

EVA LIINA ASU, PÄRTEL LIPPUS, ELLEN NIIT, HELEN TÜRK (Tartu)

THE ACOUSTIC CHARACTERISTICS
OF MONOPHTHONGS AND DIPHTHONGS
IN THE KIHNU VARIETY OF ESTONIAN

Abstract. This paper presents the first acoustic study of monophthongs and diphthongs in the Kihnu variety of Estonian. The focus is on six diphthongs which have arisen after the diphthongization of long open and mid vowels. The comparison of their components with monophthongs revealed that the target values of the diphthongs are close to the corresponding monophthongs. There was some variation due to coarticulatory influences between the diphthong components. It was shown that the duration of both target vowels in the diphthong is longer in the third quantity than in the second quantity. The analysis also revealed that there is an acoustic basis for postulating the existence of two triphthongs in Kihnu. Formant trajectory length proved to be a useful measure for comparing monophthongs, diphthongs and triphthongs.

Keywords: Estonian dialects, Kihnu, vowel quality, diphthongs, triphthongs, formant trajectory length.

1. Introduction

Regional varieties of Estonian exhibit considerable variation as to the realization of their vowel systems. One of the richest inventories is that of the variety of Kihnu, which belongs to the Insular dialects of the North Estonian dialect group, and is spoken by about 600 people on the Island of Kihnu. Due to the relative isolation of the variety several features that are rare or not present elsewhere in the North Estonian dialect area have been preserved. Kihnu is, for instance, unique among the North Estonian dialects in that it has retained vowel harmony involving four different vowels (Lonn, Niit 2002): /æ/¹ (e.g. /kylæ/ 'village'), /ø/ (e.g. /pøgø/ 'ghost'), /y/ (e.g. /sygyse/ 'in autumn') and /ɤ/ (e.g. /pɤlɤ/ 'is not'). Another characteristic which does not occur in other North Estonian dialects (although is common in South-East Estonian dialects) is the change of the front vowels /i/ and /æ/ in the first syllable into /ɤ/ and /ɑ/ respectively, with the accompanying /j/-glide (Eek 2008 : 106; Pajusalu, Hennoste, Niit, Päll, Viikberg 2009 : 133), e.g. /lina/ > /ljɤna/ 'flax', /nægu/ > /njagu/ 'face'. This phenomenon only takes place in words which have a back vowel in the second syllable.

The Kihnu vowel system is also characterized by the diphthongization of long vowels. This general feature of the North Estonian dialect group is rare among

¹ The symbols used in this paper are those of the International Phonetic Alphabet (IPA). Equivalents in the Finno-Ugric transcription system are as follows: α = a, æ = ä, ø = ö, y = ü, ɤ = ɛ, /:/ = length mark.

Insular varieties, appearing only in Eastern Saaremaa and Muhu. Diphthongization in Kihnu involves long open and mid vowels /æ, a, e, ø, ʏ, o/ resulting in diphthongs which have mostly been written down as *iä, ua, ie, üe, õe, uõ* (e.g. /sæ:r/ > /siæ:r/ 'leg', /ma:/ > /mua:/ 'land', /ke:l/ > /kie:l/ 'tongue', /tø:/ > /tye:/ 'work', /vʏ:ras/ > /vʏeras/ 'stranger', /ko:li/ > /kuy:li/ 'to school'). Long close vowels /i, y, u/ have been preserved in all words (Saar, Valmet 1997). Despite widespread diphthongization long open and mid vowels occur in certain words and word forms, in particular newer vocabulary and loan words (Saar 1958); least complete is the diphthongization of long /ø/ and /ʏ/ (Lonn, Niit 2002 : 25). Thus, the Kihnu variety has both long monophthongs and diphthongs.

Of particular interest are the diphthongs that have been formed as a result of the diphthongization of long /a/ and /o/. While for other diphthongs dialectal transcriptions are uniform, there is notable variation for these two. The diphthong /ua/ corresponding long /a/ has also been transcribed as *uä* or *ua^e* (e.g. *puät, puä^et* 'boat' sg. nom.) or *oa* (e.g. *puadi ~ poadi* 'boat' sg. gen.), and the diphthong /uʏ/ corresponding long /o/ as *ue* or *uõ^e* (e.g. *kuel, kuõ^el* 'school' sg. nom.) or *uo* (e.g. *kuoli* 'school' sg. gen.) (Saar 1958; Grigorjev, Keevallik, Niit, Paldre, Sak, Veismann 1997). The transcription *uä* was first used by Saareste (1920), and is in itself curious because this diphthong does not occur in any other Estonian variety (Pajusalu, Hennoste, Niit, Päll, Viikberg 2009 : 133). It is obviously not easy to judge the quality of a sound solely by auditory impression, but in this case an additional reason for varying transcriptions lies in palatalization (Saar 1958; Eek 2008; Sang 2009). In words where a diphthong is followed by a palatalized consonant, an epenthetic vowel is inserted before the consonant after the diphthong (Eek 2008 : 132). The result is a triphthong. Alternatively, the formation of triphthongs has been explained by consonant fission (Viitso 2003 : 191–192). According to Sang (2009) the phonological system of Kihnu contains the triphthongs /uae/ and /uʏe/. This is a generally accepted view among the speakers of Kihnu. The textbook on Kihnu morphology uses three adjacent vowel letters to spell the words with triphthongs (e.g. *puæk* 'lighthouse') (Laos 2010 : 15). In other publications in Kihnu, also variants with the superscript *e* are used. Triphthongs have also been transcribed in the dialectal texts for Muhu (Lonn, Niit 2002 : 22–23) and Leivu (Ariste 1953 : 69).

The existence of triphthongs has not been studied acoustically. In fact, the vowel system of Kihnu has only been described in the accounts of traditional dialectological studies (e.g. Saar 1958; Lonn, Niit 2002). The only acoustic analysis of Kihnu vowels is a small-scale study focusing on the quality of short vowels in stressed and unstressed syllables based on data from one elderly female subject (Türk 2010).

The present paper aims to address this gap in our knowledge and focus on the acoustic study of the six diphthongs which have formed due to the diphthongization of long open and mid vowels. We are interested in both the quality and durational characteristics of diphthong components. Our hypotheses are derived from earlier acoustic studies on Standard Estonian diphthongs, of which there are only a handful. Firstly, based on Lehiste (1970) we hypothesize that the diphthong components are similar in their quality to the corresponding monophthongs, although following Piir (1985 : 49) some variation is likely to occur. Secondly, also based on Lehiste (1970) we would expect the duration of both vowels in a diphthong to be longer in the third quantity (Q3) than in the second quantity (Q2). There are, however, contradicting views on this matter. It has also been suggested that it is the duration of the second element of the diphthong that differentiates between Q2 and Q3, whereas the first element is not influenced by the phonological quantity and is similar in both quantities (Eek 2008 : 136–137). Piir's (1985) results imply that the durations vary depending on the diphthong, but the second components are for the most part longer than the first components. Finally, we hypothesize that there is an acoustic basis for postulating the existence of two triphthongs in Kihnu.

2. Materials and method

The data set for the present analysis comprised disyllabic test words containing 9 monophthongs /a, e, i, o, u, ʏ, æ, ø, y/ represented by the letters *a, e, i, o, u, õ, ä, ö, ü* and 6 diphthongs /iæ, ua, ʏe, uʏ, ie, ye/ spelled as *iä, ua, õe, uõ, ie, üe*. The test words for monophthongs were in all three quantity degrees (short (Q1), long (Q2) and overlong (Q3)), and those for diphthongs in Q2 and Q3 (as diphthongs cannot be in the short quantity degree). In order to elicit triphthongs, monosyllabic words where the triphthongs /uae, uʏe/ were expected to occur were included in the material, and spelled with *uae* and *uõe* respectively. As monosyllabic words in Estonian are considered to be in Q3, the triphthongs in the present study are only in Q3.

All test words were embedded in utterance initial and final positions of read sentences, which were written down in the Kihnu variety, e.g. *Kuõli läksid poesid* 'To school went the boys', *Poesid läksid kuõli* 'The boys went to school'. Each sentence was printed on a separate card. In total there were 153 sentences which contained 165 test vowels.

The data was recorded from six female speakers aged 23–42 years (average age 37) who are all native speakers of the Kihnu variety. All subjects were born in Kihnu and have at least one parent from Kihnu. Five speakers have higher education and one was a university student at the time of recording. Mainly in connection with their studies they have all lived outside Kihnu. The time spent elsewhere in Estonia varied from a couple of years to 10 years.

The subjects were presented with the pack of test cards where the sentences appeared in random order. The recordings were carried out in quiet settings on the Island of Kihnu using an Edirol R-09HR digital recorder. Not all test tokens could be used in the analysis because of the background noise, whisper, or other reasons. In total, the final set of materials consisted of 939 test words, the distribution of which can be seen in Table 1.

Table 1

The distribution of the analyzed tokens according to the three quantities (Q1, Q2, Q3)

	Q1	Q2	Q3
Monophthongs	225	163	121
Diphthongs	–	176	224
Triphthongs	–	–	30

The data were analyzed with the speech analysis software Praat (Boersma, Weenink 2011). In this paper, for the first time for Estonian, the method of formant dynamics was used, i.e. rather than measuring a single point in the middle of the steady state of the vowel, equidistant temporal points were used to characterize the formant trajectory within the whole vocalic part. The boundaries of all vocalic segments were marked by hand. A script was used to calculate the total duration of each vocalic segment and divide it into ten equal intervals (see Figure 1). The frequencies of F1 and F2 were automatically measured and checked manually for inconsistencies. Thus, for each formant a contour consisting of nine values was obtained. For plotting the trajectories, the values of the first and last point were left out so as to minimize the effects of formant transitions associated with the adjacent consonants.

Additionally, in order to track more closely the formant frequency change over the course of vowels' duration in both monophthongs and diphthongs, trajectory length (TL) for each separate vowel section was calculated (cf. Jacewicz, Fox, Salmons 2011).

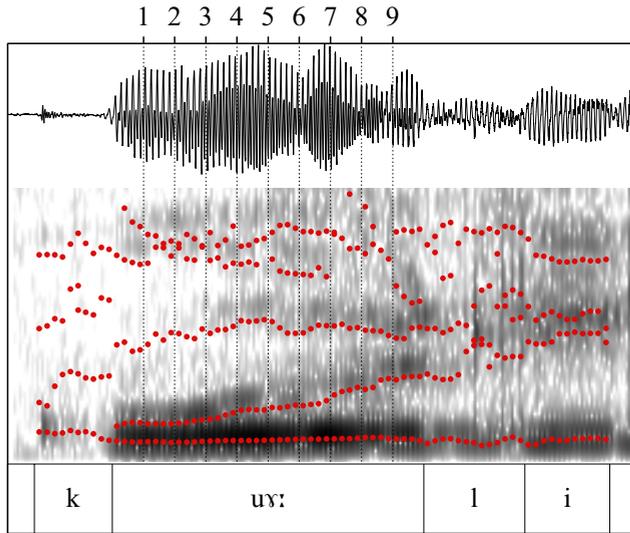


Figure 1. Spectrogram of the test word /kuɣ:li/ 'to school'.

3. Results and discussion

3.1. The acoustics of monophthongs

We will start the presentation of our results by looking at the quality of Kihnu monophthongs. We compared their values to earlier data from Standard Estonian based on spontaneous speech from 8 women in a comparable age group (Lippus 2011). Figure 2 plots all Kihnu monophthongs in the three quantity degrees in F1—F2 space. The mean values of 6 Standard Estonian monophthongs are marked with grey letters in italics. The general direction of the formant frequency change is indicated by arrows.

It can be observed that the trajectories for monophthongs are very short and are clustered around the target value of the vowel. It is perhaps striking that in Q1 the vowels /ɤ/ and /ø/ coincide in Kihnu. Our comparative Standard Estonian data did not include these vowels, but a similar result has also been shown for Standard Estonian where a short stressed /ɤ/ has shifted into the space of /ø/ (Eek, Meister 1998 : 229). Overall, Kihnu monophthongs are similar to those of Standard Estonian. In both varieties, short monophthongs are more centralized than the long ones, and there is not much difference between the monophthongs in Q2 and Q3. The two vowel systems differ only in the realization of /æ/ and /i/. In Q1 and Q3, the Kihnu /æ/ is more reduced and higher than the Standard Estonian /æ/. Long /i/ in Standard Estonian Q2 and Q3 is more peripheral than its Kihnu counterpart.

3.2. The quality of diphthong components

In order to study the quality of Kihnu diphthongs, the target values of the diphthong components were compared to the corresponding monophthongs. Figure 3 displays the six diphthongs in Q2 and Q3 in F1—F2 vowel space. The mean values of Kihnu monophthongs are marked with grey letters. It can be seen that the diphthongs have longer trajectories than the monophthongs in Figure 2. All diphthongs are acoustically clearly realized. The trajectories for diphthongs in Q2 and Q3 are very similar implying that there is not much difference between the quality of the diphthongs in the two quantities.

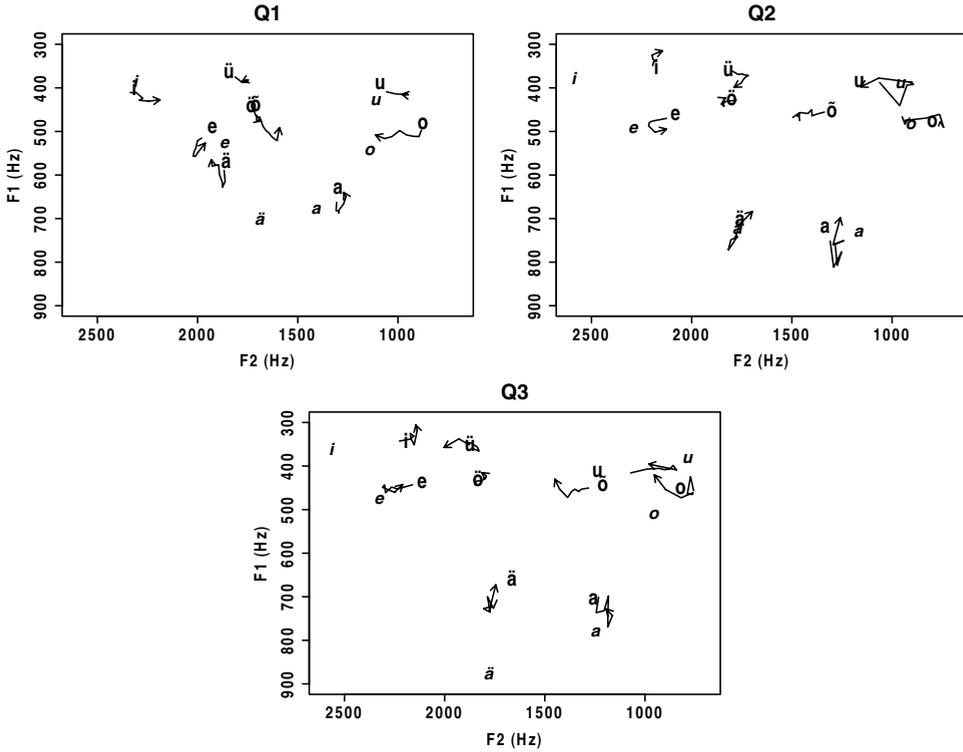


Figure 2. Monophthongs in three quantity degrees (Q1, Q2, Q3) in F1—F2 vowel space. Mean values of six Standard Estonian vowels are marked with bold letters.

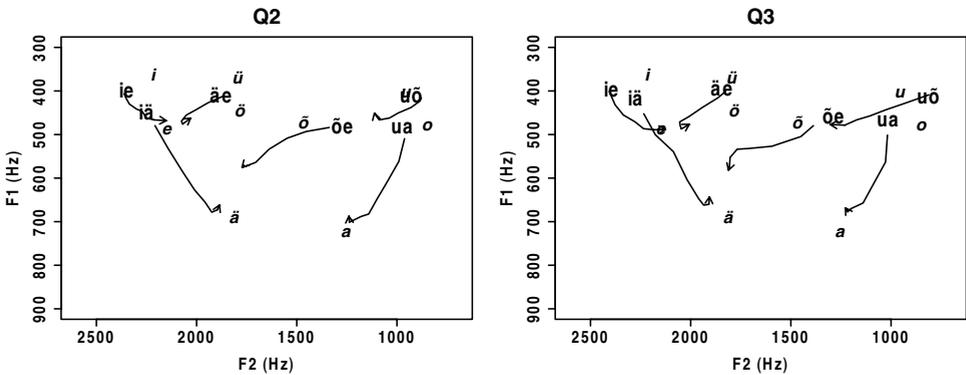


Figure 3. Diphthongs in two quantity degrees (Q2, Q3) in F1—F2 vowel space. Mean values of Kihnu monophthongs are marked with bold letters.

The comparison of the diphthong components with monophthongs shows that the target values of the diphthongs are close to the corresponding monophthong values, which was expected in our first hypothesis. The realization of similar diphthong components is, however, affected by the other component. Thus, /e/ in /ye/ and /æ/ is realized differently, being lower and further back in /æ/. Also, /i/ in /ie/ and /iæ/ differs in quality and is realized closer to /e/ in /iæ/. The first component of /uæ/ and /ua/ is closer to /o/ in the latter diphthong. Analogous results for Standard Estonian were obtained by Piir (1985) who showed that similar first or second components have different values. Also, a study of the diphthongs

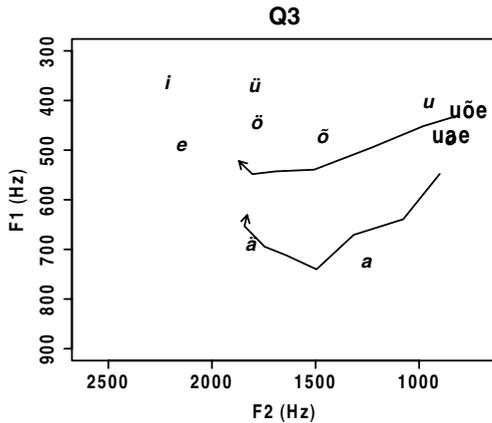


Figure 4. Triphthongs in F1–F2 vowel space. Mean values of Kihnu monophthongs are marked with bold letters.

/ai, ei, ui/ in the second syllable of disyllabic Standard Estonian words (Teras 1996) demonstrated strong coarticulatory influences between the diphthong components.

3.3. The acoustics of triphthongs

Before examining the durational characteristics of the diphthongs we will test our third hypothesis concerning the realization of triphthongs. The trajectories for the triphthongs are plotted in Figure 4. It can be seen that the two postulated triphthongs in Kihnu Estonian are indeed acoustically realized as such. They have visibly longer trajectories, traversing clearly three different vocalic qualities, as compared to the diphthongs in Figure 3. It can be seen, though, that the trajectory for /uae/ is lower than what we would have expected based on dialectal transcriptions of this triphthong. Its first component starts below /o/, moves to /a/ and ends just above /æ/ without reaching /e/. Thus, the realization of this triphthong is closer to /oaæ/ than /uae/, which means that the controversial transcription *uä* used since Saareste (1920) captures its actual quality relatively well.

3.4. Durational characteristics

3.4.1. Trajectory length

We will now look more closely at the measures trajectory length (TL) and vowel section length (VSL). For calculating TL we adjusted the formula from Jacewicz, Fox, Salmons 2011. This was done because the calculation of TL using the original formula (based on the sum of vowel sections using all 9 measurement points in the vocalic segment) did not separate monophthongs and diphthongs as clearly as would be expected from Figures 2 and 3. Long monophthongs, in particular /o/ and /u/ and sometimes /a/, have a relatively long trajectory, which spans about 2–3 Barks. For other vowels, this trajectory is within 1 Bark. The movement of the trajectory is, however, different in monophthongs and diphthongs. Diphthongs (and also triphthongs) have trajectories which move from one target value to another, but for monophthongs the trajectory is not unidirectional but zigzags around the target value of the vowel. Thus, it was considered optimal to use only two points in each vocalic segment, and the trajectory length was calculated on the basis of formant values at 20% and 80% points of the vocalic segment using the formula:

$$TL = \sqrt{(F1_{20\%} - F1_{80\%})^2 + (F2_{20\%} - F2_{80\%})^2} \quad (1)$$

Figure 5 presents the boxplots for the trajectory length of Kihnu monophthongs, diphthongs and triphthongs. The three types are clearly separated [$F(2, 238) = 302.61$; $p < 0.001$]. An ANOVA also showed a significant main effect of quantity [$F(2, 238)$

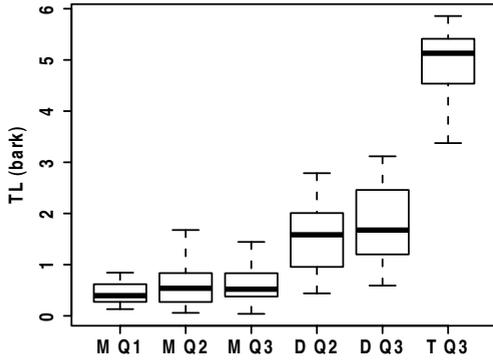


Figure 5. Boxplots for the trajectory length (TL) measured on the basis of formant values at 20% and 80% points of the vocalic segment.

= 67.249; $p < 0.001$], although this could be explained by unbalanced data: diphthongs cannot be in Q1 and triphthongs only occurred in Q3. The post-hoc test showed no significant difference between the realization of monophthongs in the three quantity degrees or diphthongs in the two quantity degrees.

The boxplots for vowel duration are presented in Figure 6. The trajectory length is not connected with the duration of the vowel, which is dependent on the phonological quantity of the word rather than whether the vowel is a monophthong, diphthong or a triphthong. An ANOVA showed significant main effects of both quantity [$F(2, 238) = 224.65$; $p < 0.001$] and vowel type [$F(2, 238) = 57.011$; $p < 0.001$], but post-hoc tests confirmed that the difference was between short (Q1) and long (Q2, Q3) vowels ($p < 0.001$) on the one hand, and between Q2 and Q3 vowels ($p < 0.05$) on the other hand.

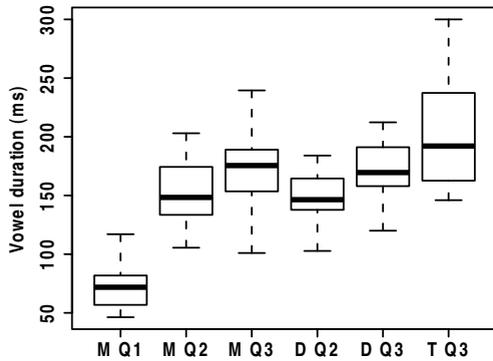


Figure 6. Boxplots for vowel duration in monophthongs, diphthongs and triphthongs.

There was no difference between monophthongs and diphthongs in the same quantity degree, but the triphthongs were longer than the monophthongs and diphthongs in Q3 words ($p < 0.05$). The latter might, however, be an effect of word structure as the triphthongs occurred only in monosyllabic words.

3.4.2 Vowel section length

The length of each vowel section (VSL) was calculated using the following formula where $n = 8$:

$$VSL_n = \sqrt{(F1_n - F1_{n+1})^2 + (F2_n - F2_{n+1})^2} \quad (2)$$

Figure 7 displays the VSL values for monophthongs (left panel), and diphthongs and triphthongs (right panel). Each contour consists of 8 points corresponding to 8 sections calculated on the basis of 9 measured segments. We can see that in the case of monophthongs the VSL contour is bowl-shaped, which implies that bigger changes

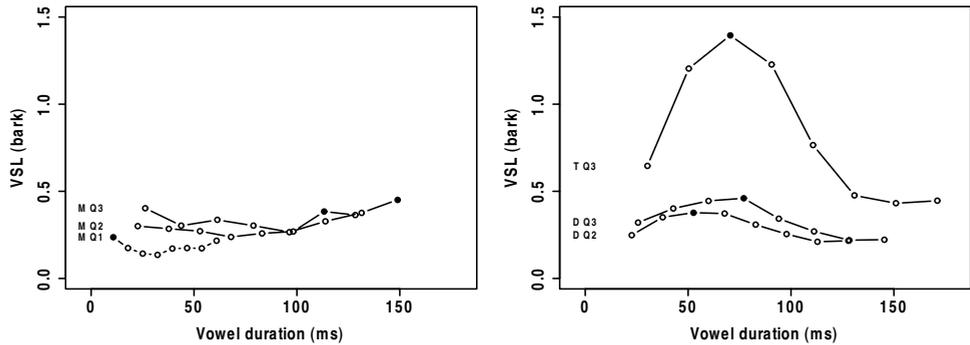


Figure 7. The VSL contours for monophthongs (left panel), and diphthongs and triphthongs (right panel). The vowel sections with the biggest change are marked with filled circles.

take place in the beginning and end of the vowel (i.e. transitions from and to the consonant), while the middle part is relatively stable. The VSL contours for diphthongs and triphthongs are hat-shaped; the more stable parts are in the beginning and in the end, and the transition from one target value to another takes place in the middle.

We were interested in the duration of diphthong components in Q2 and Q3. Based on earlier findings from Lehiste (1970) we expected the duration of both components to be longer in Q3. Rather than trying to divide the diphthongs into components by hand, we located the vowel section with maximal change in formant values, taking it to mark the most objective boundary between the two diphthong components. In Figure 7, the sections with maximal change are marked with filled circles. It can be seen that the transition from one target value to the next occurs later in Q3 diphthongs, which means that the first component of a diphthong is longer in Q3. Thus, our results lend support to Lehiste's work (1970), and are not in line with the view expressed in Eek (2008) according to which the first component is similar in both quantities. Similar to Piir (1985) our results show that both in Q2 and Q3 diphthongs, the first component has a shorter duration than the second one.

In the case of triphthongs we would expect to see a two-peaked VSL contour which would traverse three target values and two transitions. In Figure 7 we can see that this not the case, as there is only one clear peak. Nevertheless, the trajectory of the triphthongs is different from those of the diphthongs; the triphthongs are considerably more dynamic than the diphthongs.

4. Conclusions

The main aim of this paper was to study the acoustics of six Kihnu diphthongs which have arisen as a result of diphthongization of long open and mid vowels. The analysis used the method of formant dynamics. Three hypotheses were tested.

Firstly, the results showed that the diphthong components are close in their quality to the corresponding monophthongs. The realization of similar diphthong components was, however, shown to be affected by the other component. Coarticulatory influences between the diphthong components were evident in the case of /ua/ and /iæ/ where the first components were lower than the target values. Also, the second component /e/ was lower and further back in /e/ than in /ie/. A comparison of Kihnu monophthongs with Standard Estonian showed that in Q1 and Q3, the Kihnu /æ/ is more reduced and higher than the Standard Estonian /æ/. Long /i/ in Standard Estonian Q2 and Q3 is more peripheral than its Kihnu counterpart.

Secondly, using the measure vowel section length we showed that the first diphthong component is affected by the phonological quantity. Both components were longer in Q3 diphthongs than in Q2 diphthongs.

Thirdly, the results of the present study demonstrated that there is an acoustic basis for postulating two triphthongs in Kihnu. The trajectory length measure proved to be useful in separating monophthongs, diphthongs and triphthongs.

Acknowledgements

We would like to thank Külli Laos for her help with test sentences and all our subjects. Research presented in this paper was supported by 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

Pärtel Lippus

Institute of Estonian and General Linguistics, University of Tartu

E-mail: partel.lippus@ut.ee

Ellen Niit

Institute of Estonian and General Linguistics, University of Tartu

E-mail: ellen.niit@ut.ee

Helen Türk

Institute of Estonian and General Linguistics, University of Tartu

E-mail: helen.tyrk@gmail.com

REFERENCES

- Ariste, P. 1953, *Eesti keele foneetika*, Tallinn.
- Boersma, P., Weenink, D. 2011, Praat: doing phonetics by computer [Computer program]. Version 5.3. <http://www.praat.org/>.
- Eek, A. 2008, *Eesti keele foneetika I*, [Tallinn].
- Eek, A., Meister, E. 1998, Quality of Standard Estonian Vowels in Stressed and Unstressed Syllables of the Feet in Three Distinctive Quantity Degrees. — *LU XXXIV*, 226–233.
- Grigorjev, P., Keevallik, 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.
- Jacewicz, E., Fox, R. A., Salmon, J. 2011, Vowel Change across Three Age Groups of Speakers in Three Regional Varieties of American English. — *Journal of Phonetics* 39, 683–693.
- Laos, K. 2010, *Kihnlasõ emäkiel*, Pärnu.
- Lehiste, I. 1970, Diphthongs versus Vowel Sequences in Estonian. — *Proceedings of the Sixth International Congress of Phonetic Sciences, Held at Prague 7–13 September 1967*, Prague, 539–544.
- Lippus, P. 2011, The acoustic features and perception of the Estonian quantity system, Tartu (*Dissertationes Philologiae Estonicae Universitatis Tartuenssis* 29).
- Lonn, V., Niit, E. 2002, *Saarte murde tekstid*, Tallinn (*Eesti murded VII*).
- Pajusalu, K., Hennoste, T., Niit, E., Päll, P., Viikberg, J. 2009, *Eesti murded ja kohanimed*, Tallinn.
- Piir, H. 1985, Acoustics of the Estonian Diphthongs. — *Estonian Papers in Phonetics. EPP 1982–1983*, Tallinn, 5–103.

- S a a r, T. 1958, Kihnu murde häälikud (MS. Institute of the Estonian Language. The Archive of Estonian dialects and Finno-Ugric languages).
- S a a r, T., V a l m e t, A. 1997, Kihnu murrakust. — O. K õ i v a, I. R ü ü t e l, Vana kannel VII:1. Kihnu regilaulud, Tallinn, 35—40.
- S a a r e s t e, A. 1920, Häälikajalooline uurimus Kihnu murdest. Konsonandid. H-98 (MS. The University of Tartu Archives of Estonian Dialects and Kindred Languages).
- S a n g, J. 2009, Ühest fonotaktilisest kollisioonist (Kihnu näitel). — KK, 809—817.
- T e r a s, P. 1996, Eesti keele teise silbi diftongide akustikast. Lõputöö, Tartu (MS. University of Tartu).
- T ü r k, H. 2010, Kihnu murraku vokaalidest. Bakalaureusetöö, Tartu (MS. University of Tartu).
- V i i t s o, T.-R. 2003, Rise and Development of the Estonian Language. — Estonian Language, Tallinn (Linguistica Uralica, Supplementary Series 1), 130—230.

ЭВА ЛИЙНА АСУ, ПЯРТЕЛ ЛИППУС, ЭЛЛЕН НИЙТ, ХЕЛЕН ТЮРК (Тарту)

АКУСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ МОНОФТОНГОВ И ДИФТОНГОВ В КИХНУСКОМ ГОВОРЕ ЭСТОНСКОГО ЯЗЫКА

В статье представлено первое акустическое исследование монофтонгов и дифтонгов кихнуского говора эстонского языка. Основное внимание уделяется шести дифтонгам, которые возникли в результате дифтонгизации долгих гласных нижнего и среднего подъема. Сопоставление компонентов дифтонгов с соответствующими монофтонгами показало их сходство. Некоторые существующие отличия объясняются влиянием эффекта коартикуляции между компонентами дифтонгов. В статье показано, что оба целевых гласных дифтонга в третьей степени долготы длиннее, чем во второй. Проведенный анализ также продемонстрировал, что с акустической точки зрения есть все основания постулировать существование в говоре двух трифтонгов. При сравнении монофтонгов, дифтонгов и трифтонгов эффективным параметром сопоставления оказалась длина формантных траекторий.