonian Center of Excellence in Computer Science (EXCS) and the National Program for Estonian Language Technology. All subjects who participated in the study are gratefully acknowledged.

7. REFERENCES

- [1] Best, C. T., Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn, O.-S., Munro, M. J., (eds), *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins chapter 2, 13–34.
- [2] Boersma, P., Weenink, D. Praat: doing phonetics by computer [computer program]. Version 5.3.10, retrieved 10 April 2012 from <http://www.praat.org/>.
- [3] Eek, A., Meister, E. 1998. Quality of standard Estonian vowels in stressed and unstressed syllables if the feet in three distinctive quantity degrees. *Linguistica Uralica* 34(3), 226–233.
- [4] Elvin, J., Escudero, P., Vasiliev, P. 2014. Spanish is better than English for discriminating Portuguese vowels: acoustic similarity versus vowel inventory size. *Frontiers in Psychology* 5.
- [5] Escudero, P. 2005. *Linguistic Perception and Second Language Acquisition. Explaining the attainment of optimal phonological categorization*. PhD thesis Utrecht University.
- [6] Escudero, P., Boersma, P. 2002. The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. Skarabela, B., Fish, S., Do, A. H.-J., (eds), *Proceedings of the 26th Annual Boston University Conference on Language Development* 208–219.
- [7] Escudero, P., Sisinni, B., Grimaldi, M. 2014. The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *Journal of the Acoustical Society of America* 135, 1577–1584.
- [8] Escudero, P., Williams, D. 2012. Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America* 131, EL406–EL412.
- [9] Flege, J. E. 1995. Second language speech learning: Theory, Findings, and Problems. In: Strange, W., (ed), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*,. Timonium, MD: York Press chapter 8, 233–277.
- [10] Hirata, Y., Tsukada, K. 2004. The effects of speaking rate and vowel length on formant movements in Japanese. Agwuele, A., Warren, W., Park, S.-H., (eds), *Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in Speech Production and Perception* 73–85.
- [11] Iverson, P., Evans, B. G. 2007. Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America* 122(5), 2842–2854.
- [12] Iverson, P., Evans, B. G. 2009. Learning English vowels with different first-language vowel systems ii: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America* 126(2), 866–877.
- [13] Kamiyama, T. 2011. Pronunciation of French vowels by Japanese speakers learning French as a foreign language: back and front rounded vowels /u y ø/. *Phonological Studies (Phonological Society of Japan)* 97–108.
- [14] Kamiyama, T., Vaissière, J. 2009. Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. *Revue AILE...LIA* (2), 9–41.
- [15] Lambacher, S., Martens, W., Molholt, G. 2000. Comparison of identification of american english vowels by native speakers of japanese and english. *Proceedings of the Phonetic Society of Japan* 213–217.
- [16] Lehiste, I. 1960. Segmental and syllabic quantity in Estonian. In: *American Studies in Uralic Linguistics* volume 1. Bloomington: Indiana University Publications.
- [17] Lippus, P. 2010. Variation in vowel quality as a feature of Estonian quantity. Hasegawa-Johnson, M., Bradlow, A., Cole, J., Livescu, K., Pierhumbert, J., Shin, C., (eds), *Speech Prosody 2010* Chicago, USA.
- [18] Meister, L., Meister, E. 2012. Aktsendikorpus ja võõrkeele aktsendi uurimine. *Keel ja Kirjandus* 55(8–9), 696–714.
- [19] Meunier, C., Frenck-Mestre, C., Lelekov-Boissard, T., Le Besnerais, M. 2003. Production and perception of vowels: Does the density of the system play a role? *Proceedings of the 15th International Congress of Phonetic Sciences* 723–726.
- [20] Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., Trent-Brown, S. A. July 2008. Acoustic and perceptual similarity of Japanese and American English vowels. *Journal of the Acoustical Society of America* 124(1), 576–588.
- [21] Strange, W., Akane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., Jenkins, J. J. 1998. Perceptual assimilation of american english vowels by japanese listeners. *Journal of Phonetics* 26, 311–344.