

EVA LIINA ASU, NELE SALVESTE (Tartu)

THE PHONETIC AND PHONOLOGICAL ANALYSIS
OF THE FALL-RISE INTONATION PATTERN
IN THE KIHNU VARIETY OF ESTONIAN

Abstract. The paper presents an analysis of fall-rise intonation patterns in the Kihnu dialect of Estonian. The study is based on recordings of spontaneous speech from three generations of female speakers. Fall-rises occurred almost five times less frequently in the data of the speakers born around 1935 as compared to those born around 1900. The data of the speakers born after 1970 did not contain any fall-rises at all, which could be interpreted as a sign of prosodic levelling of the traditional variety. An investigation of F0 alignment revealed a systematic variation in the location of high and low tonal targets. The peaks were aligned significantly later in the case of longer phrases. The first low elbow was aligned either with the accented syllable (in monosyllabic and disyllabic words) or with the syllable following the accented syllable (in longer words and phrases), and the second low elbow with the final syllable of the phrase just before the high boundary. These findings lend support to the treatment of the fall-rise accent as H*L H% in accordance with the autosegmental-metrical analysis of Estonian intonation.

Keywords: Estonian, Kihnu, intonation, fall-rise accent, peak alignment, dialect levelling.

1. Introduction

Prosodically, one of the most interesting areas in Estonia is constituted by the islands off the west coast of the Estonian mainland. Distinctive intonational variation of insular dialects has been attributed to influences from Swedish, which used to be an important contact-language in this language convergence area (Niit 1980; Kalits 2006). The two Estonian varieties usually singled out for their intonation, and often impressionistically characterised as "sing-songy" are those of the Islands of Saaremaa and Kihnu. It has been shown that in the Saaremaa variety, the impression of singing intonation is phonetically given by systematically later peak alignment as compared to Standard Estonian (Asu 2005; 2006). The Kihnu variety, on the other hand, is noted for the frequent occurrence of rising intonation, which has already been mentioned in older dialectal descriptions (Saar 1934; Tanning 1948). On closer examination the rises actually appeared to be mostly fall-rises occurring in phrase-final position (Asu, Niit 2010). Such accents have not been encountered in the intonational inventory of Standard Estonian (Asu 2004), and are therefore particularly interesting from the typological point of view. The present paper focuses on their phonetic and phonological analysis in spontaneous speech.

Fall-rise accents occur in many other European languages, including English, Dutch and German. The most common analysis of such accents within the autosegmental-metrical theory of intonation is a fall followed by a high boundary tone: H*+L H% (Grabe, Post, Nolan 2001; Gussenhoven 2005; Uhmman 1991; Féry 1993).

In the transcription system for German (GToBI), fall-rises are generally represented using the L- phrase accent as (L+)H* L-H% (Grice, Baumann, Benz Müller 2005). As intermediate phrases are not used in the intonational transcription for Estonian (Asu 2004), the optimal phonological analysis for the fall-rise accent of Kihnu would be H*L H%. We hope to find further confirmation to the appropriateness of this label on the basis of the following analysis.

The present paper has two broader research aims. Firstly, our goal is to investigate the general occurrence of the fall-rise accent in the speech of women representing different age groups with a view to establish whether there are any differences. This research question is relevant from the point of view of dialect levelling. Already back in the 1930s, it was claimed that the characteristic quality of Kihnu speech melody was disappearing in the speech of younger islanders (Saar 1934). Levelling in the Kihnu variety has so far only been addressed with respect to some grammatical and phonological aspects (Grigorjev, Keevallik, Niit, Paldre, Sak, Veismann 1997). If we assume that the fall-rise accent is (at least) one of the manifestations of the distinct Kihnu melody, the study of the distribution of this pattern in different generations of speakers would shed some more light on this matter. As a result of prosodic dialect levelling we would expect the fall-rise to appear less frequently in the data of younger informants.

The second main goal of the paper is to investigate the phonetic realisation of the fall-rise accent. We will focus on the study of the alignment of H and L targets in nuclear fall-rises of varying length. The alignment of tonal targets has been shown to depend on a variety of factors, such as for instance, the position of the accent in the utterance (sentence-medial vs. sentence-final) (Ladd, Schepman, White, Quarmby, Stackhouse 2009), the phonological length of the accented vowel (Ladd, Mennen, Schepman 2000; Ladd, Schepman, White, Quarmby, Stackhouse 2009), or the number of unstressed syllables following the accented syllable (Silverman, Pierrehumbert 1990; Prieto, van Santen, Hirschberg 1995; Plüschke 2011). Studies on the influence of syllable structure on peak alignment have yielded contradicting results. For instance, it was shown for Spanish that peaks were aligned earlier in closed syllables than in open syllables (Prieto, Torreira 2007), while no effect of syllable structure on alignment was found in Dutch (Schepman, Lickley, Ladd 2006).

In Estonian, the F0 peak placement is additionally affected by the phonological quantity of the foot. Broadly speaking, in the short (Q1) and long (Q2) quantity degree, peaks are aligned later than in the overlong quantity degree (Q3) that is characterised by a steep F0 fall early in the accented syllable. It has been shown that these tonal differences are also present in spontaneous speech (Asu, Lippus, Teras, Tuisk 2009).

Thus, in our data we would expect H* to be aligned later in the case of longer unstressed material following the accented syllable. Despite the number of syllables in the phrase we expect the location of the peak to be dependent on the phonological quantity. Possible differences in alignment due to syllable structure will also be investigated.

2. Materials and method

For the present analysis we used recordings of spontaneous speech from three generations of Kihnu women: (1) those born around 1900, (2) around 1935, and (3) after 1970. The data was drawn from the University of Tartu Archives of Estonian Dialects and Kindred Languages based on the birth-year of the informants. The speech files consisted mainly of dialogues between a field-worker and an informant, and in the case of some newer recordings between two speakers of Kihnu Estonian. In total we analysed 10 h 47 min of speech by 21 informants. The summary of the data is presented in Table 1. It can be seen that there was least data from the youngest generation, which is due to there being on the whole less data for younger speakers

in the database. The division of speakers into three groups according their birth-year is in accordance with the speakers' age at recording. With only one exception (the 80 year old informant in the middle group) the informants in the oldest age group, as classified by their birth-year, were also oldest at the time of recording.

Table 1

Speech data used for the analysis				
Year of birth	Number of informants	Age at recording	Year of recording	Duration of the data (h:min:sec)
1889–1902	6	68–91	1961–1988	3:36:35
1930–1943	8	61–68, 80	1992–2011	5:17:39
1971–1989	7	20–38	2009	1:53:28

The recordings were made at different times during dialectal field-work on the Kihnu Island. Therefore, the quality of the sound files varies considerably depending on the recording equipment used and the nature of the recording environment. Despite this variation the data was deemed suitable for the present purposes.

We located all instances of fall-rise accents in the recordings by combined listening and visual inspection of the F0 traces. In total 282 tokens were found (275 in intonational phrase-final and 7 in non-final position). The present analysis focuses on the phrase-final (nuclear) accents. The distribution of the tokens according to the phonological quantity of the accented syllable was the following: Q1 – 84, Q2 – 49, Q3 – 142.

The fall-rise accents were transcribed on four annotation tiers: word, syllable, segment and intonation, using the speech analysis software Praat (Boersma, Weenink 2011). On the intonation tier, the F0 contour of each phrase was labelled at 5 events: (1) the beginning of the word at the start of fall-rise, which often coincided with the phrase boundary and was therefore marked as %, (2) the F0 maximum in the accented syllable (H*), (3) the first low elbow after the fall (L₁), (4) the second low elbow before the rise (L₂), and (5) the final intonational boundary (H%). We chose to identify two low pitch elbows rather than the lowest F0 point in order to render the shape of the accent contour better particularly in the case of longer phrases (cf. Lickley, Schepman, Ladd 2005). Figure 1 presents an F0 contour with the H*L H% accent, exemplifying the annotation and measurement points.

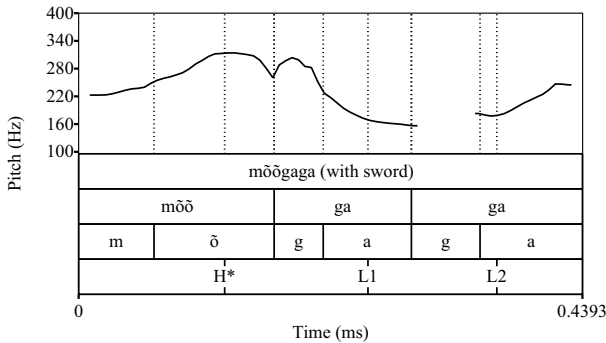


Figure 1. An F0 contour with the H*L H% accent on a trisyllabic utterance showing the annotation and measurement points.

A Praat script was used to extract the values of pitch (Hz) and time (ms) at the labelled locations, as well as to calculate the duration of each transcribed phrase and syllable. All F0 measurements were subsequently manually checked and corrected.

The H*L H% accent occurred on material of varying length, from monosyllabic words to words or phrases consisting of six syllables. Table 2 shows the distribution of the data according to the quantity and structure of the first syllable, and the number of syllables in the phrase.

In order to study the alignment of tonal targets the proportional time of H* was calculated using the formula:

$$H^* - S_0 / S_B - S_0 \quad (1)$$

where H^* denotes the time for the peak, S_0 for the beginning of the stressed syllable, S_B for the end of the stressed syllable (syllable boundary).

Table 2

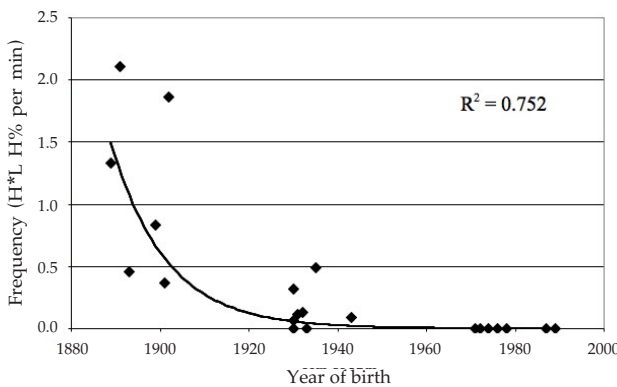
The distribution of the data according to the quantity (Q1, Q2, Q3) and structure of the first syllable, and the number of syllables in the phrase

Number of syllables in the phrase	(C)V			(C)VV			(C)VC		(C)VVC	Total
	Q1	Q2	Q3	Q2	Q3	Q3	Q3			
1	–	–	11	–	6	13			30	
2	27	7	23	5	26	–			88	
3	30	7	16	7	12	–			72	
4	11	4	11	11	12	–			49	
5	11	5	4	3	5	–			28	
6	5	–	1	–	2	–			8	
Total	84	23	66	26	63	13			275	

3. Results and discussion

3.1. Distribution of the fall-rise accent in three generations of speakers

In order to test the hypothesis about prosodic dialect levelling we looked at the distribution of the fall-rise accent in the three age groups presented above. The frequency was calculated on the basis of the number of times the accent occurred in every sound file. Admittedly, this is a relatively simplified way of analysis but nevertheless provides us with a clear answer. Figure 2 presents an exponential curve for the frequency of the accent in relation to the birth-year of the speaker. It can be seen that there is a strong correlation ($R^2 = 0.752$). In the older group, the fall-rise accent occurred on average 1.2 times/min, and in the middle group on average 0.2 times/min, or in other words, the speakers in the older group used the fall-rise accent on average every 1.3 min, whereas those in the middle group every 6 min. The accent did not occur at all in the data of the youngest informants. It can be argued that this might be due to there being least data for this group of speakers, but on the other hand the data was similar in nature and by a comparable number of informants as for the other two age groups. Therefore, we can infer that the hypothesis about the gradual disappearance of the characteristic speech melody with regard to the occurrence of the $H^*L H\%$ accent was borne out.



*Figure 2. Exponential curve for the frequency of $H^*L H\%$ vs. year of birth of the speaker.*

The lack of the fall-rise accent in the speech of younger informants can be interpreted as a reflection of the levelling of traditional Kihnu speech melody. Yet, the general infrequency of the $H^*L H\%$ pattern in the present data (only 275 phrase-final

tokens in nearly 11 hours of speech) can denote its limited discourse functions. It is likely that the accent only occurs in certain contexts. Preliminarily, we observed that fall-rises were often used in short (one-word) answers, and echo-questions. It was also clear that the fall-rise accent was much more common in phrase-final than in non-final position. A more thorough investigation of its discourse functions is needed.

3.2. Alignment of H and L targets

Peak alignment was analysed with respect to the number of syllables in the phrase, and the phonological quantity and syllable structure of the accented syllable. The phrases in the data were maximally 6 syllables long, but due to there being no tokens for Q2 for the longest material, we included maximally pentasyllabic phrases in the analysis. Firstly, we tested whether the peak alignment was influenced by syllable structure, separating open ((C)V(V)) and closed ((C)(V)VC) syllables. As we did not find any significant effect of syllable structure on peak location, the results will be presented based on the average of all syllable structures.

Figure 3 shows the proportional location of the peak relative to the beginning of the accented syllable in phrases of different length. As expected, the peak was aligned later in the case of more unstressed syllables following the stressed syllable. There was a significant effect of the number of syllables on the alignment of H* ($F(4, 254) = 15.29, p < 0.001$). This finding is line with earlier work on Estonian (Plüschke 2011) and other languages (e.g. Silverman, Pierrehumbert 1990; Ladd, Schepman, White, Quarmby, Stackhouse 2009; Schepman, Lickley, Ladd 2006).

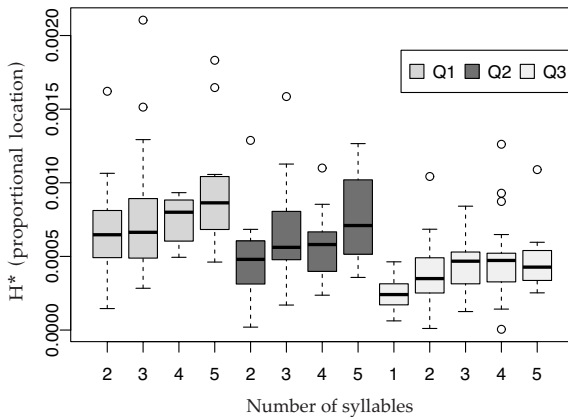


Figure 3. Boxplots showing the proportional location of H* relative to the duration of the accented syllable in three quantities (Q1, Q2, Q3) in phrases of different length (1–5 syllables).

Various explanations for the observed systematic differences in peak alignment have been offered. The most compelling reason for earlier peaks in the case of shorter phrases in the present data seems to be so called tonal crowding, where the variable spacing of tones is influenced by the preceding and following tones (cf. Silverman, Pierrehumbert 1990; Arvanti, Ladd, Mennen 2006). In the case of shorter material, there is less time for the realisation of the fall-rise contour before the phrase boundary, which means that the whole tune has to be compressed. This explains the significantly earlier location of H* in mono- and disyllabic utterances, and is a strong argument for the analysis of the contour as one pitch accent rather than composed of different tunes.

Despite the influence of the length of the post-accentual material, the quantity-dependent effect on peak alignment was also clearly evident and statistically significant ($F(2, 254) = 29.34, p < 0.001$). As expected, the F0 peak occurred earlier in Q3 than in Q1 and Q2 words. A similar result was shown for Standard Estonian in (Plüschke 2011) on the basis of tightly controlled read sentences.

Subsequently, the alignment of low targets was explored. Figure 4 presents both the timing of H* and the two low elbows in phrases varying from 1 to 5 syllables. The accented syllables are marked in dark grey. It can be seen that a systematic pattern also emerges for the alignment of low elbows. In monosyllabic and disyllabic words in Q3, the first low elbow (L₁) was located in the first syllable; in longer phrases it was always aligned with the second syllable with only one exception (pentasyllabic phrases in Q2). This relatively invariant location of L₁ in relation to H* denotes that the tones belong together, which gives support to the analysis of the first component of the fall-rise accent as a bi-tonal pitch accent, a fall (H*L). The second low elbow (L₂) was always aligned with the final syllable of the phrase (with only one exception). It can be argued that the stable alignment of L₂ might be an evidence for the existence of a low phrase accent (L-), or a bi-tonal boundary tone (LH%). It is, however, impossible to solve the issue on the basis of the present data. Therefore, until further evidence, there is no reason not to analyse the pattern as H*L H%, which is in line with the treatment of similar patterns in comparable intonational transcription systems such as IViE for British English (Grabe, Post, Nolan 2001) and ToDI for Dutch (Gussenhoven 2005).

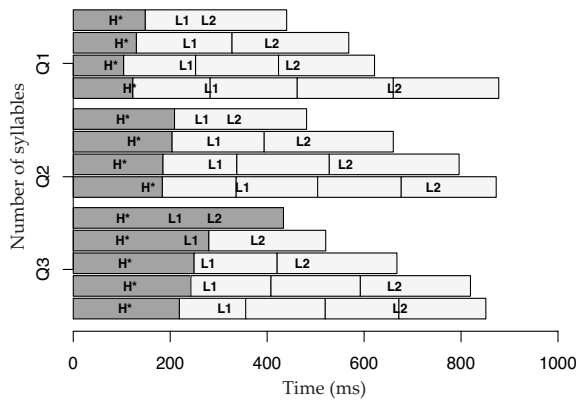


Figure 4. The alignment of H* relative to the onset of the accented syllable (dark grey) and the location of low targets (L₁ and L₂) in phrases of different length (1–5 syllables) in three quantities (Q1, Q2, Q3).

In sum, the general shape of the tonal contour is realised similarly in phrases of different length independent of the structure or quantity of the accented syllable. There is a fall in pitch immediately after the H* after which the F0 contour stays low until the rise starts just before the high boundary. Figure 5 presents boxplots showing pitch at the 5 locations in the fall-rise contour. The data has been averaged over all syllable structures and quantity degrees as well as phrases of different length, as there were no significant differences in the range of the fall from H* to L₁ and the span of the rise from L₂ to H% between these different conditions.

4. Conclusions and further work

The paper studied the fall-rise pitch accent in Kihnu Estonian. Firstly, we explored its distribution in the speech of three generations of female speakers. It was hypothesised that due to the levelling of the traditional dialect, fall-rise patterns (as one of the manifestations of the distinct intonation of the variety) would occur less frequently in the speech of younger informants. This hypothesis was borne out. The speakers born around 1935 used the fall-rise accent almost five times less frequently than those born around 1900, and the data of the speakers born after 1970 did not contain any fall-rises at all.

Secondly, we investigated the alignment of tonal targets in the fall-rise contour. The analysis revealed a systematic variation in the location of the peak depending on the number of syllables in the phrase. The H* was aligned significantly later when more unstressed material followed the accented syllable. This finding is in line with work on other languages. It was also shown that the previously attested tonal align-

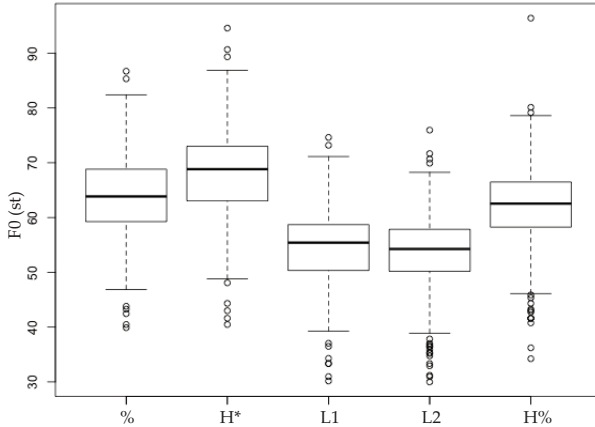


Figure 5. Boxplots of the pitch (st) at the initial boundary (%), High target (H*), Low targets (L₁, L₂) and final intonational boundary (H%).

ment characteristics of the Estonian three quantity degrees were realised in the present spontaneous data: the peak was aligned earlier in Q3 than in Q1 and Q2.

Two low pitch elbows were located in order to characterise the low valley between the end of the fall and the beginning of the rise in the fall-rise contour. It was shown that the first low elbow (L₁) was aligned after the accented syllable (except in monosyllabic and disyllabic words in Q3, where it was located in the accented syllable), and the second low elbow (L₂) with the final syllable of the phrase. Further work is needed to establish whether the stable alignment of L₂ just before the high boundary tone might be evidence of a low phrase accent (L-), or alternatively a bi-tonal boundary tone (LH%). These questions have to be tested on more tightly controlled materials.

The domain of the fall-rise varied from monosyllables to words and phrases consisting of six syllables. The realisation of the tonal contour was not dependent on the length of the material, which implies that the pattern is not composed of different tunes but constitutes one pitch accent plus a boundary tone. This finding lends support to its analysis as H*L H% following the current framework for the analysis of Estonian intonation (Asu 2004).

We are aware that the characteristic intonation of Kihnu may also have other acoustic correlates in addition to the H*L H% pattern. As has been shown for Estonian (Asu 2005) and many other languages (e.g. Atterer, Ladd 2004; Arvanti, Garding 2007; Ladd, Schepman, White, Quarmby, Stackhouse 2009), different varieties of the same language can vary in the realisation of the same F0 contour with respect to peak alignment. A comparison of Kihnu intonation with other varieties of Estonian might therefore reveal further important characteristics of the prosody of this variety.

Acknowledgements

We would like to thank Pärtel Lippus for his assistance with the statistical analysis, and Francis Nolan for discussing aspects of the paper. This research was supported by the Estonian Science Foundation grant no. 7904.

Addresses

Eva Liina Asu
 Institute of Estonian and
 General Linguistics, University of Tartu
 E-mail: eva-liina.asu@ut.ee

Nele Salveste
 Institute of Estonian and
 General Linguistics, University of Tartu
 E-mail: nele.salveste@ut.ee

REFERENCES

- Arvaniti, A., Garding, G. 2007, Dialectal Variation in the Rising Accents of American English. — *Papers in Laboratory Phonology* 9. Berlin—New York, 547—576.
- Arvaniti, A., Ladd, D. R., Mennen, I. 2006, Phonetic Effects of Focus and "Tonal Crowding" in Intonation. Evidence from Greek Polar Questions. — *Speech Communication* 48, 667—696.
- Asu, E. L. 2004, *The Phonetics and Phonology of Estonian Intonation* (Doctoral dissertation. University of Cambridge).
- 2005, Intonational Contour Alignment in Saaremaa and Standard Estonian. — *LU XLI*, 107—112.
- 2006, Tonal Alignment in Two Varieties of Estonian. — *Nordic Prosody. Proceedings of the IXth Conference, Lund 2004, Frankfurt am Main*, 29—35.
- Asu, E. L., Lippus, P., Teras, P., Tuisk, T. 2009, The Realization of Estonian Quantity Characteristics in Spontaneous Speech. — *Nordic Prosody. Proceedings of the Xth Conference, Helsinki 2008, Frankfurt*, 49—56.
- Asu, E. L., Niit, E. 2010, Prosodic Features of the West Estonian Dialect of Kihnu. — *Congressus XI Internationalis Fenno-Ugristarum, Pars III. Summaria Acroasium in Symposiis Factarum, Piliscsaba*, 7.
- Atterer, M., Ladd, D. R. 2004, On the Phonetics and Phonology of "Segmental Anchoring" of F0. Evidence from German. — *Journal of Phonetics* 32, 177—197.
- Boersma, P., Weenink, D. 2011, Praat: doing phonetics by computer [Computer program]. <http://www.praat.org/>.
- Féry, C. 1993, *German Intonational Patterns*, Tübingen.
- Grabe, E., Post, B., Nolan, F. 2001, Modelling Intonational Variation in English. The IViE System. — *Proceedings of Prosody 2000, Poznan*, 51—57.
- Grice, M., Baumann, S., Benz Müller, R. 2005, German Intonation in Autosegmental-Metrical Phonology. — *Prosodic Typology. The Phonology of Intonation and Phrasing*, Oxford, 55—83.
- Grigorjev, P., Keevalik, L., Niit, E., Paldre, L., Sak, K., Veismann, A. 1997, Kihnu murde assimileerumise mustreid Manilaiul. — *Pühendusteos Huno Rätsepale, Tartu (Tartu Ülikooli eesti keele õppetooli toimetised 7)*, 26—44.
- Gussenhoven, C. 2005, *Transcription of Dutch Intonation. — Prosodic Typology. The Phonology of Intonation and Phrasing*, Oxford, 118—145.
- Kalits, V. 2006, Kihnlaste elatusalad XIX sajandi keskpaigast XX sajandi keskpaigani, Kihnu.
- Ladd, D. R., Mennen, I., Schepman, A. 2000, Phonological Conditioning of Peak Alignment of Rising Pitch Accents in Dutch. — *Journal of the Acoustical Society of America* 107, 2685—2696.
- Ladd, D. R., Schepman, A., White, L., Quarby, L. M., Stackhouse, R. 2009, Structural and Dialectal Effects on Pitch Peak Alignment in Two Varieties of British English. — *Journal of Phonetics* 37, 145—161.
- Lickley, R., Schepman, A., Ladd, D. R. 2005, Alignment of "Phrase Accent" Lows in Dutch Falling Rising Questions. Theoretical and Methodological Implications. — *Language and Speech* 48, 157—183.
- Niit, E. 1980, *The Structure of the Baltic Prosodic Area and the Place of Estonian Dialects in It*, Tallinn (Preprint KKI-17).
- Plüschke, M. 2011, Peak Alignment in Falling Accents in Estonian. — *Proceedings of the XVII International Congress of Phonetic Sciences, Hong Kong, 17—21 August 2011, Hong Kong*, 1614—1617.
- Prieto, P., Torreira, F. 2007, The Segmental Anchoring Hypothesis Revisited. Syllable Structure and Speech Rate Effects on Peak Timing in Spanish. — *Journal of Phonetics* 35, 473—500.
- Prieto, P., van Santen, J., Hirschberg, J. 1995, Tonal Alignment Patterns in Spanish. — *Journal of Phonetics* 23, 429—451.
- Saar, T. 1934, *Murdeülevaade* (MS. Institute of the Estonian Language. Archive of Estonian dialects and Finno-Ugric languages).

- Schepman, A., Lickley, R., Ladd, D. R. 2006, Effects of Vowel Length and "Right Context" on the Alignment of Dutch Nuclear Accents. — *Journal of Phonetics* 34, 1–28.
- Silverman, K., Pierrehumbert, J. 1990, The Timing of Prenuclear High Accents in English. — *Papers in Laboratory Phonology I*, 72-106.
- Tanning, S. 1948, Murdepäevik (MS. Institute of the Estonian Language. Archive of Estonian dialects and Finno-Ugric languages).
- Uhmann, S. 1991, Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie, Tübingen.

ЭВА ЛИЙНА АСУ, НЕЛЕ САЛВЕСТЕ (Тарту)

**ФОНЕТИЧЕСКИЕ И ФОНОЛОГИЧЕСКИЕ АНАЛИЗЫ
НИСХОДЯЩЕ-ВОСХОДЯЩЕЙ ИНТОНАЦИИ
В КИХНУСКОМ ГОВОРЕ ЭСТОНСКОГО ЯЗЫКА**

В статье представлен анализ нисходяще-восходящей интонации в кихнуском говоре эстонского языка. Работа основана на записях спонтанной речи, полученных от носителей языка женского пола трех поколений. Если сравнить носителей языка, родившихся около 1900 года и около 1935 года, то в речи последних нисходяще-восходящая интонация встречается почти в пять раз реже. Речь носителей, родившихся после 1970 года, вообще не содержит нисходяще-восходящих интонационных контуров, что может быть интерпретировано как просодическое выравнивание исходных диалектных признаков. Исследование контура основного тона показало систематическое варьирование в расположении высоких и низких тональных пиков. В более длинных фразах пики расположены значительно правее. Первый нижний пик соотносится либо с ударным слогом (в односложных и двусложных словах), либо со слогом, следующим за ударным (в более длинных словах и фразах). Второй нижний пик соотносится с последним слогом фразы непосредственно перед высокой границей. Исследование подтвердило интерпретацию нисходяще-восходящей интонации как H*L H% в соответствии с автосегментным метрическим анализом эстонской интонации.