Vowel quantity contrasts in sung Estonian

Kaili Vesik - University of British Columbia

Estonian contrasts length in both vowels and consonants (Fox & Lehiste, 1989; Prillop, 2013; Lippus, Pajusalu, & Allik, 2009). In particular, it has been described as having a three-way length contrast (Ross & Lehiste, 1998; Eek, 1983). The three-way contrast is available only in the initial syllable of polysyllabic words (Asu & Teras, 2009), which is the location of primary stress (Lehiste, 1960). These contrasting values are referred to as Quantity 1 (associated with short phonological length), Quantity 2 (long), and Quantity 3 (overlong). For example (Spahr, 2014):

(1)

Q1 (short)	Q2 (long)	Q3 (overlong)
vina	*viina*	*viina*
[vi.nɑ]	[vii.nɑ]	[vii:.nɑ]
vapour.NOM	vodka.GEN	vodka.PART

Estonian quantities, assigned to word-initial syllables, are correlated perceptually with particular values for the ratios of first-syllable to second-syllable durations, rather than the raw durations of the stressed syllables themselves (Eek & Meister, 1997). Lehiste (1960) showed that the average $\sigma 1/\sigma 2$ duration ratio of Q1 words is 2:3, that of Q2 words is 3:2, and of Q3 words is 2:1. However, Lippus et al. (2009) showed that native speakers required the addition of falling pitch on the first syllable in order to successfully identify targets as Q3 rather than Q2 at levels better than chance. Thus only the distinction between Q1 and Q2 is available independent of pitch.

In some settings, typical temporal patterns are disrupted; for example, in music. This is particularly relevant to Estonian in that joint singing (e.g. choral music) is a large part of national identity and cultural transmission in Estonia (Raudsepp & Vikat, 2009, 2011). This study investigates whether Estonian listeners are able to adjust to a greater degree of length variation in sung vs spoken language. More specifically, via an initial corpus exploration that informs a subsequent perception experiment, I show that (a) in a context that provides lexical clues to vowel quantity, the measures that determine Estonian vowel quantity contrasts in speech are not entirely preserved in natively composed choral music, and (b) in a context where those lexical clues are not applicable, native listeners perceive vowel quantity contrasts in sung Estonian similarly to the way they do in speech.

Due to the necessity of f0 cues for distinguishing Q2 from Q3, the focus of this study is simply on the perception of Q1 vs Q2 (i.e., short vs long). In sung Western music, text is often set within a metrical framework in which $\sigma 1/\sigma 2$ ratios of 1:1, 2:1, or 3:1 (or their reciprocals) are much more likely than those of 2:3 or 3:2. These Western conventions would prevent composers from setting text in a way that faithfully follows Lehiste's ratios, but it should still be expected that Q1 syllable pairs would tend to be assigned to shorter:longer note-length ratios while Q2 syllable pairs would tend to be assigned to longer:shorter note-length ratios. However, given a small collection of Estonian choral music, I show that there is no significant effect of vowel quantity on the note-length ratios used by composers in setting Estonian text to music; that is, that text is not set to music according to the same ratios found in speech. This suggests that preserving speech-based durational ratios is not of primary importance to composers of Estonian music, who instead choose text settings that allow for a greater degree of musical expression. The ambiguity in how individual words are

set rhythmically could potentially be accommodated by listeners due to the availability of the lexical information in the surrounding text. In the presence of lexical clues, duration need not be a strong cue for vowel quantity.

Informed by the corpus exploration, I analyze native listeners' perception of vowel quantity in sung Estonian as compared to spoken Estonian, investigating how native listeners perceive sung Estonian vowel quantities over a continuum of $\sigma 1/\sigma 2$ duration ratios. I show that there is a significant positive effect of note-length ratio on the probability that a bisyllabic word is identified as being in Q2, and that Lehiste's (1960) ratios for Q1 (2:3) and Q2 (3:2) in speech are identified similarly in sung Estonian (with ratios $\leq$ 2:3 perceived as Q1, those $\geq$ 3:2 as Q2, and with the crossover point of the categorization curve having approximately the same value in both contexts. The significant effect of note-length ratio on listeners' perception of vowel quantity shows that without the advantage of lexical clues from lyrics, listeners use durational cues to identify sung words as being in Q1 vs Q2, with the results at the endpoints of the note-ratio continuum being quite well-aligned with Lehiste's results for speech.

References

Asu, E. L., & Teras, P. (2009). Estonian. *Journal of the International Phonetic Association*, *39*(3), 367–372. doi: 10.1017/S002510030999017X

Eek, A. (1983). Kvantiteet ja rõhk eesti keeles (I). *Keel ja kirjandus*, *9*, 481-489.

Eek, A., & Meister, E. (1997, 01). Simple perception experiments on estonian word prosody. In *Estonian prosody: Papers from a symposium* (p. 71-99).

Fox, R. A., & Lehiste, I. (1989). Discrimination of duration ratios in bisyllabic tokens by native English and Estonian listeners. *Journal of Phonetics*, *17*(3), 167-174.

Lehiste, I. (1960). Segmental and syllabic quantity in estonian. In *American studies in uralic linguistics* (p. 21-82). Bloomington: Indiana University.

Lippus, P., Pajusalu, K., & Allik, J. (2009). The tonal component of Estonian quantity in native and non-native perception. *Journal of Phonetics*, *37*(4), 388-396.

Prillop, K. (2013). Feet, syllables, moras and the Estonian quantity system. *Linguistica Uralica*, *49*(1), 1-29.

Raudsepp, I., & Vikat, M. (2009). Joint singing as a phenomenon of Estonian cultural transmission. *Problems of Education in the 21st Century*, *13*, 103-109.

Raudsepp, I., & Vikat, M. (2011). The role of the phenomenon of joint singing in the development of national identity in Estonia. *Procedia Social and Behavioral Sciences*, *29*, 1312-1319.

Ross, J., & Lehiste, I. (1998). Timing in Estonian folk songs as interaction between speech prosody, meter, and musical rhythm. *Music Perception: An Interdisciplinary Journal*, *15*(4), 319-333.

Spahr, C. (2014, 01). Prosodic epenthesis and floating vowels in estonian quantity [conference handout]. In *Eleventh old world conference on phonology.* Retrieved from `http://individual.utoronto.ca/spahr/spahr_ocp11_handout.pdf`