doi:10.5128/ERYa21.09

MAPPING BOUNDARIES: MICROPROSODIC AND DIALECTAL VARIABILITY IN THE PERCEPTION OF ESTONIAN LONG AND OVERLONG QUANTITY

Pärtel Lippus, Liis Kask, Sofia Lutter,
Nele Põldver, Kairi Kreegipuu

Abstract. This paper presents the results of a web-based perception experiment that tested the distinction of long (Q2) and overlong (Q3) quantity degrees in Estonian. Firstly, we observe an effect of segmental quality on the perception of quantity. Most of the quantity experiments have used one or two minimal triplets, typically with open back vowels, not considering variation due to intrinsic properties. The stimuli were created from words with 8 different segmental combinations. The second aim of this study is to map the dialectal variability in Estonian quantity perception. There has been some evidence that listeners from East and South dialect areas are less sensitive to pitch cue than those from North and West. The current study was carried out with 290 participants of different regional backgrounds. The results showed that different segmental quality sets have slightly different Q2–Q3 category boundaries. Also, the precision of quantity identification was considerably lower in the case of a nonsense word set. Unexpectedly, we were not able to find a clear dialectal background effect.*

Keywords: quantity, duration, pitch, microprosody, perception, dialectal variability, Estonian

1. Introduction

The three-way quantity system is the most studied feature of Estonian phonology, yet there are still many un- or understudied aspects. This paper attempts to shed more light on a few of them.

Estonian is a quantity language, meaning that it has a phonological length category. The distinction of the three quantity degrees (Q1 – short, Q2 – long, and Q3 – overlong) is a main feature of word prosody and is productive for both vowels

* The work was supported by the Estonian Research Council grant "Pre-attentive information processing in the brain: Relationship to state and trait variables, and behaviour" (PRG1151) and the National Program: Estonian Language and Culture in the Digital Age project "From gesture to speech: Multimodal analysis of spoken Estonian" (EKKD-TA6).

and consonants. Earlier studies of quantity perception have shown that the stressed syllables of Q2 and Q3 cannot always be distinguished from each other when they are isolated from the second syllable. The length of the second syllable affects the identification of the quantity in the first syllable (Eek, Meister 2003, Krull et al. 2003, Lehiste 1997).

In disyllabic words, the duration ratios of stressed (S1) and unstressed (S2) syllables can be described, varying within the following limits: 2:3 (0.5–0.8) in Q1, 3:2 (1.2–2.3) in Q2, and 2:1 (2.0–3.9) in Q3, although vowel quantities cannot be determined solely based on these proportions due to considerably large variability in production (Asu et al. 2016: 134–136). One of the main factors that helps distinguishing Q2 words from Q3 words is the pitch cue. In Eek's (1980a) experiment, the stimuli formed from a Q1 word was perceived as Q2 when the duration ratio of the original word had been modified to be characteristic of Q2 words. However, it was not possible to get the perception of Q3 from a Q1 or Q2 word by lengthening the duration without changing the pitch cue. In a similar experiment (Lippus et al. 2009), the stimuli were re-synthesized from six words (Q1 [sɑtɑ] – Q2 [sɑːtɑ] – Q3 [sɑːːtɑ] with vowel lengthening and Q1 [kɑtɑ] – Q2 [kɑttɑ] – Q3 [kɑtːtɑ] with consonant lengthening), changing only the duration ratios by manipulating the V1 duration. For words with consonant lengthening, respondents distinguished all three quantities by duration ratios, but for words with vowel lengthening, they did not distinguish Q2 and Q3 from words synthesized from Q1 and Q2 words with the duration ratio corresponding to Q3 words. The importance of both pitch and temporal cues for distinguishing Q2 and Q3 has also been shown in an EEG study (Kask et al. 2021).

Most of the perception experiments exploring the quantity system have used one or two minimal triplets (i.e. a series of three words with identical segmental sequence that only differ by the length of the segments), typically including only an open back vowel [ɑ]: e.g. [sɑtɑ] – [sɑːtɑ] – [sɑːːtɑ] (Lehiste 1997, Lippus et al. 2009, Lippus et al. 2011) or [jɑmɑ] – [jɑːmɑ] – [jɑːːmɑ] (Eek 1980a: 19, Eek 1980b). The intrinsic properties of segmental quality (the effects of place of articulation on the segmental duration, pitch and intensity) are lesser studied in Estonian prosodic research in general. The few studies observing intrinsic duration effects in Estonian have shown that the short-long category boundary is perceived at shorter duration for closed vowels than for open vowels (Meister et al. 2011, Meister, Werner 2009). This intrinsic duration can dramatically affect the syllable duration ratios, as was shown in an acoustic study by Lippus and Šimko (2015). As the nature of the microprosodic variation is mainly related to articulatory effort and is a psychoacoustical universal, we expect that the listener can correctly evaluate the quantity of words with different segmental quality. However, these intrinsic properties need to be accounted for in the description of the quantity system.

There are a few accounts of dialectal variability in the Estonian quantity system. In the 19th century dialect mapping, the three-way quantity distinction was missing in the North-East Coastal area. In contemporary Estonian the speakers from this region do produce the distinction of long and overlong quantities, but the variation is great, and the pitch peak does not distinguish Q2 and Q3 (Kalvik 2005). The earlier descriptions of Estonian dialects based on auditory impressions have also claimed that while the three-way system was more regular in Northern Estonian dialects, in some Southern dialects Q2 has been said to be "strengthened", meaning that it is produced more with the features of Q3. In contemporary South Estonian no significant temporal differences were found in comparison with the standard variety, but the pitch peak occurred earlier (Parve 2003). Earlier peak alignment in South Estonia compared to the later peak in Saaremaa was also noticed by Lippus and Ross (2011). A recent perception study by Piits and Mihkla (2024) showed that when asked to categorise the quantity of difficult cases, the participants from South Estonian and North-East Costal areas preferred to rate them as Q3 more frequently than the rest of the participants.

While previous perception studies (Lehiste 1997, Eek 1980b) had shown that the early falling pitch is the vital cue for perceiving the overlong quantity for Estonian listeners, Lippus and Pajusalu (2009) found that this was so only for half of the study group of 38 participants, who were mostly from Northern Estonia, while the other half of the participants, mostly from Eastern and Southern Estonia, categorised the stimuli with level pitch as Q3 if the duration of the stressed vowel was lengthened.

This paper addresses two main issues. Firstly, we aim to analyse the segmental quality (i.e. microprosody) effect and how it influences quantity perception. Secondly, we want to map the possible dialectal variability in Estonian quantity perception.

2. Material and method

2.1. Target words

The stimuli were created using re-synthesis of natural speech, manipulating ether the vowel duration or the pitch contour. As base words a set of 8 minimal triplets of quantity with different segmental phoneme combinations were used (see Table 1). Seven of the triplets (or at least their Q2 and Q3 forms) were meaningful words. The eight was a triplet of nonce words. The target words were embedded in a carrier sentence *Ütle ... kõvasti* ‘say ... out loud’ prompted on a computer screen using the SpeechRecorder software (Draxler, Jänsch 2004). The sentences were read by a native Estonian speaker (female, age 29, from Harjumaa, a speaker of the standard variety of Estonian). The recordings were done in a sound-attenuated recording booth at the University of Tartu Phonetics Lab.

Table 1. Stimulus base words in IPA transcription with glosses. The nonce words are indicated with a dash

Q1		Q2		Q3	
sɑtɑ	‘hundred’	sɑːtɑ	‘send’ SG2.IMP	sɑːːtɑ	‘get’ INF
sɑke	‘frequent’	sɑːke	‘get’ PL2.IMP	sɑːːke	‘saw’ PL.PART
sɑki	‘bustle’ SG2.IMP	sɑːki	‘harvest’ SG.GEN	sɑːːki	‘get’
lɑtʲi	–	lɑːtʲi	‘load’ SG2.IMP	lɑːːtʲi	‘style’ SG.PART
likɑ	‘slime’	liːkɑ	‘league’	liːːkɑ	‘too’
vite	–	viːte	‘reference’	viːːte	‘five’ SG.ILL
viki	–	viːki	‘even score’ SG.GEN	viːːki	‘take’ SG2.IMP
tɑtɑ	–	tɑːtɑ	–	tɑːːtɑ	–

While the Q1 word was also recorded, we here describe only the Q2 and Q3 tokens which were used for the test stimuli in this study. The vowel durations and pitch contours of the recoded target words were observed to make sure that the samples are typical representations of Estonian Q2 and Q3 words. In Q2 the mean V1 = 214 ms, V2 = 124 ms; in Q3 V1 = 272 ms and V2 = 69 ms (see Figure 1).

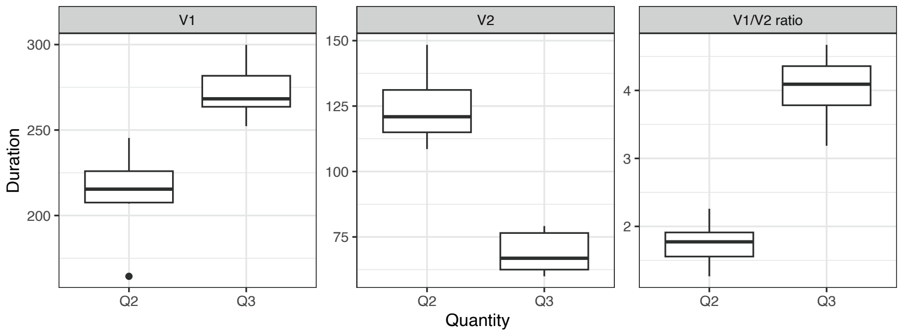

Figure 1. The duration (in milliseconds) of the stressed and unstressed vowels and the V1/V2 duration ratio of the base words

The target words were in a focal position in the carrier phrase and read with an H*L accent (see Figure 2). As the intervocalic consonant was always an unvoiced stop ([t] or [k]), the syllable boundary is clearly visible in the pitch contours as a discontinuation of the curve. The pitch in Q2 tokens was relatively level or gradually rising during the stressed syllable. The average peak in Q2 words was 189 Hz and by the end of the first syllable it was on the average at 181 Hz. Most of the fall took place during the unstressed syllable, finally reaching 149 Hz in the end of V2. In Q3 tokens, the peak was 198 Hz and took place already in the first half of the first vowel, falling to 157 Hz by the end of stressed syllable and continuing at the same level until the end of the unstressed syllable.

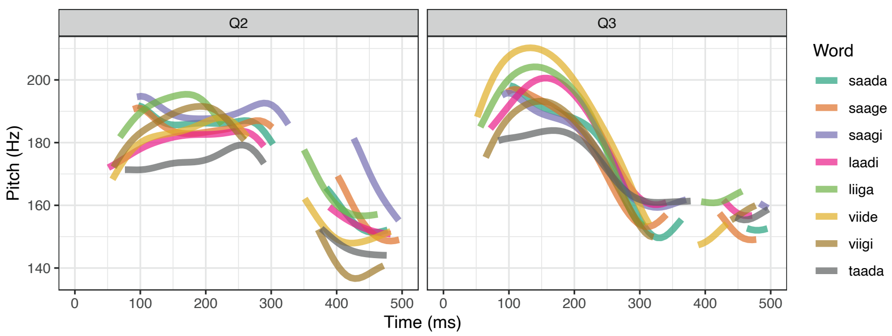

Figure 2. Pitch contours of the base words. The left panel shows the pitch contours of the words in Q2 and the right panel the words in Q3

By both the temporal and tonal patterns, the recoded samples are typical representations of Estonian Q2 and Q3 words matching the descriptions in the previous literature (see a detailed overview in Asu et al. 2016: 134–136, 140–143).

2.2. Stimuli

The re-synthesis was carried out in Praat (Boersma, Weenink 2019). The target words were extracted from the carrier sentences, making the cuts at zero crossings. From each target word three sets of stimuli were re-synthesised.

In Set 1, five stimuli were created from each Q2 and Q3 base word, where the stressed vowel duration was manipulated in 30 ms steps ranging 170–290 ms, following the same procedure that was used in Lippus et al. (2009). The durations of the other segments and the pitch contour remained unchanged. This resulted in 2 (quantity levels) x 8 (phonemic sequences) x 5 (V1 manipulation steps) = 80 different stimuli.

In Set 2, the V1 duration was altered by steps of 30 ms similarly to Set 1, but additionally also the duration of the unstressed vowel was changed by –30 ms in the case of Q2 and by +30 ms in the case of Q3 base words. This resulted in a similar set of five stimuli, where the V1/V2 ratio was gradually shifted. Set 2 again included 2 (quantity levels) x 8 (phonemic sequences) x 5 (V1 manipulation steps) = 80 different stimuli.

In Set 3 the f0 contours were swapped between Q2 and Q3 within the word pairs: the pitch tracks of the base words with same segmental sequence (e.g. [sɑːtɑ] and [sɑːːtɑ]) were extracted, rescaled to fit the total duration of the counterpart and mapped on that word. No manipulations in duration of the segments were made, creating a mismatch between the temporal and tonal pattern. Set 3 consisted of 2 (quantity levels) x 8 (phonemic sequences) = 16 stimuli.

Additionally, a control set of four stimuli were created from Q1 tokens *sada*, *sagi*, *liga*, and *tada*, where V1 duration was set to 110 ms. Those stimuli were included in the experiment to check whether the listener is paying attention, is following the instructions and has understood the concept of the quantity categories.

2.3. Perception test

An online listening experiment was carried out in Kaemus (https://kaemus.psych.ut.ee/), an experiment environment developed at the University of Tartu Institute of Psychology. The experiment was launched in two rounds with different subsets of stimuli in the spring of 2020 and again in the spring of 2021. Different subsets were used to limit the total duration of the experiment to approximately 30 minutes.

The experiment started with a short instruction explaining the notion of the three quantity categories of Estonian. The categorisation of words into three quantity degrees should be familiar to Estonian speakers from the secondary school Estonian language program. The task was to listen to the stimuli and categorise them clicking the buttons labelled "I välde", "II välde", "III välde". During the first data collection (2020) each participant was presented a main block of 128 tokens (each stimulus represented twice) and an optional extra block of 96 tokens (each stimulus represented once). The second round of data collection (2021) consisted of a main block of 126 tokens (represented once) and an extra block of 90 tokens (represented once). After every 20–30 tokens (6 times in the main block and 3 times in the extra block) the participants were asked a control question about the last heard stimulus. They had to choose which of two example sentences fit the best with the meaning of the stimulus. An extra option "The sound is meaningless" was also included (and correct in the case of a nonce word).

The participants also filled in a background questionnaire about their current and previous locations of habitation and those of their parents, basic demographic information (education, field of study or work), relevant medical information (the presence of hearing loss or neurological disorders) and various questions about language and music skills. The participants were asked to self-assess their sense of rhythm (5 levels from very bad to very good), musical hearing, language skills (various questions about reading habits, the number of foreign languages they know, etc.), knowledge of Estonian dialects, and assessment of their usual speech rate.

The purpose of this information was to exploratively explain a certain part of the variability in quantity perception. It is known, for example, that musical training fine-tunes the brain's perception of the features important for music (e.g., duration, pitch, see a recent example on language perception among musicians with Estonian and Chinese as their native languages, Lyu et al. 2024). These prosodic features are also important in the Estonian quantity distinctions and may play a role in the current experiment.

2.4. Subjects

A total of 317 participants completed the online perception test: 219 participants in 2020 and 98 participants in 2021. The participants were recruited by sharing the invitation to participate in the experiment in the university email lists and in social media.

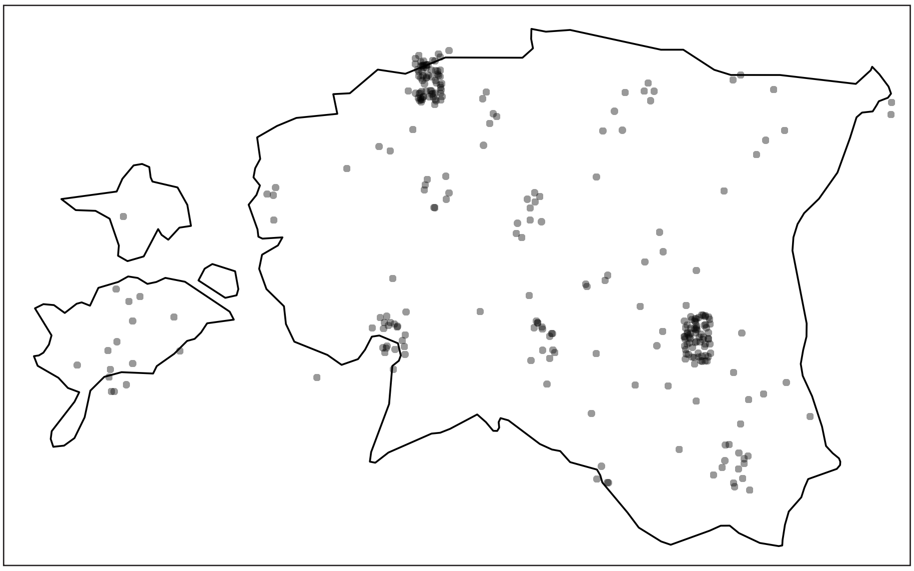

Figure 3. The place of habitation of the participants[1]

[1] The latitude and longitude values of the locations are plotted here and in the following plots with a small random variation providing better visibility of different points with identical values and at the same time providing an additional layer of anonymisation of the participants.

From the current analysis, 12 of the subjects who were not native Estonian speakers were left out. The participants who failed to answer correctly to the control question or did not categorise the control stimuli correctly as Q1 were also left out of the analysis, as they were considered to misinterpret the instructions or not paying attention during the experiment.

The set of participants analysed in this paper were 290 native Estonians: 250 female and 40 male participants. The mean age was 33.9 (ranging from 18 to 77 years). 86 of the participants had obtained secondary education and 16 vocational education, 66 had a BA, 98 an MA, and 21 a PhD degree. Figure 3 maps the reported places of habitation of the participants.

2.5. Data analysis and availability

Data analysis was carried out using R (R Core Team 2022). The Random Forest model was fitted using the *randomForest* package (Liaw, Wiener 2002).

The data is available in the DataDOI repository (Lippus et al. 2025). The repository includes the recordings of the base words with TextGrid annotations, the stimuli and the Praat scripts used for creating the stimuli, the data set with the perception test results and the R code used for analysing the results.

3. Results

Among the 290 participants, there were 23 subjects who rated all stimuli with the same quantity degree that was the quantity of the original base word independent of the manipulations in the temporal domain. These participants most probably based their rating on the pitch cue. In order to evaluate whether the listener was more sensitive to the pitch or to the temporal cues we calculated a pitch vs. duration sensitivity score (henceforth *Pitch Sensitivity Score*) based on the responses to the stimuli in Set 1 and 2. The score was calculated as follows: for each speaker we counted the number of Q3 responses to the stimuli with level pitch and the number of Q2 responses to the stimuli with falling pitch, and then divided the sum of these two by the total number of responses given by the speaker. This score ranges between 0 – 0.694. The score is zero if the participant rated all stimuli with level pitch as Q2 and all stimuli with falling pitch as Q3 regardless of the segmental durations. The larger the score is, the more the participant rated the stimuli based on the temporal cues.

In Set 3 the pitch contours were swapped between Q2 and Q3 within the word pairs and no temporal manipulations were carried out, creating a clear mismatch between the tonal and temporal cues. In this set 90 (36%) of the participants based their decision only on the pitch cue and only 29 (12%) of the participants mainly relied on the temporal cues. The performance in Set 3 was highly correlated with the Pitch Sensitivity Score (which showed the performance in Set 1 and 2): Spearman's $\rho = 0.714$ ($S = 725711$, $p < 0.001$).

In the first part of the results, we will look at the microprosodic effects on the perception of the long vs. overlong category boundary. We will analyse the temporal domain of the stimuli by looking at the results of the participants with the Pitch Sensitivity Score higher than 0.1, meaning that they rated at least 10% of the stimuli with a different quantity degree than would have been expected based solely on the pitch cue.

In the second part of the results, we will attempt to map the Pitch Sensitivity Score with the regional background of the participants. In the third section we look at how the background information of the speakers explains the variability of the Pitch Sensitivity Score.

3.1. Effects of microprosody on the boundary of long and overlong quantity

In this section we look at the category boundary between long (Q2) and overlong (Q3) quantity degrees in V1 duration. The analysis is based on the results of the 137 listeners who showed temporal sensitivity in stimuli sets 1 and 2, receiving the Pitch Sensitivity Score higher than 0.1.

Figure 4 presents the proportion of the Q3 responses to each stimulus that was created by resynthesis manipulating the duration of the stressed vowel in Set 1 and Set 2. As there are only Q2 and Q3 responses in the analysed data, the Q3 response can be interpreted as binary, meaning that if the stimulus is rated <50% as Q3, it is perceived as Q2, and if the Q3 rating is >50%, it is perceived as Q3. The confidence intervals were set to 99%. The grey shading on the graphs shows the area where the confidence interval crosses the 50% line, which is interpreted as an indication of the category boundary.

The category boundary had a different temporal location depending on V1 quality. With level pitch, the boundary for Q2 is earlier if V1 is /i/ and later when V1 is /ɑ/: for *liiga* the Q2 boundary was reached at V1 = 245 ms, *viide* V1 = 272 ms and *viigi* V1 = 275 ms vs. for *saada* and *laadi* it was at V1 = 287 ms and *saage* V1 = 268 ms. In the case of *saagi* it never reached the category boundary and all stimuli with level pitch were perceived as Q2.

With falling pitch, the pattern was less clear, but the category boundary for Q3 was earlier if V1 was /i/ and later if V1 was /ɑ/: for *viide* the Q3 boundary was at V1 = 207 ms, *viigi* V1 = 214 ms and *liiga* V1 = 236 vs. *saada* V1 = 209 ms, *laadi* V1 = 216 ms, *saage* V1 = 223 ms and *saagi* V1 = 234 ms.

The nonce word *taada* was perceived as Q2 with V1 duration up to 283 ms but there was no clear distinction between Q2 and Q3 when V1 was longer regardless of the pitch cue.

3.2. Evaluating the pitch cue and regional background

In this section we map the participants' regional background in the light of the Pitch Sensitivity Score that was calculated based on the responses to the stimuli in Sets 1 and 2. In this section we use the participants' mother's place of habitation as a reference. While the participants' current places of habitation are mostly clustered into the larger towns (mainly Tallinn and Tartu), their mother's side

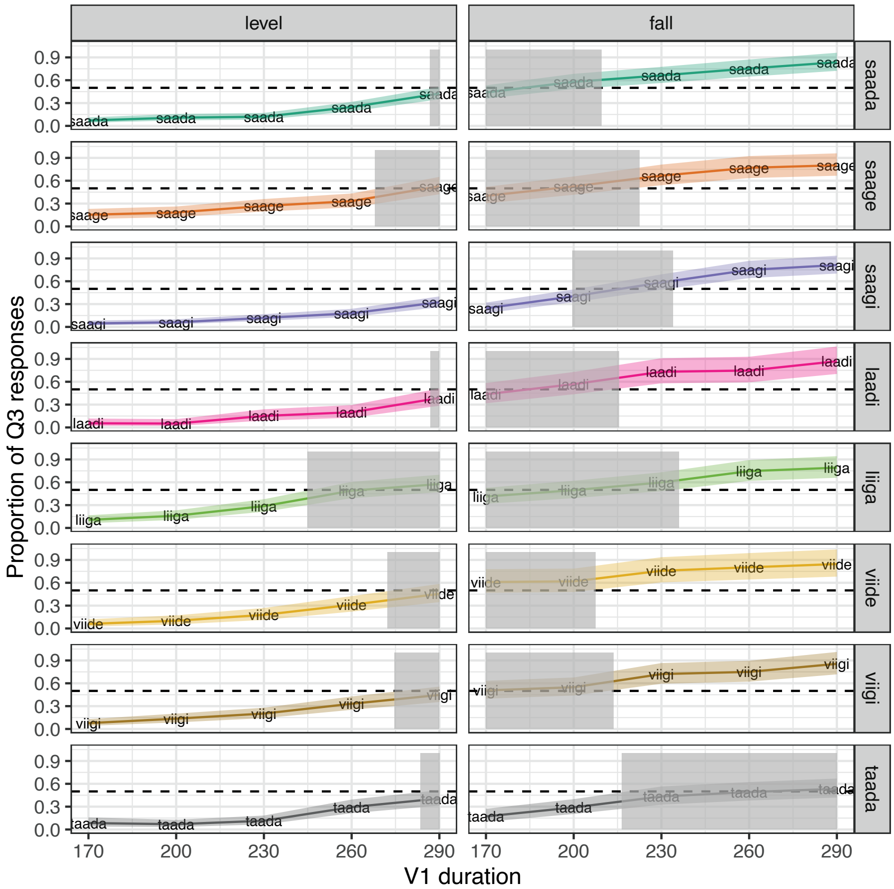

Figure 4. The proportion of Q3 responses to the stimuli with manipulated V1 duration. The panels are grouped by the base word segmental quality (in rows) and pitch contour in the stressed syllable (in columns). The grey shaded squares mark the area where the confidence interval crosses the 0.5 level (indicated with black dashed line)

may be the most relevant point in our data set to reflect the participant's dialectal background.

For 23 of the participants (which is 8% of the total group) the Pitch Sensitivity Score was 0 meaning that they only based their decisions on the pitch cue regardless of the temporal manipulations. Setting the Pitch Sensitivity Score threshold to 0.1 divides the participants into two roughly equal groups. 153 (53%) participants based their decisions mainly on the pitch cue (>90% of their responses) and can be called the "pitch sensitive" group, while 137 (47%) of the participants gave responses that correlated at least 10% with the temporal manipulations and can be called the "duration sensitive" group. Figure 5 plots this grouping as a function of the participant's regional background on the map of Estonia.

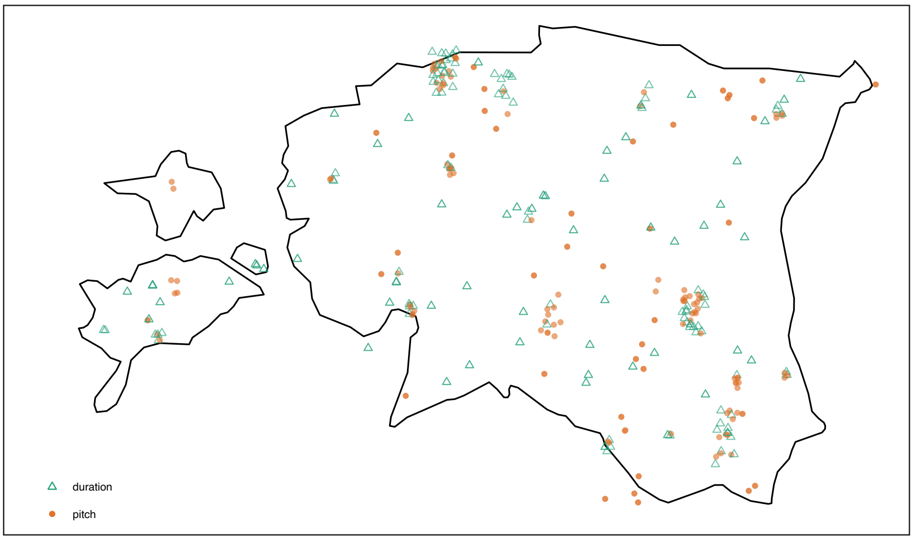

Figure 5. The "pitch sensitive" (marked with filled circles) vs. the "duration sensitive" (marked with triangles) participants mapped based on the mother's place of habitation

3.3. Evaluating the effect of participant's background factors on the pitch sensitivity

To further test the background of the variability in the Pitch Sensitivity Score, we tested it by fitting a Random Forest model with the self-reported variables from the questionnaire. The following variables were included as independent factors: participant's age and gender, education, mother's place of habitation (*mother_lat* and *mother_lon* referring to the latitude and longitude of the geographic point), sense of rhythm, musical hearing, language skills, being a speaker of some specific dialectal variety of Estonian, and being perceived as a fast or slow speaker by self (*speechrate_self*) and others (*speechrate*).

The fitted Random Forest model explains 13% of the variance of the Pitch Sensitivity Score. Figure 6 presents the mean decrease accuracy of the tested factors, evaluating each factor's contribution to the model.

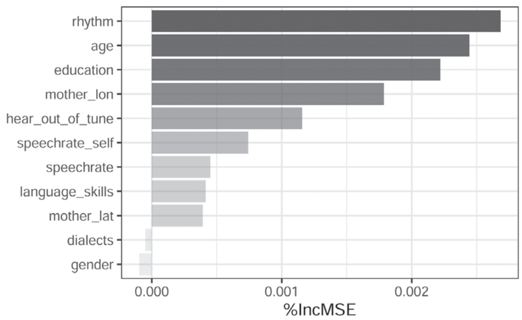

Figure 6. The Mean Decrease Accuracy of the tested factors in the Random Forest model evaluating the participants' Pitch Sensitivity Score

According to the Random Forest model, the strongest predictors for explaining the participant's Pitch Sensitivity Score are the self-assessed sense of rhythm, age, education and the longitude of the mother's place of habitation (i.e. the location in the East-West dimension). Less important factors are the self-assessed speech rate, language skills and the latitude of the mother's place of habitation (i.e. the North-South dimension). The least significant factors are the knowledge of dialects and the participant's gender.

4. Discussion

The first aim of this study was to test the interaction between prosody and microprosody in the Estonian quantity system. The results show that the boundary between the long and overlong quantity degrees is perceived at a shorter duration if the vowel in the stressed syllable is a closed vowel [i] and at a longer duration if the vowel in the stressed syllable is an open vowel [ɑ]. The effect of intrinsic duration on the foot level quantity has been observed in acoustic studies (Lippus, Šimko 2015) and the current study confirms that this is also accounted for in perception.

The results also show that it was difficult to categorise the stimuli created from nonce words. While the segmentally similar stimuli from the meaningful word pair *saada* were perceived as Q2 with level pitch and shorter duration of V1 and as Q3 with the falling pitch and longer duration of V1, the stimuli from the nonce words *taada* with falling pitch were categorised virtually randomly. The result may be surprising if we expect the quantity distinction of disyllabic speech units to be a generic task that is automatically applied to new content. However, it is commonly known that processing more frequent words is faster and more precise than the processing of less frequent words (Balota, Chumbley 1984, Brysbaert et al. 2018, Lõo et al. 2018). The difficulty of categorising nonce words may be also linked with the results from Piits and Mihkla (2024) showing that it may be difficult to categorise the quantity of words where the quantity alteration does not change the meaning (while they were also testing stimuli with category-straddling acoustic features, which was probably the main reason for the difficulty of categorisation in their study).

The second aim was to test the regional variability of the listeners' preference of using pitch vs. duration cues in Q2 vs. Q3 judgements. The previous study by Lippus and Pajusalu (2009) suggested that listeners from North and West Estonia strongly depend on the early falling pitch cue for Q3 perception, while participants from Eastern and Southern Estonia make the quantity judgements mainly based on temporal cues. The division in Lippus and Pajusalu's study was not absolute and the line separating the 38 participants into two groups was rather creatively placed on the major dialect group borders. The current study hoped to engage a larger group of participants, covering the area more densely and allowing to map the border more precisely.

Contrary to our expectations, mapping the listeners based on their own and their parents' place of habitation shows no clear patterns related to the pitch vs. duration perception. The majority of the participants based their decisions more on the pitch cue than we would have predicted when designing the stimuli from the

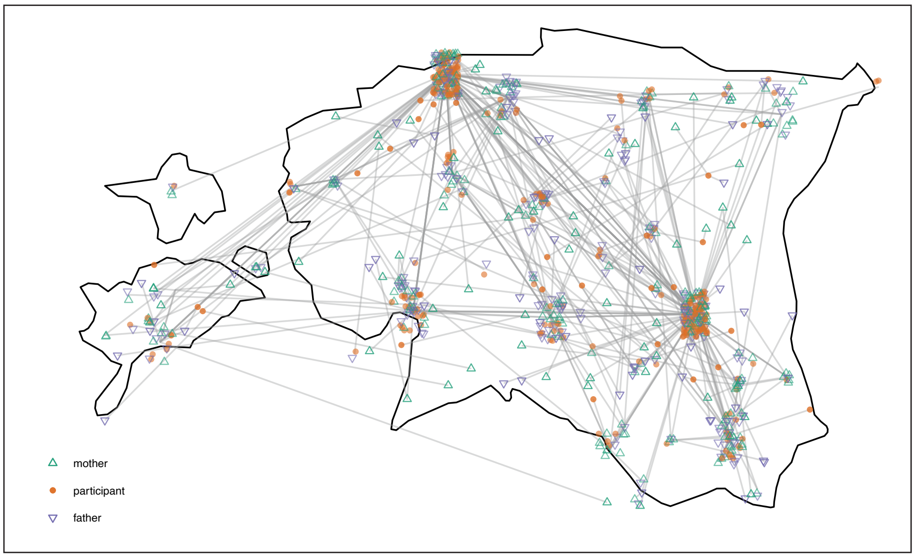

Figure 7. The place of habitation of the participants connected with the place of habitation of their mother and father

knowledge from the previous studies (Lippus et al. 2009, Lippus, Pajusalu 2009). Mapping the results by the participants' or their parents' place of habitation did not show a pattern that could be directly related with their Pitch Sensitivity Score.

A possible reason for this is the migration within Estonia that has taken place in the Estonian population during the past centuries: not only the participants of this study but also their parents have moved from their birthplace to a different town to study and then again to a different location to work. While it was difficult to identify a stable place of habitation for most of the participants, linking their place of habitation with those of their parents (see Figure 7) shows firstly migration to the major towns and secondly that the two parents are often from very different areas of Estonia.

The Random Forest model showed that the participants' background factors explained only 13% of the variability in the performance in the perception experiment. Among the studied features the mother's place of habitation was still among the five more significant factors, while the strongest factor was the listener's sense of rhythm, followed by age and education level, and the fifth was the participant's musical skills (namely the person's ability to hear if the voice is out of tune).

As a limitation of this study, having limited repetitions of the stimuli in the experiment presented some challenges, especially when we were interested in individual category boundaries and identifying random responses. Currently, it was not possible to get a better overview of individual perception of the used stimuli. With only a few data points per person, it was difficult to distinguish random responses from genuine variations in their categorization. Also, it is possible that a participant might have answered differently due to tiredness or distraction (especially in online experiments), and this variability may have been amplified with fewer repetitions.

The previous results that showed dialectal background effects had a somewhat different design: forced choice with no repeated listening option. In the current study the online perception test environment allowed repeated listening. Forcing the participants to make their decision based on the first impression and looking at the reaction times may provide more detailed data on the category boundaries and individual weights of the different cues.

5. Conclusions

This paper analysed the results of an online perception study with participants of various Estonian regional backgrounds. Analysing the microprosodic effects showed that the long vs. overlong category boundary is located differently depending on the segmental quality of the stressed vowel. The majority of listeners depend strongly on the pitch cue, while the preference of relying more on pitch or temporal cues showed no clear patterns in relation to the participants' regional background.

References

Asu, Eva Liina; Lippus, Pärtel; Pajusalu, Karl; Teras, Pire 2016. Eesti keele hääldus. Eesti keele varamu, 2. Tartu: Tartu Ülikooli kirjastus. http://hdl.handle.net/10062/57960

Balota, David A.; Chumbley, James I. 1984. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. – Journal of Experimental Psychology: Human Perception and Performance, 10 (3), 340–357. https://doi.org/10.1037/0096-1523.10.3.340

Boersma, Paul; Weenink, David 2019. Praat: Doing Phonetics by Computer. http://www.praat.org

Brysbaert, Marc; Mandera, Paweł; Keuleers, Emmanuel 2018. The word frequency effect in word processing: An updated review. – Current Directions in Psychological Science, 27 (1), 45–50. https://doi.org/10.1177/0963721417727521

Draxler, Christoph; Jänsch, Klaus 2004. Speechrecorder: An universal platform independent multi-channel audio recording software. – Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa, Raquel Silva (Eds.), Proceedings of 4th International Conference on Language Resources and Evaluation (LREC'04). Lisbon: European Language Resources Association, 559–562.

Eek, Arvo 1980a. Estonian quantity: Notes on the perception of duration. – Arvo Eek (Ed.), Estonian Papers in Phonetics 1979. Tallinn: Academy of Sciences of the Estonian S.S.R. Institute of Language and Literature, 5–30.

Eek, Arvo 1980b. Further information on the perception of Estonian quantity. – Arvo Eek (Ed.), Estonian Papers in Phonetics 1979. Tallinn: Academy of Sciences of the Estonian S.S.R. Institute of Language and Literature, 31–57.

Eek, Arvo; Meister, Einar 2003. Foneetilisi katseid ja arutlusi kvantiteedi alalt (I): Häälikukestusi muutvad kontekstid ja välde. – Keel ja Kirjandus, 46 (11), 815–837.

Kalvik, Mari-Liis 2005. Välted muutuvas rannikumurdes. – Keel ja Kirjandus, 48 (3), 209–222.

Kask, Liis; Põldver, Nele; Lippus Pärtel; Kreegipuu, Kairi 2021. Perceptual asymmetries and auditory processing of Estonian quantities. – Frontiers in Human Neuroscience, 15, 194. https://doi.org/10.3389/fnhum.2021.612617

Krull, Diana; Traunmüller, Hartmut; van Dommelen, Wim A. 2003. The effect of local speaking rate on perceived quantity: A comparison between three languages. – Maria-Josep

Solé, Daniel Recasens, Joaqúin Romero (Eds.), Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain, 3–9 August 2003. Barcelona, 1739–1742.

Lehiste, Ilse 1997. Search for phonetic correlates in Estonian prosody. – Ilse Lehiste, Jaan Ross (Eds.), Estonian Prosody: Papers from a Symposium. Tallinn: Institute of Estonian Language, 11–35.

Liaw, Andy; Wiener, Matthew 2002. Classification and regression by randomForest. – R News, 2 (3), 18–22.

Lippus, Pärtel; Karl Pajusalu 2009. Regional variation in the perception of Estonian quantity. – Martti Vainio, Reijo Aulanko, Olli Aaltonen (Eds.), Nordic Prosody. Proceedings of the Xth Conference, Helsinki 2008. Frankfurt, Berlin, New York: Peter Lang, 151–157.

Lippus, Pärtel; Pajusalu, Karl; Allik, Jüri 2009. The tonal component of Estonian quantity in native and non-native perception. – Journal of Phonetics, 37 (4), 388–396. https://doi.org/10.1016/j.wocn.2009.07.002

Lippus, Pärtel; Pajusalu, Karl; Allik, Jüri 2011. The role of pitch cue in the perception of the Estonian long quantity. – Sónia Frota, Gorka Elordieta, Pilar Prieto (Eds.), Prosodic Categories: Production, Perception and Comprehension. Dordrecht: Springer Netherlands, 231–242. https://doi.org/10.1007/978-94-007-0137-3_10

Lippus, Pärtel; Ross, Jaan 2011. Has Estonian quantity system changed in a century? Comparison of historical and contemporary data. – Wai-Sum Lee, Eric Zee (Eds.), Proceedings of the 17th International Congress of Phonetic Sciences. Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong, 1262–1265. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Lippus/Lippus.pdf (31.3.2025).

Lippus, Pärtel; Šimko, Juraj 2015. Segmental context effects on temporal realization of Estonian quantity. – Maria Wolters, Judy Livingstone, Bernie Beattie, Rachel Smith, Mike MacMahon, Jane Stuart-Smith, Jim Scobbie (Eds.), Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow: University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0318.pdf (31.3.2025).

Lippus, Pärtel; Kask, Liis; Lutter, Sofia; Põldver, Nele; Kreegipuu, Kairi 2025. Mapping boundaries: Microprosodic and dialectal variability in the perception of Estonian long and overlong quantity (DATA). DataDOI. https://doi.org/10.23673/re-497

Lõo, Kaidi; Järvikivi, Juhani; Baayen, R. Harald 2018. Whole-word frequency and inflectional paradigm size facilitate Estonian case-inflected noun processing. – Cognition, 175, 20–25. https://doi.org/10.1016/j.cognition.2018.02.002

Lyu, Siqi; Põldver, Nele; Kask, Liis; Wang, Luming; Kreegipuu, Kairi 2024. Effect of musical expertise on the perception of duration and pitch in language: A cross-linguistic study. – Acta Psychologica 244, a104195. https://doi.org/10.1016/j.actpsy.2024.104195

Meister, Einar; Werner, Stefan 2009. Duration affects vowel perception in Estonian and Finnish. – Linguistica Uralica, 45 (3), 161–177. https://doi.org/10.3176/lu.2009.3.01

Meister, Einar; Werner, Stefan; Meister, Lya 2011. Short vs. long category perception affected by vowel quality. – Wai-Sum Lee, Eric Zee (Eds.), Proceedings of the 17th International Congress of Phonetic Sciences. Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong, 1362–1365.

Parve, Merike 2003. Setu välted. – Karl Pajusalu, Pire Teras (Eds.), Lõunaeesti häälikud II. Publications of the Department of Estonain of the University of Tartu, 27. Tartu, 17–47. http://hdl.handle.net/10062/55279

Piits, Liisi; Mihkla, Meelis 2024. Varieeruva vältega sõnade taju ja akustika. – Emakeele Seltsi aastaraamat, 69, 178–203. https://doi.org/10.3176/esa69.08

R Core Team 2022. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/

PIIRIDE KAARDISTAMINE: MIKROPROSOODIA JA MURDETAUSTA VARIEERUMINE EESTI PIKA JA ÜLIPIKA VÄLTE TAJUMISEL

Pärtel Lippus, Liis Kask, Sofia Lutter, Nele Põldver, Kairi Kreegipuu
University of Tartu

Käesolev artikkel esitab tulemusi suuremast veebipõhisest tajukatsest, millega uuriti teise ja kolmanda välte eristamist eesti keeles. Uurimusel on kaks peamist huvipunkti. Esiteks soovime uurida häälikukvaliteedi efekti vältetajule. Enamik varasemaid vältetaju katseid on kasutanud mõnda üksikut minimaalkolmikut, tüüpiliselt on selleks olnud [sɑtɑ] – [sɑːtɑ] – [sɑːːtɑ]. Samas mõjutab hääliku omakestus oluliselt selle pikkuse tajumist ning väldet iseloomustavad kestussuhted võivad olulisel määral varieeruda sõna häälikulise koostise kombinatsioonist lähtuvalt. Siin uurimuses genereeriti stiimulid kaheksast erinevast häälikujärjendist. Stiimulid resünteesiti manipuleerides vokaalide kestust ja põhitoonikontuuri.

Uurimuse teine eesmärk oli kaardistada vältetaju võimalikku murdetaustast tingitud varieerumist. Varasemad uurimused on näidanud, et Põhja- ja Lääne-Eesti murdetaustaga katseisikud on toonitundlikumad võrreldes Ida- ja Lõuna-Eesti taustaga katseisikutega, kes tuginevad vältehinnangutes rohkem kestuse varieerumisele. Käesolevas uuringus osales 290 eesti emakeelega katseisikut üle kogu Eesti.

Tulemused näitasid, et esisilbis kõrge vokaaliga stiimuliseeriates oli teise ja kolmanda välte piir varasem kui madala vokaaliga stiimuliseeriates. Samuti oli tähenduseta sõnade välte kategoriseerimise täpsus oluliselt madalam kui tähendusega sõnade puhul. Vastu ootusi ei õnnestunud leida selget seost toonitundlikuse ja osalejate murdetausta vahel.

Märksõnad: välde, kestus, põhitoon, mikroprosoodia, taju, dialektoloogiline varieerumine, eesti keel

Pärtel Lippus (University of Tartu) is the Professor of Estonian Phonetics at the Institute of Estonian and General Linguistics. His main research area is Estonian prosody, namely the quantity system, but also stress and prominence, intonation, and sociophonetic variation of Estonian.
Jakobi 2, 51005 Tartu, Estonia
partel.lippus@ut.ee

Liis Kask (University of Tartu) is a Junior Research Fellow in Experimental Psychology at the Institute of Psychology. Her research involves language and auditory information proccessing and brain mechanisms in language acquisition.
Näituse 2, 50409 Tartu, Estonia
liis.kask@ut.ee

Sofia Lutter (University of Tartu) is a MA student at the Institute of Psychology. In her master's research she explores bilinguals' perception of Estonian quantity.
Näituse 2, 50409 Tartu, Estonia
sofia.lutter@ut.ee

Nele Põldver (University of Tartu) is a Research Fellow of Experimental Psychology at the Institute of Psychology. Her research interests include visual and auditory perception, including speech and music perception, and their underlying processes in the brain.
Näituse 2, 50409 Tartu, Estonia
nele.poldver@ut.ee

Kairi Kreegipuu (University of Tartu) is the Professor of Experimental Psychology at the Institute of Psychology. Her interest is in cognitive information processing and its relation to states (including emotions and fatigue), as well as the study of markers and correlates of cognition.
Näituse 2, 50409 Tartu, Estonia
kairi.kreegipuu@ut.ee