Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Quantity contrast in Inari Saami: the role of pitch and intensity

Helen Türk, Pärtel Lippus, Karl Pajusalu, Pire Teras

Institute of Estonian and General Linguistics, University of Tartu, Estonia
{helen.turk, partel.lippus, karl.pajusalu, pire.teras} @ut.ee

Abstract

The paper investigates the variation of pitch and intensity as a feature of the quantity contrast in Inari Saami. Previous studies [1, 2] have shown that in the context of short vowels, an intervocalic consonant has three degrees of length, while after a long vowel the distinction is binary, and that there is also a compensatory lengthening mechanism between the segments within the foot. The research here is based on the same data that was used for studying the temporal characteristics of quantity. We analyze fundamental frequency of the whole disyllabic foot as well as the intensity of the intervocalic consonant and surrounding vowels. In other Finno-Ugric languages such as Estonian, Livonian and Lule Saami, these characteristics have been shown to be important in the realization of the quantity contrast. Our results indicate that in Inari Saami the interaction between the overall falling pitch contour and quantity is not as strong as it is in Estonian. However, in words with long geminates the pitch fall is shallower than in other word structures. Similar to Lule Saami, in the stressed first vowel the intensity increases with consonant quantity, while in the unstressed second vowel it decreases.

Index Terms: ternary quantity contrast, Inari Saami, word prosody, fundamental frequency, intensity

1. Introduction

Inari Saami is a Finno-Ugric language that is spoken around Lake Inari in Nothern Finland by approximately 250 mostly elderly native speakers [3]. Its sound system has a relatively rare phonological feature – the three-way quantity contrast in consonant duration, which is realized by the distinction of single consonants (Q1), half-long (Q2) and long (Q3) geminates or consonant clusters in the intervocalic position in left-headed disyllabic feet [1, 4]. For vowels the distinction is binary [5], i.e. in stressed syllables there are both short and long vowels, but in unstressed syllables there is no phonological length distinction [6]. In an initial syllable Inari Saami has also short and long diphthongs [5, 7].

Only a few studies about quantity in Inari Saami have been carried out and they focus on temporal aspects, e.g., [2, 8]. In our recent study [2] it was revealed that in disyllabic words an intervocalic consonant preceded by a phonologically short vowel has three degrees of length: short, half-long or long. Following a long vowel the consonant is short or half-long. Our results also showed that the durations of segments in disyllabic feet are interrelated indicating that quantity in Inari Saami is a foot-level phenomenon. With the increase in consonant duration after a short first syllable vowel the durations of surrounding vowels decrease. The second syllable vowel is also shortened in words where there is a long first syllable vowel followed by a short and a half-long consonant.

The three-way quantity distinction is more widely known and studied in Estonian, a Finno-Ugric language that has a complicated quantity system showing a ternary opposition for both vowels and consonants. The domain of quantity in Estonian is a disyllabic foot and while duration has been found to be the main correlate of quantity, other prosodic features, such as pitch and intensity have also been shown to be important. It is claimed that in Q1 and Q2 words the peak of the pitch contour is in the end of the first syllable, in Q3 it falls to the beginning of the first syllable [9, 10, 11]. The intensity movement of whispered speech roughly follows the F0 contour [12].

According to Suomi et al. [13] the fundamental frequency contour in Finnish is similar in different word structures, and for keeping this uniformity the durations of segments are adjusted. Tuisk [14] has also found the same patterns in Livonian.

In Saami languages the nature of non-temporal prosodic variables is less studied. McRobbie Utasi [15] has found that in Skolt Saami quantity also involves pitch, stress and intensity. Fangel-Gustavson et al. [16] show that in Lule Saami quantity affects vowel intensity but not consonant intensity, and that there is no significant effect of tonal component. In Inari Saami fundamental frequency and intensity have not been studied before. Thus, the aim of this research is to investigate tonal patterns of quantity in disyllabic foot structures. While temporal characteristics of quantity in Inari Saami are somewhat similar to Estonian (cf. [1, 2, 4]) we would expect some similarity in pitch movement as well. Secondly, we study the intensity of stressed and unstressed vowels and intervocalic consonants in relation to quantity.

2. Materials and method

In order to expand on the previous research of Inari Saami quantity contrast, for the current study we used the same data as in [2] where the temporal characteristics were discussed. The data comes from four male native speakers (aged 62, 68, 76, 77). It was collected in 2013 in Inari area using an Edirol R-09 digital recorder at 48 kHz. The subjects were asked to read sentences that contained test words which occurred in the phrase-medial (e.g., *Ohtâ **maņe** lii taa, mut ohtâ **lodde** lii tobbeen* 'Here is one **egg**, but there is one **bird**') or phrase-final position (e.g., *Mist láá kuttâ **sare**, mut sist láá kuttâ **juuŋâ*** 'We have six **blueberries**, but they have six **lingonberries**').

This paper focuses on disyllabic words with a phonologically short and long vowel in the first syllable followed by a short consonant, or a half-long or long geminate. Unstressed syllables contained a phonologically short vowel and were open or closed by a voiced consonant. Thus, the analyzed word structures were CVCV, CVCCV, CVC:CV(C), CVVCV(C), CVVCCV(C). Following are some examples of test words for each word structure: 1) CVCV – a

 doi: 10.21437/SpeechProsody.2016-224

short vowel followed by a short consonant (Q1), e.g. *sare* 'blueberry, acc/gen.sg', *kye'le* 'fish, acc/gen.sg'; 2) CVCCV – a short vowel followed by a half-long geminate (Q2) (in the orthography it is marked with a dot under a single letter), e.g. *saṛe* 'blueberry, nom.sg'; 3) CVC:CV(C) – a short vowel followed by a long geminate (Q3), e.g. *pállu* 'ball, nom.sg', *uáb'bi* 'sister, nom.sg'; 4) CVVCV(C) – a long vowel followed by a short consonant (Q1), e.g. *määli* 'soup, acc/gen.sg', *muorâ* 'tree, acc/gen.sg'; 5) CVVCCV(C) – a long vowel followed by a half-long geminate (Q2), e.g. *määḷi* 'soup, nom.sg', *muoṛâ* 'tree, nom.sg'. In total 554 words were used for this study (see Table 1).

Table 1. *The distribution of analyzed tokens.*

Foot structure	Position	
	Phrase-medial	Phrase-final
CVCV	16	13
CVCCV	31	31
CVC:CV(C)	108	66
CVVCV(C)	60	51
CVVCCV(C)	60	118

Using Praat [17] the data were segmented and annotated. A Praat script was used to obtain the durations of all segments of the test words, and values for fundamental frequency were retrieved from 20 equally spaced points from the beginning of the vowel (V1) in the first syllable to the end of the vowel (V2) in the second syllable enabling to get pitch contours which were interpolated and smoothed. For each test word the pitch contours were manually checked for discrepancies and necessary corrections were made. The values measured in Hertz were converted into semitones (st) in reference to 50 Hz. The fundamental frequency turning point was found by calculating the time (in milliseconds) from the beginning of V1 to the next point after F0 maximum where the F0 was falling. Also, the linear regression slope of the fundamental frequency in relation to the turning point in V1 and time normalized in relation to the whole word was found. For intensity the mean values of V1 and V2 and the intervocalic consonant were extracted in decibels.

Statistical analysis was carried out using LME4 package in R [18]. F0 turning point and the difference between V1 maximum and V2 minimum as well as the mean values of V1, V2 and C2 intensity were tested with a mixed effects model for three factors: 1) the length of an intervocalic consonant (C2 Length, levels: short, half-long, long), 2) the length of the first syllable vowel (V1 Length, levels: short, long), 3) position in a sentence (Position, levels: phrase-medial, phrase-final). The speaker and the word were considered as random factors.

3. Results

3.1. Fundamental frequency

The results of the fundamental frequency analysis are summarized in Table 1. Figure 1 shows the F0 contours in two sentence positions (phrase-medial and phrase-final) and in word structures with a short (V) and long (VV) first syllable vowel.

The fundamental frequency turning point is shown with a vertical line in Figure 1. As it can be seen, there are no great differences between word structures. The main effects of V1 Length and C2 Length on the turning point are not statistically significant, but a mixed effects model shows a significant effect of Position [$\chi^2(df=8, N=554)=11.87, p<0.001$]. In the phrase-medial position the turning point is located earlier (on average, at 86 ms with a short V1 and at 87 ms with a long V1) than in the phrase-final position (at 108 ms with a short V1 and at 106 ms with a long V1).

The overall fundamental frequency range, i.e. the difference between the F0 maximum of V1 and F0 minimum of V2, is also similar in different word structures, but again, the main effect of Position is statistically significant [$\chi^2(df=8, N=554)=31.78, p<0.001$]. The range is smaller in the phrase-medial position (with a short V1 it is 5.3 st and with a long V1 5.2 st) than in the phrase-final position (6.4 st with a short V1 and 6.6 st with a long V1).

Table 2. *Average values of F0 turning point (TP, time from V1 beginning in milliseconds), maximum F0 of the first (V1 F0 max) and minimum F0 of the second (V2 F0 min) syllable vowel in semitones.*

Pos.	Foot structure	TP	V1 F0 max	V2 F0 min
Phrase-medial	CVCV	73	17.5	11.8
	CVCCV	94	17.7	12.0
	CVC:CV(C)	85	17.7	12.6
	CVVCV(C)	77	17.1	12.0
	CVVCCV(C)	97	17.7	12.5
Phrase-final	CVCV	103	17.1	10.5
	CVCCV	109	18.8	11.9
	CVC:CV(C)	109	18.6	12.4
	CVVCV(C)	105	17.8	11.8
	CVVCCV(C)	107	18.7	11.9

In order to describe the fall of fundamental frequency during the word the slope of fundamental frequency was calculated by fitting a linear regression line on time-normalized pitch measurements. The F0 measurements taken from 20 equally spaced points during each word were normalized in reference to the F0 maximum and the time was normalized in reference to the duration from the turning point to the end of a word.

The results show significant main effects of C2 Length [$\chi^2(df=7, N=554)=7.62, p<0.05$] and Position [$\chi^2(df=7, N=554)=9.86, p<0.01$] on the slope, but there is no interaction. Post-hoc testing confirms the difference between Q1 and Q2 vs. Q3 consonant ($p<0.005$). In the words with a short V1 the slope is decreasing with the increasing quantity of the intervocalic consonant. The slope of F0 is the biggest with a short consonant (-5.4 st in the phrase-medial and -6.1 st in the phrase-final position), it is smaller in the words with a half-long geminate (-4.9 st in the phrase-medial and -5.8 st in the phrase-final position) and the smallest in words with a long geminate (-4.4 st in the phrase-medial and -5.0 st in the phrase-final position). Similar tendencies can be seen in the phrase-medial words with a long V1, where the slope of F0 is bigger with a short consonant (-4.9 st) than it is with a half-long geminate (-4.7 st). In the phrase-final words with a long V1 the situation is vice versa: the slope is smaller in the words with a short consonant (-5.1 st) than with a half-long geminate (-6.0 st).

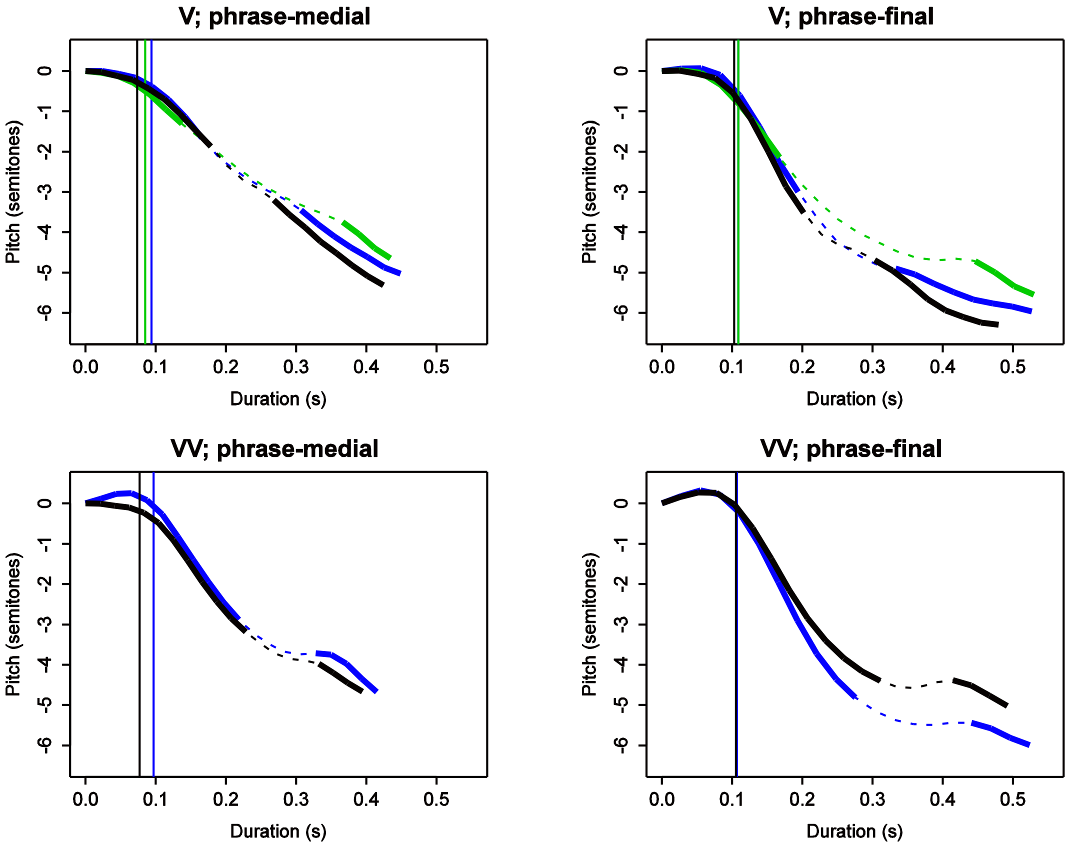

Figure 1. *Contours of fundamental frequency (in semitones) in two positions of a test sentence (phrase-medial and phrase-final) and in word structures with a short (V) and long (VV) first syllable vowel. Black contour marks the foot structures with a short intervocalic consonant, blue is for a half-long and green for a long geminate. Lines in bold denote the durations of first and second syllable vowels and dashed line shows the duration of intervocalic consonant. Vertical line is for the F0 turning point.*

Thus, it can be concluded that in words with a long (Q3) geminate the slope of fundamental frequency is smaller than with a short (Q1) consonant and a half-long (Q2) geminate. This shows that the overall pitch fall is similarly steep with Q1 consonant and Q2 geminate, but considerably shallower with a Q3 geminate. Position-wise the slope in the phrase-final position is steeper than in the phrase-medial position.

3.2. Intensity

The results of intensity analysis are presented in Table 3. In general, the average intensity of V1 is higher than the intensity of V2. There is a main effect of C2 Length [$\chi^2(df=7, N=554)=10.20$, $p<0.01$] on V2 intensity, which is decreasing with quantity. Post-hoc testing showed the difference in the case of Q1 and Q2 vs. Q3 consonant ($p<0.05$). Also, V2 intensity is significantly lower in the phrase-final position [$\chi^2(df=7, N=554)=15.11$, $p<0.001$]. The intensity of an intervocalic consonant is marginally affected by Position [$\chi^2(df=7, N=554)=5.96$, $p<0.05$], but not by C2 Length. In the phrase-medial position the average intensity of the consonant is higher (57.8 dB) than in the phrase-final position (57.1 dB).

For observing the intensity differences between the first and the second vowel, the difference between the average intensity of V1 and V2 was calculated by subtracting the intensity of V1 from the intensity of V2. In Figure 2 the results are shown comparatively in the two sentence positions (phrase-medial, phrase-final). There are significant main effects of Position [$\chi^2(df=7, N=554)=21.65$, $p<0.001$] and C2 Length [$\chi^2(df=7, N=554)=18.10$, $p<0.001$] on the difference of the intensity between V1 and V2. Post-hoc testing showed a significant difference between Q1 and Q3 consonant ($p<0.001$) and Q2 and Q3 consonant ($p<0.05$).

Table 3. *The average intensity (in dB) of the first (V1) and the second (V2) syllable vowel and of an intervocalic consonant (C2).*

Pos.	Foot structure	V1 intensity	V2 intensity	C2 intensity
Phrase-medial	CVCV	65.6	61.5	60.1
	CVCCV	66.3	60.4	58.7
	CVC:CV(C)	66.1	58.7	58.3
	CVVCV(C)	65.3	59.3	58.2
	CVVCCV(C)	66.1	60.2	58.1
Phrase-final	CVCV	65.3	58.4	57.8
	CVCCV	66.7	59.3	59.0
	CVC:CV(C)	66.1	56.3	57.8
	CVVCV(C)	65.2	59.0	57.8
	CVVCCV(C)	66.5	58.9	57.3

In the case of a short V1 in phrase-medial words the intensity difference between V1 and V2 is the smallest with a short consonant (the difference is 4.1 dB), it is bigger in the case of a half-long geminate (5.9 dB) and the biggest when there is a long geminate (7.4 dB). In the phrase-final position similar tendencies can be found as the intensity difference is increasing with the intervocalic consonant's quantity: with a short consonant 6.9 dB, 7.4 dB with a half-long and 9.8 dB with a long geminate.

In the words with a long V1 the difference between V1 and V2 intensity in the phrase-medial position is 5.7 dB in the case of a short consonant and 5.9 dB in the case of a half-long geminate. In the phrase-final position the differences are 6.2 dB and 7.6 dB, respectively. Thus, the longer the consonant is after both short and long vowel, the bigger is the difference between V1 and V2 intensity.

Also, analogically the difference between V1 and C2 intensity was calculated. It was revealed that the difference is increasing with the increasing consonant quantity (Q1 vs. Q3, $p<0.001$; Q2 vs. Q3, $p<0.05$). There is also a main effect of Position [$\chi^2(df=7, N=554)=8.76$, $p<0.01$]. The difference is smaller in phrase-medial position than in phrase-final position.

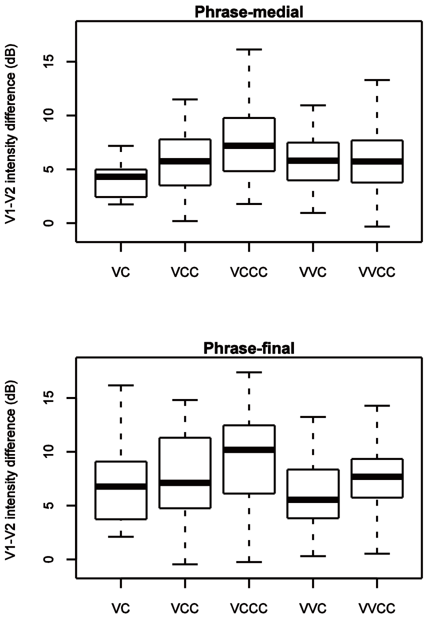

Figure 2. *The difference between the average intensity of V1 and V2 in the analyzed foot structures.*

4. Discussion

For fundamental frequency our results showed that the turning point of F0 occurs roughly at the same location in all of the analyzed word structures. Similarly, the range of fundamental frequency during the whole word is not significantly varying depending on the quantity of a word. Our previous studies about durational characteristics of quantity in Inari Saami [1, 2] have indicated that in order to keep the feet the same length there is a compensatory relation between segment durations. This implies that fundamental frequency has a secondary role in the realization of quantity contrast in Inari Saami and it is connected to the primary durational contrasts. These results can be compared to the findings about Finnish [13] where irrespective of a different syllabic structure of a word there is no variation in the movement of F0. Similar tendencies have been found in Livonian [14] which is also a Finno-Ugric language where the three-way quantity system exists. In addition, in Lule Saami [16] the small differences in the F0 movement are claimed to be less important in the realization of the ternary quantity contrast.

We also studied the slope of F0 that describes the overall fall of the pitch movement. In the word structures with a Q1 and Q2 consonant the fall is steeper than in the words with a Q3 consonant. In the case of the latter it was found that the fundamental frequency in the second syllable is clearly higher. This is different from Estonian where the F0 has fallen low in the second syllable of Q3 words [9]. Therefore, this suggests that in Inari Saami the unstressed vowels of words with a Q3 consonant may not be reduced as much as in Estonian.

Regarding temporal characteristics, Inari Saami is similar to Estonian where there is a three-way quantity opposition distinguishing between the meanings of words, but in the sense of tonal features it is different. However, similarities can be found with other Finno-Ugric languages which have three quantity degrees, such as Lule Saami and Livonian.

Our results about intensity show a correlation between quantity and the intensity of V1 and V2, but there is no impact of quantity on the intensity of an intervocalic consonant. These findings are in line with the results of Fangel-Gustavson et al. [16], who studied the intensity in Lule Saami. The intensity patterns of vowels in Inari Saami are roughly the same as in Lule Saami where the intensity of the first syllable vowel increases with consonant quantity and in the second vowel it decreases. The fact that the intensity of an intervocalic consonant which carries the quantity contrast is kept unaffected shows that intensity could also be considered as a secondary feature in the realization of quantity.

5. Conclusions

In this paper we studied the fundamental frequency and intensity characteristics of quantity contrast realized in different disyllabic foot structures of Inari Saami. Our results show that regarding to fundamental frequency there is no strong interaction between quantity and the overall falling pitch. Albeit, in the case of words with a long geminate in the intervocalic position the F0 fall is shallower compared to words with a short consonant and a half-long geminate, that have a steeper fall in V1. Similar to Lule Saami and Livonian, while the durations of segments in the different disyllabic foot structures are adjusted, the tonal movement stays the same. This shows that fundamental frequency has a secondary role in the realization of the quantity contrast. A connection between the intensity of vowels and quantity was found. It was revealed that with the increasing quantity the intensity in the stressed syllable is also increasing, but in the unstressed syllable it is decreasing. The intensity of an intervocalic consonant is not affected by the quantity. Our results suggest that fundamental frequency and intensity as the secondary features are related to the primary temporal characteristics of quantity contrast, but their phonological significance is not yet clear. Therefore, further research could be undertaken to expand on this issue, for example, by also carrying out a perception test.

6. Acknowledgements

We would like to thank all our Inari Saami speakers. We are also grateful to Eva Liina Asu for proofreading the paper. This study was supported by the Estonian Research Agency grant No. IUT2-37.

7. References

[1] H. Türk, P. Lippus, K. Pajusalu, P. Teras, "The ternary contrast of consonant duration in Inari Saami," in *Proceedings of the 7th Speech Prosody, May 20–23, Dublin, Ireland*, pp. 361–364, 2014.

[2] H. Türk, P. Lippus, K. Pajusalu, P. Teras, "Temporal patterns of quantity in Inari Saami," in *The Scottish Consortium for ICPhS 2015 – Proceedings of the 18th International Congress of Phonetic Sciences, August 10–14, Glasgow, UK,* 2015.

[3] M.M.J. Fernandez-Vest, *Sami: an introduction to the language and culture*. Helsinki: Finn Lectura, 2012.

[4] E. Markus, P. Lippus, K. Pajusalu, P. Teras, "Three-way opposition of consonant quantity in Finnic and Saamic languages," in *Nordic Prosody – Proceedings of the XIth conference, August 15–17, Tartu, Estonia*, pp. 225–234, 2012.

[5] P. Sammallahti, "New Developments in Inari Lappish Phonology," in P. Hajdú, L. Honti (eds), *Studien zur Phonologischen beschreibung Uralischer Sprachen.* Budapest: Akadémiai Kiadó, pp. 303–310, 1984.

[6] E. Itkonen, "Ehdotus inarilapin fonemaattiseksi transkriptioksi," in E. Itkonen, T. Itkonen, M. Korhonen, P. Sammallahti, *Lapin murteiden fonologiaa*. Helsinki: Castrenianumin toimitteita 1, pp. 43–68, 1971.

[7] P. Bye, "Grade alternation in Inari Saami and abstract declarative phonology," in I. Toivonen, D. Nelson (eds), *Saami Linguistics*. Amsterdam/Philadelphia: John Benjamins B.V., pp. 53–90, 2007.

[8] P. Bye, E. Sagulin, I. Toivonen, "Phonetic Duration, Phonological Quantity and Prosodic Structure in Inari Saami," *Phonetica*, vol. 66, no. 4, pp. 199–221, 2009.

[9] I. Lehiste, "Segmental and syllabic quantity in Estonian," *American Studies in Uralic Linguistics* 1, pp. 21–82, 1960.

[10] G. Liiv, "Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid," *Keel ja Kirjandus*, vol. 7–8, pp. 412–424, 480–490, 1961.

[11] P. Lippus, E. L. Asu, P. Teras, T. Tuisk, "Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian," *Journal of Phonetics*, vol 41, no. 1, pp. 17–28, 2013.

[12] D. Krull, "Perception of Estonian word prosody in whispered speech," in *Nordic Prosody – Proceedings of the VIII Conference, Trondheim*, pp. 153–164, 2001.

[13] K. Suomi, J. Toivanen, R. Ylitalo, "Durational and tonal correlates of accent in Finnish," *Journal of Phonetics*, vol. 31, pp. 113–138, 2003.

[14] T. Tuisk, "Tonal and temporal characteristics of disyllabic words in spontaneous Livonian," *Linguistica Uralica*, vol. XLVIII, no. 1, pp. 1–11, 2012.

[15] Z. McRobbie-Utasi, "The instability of systems with ternary length distinctions: the Skolt Saami evidence," in I. Toivonen, D. Nelson (eds), *Saami Linguistics*. Amsterdam/Philadelphia: John Benjamins B.V., pp. 167–206, 2007.

[16] N. Fangel-Gustavson, R. Ridouane, B. Moren-Duollja, "Quantity contrast in Lule Saami: a three-way system," in *Proceeding of the 10th ISSP, Cologne, Germany*, pp. 106–109, 2014.

[17] P. Boersma, D. Weenink, *Praat: doing phonetics by computer*. Computer program, version 5.4.21, 2015.

[18] D. Bates, M. Maechler, B. Bolker, S. Walker, *lme4: Linear mixed-effects models using Eigen and S4*, R package, 2015.