



# The role of duration and pitch in signaling quantity in Finnmark North Sámi

Katri Hiovain, Atte Asikainen and Juraj Šimko

University of Helsinki, Finland

firstname.lastname@helsinki.fi

## Abstract

Ternary quantity opposition is a cross-linguistically extremely rare typological feature. One of the languages using ternary opposition of consonants to signal linguistic contrasts is North Sámi, an endangered language spoken in several countries in the northernmost Scandinavia. Previous studies have shown that while the contrast between the two shorter quantity degrees is phonetically robustly realized using segmental durations, phonetic differences between the two longer degrees are much more subtle and show a considerable regional variation.

In this work we investigate other prosodic means that might be used to mark the contrast alongside duration, namely  $f_0$  movement and range. We show that the North Sámi speakers that are also native speakers of Norwegian use pitch to co-signal the differences between the two higher quantity degrees, while speakers that are Finnish-North Sámi bilinguals use primarily durational cues.

Documenting and interpreting these findings in the light of prosodic characteristics of the majority languages (Finnish and Norwegian) we argue that these regional differences reflect the majority language influence which can be a source of the ongoing dialectal divergence, and potential language change.

**Index Terms:** ternary quantity contrast, North Sámi, prosody,  $f_0$ , duration

## 1. Introduction

The North Sámi language is a part of the Sámi language continuum, the traditional speaking areas of which are spanning from the central Sweden to the tip of Kola peninsula in north-west Russia (see Fig. 1). All of the currently spoken nine Sámi languages, are endangered at different levels, ranging from definitely endangered to almost extinct [1]. North Sámi has approximately 20,000 – 24,000 speakers, which makes it the biggest and the least endangered Sámi language and it is still a predominant language in a few municipalities situated in the traditional Sámi speaking area (e.g. Guovdageaidnu). The language is nevertheless a minority language in the respective countries where it is spoken, meaning that all its adult speakers are bilingual, and conduct many of their daily activities in the official state language. As can be expected, this bilinguality has an influence on many aspects of the language, including several prosodic phenomena (see [2, 3]).

The North Sámi language is traditionally divided into four main dialect groups and areas: Western and Eastern Finnmark dialects, Torne dialects and Sea dialects. The dialects have a number of (morpho)phonological and lexical differences but are still mutually intelligible. In this paper we focus on the two dialects of the Finnmark North Sámi spoken in the northernmost areas of Norway and Finland.

Like many other languages from the Finnic branch of the Uralic language family, North Sámi uses quantity (phonological length) to mark lexical and grammatical contrasts. Similarly

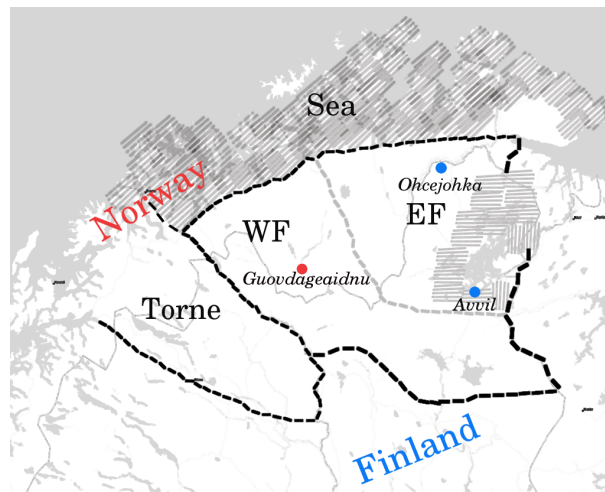


Figure 1: *The North Sámi varieties. The dashed lines separate the four main dialects of the North Sámi dialects and their traditional speaking areas (see [4, 5]). WF = Western Finnmark, EF = Eastern Finnmark.*

to, e.g., Estonian, the relatively complex system of gradation in North Sámi results in frequent consonant gemination with three possible quantities of the intervocalic consonant: short (Q1), long (Q2), and overlong (Q3). While the gradation system (described in more detail in, e.g., [5]) results in three phonological quantities, only two of them usually occur in the paradigm of a single word root. Therefore, the quantity opposition contrast is mainly realized in the form of Q2–Q1 or Q3–Q2, but usually not as Q1–Q3, see Table 1 for examples. This is the traditional way of describing the North Sámi quantity, and for consistency with earlier research, we are following this paradigm in our paper.

Table 1: *Examples of length oppositions in North Sámi.*

Q1 (short)	Q2 (long)	Q3 (overlong)
<i>gova</i>	<i>govva</i>	
‘picture’ GEN.SG	‘picture’ NOM.SG	
	<i>duvvá</i>	<i>duvvá</i>
	‘dove’ GEN.SG	‘dove’ NOM.SG

These quantity oppositions are assumed to be phonetically manifested primarily by durational variation. As shown in our previous work [3, 6], the Q1–Q2 contrast is indeed robustly signalled by the intervocalic consonant duration (shorter in Q1 than in Q2), and also by a compensatory shortening of the previous vowel in the longer quantity. The situation is somewhat more subtle for Q2–Q3 contrast, and the phonetic realization of this opposition shows interesting dialectal differences. While speakers of the Eastern dialect (and those from Finland) signal the contrast by durational means both for the consonant and the preceding syllable vowel, in the Western dialect (and

Norway), the durational differences are considerably less pronounced. The complementary effects affecting the durations of the consonant and the preceding vowel also suggest that quantity might at least partly be signalled in relative terms which can be quantified as a ratio of the consonant’s duration to that of the vowel.

In the present paper, we investigate whether this quantity opposition is co-signalled by additional prosodic means, namely fundamental frequency ( $f_0$ ). Given the differences mentioned above, we compare the realization of quantity opposition between the Western and Eastern dialects of the Finnmark North Sámi. As the Western and Eastern dialects are spoken primarily (although not exclusively) in Norway and Finland, respectively, we conceptualize the potential differences between the dialects in terms of the majority language influence on word prosody. Our speech material is thus divided by the origin of the speakers: Norway (Western dialect speakers) and Finland (Eastern dialect speakers).

In North Sámi, the primary word-level stress is assumed to always fall on the first syllable of each word, like in Finnish and other Finnic languages. The word stress is claimed to be realized by intonational patterns, with the stressed syllable having the highest point of pitch amongst the syllables in the word [4]. Thus, like in Finnish, the general word intonation in North Sámi is falling.

Unlike Finnish and North Sámi, Norwegian (a Germanic Indo-European language) does not have fixed stress on first syllable, but a lexically alternating stress position, often falling on the second syllable. The northernmost dialects of the Bokmål Norwegian that are in contact with North Sámi are characterized as “high-pitched” or *høytone* [4], meaning that the syllable stress is marked by high pitch; this means that in Norwegian the word-level intonation is often rising. In addition, Norwegian has a lexical pitch accent (or *lexical tone*), which means that many segmentally homonymous bisyllabic word pairs differ only by the tonal contour: *falling* or *rising-falling* [7].

It has been shown that, alongside duration, pitch cues participate in signalling quantity opposition in several languages. In Finnish, the phonologically long (heavy) syllables are associated with a dynamic (falling) tonal target whereas lighter syllables are connected with a static (high) tonal target [8]. In Estonian, a Finnic language with ternary quantity opposition, certain intonational patterns are connected to different quantity degrees [9, 10, 11]. For example, the contrast between the two shorter quantity degrees and the overlong degree is co-signalled by the position of the  $f_0$  turning point, located relatively earlier in Q3 words than in Q1 and Q2 ones. Some distinct patterns for different quantity degrees were also found in Skolt and Inari Sámi [12, 13].

The present work introduces an investigation of possible tonal elements of signalling quantity contrast in Finnmark North Sámi. Namely, we look at the  $f_0$  movement and its range over the (first syllable) vowel that participates in quantity patterns primarily associated with the following intervocalic consonant. Given the differences between prosodic characteristics of the majority languages (Finnish and Norwegian) spoken in the Finnmark North Sámi area (see Fig. 1), we concentrate on describing possible differences between speakers based on their majority language background. In particular, we hypothesize that if there are any dialectal differences, the North Sámi–Norwegian bilinguals use more tonal features than the speakers from Finland, which could possibly be a sign of majority language influence and language change.

## 2. Material and methods

### 2.1. Data collection

The speech material analyzed for this paper was collected during 2018 at two sites, Oulu (Finland) and Guovdageaidnu (Norway), and consists of utterances from four native and professional North Sámi speakers (age range 25–45), two from each site. The two speakers recorded in Oulu, originally from Ohcejohka and Avvil, F1 (female) and F2 (male), represent the Eastern Finnmark North Sámi dialect and are Finnish–North Sámi bilinguals. The speakers recorded in Guovdageaidnu, N1 and N2 (both female), also originally from there, are Western dialect speakers and Norwegian–North Sámi bilinguals. As we are primarily interested in the possible links between the majority language and quantity signalling in North Sámi. The speakers from Finland are marked in blue and Norwegian ones in red.

The recordings were collected in small office rooms using a Zoom H2n portable recording device. The instructions were given in Sámi and (when appropriate) Finnish, and a consent form was filled in before the recording. The subjects were asked to read aloud carrier sentences with target words embedded in a random order, presented on a computer screen, for example:

*Govas lea okta **happjá** (Q3) gurutbealde.*

‘There is one **long-tailed duck** on the left.’

*Govas leat guokte **happjá** (Q2) olgesbealde.*

‘There are two **long-tailed ducks** on the right.’

In order to maintain natural meanings of utterances, the carrier sentences somewhat differed between target words but all of them followed the same pattern with the target word in sentence-medial position.

In North Sámi, the numeral *okta*, ‘one’ elicits a singular nominative form of a noun, while the numeral *guokte*, ‘two’, is accompanied with the singular genitive/accusative case. As mentioned earlier (see Table 1), this grammatical opposition is commonly signalled by quantity degree of the intervocalic consonant of the noun as shown in the table.

Note that, as common in the North Sámi writing system, the difference between Q3 and Q2 geminates is not indicated orthographically, but the correct form is elicited by the grammatical relationships. The Q1–Q2 contrast is orthographically marked.

The data set contains 36 different two-syllable (4-segment) target words, structured as C1V1C2V2. The consonant C2, the main locus of quantity contrast, was one of /f, l, m, n, ŋ, r, s, ʃ, v/, most of them elicited in all three quantity degrees. The preceding vowel V1, known to participate in signalling quantity contrast, was one of /a, i, o, u/, each occurring in all quantity degrees.

The speakers were told to pronounce the carrier sentence as they found most suitable in their own dialect. In the case of mistakes or corrections, only the last production of each token was taken into analysis.

Every speaker uttered each target word three times, resulting in altogether 577 tokens (132 in Q1, 288 in Q2, and 157 in Q3).

### 2.2. Data processing

Recordings were automatically transcribed and segmented on word and phoneme levels using WebMAUS Basic forced-aligner [14, 15]. The resulting annotations were subsequently manually checked and corrected, and durations of the target word segments were extracted for statistic processing.

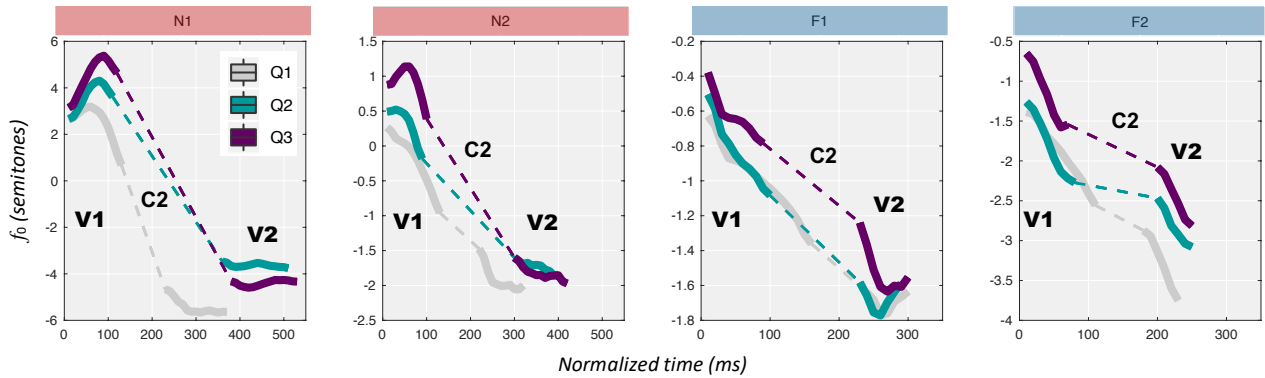


Figure 2: Time-normalized  $f_0$  trajectories over the V1C2V2 portion of the test words, averaged for each of the three quantity degrees, separately for the four speakers. The contours for the consonantal segments are shown as interpolations between the vowels.

The  $f_0$  contours of the target words were extracted using a customized Praat script, utilizing the Praat Pitch tier (pitch range 50-400Hz). The contours of each word were manually checked for octave jumps and other artefacts, after which the  $f_0$  values were collected and upsampled to 100 Hz sampling rate. For unvoiced (or creaky) portions of the vocalic segments, the  $f_0$  contours were obtained by interpolation and/or extrapolation using the surrounding material; when this was not possible (as was often the case for V2), the given data point was excluded from subsequent evaluation.

### 3. Results

Fig. 2 depicts average time-normalized  $f_0$  trajectories over the V1C2V2 portion of the test words in all three quantity degrees. For every speaker, mean durations of each of these three segments was calculated, separately for each quantity degree. Individual  $f_0$  contours for the two vocalic segments, V1 and V2, where then time-normalized to their respective mean durations, and the resulting time-normalized contours were then averaged. (Please note that the time-normalization was only used for the plotting purposes, statistical analyses operate on non-normalized values.) For intervocalic consonants, an interpolation between the vocalic portions of  $f_0$  is plotted.

The resulting figure suggests several observations regarding duration patterns and  $f_0$  contours over the individual segments:

While C2 is on average considerably shorter in Q1 compared to the two longer quantity degrees, the difference in C2 duration between Q2 and Q3 is less obvious for all four speakers. Similar but complementary durational patterns can be seen for the preceding vowel V1, that is longer in Q1 compared to the other two degrees (more so for speakers F1 and F2); for speakers from Finland V1 is on average slightly longer in Q2 than in Q3. The small lengthening of C2 and the compensatory shortening of V1 in Q3 compared to Q2 may be result in larger differences in relative durations captured by, for example, a durational ratio between C2 and V1 (cf. [6]).

Interestingly, the Fig. 2 shows some differences between the speakers from Norway and those from Finland. First, the V1C2V2 sequences are considerably shorter overall for F1 and F2 than for N1 and N2. Also, the more robust compensation for F1 and F2 lead to relatively equal average duration of the entire sequences for the three quantities, while for N1 and N2 the sequences are somewhat shorter for Q1 than for Q2 and Q3.

The two groups of speakers also seem to exhibit different tonal patterns in terms of the average  $f_0$  contours for the vowel V1. While the contours for F1 and F2 are generally falling,

N1 and N2 seem to generally exhibit more hat-like patterns. Also, the overall amount of  $f_0$  movement seems to be greater for Q1 than for Q2, and for Q3 than for Q2 (at least for the speakers from Norway). These differences suggest the presence of tonal phenomena participating in signalling Q1-Q2 and Q2-Q3 oppositions.

To evaluate these observations statistically, mixed effect models with quantity degree as a fixed independent factor and the phonetic segments of V1, C2, and V2 as random effects (to counteract inherent durational and tonal properties of different segments) were fitted separately for each speaker. The dependent variables were (1) **V1 duration**; (2) **V1  $f_0$  range** ( $f_0$  maximum -  $f_0$  minimum over V1); (3) the overall  **$f_0$  movement over V1** (computed as a sum of absolute values of differences between adjacent  $f_0$  samples); (4) **C2 duration**; (5) **ratio of C2 and V1 durations** (C2 duration divided by V1 duration); and (6) the **foot duration** (duration of the entire sequence V1C2V2).

After visual inspection of the dependent variable distributions, all dependent variables were log2-transformed.

Table 2: The  $t$ -values of the quantity effects (relative to Q2) from the fitted mixed effect models. The values considered significant ( $t$ -value > 2) are shaded in a speaker dependent color. dur = duration;  $f_0$ -r =  $f_0$  range;  $f_0$ -m =  $f_0$  movement

		V1		C2	C2:V1	foot	
		dur	$f_0$ -r	dur	ratio	dur	
Q1-Q2	N1	5.06	2.68	2.49	-31.5	-22.9	-13.6
	N2	8.97	3.70	2.69	-29.2	-29.1	-7.54
	F1	8.87	0.89	2.23	-12.9	-14.9	0.40
	F2	6.97	1.97	2.19	-14.8	-13.9	-2.14
Q3-Q2	N1	1.71	2.10	2.20	1.30	-0.35	2.24
	N2	2.43	2.79	3.54	-0.37	-2.84	2.58
	F1	-1.66	-0.56	-1.02	1.77	2.39	0.66
	F2	-2.15	-1.17	-1.42	0.73	2.07	-0.57

Table 2 lists  $t$ -values for the main effects of quantity on the dependent variables as obtained by the fitted models. The effects with absolute  $t$ -values greater than 2 are marked as significant by shading in the majority language dependent color (the grey shading marks significant differences going the ‘wrong way round’, e.g., N2 produced V1 significantly longer in Q2 than in Q3). As the independent factor variable of quantity was centered on Q2, the effects correspond to the sizes of difference in the respective dependent variable in Q2 versus the other two quantity degrees. The rows marked ‘Q1-Q2’ and ‘Q3-Q2’ thus refer to the effect in terms of these quantity contrasts.

For Q1–Q2 opposition, most effects of quantity on the dependent variables are significant. For all speakers, C2 was significantly longer in Q2 than in Q1, and, in compensatory manner, V1 was significantly shorter in the former than in the latter. This effect is also captured by the C2:V1 duration ratio. Probably reflecting these durational patterns, the  $f_0$  movement over V1 was significantly greater in Q1 than in Q2 for all speakers. The  $f_0$  range within V1 was significantly greater in Q1 than Q2 only for the speakers from Norway. With an exception of F1, the foot was significantly longer in Q2 than in Q1.

The Q3–Q2 opposition shows more subtle patterns. The effect of quantity on the duration of C2 was not significant for any of the the speakers. Speaker F2 produced V1 significantly shorter in Q3 than in Q2, but this compensatory behaviour was not significant for any other speaker; speaker N1 actually produced significantly *longer* V1s in Q3 than in Q2. Both speakers from Finland, however, signalled the opposition by significantly greater C2:V1 ratio in Q3 than in Q2; again, for N2 the ratio was significantly smaller in Q3 than in Q2. In terms of intonation, only the speakers from Norway had significantly greater  $f_0$  range and overall  $f_0$  movement over V1 in Q3 than in Q2; this effect was not significant for the speakers from Finland. Finally, the overall foot duration was significantly longer in Q3 than in Q2 only for the speakers N1 and N2; the speakers from Finland show no significant differences.

To further illustrate the key findings, the boxplots in Fig. 3 show the quantity effects for three of the six dependent variables depicting differences between the two speaker groups: C2:V1 ratio, and  $f_0$  range and movement over V1. To partly simulate the mixed effect modelling, the data were normalized by dividing the raw values by the mean value for the same segment in Q1 (the mean duration for same consonant phoneme in Q2 for C2 duration). The significance marking in the figure is from the models listed in Table 2.

#### 4. Discussion and conclusions

As attested previously [16, 3, 6], the opposition between the two shorter quantity degrees, Q1 and Q2, is robustly signalled by durations of both the consonant and preceding vowel, with the consonant being considerably longer in Q2 than in Q1, and the vowel shorter in the former than in the latter, in a compensatory manner. Naturally, this inverse relationship is further magnified when considering durations in relative terms, such as as a C2:V1 duration ratio.

Intonational clues considered here seem to contribute to signalling the Q1–Q2 contrast for V1. As indicated in Fig. 2 (see, e.g., speaker F1), the differences in  $f_0$  excursions might, however, simply reflect the differences in duration between the two quantity degrees.

Our data reveal a more interesting yet durationally subtler patterns concerning the Q2–Q3 contrast. On the one hand, the two Finnish-North Sámi bilinguals significantly mark the contrast using durational means, at least in the relative terms of C2:V1 ratio, but do not show any significant differences in the intonational measures. On the other hand, the speakers from Norway do not adhere to the expected durational pattern; speaker N2 in fact produced significantly *longer* V1s in Q2 than Q3 leading to a significantly *smaller* C2:V1 ratio in the former than in the latter quantity. Instead, these two speakers seem to “magnify” the pattern in Q3 compared to Q2 by producing the vowel with significantly greater  $f_0$  movement and range.

The lack of durational compensations between the consonant and the preceding vowel by the two Norwegian speakers

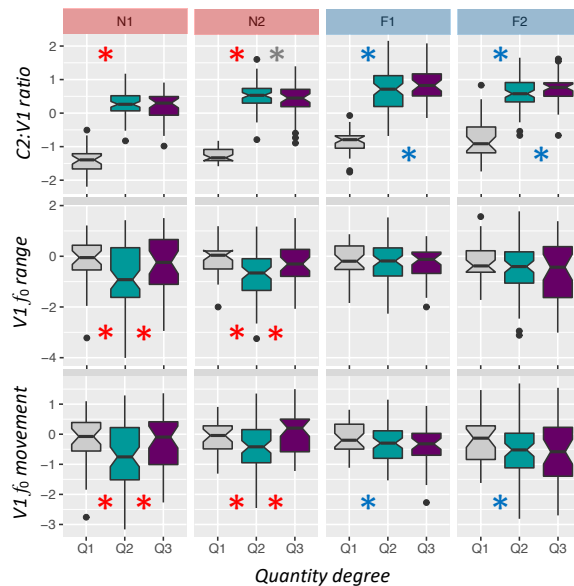


Figure 3: Boxplots of normalized distributions of three dependent variables. The significance markings as in Table 2.

leads to the significant dependency of the ‘foot’ duration on quantity, in particular for Q1–Q2 contrast. With an exception of a small effect for Q1–Q2 opposition for F2, the foot durations did not significantly depend on quantity for the Finnish bilinguals, the behavior consistent with compensations leading to a foot isochrony [17, 6].

We suggest that these differences between the areal varieties could be attributed to the majority language influence on the bilingual speakers of North Sámi. Of course, our own data set falls well short of providing a strong justification for this claim. First, its small size does not yield sufficient support for any generalizations over population groups (that is also why we performed statistical analyses on a by-speaker basis). A considerably larger and more varied data sample will be needed to present a more convincing argument for the claim.

Second, the division of our speakers by the majority language is somewhat arbitrary as both speakers from Finland are Eastern dialect speakers and both Norwegian subjects speak Western dialect of Finnmark North Sámi. This confound, however, also reflects the geographical distribution of the dialects: the Eastern dialect is spoken predominantly in Finland, while the Western one is more wide-spread in Norway, see Fig. 1. It is therefore plausible that the influence of majority language that, as in this case, has been influencing the language for many generations, can shape the dialects themselves.

Importantly, the reported differences between speakers meaningfully reflect the differences between the majority languages. Finnish is a quantity language that uses predominantly duration to contrast quantity degrees; and so do our speakers from Finland. Norwegian uses word-level intonation to mark contrasts, and our results suggest that the speakers from Norway use intonational means to signal Q2–Q3 contrast.

#### 5. Acknowledgements

We are very grateful to Štefan Beňuš, Pärtel Lippus, Jussi Ylikoski and to all the Sámi speakers who participated in the research.

## 6. References

- [1] C. Moseley, *Atlas of the World's Languages in Danger*. Unesco, 2010.
- [2] A. Aikio, L. Arola, and N. Kunnas, "Variation in North Saami," *Globalising sociolinguistics: Challenging and expanding theory*, pp. 243–255, 2015.
- [3] K. Hiiovain and J. Šimko, "Duration patterns in finnmork North Sámi quantity," in *International Congress of Phonetic Sciences*. Australasian Speech Science and Technology Association Inc., 2019, pp. 1560–1564.
- [4] K. P. Nickel and P. Sammallahti, *Nordsamisk grammatikk*. Davvi Girji AS, 2011.
- [5] P. Sammallahti, *The Saami languages: an introduction*. Davvi girji, 1998.
- [6] K. Hiiovain, J. Šimko, and M. Vainio, "Dialectal variation of duration patterns in finnmork North Sámi quantity," *Journal of the Acoustical Society of America*, (In press).
- [7] A. Wetterlin, *Tonal accents in Norwegian: phonology, morphology and lexical specification*. Walter de Gruyter, 2010, vol. 535.
- [8] M. Vainio, J. Järvikivi, D. Aalto, and A. Suni, "Phonetic tone signals phonological quantity and word structure," *The Journal of the Acoustical Society of America*, vol. 128, no. 3, pp. 1313–1321, 2010.
- [9] P. Lippus, K. Pajusalu, and J. Allik, "The tonal component of Estonian quantity in native and non-native perception," *Journal of Phonetics*, vol. 37, no. 4, pp. 388–396, 2009.
- [10] P. Lippus, E. L. Asu, P. Teras, and 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.
- [11] P. Lippus, "The acoustic features and perception of the Estonian quantity system," Ph.D. dissertation, 2011.
- [12] Z. McRobbie-Utasi, "The instability of systems with ternary length distinctions: The Skolt Saami evidence," *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, vol. 288, p. 167, 2007.
- [13] H. Türk, P. Lippus, K. Pajusalu, and P. Teras, "The acoustic correlates of quantity in Inari Saami," *Journal of Phonetics*, vol. 72, pp. 35–51, 2019.
- [14] T. Kisler, U. Reichel, and F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, vol. 45, pp. 326–347, 2017.
- [15] F. Schiel, "Automatic phonetic transcription of non-prompted speech," in *Proceedings of the XIVth International Congress of Phonetic Sciences*, 1999, pp. 607–610.
- [16] T. Magga, *Duration in the quantity of bisyllabics in the Guovdageaidnu dialect of North Lappish*. University of Oulu, 1984, vol. 11.
- [17] D. Krull, "Foot isochrony in Estonian conversational speech," *Gothenburg papers in theoretical linguistics*, no. 81, pp. 97–100, 1999.