The production of Swedish sibilant fricatives by native speakers of Estonian – an acoustic analysis

Merle Põdra, Eva Liina Asu
Institute of Estonian and General Linguistics, University of Tartu, Estonia
merlepodra@gmail.com, eva-liina.asu@ut.ee

Abstract

In Estonian, there are two word-initial sibilant fricatives /s/ and /ʃ/, while Swedish distinguishes between three: /s/, *tj*-sound and *sj*-sound. The latter two lack direct counterparts in Estonian, and can therefore be difficult for the Estonian learners to acquire. An experiment was conducted with 11 Estonian L1 speakers with an advanced level of Swedish with the aim to test the production of their Swedish fricatives. The means of duration, intensity and center of gravity were calculated. The results revealed a large allophonic variation in the Estonian speakers' data. On the whole, the advanced Estonian speakers of Swedish seem to have acquired the acoustic characteristics of the Swedish fricatives, but there was some confusion in choosing a suitable allophone, e.g. [ɕ] was often used for both the *sj*- and *tj*-sound.

Introduction

Estonian and Swedish have a different set of sibilant fricatives in the word-initial position. There are only /s/ and /ʃ/ in Estonian (Ariste, 1981) but in Swedish, a distinction is made between *tj*-sound, *sj*-sound and /s/ (Lindblad, 1980; Garlén, 1988). *Sj*- and *tj*-sound do not have direct counterparts in Estonian and their pronunciation varies a lot geographically and socially (Lindblad, 1980). Therefore, they may be difficult for the Estonian learners to acquire.

While the standard version of the *tj*-sound is considered to be [ɕ] (Elert, 2000), also [ç] (Lindblad, 1980), [t͡ɕ] and [t͡ʃ] are used (Garlén, 1988). The variation is even more complex in the case of *sj*-sound. It can either be a sibilant or a non-sibilant (Garlén, 1988) and its many allophones can be divided into 'back' (velar, dorsal) or 'front' (apical, predorsal) variants (Elert, 2000, Malmberg, 1971). Additionally, a distinction of 'lighter' (*ljusa*) or 'darker' (*mörka*) allophones is made depending on the place of articulation. A foreign learner of Swedish is expected to distinguish between the *sj*- and *tj*-sounds, but there are no fixed versions to be learned (Garlén, 1988).

There is one grapheme *s* representing /s/ in Swedish, but for the *sj*- and *tj*-sounds various grapheme combinations are used (for a detailed list see Bolander, 2003: 50). This can additionally complicate the acquisition process of the latter sounds.

It is assumed that only 0.3–10% of adult learners are capable of acquiring a second language without accent (Markham, 1997). The acquisition of L2 sounds is highly connected to perception, and the similarities between L1 and L2 sounds. Both the Perceptual Assimilation Model for L2 learners (PAM-L2) (Best & Tyler, 2007) and the Speech Learning Model (Flege, 1995) place L1 and L2 sounds in one continuum and connect the precision of acquisition to the language experience of the learner. According to PAM-L2, it is the perceptual similarities between L1 and L2 sounds that determine whether a new category will be created or the L2 sound in question is assimilated with the most similar L1 category. The Speech Learning Model (Flege, 1995) claims that the contrasts between sounds are created through similarities or differences between the phonetic features of the L1 and L2

sounds. L2 sounds are acquired when the L2 learner's phonetic categories match with the native speaker's categories (Flege, 1995; Meister, 2011).

This article focuses on the production of Swedish /s/, *sj*-sound and *tj*-sound by native speakers of Estonian by analyzing the allophonic variation as well as duration, intensity and center of gravity of the fricatives. Because both Estonian and Swedish have /s/, we expect, in line with PAM-L2 and the Speech Learning Model, the Estonian speakers to use their L1 categories for the Swedish /s/. We assume that at least one new category is formed for the Swedish *sj*- and *tj*-sound in order to differentiate the Swedish sounds from the Estonian /ʃ/.

Materials

An experiment was conducted to test the production of the Swedish sibilants by Estonian L1 speakers. The experiment consisted of reading a test word in a carrier sentence *Jag sa ... igen* ('I said ... again') that appeared on a PowerPoint slide together with an illustrative picture. The test words (*särbo, sed, syl, sol, sal, kärna, kedja, kyl, kjol, tjat, stjärna, sked, sky, sjok* and *sjal*) were partly taken from Lindblad (1980). The fricatives were always word-initial and were followed by the vowels /æː/, /eː/, /yː/, /uː/ and /ɑː/. Each slide appeared three times in a random order.

Eleven Estonian subjects (6 women and 5 men) with an advanced level of Swedish (B1-C1, according to CEFR) participated in the experiment. All have learned Swedish after the *critical period* (Scovel, 1969; Meister, 2011) but use it daily as their working language. The Estonian subjects were also asked to read a set of test sentences in Estonian including six test words *(seen, suu saal, šeff, šunt, šaht*) which appeared in a carrier sentence *Anna ... edasi* ('Pass ... on').

In order to compare the Estonian speakers' L2 data with Swedish, 5 native Swedish speakers were recorded reading the same Swedish sentences.

Method

The recordings were made in the Phonetics Lab of the University of Tartu using a Beyerdynamic MC 930 microphone and a Sound Devices USBPre sound card. The data was annotated in Praat (Boersma & Weenink, 2014) and the measurements were obtained automatically using Praat scripts. Intensity difference between the fricative and the following vowel was calculated manually. For the statistical analysis in R, ANOVA and *Post-hoc Tukey**HSD* tests were used (R Core Team, 2015).

Results

Allophonic variation

The Estonian subjects' allophones for the *sj*-sound can be divided into three groups: voiceless palatal predorsal-alveolar fricative [ɕ] (28%), voiceless velar fricative [x] (31%), and voiceless dorso-velar fricative [ɧ] (39%). [x] was used as the only allophone by two speakers, [ɧ] by two speakers, and [ɕ] by one speaker. The others produced at least two of the allophones of which one was always [ɕ]. The 'darker' allophones [x] and [ɧ] were never used together. One informant also used consonant clusters /sj/ (1%) and /hʃ/ (1%).

The palatal fricative [ɕ] was the most common pronunciation variant for the *tj*-sound (92%). Five Estonian subjects used only [ɕ], but the others produced [ɕ] together with one or more of the following allophones: [x] (1%), [s'] (2%), [sj] (<1%), [k] (<1%), [hʃ] (<1%), [stɕ] (<1%), [sts'] (<1%) and [tɕ] (<1%).

/s/ was the most homogenic sound: [s] was used in 96% of the cases, but a few other allophones also appeared: [ʃ] (1%), [ɕ] (2%) and [h] (1%).

The native Swedish speakers used [ʂ] (21%), [x] (67%) and [ɧ] (12%) for the *sj*-sound. Of these allophones, [ʂ]

did not appear in the pronunciation of the Estonian speakers. Only [ɕ] was used for the *tj*-sound and [s] for /s/.

By way of comparison, only [s] was used for the Estonian /s/ and only [ʃ] for the Estonian /ʃ/ by the Estonian informants in the Estonian test words.

Acoustic analysis

Duration

The durational variation (and the medians) of the main allophones of the Swedish sounds as produced by the Estonian subjects can be found in Figure 1. The mean durations of the main allophones were the following: [s] 167 ms, [ɕ] /*tj*-sound/ 162 ms, [ɕ] /*sj*-sound/ 162 ms, [ɧ] /*sj*-sound/ 149 ms and [x] /*sj*-sound/ 166 ms. It can be seen from the data that the allophone [ɕ] of the *tj*-sound and the allophone [ɕ] of the *sj*-sound have the same mean duration. There was no significant effect of sibilant type on the mean duration ($F(4, 163) = 1.862$; $p>0.05$).

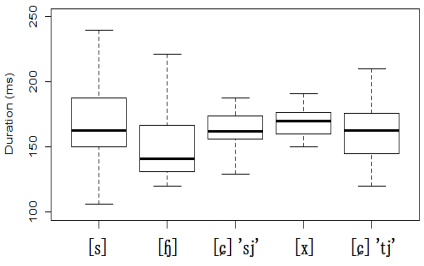

Figure 1. The variation in the duration of the allophones of the Swedish /s/, *tj*-sound and *sj*-sound by the Estonian subjects.

The mean duration of [ɧ] was slightly shorter than that of the other fricatives (149 ms). This can be explained by the even shorter (135 ms) mean duration of [ɧ] produced by the female informants while the mean duration of [ɧ] by male informants was 174 ms. *Post-hoc TukeyHSD* test showed a significant difference between the duration of [ɧ] ($p<0.01$) produced by male and female speakers of Estonian.

By comparison, the mean durations of the native Swedish speakers' sibilants were the following: [s] 180 ms, [ɕ] 166 ms, [ʂ] 163 ms, [ɧ] 194 ms and [x] 150 ms. There was a significant effect of sibilant type on the mean duration of the native Swedish speakers' fricatives ($F(4, 70) = 7.122$; $p<0.001$).

The average durations of the Estonian sibilants by the Estonian subjects were 123 ms for /s/ and 125 ms for /ʃ/ which is considerably shorter than the average durations of their Swedish sibilants.

Intensity

Figure 2 presents the variation and the medians for the intensity differences between the Swedish fricatives and the following vowel as produced by the Estonian subjects. The mean differences were the following: [s] 10 dB, [ɕ] /*tj*-sound/ 10 dB, [ɕ] /*sj*-sound/ 11 dB, [ɧ] /*sj*-sound/ 14 dB and [x] /*sj*-sound/ 16 dB. There was a significant effect of sibilant type on intensity ($F(4, 163) = 28.02$; $p<0.001$).

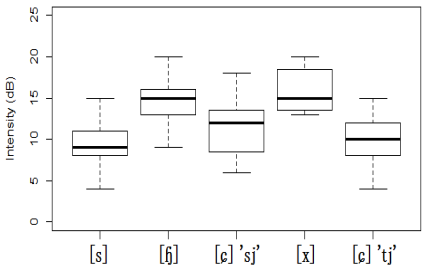

Figure 2. The variation in the mean intensity differences between the allophones of the Swedish /s/, *tj*-sound and *sj*-sound and the following vowel by the Estonian subjects.

It can be seen in Figure 2 that [ɧ] and [x] had a significantly ($p<0.001$) higher intensity differences, and therefore lower fricative intensity, as compared to /s/ and the allophone [ɕ] of the *tj*-sound. A *Post-hoc TukeyHSD* test showed no significant differences between the allophone [ɕ] of the *sj*-sound and the [ɕ] of the *tj*-sound ($p>0.05$).

The Estonian /s/ and /ʃ/ had the intensity differences of 9 dB and 11 dB, which are closer to the Swedish /s/ and

[ɕ] than to the allophones [x] and [ɧ] of the Swedish *sj*-sound.

The results of the native Swedish speakers also showed a significant effect of sibilant type on intensity ($F(4,70) = 9.494$; $p<0.001$). The intensity difference for [ɧ] was 18 dB and for [x] 16 dB, while [ɕ] /*tj*-sound/ had 9 dB, [ʂ] 10 dB, and /s/ 10 dB. Similarly to the Estonian speakers, the /sj/ allophones [ɧ] and [x] of the native Swedish speakers had significantly higher intensity differences as compared to [s] and [ɕ] ($p<0.01$). There was no significant difference between the results of the Estonian and the native Swedish speakers' Swedish sibilants ($p > 0.05$).

Center of gravity

The results for the Swedish sibilants by the Estonian subjects are in line with those of earlier studies (e.g. Gordon et al. 2002; Gordon & Appelbaum, 2006; Padgett & Zygis, 2003). The variation and the medians of the center of gravity of all the allophones can be seen in Figure 3. There was a significant effect of sibilant type on the mean center of gravity ($F(4,163) = 174.5$; $p<0.001$); /s/ had the highest mean center of gravity (7141 Hz), followed by [ɕ] /*sj*-sound/ 4455 Hz, [ɕ] /*tj*-sound/ 4370 Hz, [ɧ] /*sj*-sound/ 2422 Hz and [x] /*sj*-sound/ 1064 Hz. *Post-hoc TukeyHSD* tests showed a highly significant difference ($p<0.001$) between all sound pairs except between the allophone [ɕ] of the *tj*-sound and *sj*-sound ($p>0.05$), which refers to their similar place of articulation. It could also be seen that the center of gravity of the sounds produced by female speakers had somewhat higher values. For example, the mean center of gravity of the Swedish /s/ produced by the male Estonian subjects was 6237 Hz while it was 7906 Hz when produced by the female subjects ($p<0.001$).

There was also a significant effect of sibilant type on the mean center of gravity as produced by the native Swedish speakers ($F(4, 70) = 339.1$; $p<0.001$). The values for the center of gravity in the native Swedish speakers' sibilants were the following: /s/ 8051 Hz, [ɕ] 4583 Hz, [ʂ] 3664 Hz, [ɧ] 1862 Hz, and [x] 1104 Hz. This order is similar to the Estonian speakers' results where /s/ had the highest center of gravity.

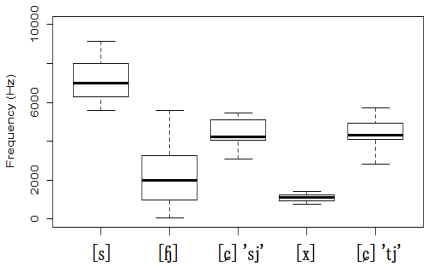

Figure 3. The variation in the center of gravity of the allophones of the Swedish /s/, *tj*-sound and *sj*-sound by the Estonian subjects.

As compared to the /s/ of the native Swedish speakers (8051 Hz), the Estonian speakers' Swedish /s/ had a lower mean center of gravity (7141 Hz) but it was still higher than the center of gravity of the Estonian speakers' /s/ in Estonian (6405 Hz). The mean center of gravity of the Estonian /ʃ/ was 4148 Hz, which is more similar to that of the Swedish *tj*-sound than to the *sj*-sound.

Discussion

As can be seen from the results, the Estonian subjects used various allophones for the Swedish *sj*- and *tj*-sound. According to PAM-L2 (Best & Tyler, 2007), a new category may be formed when an L2 sound is considered to be different from an L1 sound. It can be seen that in general Estonian subjects distinguish these Swedish sounds well from the Estonian /ʃ/ because a new category – either a palatal or velar fricative – has been created for both the *sj*-sound and *tj*-sound. Acoustically, the Swedish sibilants produced by the Estonian speakers' are more similar to those of the native Swedish speakers than the Estonian sounds.

The analysis also revealed a large interspeaker variation, but the common pattern was that only one of the 'darker' fricatives was used by each speaker. The reason for this could be that Estonian does not have fricatives with back articulation, and thus it is easier to maintain a greater contrast with the palatal [ɕ]. Although most of the subjects distinguished between the two Swedish sounds, there was one Estonian subject who used only one allophone [ɕ] for both sounds. A possible reason for using [ɕ] for both *sj*- and *tj*-sound can be a lexical influence. There was a tendency, also by some other subjects to pronounce some *sj*-words (e.g. *stjärna* and *sjok*) with an [ɕ]. The fact that /tj/ is chosen instead of /sj/ is confirmed by the acoustic results which are almost identical in case of [ɕ] of the *sj*-sound and [ɕ] of the *tj*-sound. The native Swedish speakers did not use this allophone for the *sj*-sound at all.

Duration of the sibilants differs a lot in Estonian and in Swedish, but the duration of the Swedish sibilants pronounced by the Estonian subjects was more similar to that of the native Swedish speakers than to Estonian /s/ and /ʃ/. The Swedish /s/ produced by Estonian speakers had a shorter duration than the /s/ of the native Swedish speakers. This may imply L1 Estonian influence on Swedish. The same was true for the center of gravity of the Swedish /s/ as produced by the Estonian subjects.

The results indicate that the intensity of the Estonian /s/ and /ʃ/ is similar to some of the Swedish sounds. Also, the intensity of the Swedish [ɧ] and [x] produced by the Estonian speakers was lower as was the case in the native Swedish speakers' pronunciation. This means that the Estonian speakers Swedish sibilants are more similar to those of the native Swedish speakers than their own Estonian productions.

The acoustic patterns of the center of gravity of the Swedish sounds as produced by the Estonian and native Swedish speakers were similar, and even here the Estonian speakers managed to follow the patterns of the Swedish sounds. The center of gravity of the Estonian /ʃ/ is more similar to the *tj*-sound than to the non-palatal allophones of the *sj*-sound, which means that the characteristics of the /ʃ/ are not transferred over to the *sj*-sound.

Conclusion

Although the Swedish *sj*-sound and the *tj*-sound are considered difficult for the Estonian learners, our experiment showed that the Estonian subjects had acquired the acoustic characteristics of these sounds. Large allophonic variation gives evidence of some confusion in choosing a suitable allophone. Curiously, most L1 influence was found in the pronunciation of /s/.

It is important to remember that the Estonian subjects who participated in the experiment are all experienced Swedish users, which undoubtedly contributed to their very similar results as compared to the native speakers. To get a better overview of the articulation boundaries between the different allophones and of the allophonic choices between Swedish sibilants, perception tests are needed. Also, a similar experiment could be repeated with subjects with a lower level of Swedish.

Acknowledgements

We are grateful to our subjects for participating in the experiment, and to Pärtel Lippus for his help with writing the Praat and R scripts for the analysis. The second author was supported by Estonian Research Council grant IUT2-37.

References

Ariste, P. (1981). *Eesti keele foneetika*. 1. osa. Tartu: Tartu Riiklik Ülikool

Best, C. T., & Tyler, M. D. (2007). 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.* (pp. 13–34). Amsterdam: John Benjamins.

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer. [Computer program]. (Version 5.4.) Retrieved 5 October 2014 from http://www.praat.org/.

Bolander, M. (2003). *Funktionell svensk grammatik*. Stockholm: Liber.

CEFR = *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Language Policy Unit, Strasbourg. Retrieved from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf

Elert, C.-C. (2000). *Allmän och svensk fonetik*. (8. omarb. uppl.) Stockholm: Norstedts.

Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language research*. (pp. 229–273). Timonium, MD: York Press.

Garlén, C. (1988). *Svenskans fonologi*. Lund: Studentlitteratur.

Gordon, M., Barthmaier, P. & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association 32*, 141–174.

Gordon, M., & Applebaum, A. (2006). Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association 36 (2)*, 159–186.

Lindblad, P. (1980). *Svenskans sje- och tje-ljud i ett allmänfonetiskt perspektiv*. Travaux de l'Institut de Linguistique de Lund 16, Lund: Gleerup.

Malmberg, B. (1971). *Svensk fonetik.* Lund: Gleerups.

Markham, D. (1997). *Phonetic Imitation, Accent, and the Learner*. Lund: Lund University Press.

Meister, L. (2011). *Eesti vokaali- ja kestuskategooriad vene emakeelega keelejuhtide tajus ja hääldus*es. *Eksperimentaalfoneetiline uurimus.* Doktoriväitekiri. Tartu: Tartu Ülikool.

Padgett, J., & Zygis, M. (2003). The Evolution of Sibilants in Polish and Russian. In T. A. Hall, & S. Hamann (Eds.), *ZAS Papers in Linguistics 32*, 155–174.

R Core Team (2015). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Retrieved from http://www.R-project.org

Scovel, T. (1969). Foreign Accents, Language Acquisition and Cerebral Dominance. *Language Learning 19*, 245–253.