VILNIUS UNIVERSITY

EVALDAS ŠVAGERIS

ACOUSTIC FEATURES OF MONOPHTHONG TONES IN LITHUANIAN AND LATVIAN DIALECTS: COMPARATIVE ANALYSIS

Summary of Doctoral Dissertation

Humanities, Philology (04 H)

Vilnius, 2015

Dissertation was prepared at Vilnius University in 2011 – 2015.

Research supervisor:

Prof. Habil. Dr. **Bonifacas Stundžia** (Vilnius University, Humanities, Philology – 04H).

The dissertation will be defended at the Academic Council of Philology, Vilnius University:

Chair – Prof. habil. dr. **Danguolė Mikulėnienė** (Institute of Lithuanian Language, Humanities, Philology – 04H)

Members:

Doc. Dr. **Asta Leskauskaitė** (Institute of Lithuanian Language, Humanities, Philology – 04H)

Doc. Dr. **Jolita Urbanavičienė** (Institute of Lithuanian Language, Humanities, Philology – 04H)

Doc. Dr. **Lina Murinienė** (Lithuanian University of Educational Sciences, Humanities, Philology – 04H)

Doc. Dr. **Edmundas Trumpa** (University of Latvia, Humanities, Philology – 04H)

The official defense of the dissertation will be held in the public meeting of the Council of Philological Sciences at 4 p.m. on January 14, 2016 in the .Kristijonas Donelaitis Auditorium, Faculty of Philology, Vilnius University.

Adress: 5 Universiteto st., Vilnius, Lithuania.

The summary of the doctoral dissertation was sent out on 14th December 2015.

The text of the dissertation is available at Vilnius University Library and on Vilnius University web page: *www.vu.lt/lt/naujienos/ivykiu-kalendorius*

VILNIAUS UNIVERSITETAS

EVALDAS ŠVAGERIS

LIETUVIŲ IR LATVIŲ TARMIŲ MONOFTONGŲ PRIEGAIDŽIŲ AKUSTINIAI POŽYMIAI: LYGINAMOJI ANALIZĖ

Daktaro disertacijos santrauka

Humanitariniai mokslai, filologija (04 H)

Vilnius, 2015

Disertacija rengta 2011 – 2015 metais Vilniaus universitete.

Mokslinis vadovas:

Prof. habil. dr. **Bonifacas Stundžia** (Vilniaus universitetas, humanitariniai mokslai, filologija – 04H)

Disertacija ginama Vilniaus universiteto Filologijos mokslo krypties taryboje:

Pirmininkė – Prof. habil. dr. **Danguolė Mikulėnienė** (Lietuvių kalbos institutas, humanitariniai mokslai, filologija – 04H)

Nariai:

Doc. dr. **Asta Leskauskaitė** (Lietuvių kalbos institutas, humanitariniai mokslai, filologija – 04H)

Doc. dr. **Jolita Urbanavičienė** (Lietuvių kalbos institutas, humanitariniai mokslai, filologija – 04H)

Doc. dr. **Lina Murinienė** (Lietuvos edukologijos universitetas, humanitariniai mokslai, filologija – 04H)

Doc. dr. **Edmundas Trumpa** (Latvijos universitetas, humanitariniai mokslai, filologija – 04H)

Disertacija bus ginama viešame Filologijos mokslo krypties tarybos posėdyje 2016 m. sausio mėn. 14 d. 16 val. Filologijos fakulteto K. Donelaičio auditorijoje.

Adresas: Universiteto g. 5, Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2015 m. gruodžio mėn. 14 d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: *www.vu.lt/lt/naujienos/ivykiu-kalendorius*

INTRODUCTION

Dissertation object and relevance of the topic

The **object** of the present dissertation is monophthong tones in Lithuanian and Latvian dialects and their acoustic features. The initial intention was to encompass a larger number of dialects, the abundance of experimental material made me revise it and limit the scope of the work to two of them, namely that of the Telšiai subdialect, north-central Samogitia, and the Vidzeme range of the central Latvian dialect (the Valmiera subdialect). The choice of dialects is due to several reasons. First, the preference was to research and compare those dialects in which, according to the work of other researchers in the field to date, the tones are well distinguished. We can expect an analysis to be productive only if we fulfil this criterion (otherwise there are no objective grounds for talking about acoustic similarities and differences of tones in Lithuanian and Latvian dialects). Second, the Latvian Valmiera subdialect belongs to an area with the so-called threefold tonal distinction system. This threefold system (the broken, sustained and falling tones) has become the conventional reference point for describing the tonal systems of other Latvian dialects (many of which have undergone a transformation form the threefold to a twofold system). This system has also been traditionally used for identifying the relations between tones in Lithuanian vs. Latvian (according to Endzelin's law; see Endzelīns 1951: 35–41)[1].

There is indeed no lack of experimental research for determining prosodic features of Latvian and, especially, Lithuanian dialects; however, there have been no comparative works to date which, using a single methodology, analyse and compare the acoustic tonal base of living Baltic languages, identifying the similarities and differences between them. The multitude of interpretations, which has accumulated through a considerable amount of time, of the acoustic nature of tones serves as a good indicator of the relevance of the issue in question, i.e. what are the most universal features of tones. Once this work has been done, one can compare the tones of Baltic languages with respective prosodic units in other languages. One problematic issue at hand is the distinction of tones on long vowels (part of the charge of distinguishing mixed diphthongs and diphthong tones rests on qualitative features), and the attention

[1] For exceptions to this law see Stundžia 1985: 142–151.

here is therefore directed exclusively at this subgroup of vocalism. Georg Gerullis, as far back as the middle of the previous century, mentioned a "wave" of gradual decrease in differences between tones (primarily on the long vowels), spreading from eastern to western Lithuania (Gerullis 1930), and it is therefore important to determine the extent and the nature of the progression of this linguistic phenomenon. We should also verify the issue of the functioning of the threefold tonal system in Valmiera today (i.e. whether or not we can say that all three tones are still distinguished). Furthermore, in an age of continuously evolving technology, the data gathered in this work could also be both relevant and useful. Once the universal features of tones are identified, their acoustic specifications could be of use in improving speech recognition and synthesising technologies.

Dissertation aim and objectives

The **aim** of the present work is to analyse, compare, and describe the acoustic features of monophthong tones in Latvian (central) and Lithuanian (Telšiai subdialect, northern Samogitia) dialects.

Objectives:

1. To provide a detailed review of the existing research into tones of northern Samogitian and central Latvian dialects;
2. To collect experimental material, process it, and prepare for analysis (to arrange dialectological field trips, record data, etc.);
3. To propose, motivate, and apply a technical data analysis method (for automated parameter generation) that is up to the standard of today's experimental phonetics;
4. To analyse parameter groups of the duration of long vowels, pitch, and intensity, also identifying their degree of correlation and their relationship with tones;
5. To graphically illustrate all distinctive features of long stressed vowels possibly implicated by their tones, also verifying the probability of their differences using statistical criteria;
6. To identify the acoustic similarities and differences between tones in Lithuanian and Latvian dialects.

Defended theses

1. Tones in Lithuanian and Latvian languages differ acoustically in combinations of pitch and length.
2. There is a relationship of inverse proportionality between the rate of pitch change and vowel duration. The shorter the vowel, the more abrupt the (relative) change of pitch; conversely, the longer the vowel, the slower the relative change of pitch.
3. The most reliable technical-mathematical parameter for identifying tones is the pitch sustainability coefficient (the value of relative pitch change multiplied by the duration of the vowel);
4. The pitch contour of tones, first and foremost, indicates differences in pitch sustainability levels of the vowel rather than melodic modulations.

Dissertation structure

The dissertation comprises:

1. Introduction. It identifies the object of the present dissertation, while also defining its aim and objectives and emphasizing the relevance of the topic.
2. Dissertation methodology section. It introduces the methodology chosen and the technical tools for data analysis.
3. Empirical part. It comprises an overview of the existing research (separate parts for each dialect) and three main sections corresponding to the groups of prosodic features, namely duration, pitch, and intensity.
4. Conclusions.
5. List of references and appendices.

Approbation of dissertation

Published **articles** on the topic of the dissertation:

1. Švageris, Evaldas 2013, Monoftongų priegaidžių skirtumai jaunimo kalboje, *Jaunimo kirčiavimo polinkiai: bendrinės kalbos normos ir vartosenos tendencijos* (straipsnių rinkinys), 77–88: VDU.
2. Švageris, Evaldas 2014, Tone Features in the Latvian Dialect in Lithuania, (*I W o B A V I I I (2 0 1 2) Proceedings of the 8 th international workshop on Balto-Slavic accentology*) (straipsnių rinkinys), 355-366.
3. Švageris, Evaldas 2015, Akustinių požymių koreliacijos svarba lietuvių kalbos priegaidžių skyrimui, *Baltistica* 50 (1), 89–111.

Talks given at international conferences:

1. *Acoustic features of tone in Latvian dialect of Lithuania*, IWoBA VIII (International Workshop on Balto-Slavic Accentology), 6–8 July 2012, Novi Sad, Serbia.
2. *Eksperimentinių prozodijos tyrimų naujovės: šiaurinių žemaičių telšiškių priegaidžių analizė*, Šiuolaikinės kalbos tyrimai ir problemos (21-oji tarptautinė Jono Jablonskio konferencija), 3 October 2014, Vilnius.
3. *Akustinių požymių koreliacijos svarba lietuvių kalbos priegaidžių skyrimui*, Baltų kalbų fonetika ir fonologija (tarptautinė mokslinė konferencija profesoriui Aleksui Girdeniui (1936–2011) atminti), 30–31 October 2014, Vilnius.
4. *Lietuvių, latvių ir slovėnų kalbų priegaidžių akustiniai panašumai,* 12th International Congress of Balticists, 28–31 October 2015, Vilnius.

The phonological status of tone

Typologically, Lithuanian and Latvian are tone languages[2] (Pike 1948: 8; Jakobson, Fant, Halle 1962: 13; Halle, Vergnaud 1987: 190–196, Blevins 1993: 237–273, Yip 2007: 230 et. al.), more precisely they are contour tone languages (Pike 1948: 5–17, Zhang 2002: 3–8, Rinkevičius 2015: 17–20). A given language is considered a tone language when lexical meanings are differentiated through features of pitch (at least on the morphemic level) (Hyman 2001)[3], and it is a contour tone language if pitch differentiates grammatical or lexical meanings on the syllabic level (Zang 2002: 3)[4]. The present dissertation is focused exclusively on the acoustic (phonetic) base of long vowels, and their phonological status (i.e. the power, *ceteris paribus*, to differentiate between words) is not questioned in the dialects under analysis (even though we could do so when describing e.g. northern Samogitian; see Remenytė 1990: 60–78; Mažiulienė 1996: 110). The dephonologization of monophthong tones in eastern Lithuanian dialects has long attracted scholarly attention (perhaps the clearest account of it is found in the work of Jonas Kazlauskas; see Kazlauskas 1966: 60–78; similar ideas are also found in the most recent scholarship, see Rinkevičius 2015: 23–24), with there also being some experimental research that attempted to rehabilitate the functionality of the monophthong tone system in this particular area (Kosienė 1982: 61–71; see also Garšva 2003, 2008). The phonological conception of tone in Lithuanian vs. Latvian is almost identical. It encompasses pitch modulations, with a contrastive function, intensity and duration (and, in the case of diphthongs, also of quality of sounds) of the long syllable with the main stress (Pakerys 1982: 147, Rudzīte 1993: 99, Girdenis 2003: 50; for an overview of the research in the field, see Bacevičiūtė 2009: 17–29).

---

[2]According to various data, tone languages make up 70 % (Yip 2007: 229) or 50 % of the world's languages (Hyman 2011: 198). It should, however, be noted that there are no data in"The World Atlas of Language Structures" (wals.info) about Lithuanian tones.

[3]The exact definition is as follows: 'A language with tone is one in which an indication of pitch enters into the lexical realization of at least some morphemes' (Hyman 2001).

[4]Original definition: 'In some tone languages, the contrastive functions of pitch are sometimes played by pitch *changes* within a syllable. Pitch changes of this kind are called contour tones' (Zhang 2002: 3).

Computer-assisted processing of data

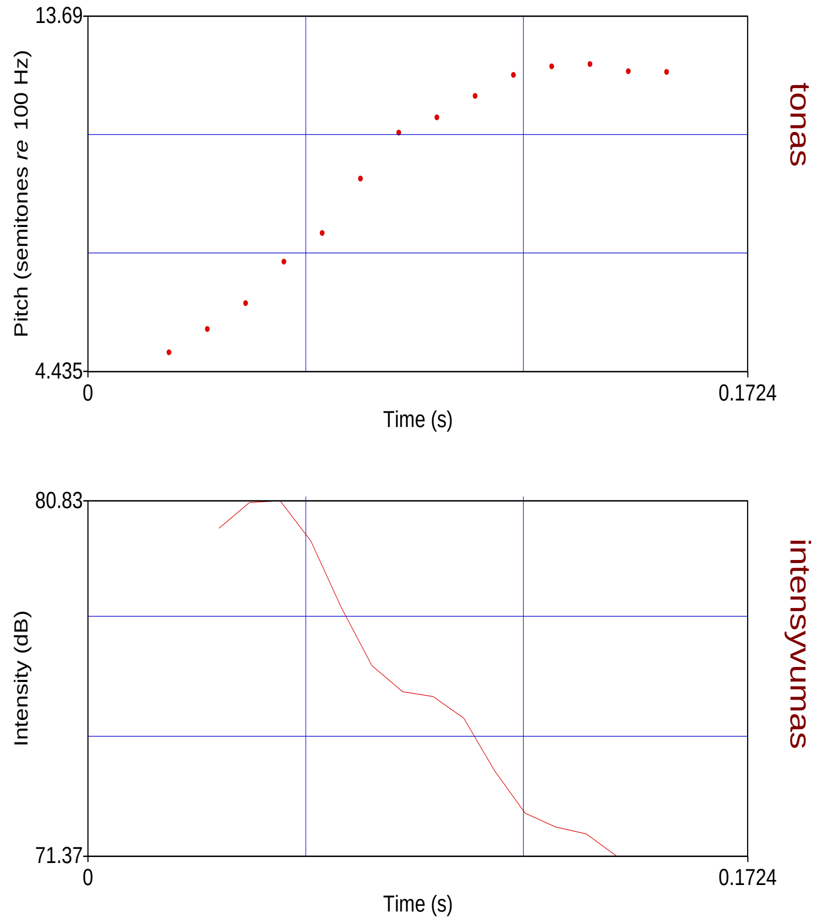
The main tool for data generation used in the research was Praat, a software package for acoustic analysis. This software is well-known and widely used by phoneticians worldwide (including Lithuania).[5] The suggestion to create a script, which is a set of specifically selected functions of *Praat* software, and thus to optimise the calculation of parameters and save a lot of time normally spent on routine work came from Marius Tverijonas, an IT specialist at Vilnius University. By working hand in hand (with me choosing the required groups of parameters and Marius Tverijonas doing the programming), we successfully implemented this idea. Once we delimit the vowel analysed (the example above concerns the acute front vowel in *dýgs*)[6], this script generates the data window shown in Figure 1 (see Figure 1). This window also contains the parameters of all three prosodic features displayed in separate rows, namely duration (row 1), pitch (row 2), and intensity (row 3). Such a display of data is convenient for seeking out any correlations between parameters (i.e. their interdependence). The pitch (here the Hertz scale has been converted to a logarithmic semitone scale) and intensity drop graphs are displayed separately (with the upper rectangle containing the pitch graph, and the lower, the intensity drop graph). The parallel lines divide the curves into three equal parts (along both vertical and horizontal axes). Each of these parts bears the following labels: A1 (initial part of the vowel), A2 (medial part of the vowel), and A3 (final part of the vowel). The first row of data displays a single parameter, namely the absolute duration of the vowel (measured in seconds; in the example above, the duration of the acute vowel is 0.1724 s).

[5]It can be downloaded free of charge from here: http://www.fon.hum.uva.nl/praat/.

[6] It should be noted that the data window displays the parameters of the long vowel only, rather than those of the whole word (the filename 'Acoustic analysis of file dygs3akūtas.wav might be somewhat misleading).

Figure 1. Praat script window[7]

Failo dygs3akūtas.wav akustinė analizė



Trukmė

Absoliučioji trukmė

0.1724

Pagrindinis tonas

Diapazonas(pt)	Maks. taškas(pt)	Maks. taško vieta balsyje	Vid. lygis(pt)	Staigumas(pt/s)	Viršūnių skaičius	A1	A2	A3	A1L	A2L	A3L	A1S	A2S	A3S
7.51	12.44	0.75	9.93	59.20	1	▲ /	▲ /	▲▼	6.07	10.49	12.31	74.60	82.04	13.41

Intensyvumas

Diapazonas(dB)	Maks. taškas(dB)	Maks. taško vieta balsyje	Vid. lygis(dB)	Staigumas(dB/s)	Viršūnių skaičius	A1	A2	A3	A1L	A2L	A3L	A1S	A2S	A3S
9.57	80.94	0.29	76.00	96.88	1	▲ /	▼ /	▼	80.48	75.92	71.94	23.43	128.55	58.05

Koreliacija

Maksimumo taškų koreliacija (proc.)

53.60

* A1L - atkarpos A1 vidutinis lygis, A1S - atkarpos A1 staigumas ir t.t.

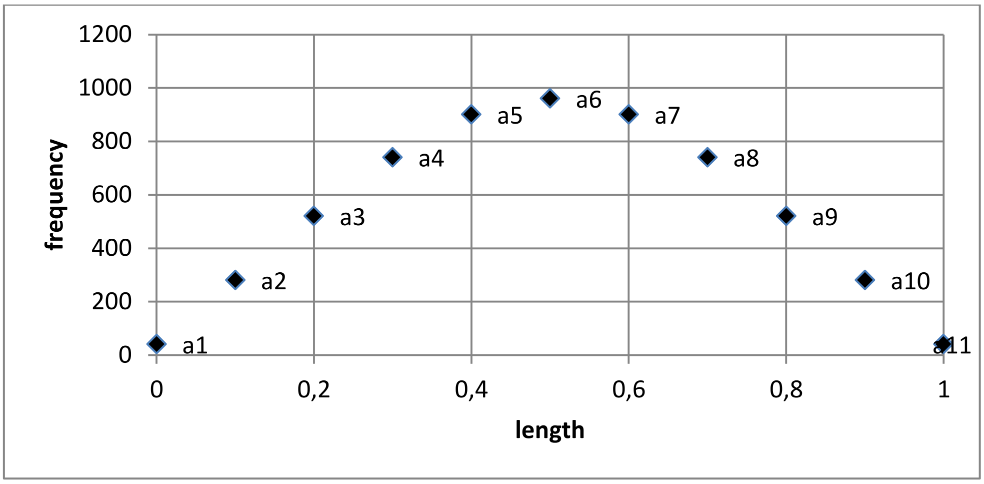
[7] Table symbols: Failo *** akustine analize – Acoustic analysis of file ***, Tonas – Pitch, Intensyvumas – Intensity, Trukme – Duration, Absoliuti trukme – Absolute duration, Pagrindinis tonas – pitch, Diapazonas – Range, Maks. Tasko vieta balsyje – Pitch peak, Vid. Lygis – mean frequency/intensity, Staigumas – rate of change, Virsuniu skaicius – Number of peaks, Koreliacija – Correlation, Maksimumo tasku koreliacija – Peak correlation, Proc. – per cent, Atkarpos *** vidutinis ilgis – Average length of segment ***, Atkarpos *** staigumas – rate of change of segment ***.

This dissertation also makes use of relative duration measurements. A formula for calculating the ratio between the duration of the vowel and the duration of the word could not be included in the script[8] (due to impossibility to include in to the script duration of entire word), so this parameter had to be calculated separately. The second (pitch) and third (intensity) rows contain the same parameters (from left to right): range (i.e. the interval between the highest and the lowest frequency of the vowel, and, respectively, between intensity levels), peak (the highest recorded frequency of the vowel in question (ht) and the level of intensity (dB)), its position in the vowel (the beginning of the vowel is conventionally given the value of 0, and the end 1),[9] mean frequency/intensity (ht and dB), average rate of pitch change (ht/s and dB/s), number of peaks, and pitch contours. The mean frequency and mean intensity of all three parts of the vowel (A1L, A2L, and A3L) and the average rate of pitch change (A1S, A2S, and A3S) were calculated separately. The formulae for calculating the rate of pitch change, peaks, and modelling pitch contours require a more detailed explanation. To facilitate the understanding of the formula for the first of these, namely the average rate of pitch change, see the diagram below (Diagram 1). The vertical axis corresponds to the scale of frequencies (of the pitch), while the horizontal one the scale of duration (in this case, please disregard the exact values). The diagram contains an ordinary curve that illustrates the change of pitch. Each of its points is assigned a label, starting from the beginning of the sound: a1, a2, a3, a4, etc. Distances between these points with respect to the x-axis are identical, because the *Praat* script measures both the frequencies of the pitch and the levels of intensity in identical 0.01 s intervals. It should be noted that the duration and pitch range of the entire vowel do *not* figure in the calculation of this parameter (these values are not included in the formula).

[8]

[9] The closer the peak is to zero, the more close it is to the beginning of the vowel; the closer it is to one, the more close it is to the vowel's end. In the example above (see Figure 1), the pitch and intensity peaks are found in different parts of the vowel: the peaks of the pitch graph are within the final part of the vowel (0.75); whereas those of the voice intensity are within the initial part of the vowel (0.29).

Diagram 1. Calculating the average rate of pitch change



The formula used for calculating this parameter is as follows:

$$\frac{|a2 - a1| + \ldots + |a_n - a_{(n-1)}|}{a_n - 1} \Big/ 0.01s$$

This formula is primarily used for calculating the change between consecutive points on the contour (of sound frequency or intensity respectively). This is obtained from a sequence of subtractions, e.g. a2 – a1, a3 – a2, a4 – a3, etc. When the difference between the points is negative (which corresponds to the difference between two points on a downward curve, e.g. a8 – a7, a9 – a8, etc.), the modulus symbol (|) converts the negative value into a positive one, or, simply put, removes the minus sign (more important for our purposes is the rate of the curve's rise or fall, rather than its actual direction). The second step is to divide all these differences (between consecutive points with respect to the y-axis) by the number of intervals (where the first interval is a2 – a1, the second a3 – a2, etc.). The final mathematical operation is to derive the relationship between the obtained value and the momentary duration (0.01 s). Once this last calculation is completed, we learn what the curve's average rate of pitch change was with respect to the momentary duration (i.e. /0,01). In the example above (see Figure 1), it is stated (in the rate of change section) that the rate of the acute $\hat{r}$ pitch curve change was, on average, 59.2 ht/s (semitones per second), while the intensity change amounted to 96.88 dB/s (decibels per second). This dissertation also employs the shortened version of the formula (which skips the ratio with the momentary duration):

$$\frac{|a2 - a1| + \ldots + |a_n - a_{(n-1)}|}{a_n - 1}$$

The value obtained by this shortened version of the formula is to be called the simple average change between points of the curve (it had to be used in some cases for mathematical purposes as it was necessary to simplify the units of measurement). Both of these parameters, namely the average rate of pitch change with and without ratio with momentary duration are not that different after all (to put it bluntly, the decimal points is moved a few positions). If we leave out the calculation of the ratio with the momentary duration, we get a simple average change between points of the curve (in semitones or decibels). So if the average rate of the pitch change was at 59.2 ht/s, then, in order to obtain the simple average pitch change parameter (without ratio with duration), we have to move the decimal point two places to the left (i.e. to undo the mathematical calculation ~~0.592 ht / 0.01 s~~), obtaining 0.592 ht (we learn that the consecutive points of the curve differed, on average, by 0.592 semitones).

The *Praat* script calculated the peak of the curve if two conditions were met, namely if the direction of the curve had changed (from a rise to a fall) and the difference between two consecutive points had been greater than one semitone (i.e. the change was required to be > 1 ht). If the curves were rising or falling for the entire duration of the vowel, the generated number of peaks always equalled to 1. If only one of the conditions was fulfilled, the script did not calculate any peaks (this allowed us to prevent the programme from reacting to small fluctuations in the curve). Pitch contours were generated following exactly the same rules. Two symbols were employed for this purpose, ▲ and ▼. The first of them represents the rise of the pitch, and the second the fall. Forward slashes (see Figure 1) separate the parts of the vowel, namely the initial, medial, and final (i.e. A1 / A2 / A3). The row of symbols in the example above, ▲ / ▲ / ▲▼ (for the acute *r̂*), enables us to see that the pitch was rising for almost the entire vowel, with the direction changing only in the final part of it (in that latter section of the vowel the pitch was of the rising-falling variety). This way of depicting the pitch contour (using symbols) seems to be both more convenient and more universal. If one attempted to simply move all the curves from the *Praat* script to a single diagram, differences between ranges and durations would get in the way, making it much harder to discern any possible trends. Combinations of symbols ▲ and ▼ do not reflect the range and duration parameters and enable one to focus exclusively on the trends of the horizontal movement of the contour (rise, fall, rise-fall, etc.).

A separate row in the Praat script data window is given for another parameter, namely the correlation between pitch and intensity peaks (the *Correlation* section). This parameter, just as some others (such as peaks, mean frequency/intensity, number of peaks, parameters of separate sections of the vowel), did not undergo a detailed analysis in the dissertation (due to limitations of space). Suffice it to say, however, that the correlation parameter (in per cent) reflects the distance between peaks of the pitch contour on the one hand and those of the intensity contour on the other. The smaller the interval between the peaks, the closer the correlation parameter value is to 100 per cent.

This dissertation also employed another, derivative parameter (and it is therefore not found in the *Praat* script window), namely the sustainability coefficient. Very briefly, it is the ratio between range and average pitch change (not rate of pitch change / duration 0,01 s!) multiplied by the entire duration of the sound under analysis. This step was taken upon noticing correlation trends between certain parameters.

Statistical data analysis

The primary source of expertise for this analysis was theoretical scholarship in statistics (Kruopis 1993; Čekanavičius, Murauskas 2002; Johnson 2008; Gries 2009[10]). The distinctiveness of data under comparison was calculated using two criteria, namely the Wald–Wolfowitz and Student's test (or t-test). The former was offered by Vilnius University assistant professor and statistics expert Petras Skirmantas.[11] The Wald–Wolfowitz test was used to calculate the probability of difference between the sequences of data under comparison in cases when the samples comprised twenty or less variables (i.e. $n \leqslant 20$), and the variables were absolute (rather than mean) values. This test suited perfectly the small samples, which comprised vowel durations (both

[10] The latter two works were primarily intended for linguists.

[11] I am very grateful to Petras Skirmantas for valuable and in-depth consultation sessions.

relative and absolute) and pitch ranges. The calculation of the probability of finding sample differences (rejection of the zero hypothesis (H_0) and/or acceptance of the alternative hypothesis (H_1)) was based on the ratio of the numbers of sequences (both calculated one and critical). First, all variables are arranged in the ascending order (from lowest to highest) in the so-called line of variation (see Table 1).[12]

Table 1. Line of variation

A	A	A	A	C	A	C	C	C	A	C	C	C	C
0,1671	0,1691	0,1692	0,1724	0,1848	0,1867	0,1966	0,2072	0,2303	0,2408	0,2488	0,3354	0,3452	0,38855

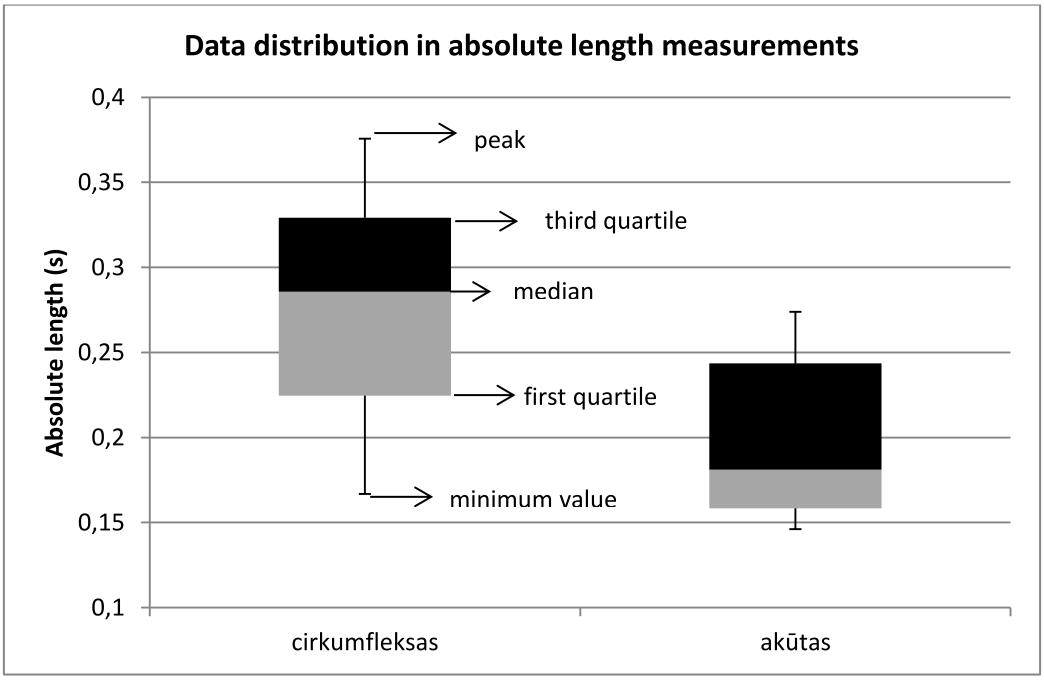
A sequence is a row of members of the same sample (prior to which and after which one finds members of another sample) (see Čekanavičius, Murauskas 2002: 25–26). In the illustrative table above, the sample of vowels representing the acute tone is marked A, and C stands for the sample of vowels representing the circumflex tone. The critical number of sequences, the exceedance of which would suggest that the risk of rejection of the zero hypothesis (H_0) (which supposes that the variables are part of the same general sample) is too high, is to be specified in the critical sequences data table (Ibid: 265). The exact critical number of sequences depends on the number of variables in each of the samples under comparison (i.e., the smaller the samples, the lower the critical number of sequences). The line of variation above contains calculations for 5 sequences, and the critical number of sequences for samples of such size is 3 (the sample of acute vowels comprises 6 variables, while that of circumflex vowels contains 8). Based on this ratio (5 > 3), we can say that the risk of the rejection of the zero hypothesis is too high (and we therefore have good reason to accept its alternative, namely that the data analysed *do not* belong to the same general sample).

The t-test was used for comparing larger samples ($n \geqslant 30$) which comprised mean values (such were the samples of average rate of pitch change and sustainability coefficient parameters). All the tools required for statistical calculations of this kind were found in the Excel 2010 software package (category: *Data analysis,* function: *t-test: two-sample assuming unequal variances*).

[12] The table contains samples comprising the minimal (absolute) vowel length parameters in the *dýgs* and *dỹks* minimal pair.

Some of the parameters (namely, both kinds of the duration and peaks location in the vowel and ranges of the pitch) were also compared graphically. Bar charts were used for this purpose (see Chart 1).[13] The coloured bars with little 'legs' on them illustrate the very same samples as in the line of variation (see Table 1). In this particular case, the chart illustrates the minimal samples for the *rũkst* and *rûkst* pair (these words were pronounced by one of the study participants from the Telšiai subdialect, northern Samogitia), comprising the calculations of the absolute duration. The rationale behind this choice is that it enables a more detailed analysis of the contrast between the data under comparison. Very often the differences between the data are only partial (just as the illustrative graph contains values that fall within the common interval). To be able to identify to what extent there is a contrast between the data, the samples had to be divided by means of their limit values, namely their peaks and minimum values, first and third quartiles, and medians.

Chart 1. A bar chart



The peak and the minimum value determine the so-called sample width (e.g., in Chart 1, the sample width of circumflex vowels is 0.1669–0.3756 s, while that of acute vowels is 0.1461–0.2739 s). The adjacent points (i.e. the minimum value and the first

[13] Their use in experimental work is not a new practice (see, e.g., Prehn 2011).

quartile, the first quartile and the median, etc.) encompass exactly one quarter of all sample values. The first quartile (also called the lower quartile) splits off the lowest 25% of data from the highest 75%, the median (or the second quartile) splits the sample in half (half of the variables in the sample are lower than or equal to the median, and the other half higher than or equal to the median). If the number of variables in the sample is odd, then the median, simply put, is the middle value of the sample.[14] The third quartile (or the upper quartile), in turn, splits off the lowest 75 % of data from the highest 25 %. So the 'legs', as well as the brighter and the darker parts of the bar each encompass 25 % of all the variables in the sample. The smaller these quarters of the data are, the narrower the interval for the variables to squeeze themselves in. The chart above shows that the first two quarters (encompassing the section between the minimum value and the median) of the acute vowels are narrower than those of the circumflex vowel. Their interval is relatively low, comprising 0.1461–0.1811 s. One should bear in mind that although these quarters are visually narrow, they nevertheless encompass one quarter of all variables in the sample (so one time out of two the acute vowels fell within the 0.1461–0.1811 s interval). Although some of the circumflex vowels also fall within this interval of duration (the 'leg' of their bar reaches the second quarter of the acute vowel sample), we have sufficient grounds to argue that the acute vowel tends to obtain shorter duration. The more fine-grained and subtle aspects of the graphic data analysis come to light, of course, only when we compare specific cases.

Acoustic features of monophthong tones in the Telšiai subdialect, northern Samogitia

Duration

The analysis of both types of duration (namely the absolute and the relative) enabled us to identify several key features of this acoustic parameter's range and contrastivity, which seem to be related to distinction of tones. First, we should note the

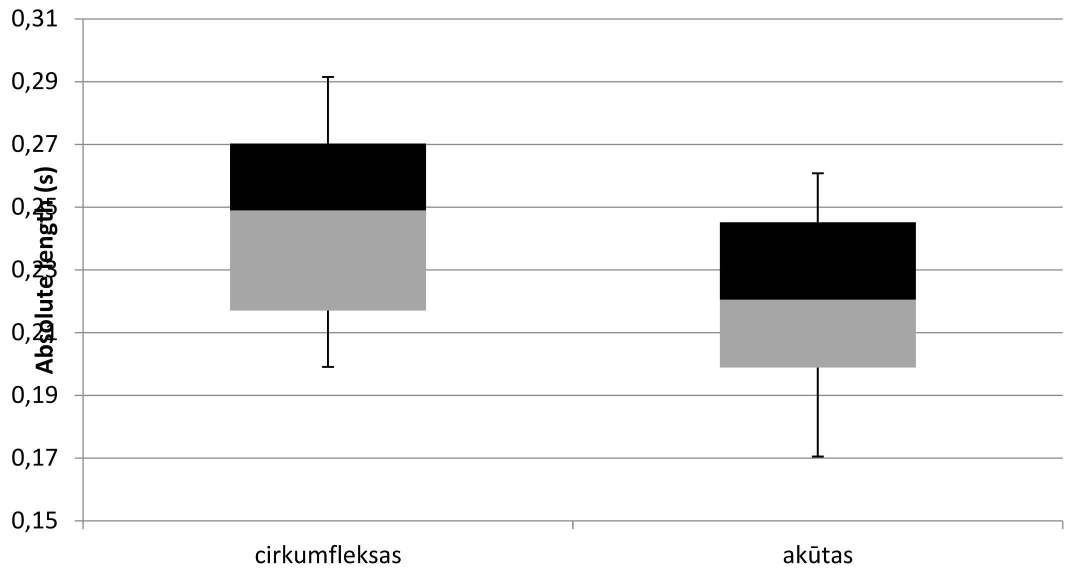
[14] The median, in contrast to the arithmetic mean, is more immune to maximum and minimum values of the sample (i.e. the disproportionately high or low value variables in the sample). The difference between the mean and the median in the same sample can sometimes be very considerable. E.g., if the sample, say comprises values 10, 20, 30, 40, 50, 60, 1000, then its mean would be 172 (the sum of variables divided by their number), while the median would equal 40 (10, 20, 30 < 40 < 50, 60, 1000, with three variables to the left and three to the right from the median).

contrast in the duration of long monophthongs between different speakers of the same dialect. Although a larger amount of data is necessary in order to fully establish this claim (i.e. a larger number of study participants), the results obtained strongly suggest that the tendency of one of the speakers to reduce the long vowels is not accidental (based on bar charts that illustrate data samples, and the results of statistical analysis). This viewpoint logically accords with the previous results obtained by scholars who researched the central north Samogitian dialect (for one, we should mention the experimental-sociolinguistic research by Remenytė–Mažiulienė)[15]. The tendency to reduce circumflex vowels is also suggested by the disappearing relationship between the vowel duration and its particular position in the word (i.e. its dependence on the position in the word, as well as the number and length of syllables and accents in it). The best indicator of this characteristic are three-syllable words (in one of the speakers, the vowels in question in three-syllable words were noticeably shorter than those in two-syllable words, while in another speaker speaking the same dialect this contract was absent). It should be stated unambiguously that the difference between circumflex and acute vowels relies on the data obtained from a single speaker (who was slightly older than most other speakers). The best indicator of this contrast is the consistent asymmetry between the limit values of data samples (minimal values and peaks, medians and first and third quartiles). This general trend is illustrated by Chart 2 (data of the minimal *kũor*ᵉ and *kûor*ᵉ pair. The central parts of the bar chart (i.e. 25 per cent of the values around each median) also serve to indicate that the acute vowels belong within the shorter-duration zone. Oftentimes these zones do not reach each other or the overlap is negligible. None of these features are shown in the data obtained from the second speaker. Presumably, the narrow intervals of duration do not provide necessary space for the quantitative features of vowels to be implemented.

---

[15] See Remenytė 1996: 49.

Chart 2. Distribution chart of the absolute duration data for the minimal pair *kũor*^{e} : *kûor*^{e}.

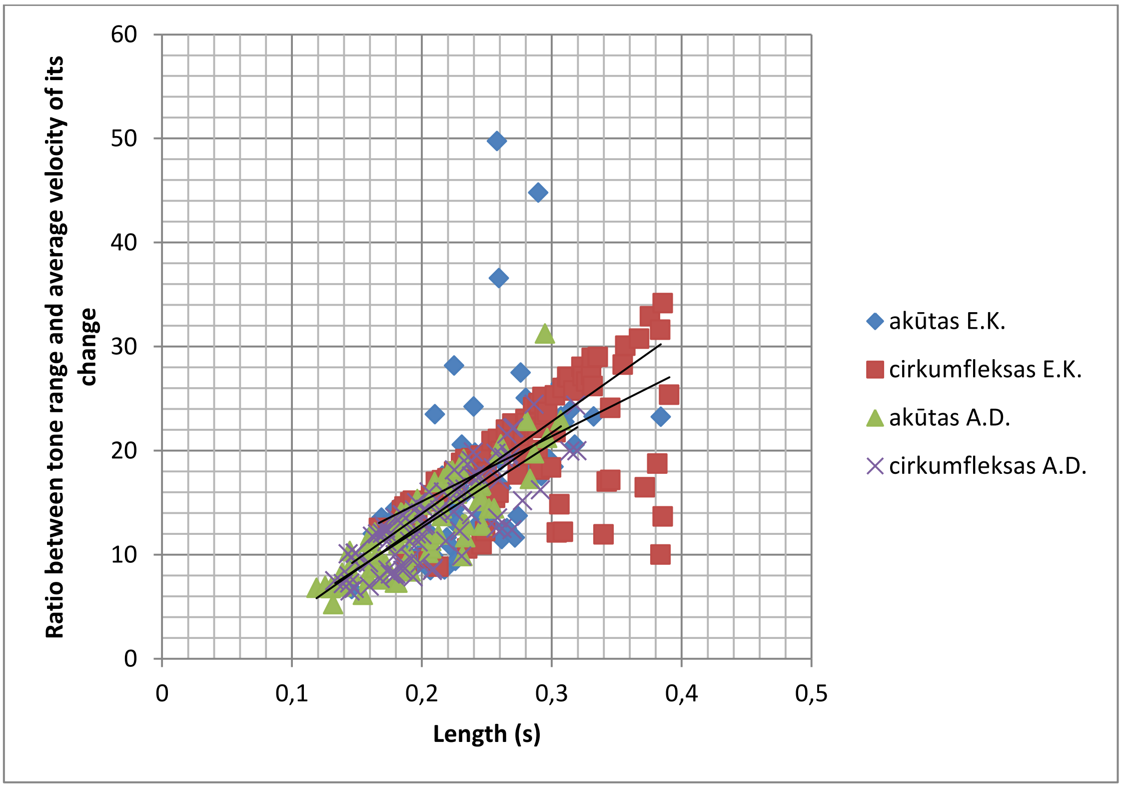


A different view is called for when one assesses the differences between the duration of circumflex and acute vowels based solely on the results of a statistical analysis. The threshold 95 % probability of difference was reached for only three pairs out of forty two pairs compared. The statistical basis, however, is insufficient to establish that members of all remaining pairs do not belong to a single general sample. One would have to say that there is no contrast between the duration of acute vs. circumflex vowels; nevertheless, the distribution of quantitative data and their insufficient statistical significance do not contradict each other (they indicate different aspects of the contrastivity of data). The smallest duration of circumflex vowels and the greatest ones of acute vowels are often found within the same interval (as shown by the Wald–Wolfowitz criterion), but this does not eliminate the asymmetry between limit values as something uninformative and insignificant. We thus seem to arrive at the following conclusion: The circumflex vowels are more often than not (though not always) pronounced as with longer duration, and the acute vowels – with shorter.

Pitch

The analysed prosodic data of the northern Samogitian Telšiai dialect forces one to conclude that the contribution of the pitch to the contrastivity of tone is on a par with that that of the duration of the vowel. The strongest argument in favour of this claim is the obvious correlation between pitch (the ratio between its range and the average pitch change) and duration (see Chart 3).

Chart 3. Correlation between the relative velocity of pitch and length



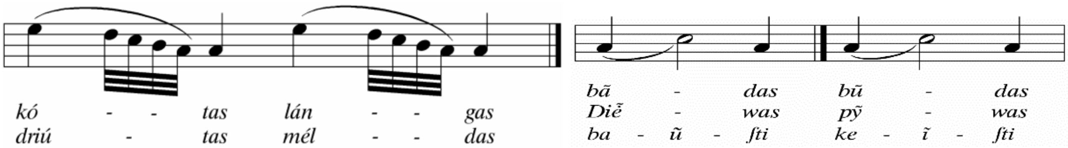
The greater the duration of the sound analysed, the slower the change of pitch (in relation to the range). There is no regularity in the differences of the horizontal pitch change (rise, fall, and rise-fall), the sole exception being the tendency of one speaker to push the pitch peaks of circumflex vowels to peripheries of it (i.e. the initial and the final parts). The least informative are the ranges of pitch. The latter parameter does not contribute to the contrastivity of tones (in some cases, acute vowels have wider ranges, but in others, circumflex). The average rate of pitch change parameter (i.e. the average change between points of the pitch per 0.01 s) contributes more to the contrastivity of tones, but it should only be assessed in relation to other parameters. As shown by one of the observed correlations, the abruptness of pitch (and also the average change between curve points) reacts at the same time to changes in range (i.e. the increasing range leads to the increasing rate of pitch change) and the duration (the latter affects the progression which describes the rise of average rate of pitch change as range increases: the greater the duration, the greater this progression is, and vice versa). Circumflex vowels tend to exhibit longer duration than the acute ones, which makes their mean rate of pitch change smaller in each range group (i.e. their pitch change seems to be smaller). The same conclusions are also implied when we include pitch sustainability parameters in our analysis. The probability of tone contrastivity according to the latter feature is also supported by statistic parameters (see Table 2).

Table 2. Sustainability coefficient assessment according to the Student test

Participant initials	E.K.		A.D.	
Tone	acute	circumflex	acute	circumflex
Mean	4.398091949	5,867294872	2.798916558	2.993815369
Dispersion	5.960585474	6,691961378	2.356389905	2.419437194
Number of trials	81	89	84	86
T value	3.809012074		0.822238912	
p value (*T*<=*t*) (two-sided alternative)	**0.00019539**		0.412106396	
t critical value (two-sided alternative)	1.974185191		1.974185191	

The totality of observed features provides both the factual material and the argumentative grounds for distinguishing two, potentially contrastive, acoustic models. The first of them could be recognized from the tendency to combine the acoustic parameters in such a way that the relative rate of pitch change of the vowels would be approaching the maximum possible level (as allowed by the sentence intonation at hand). The acoustic effect caused by this model seems to result in a pitch that is more level, more continuous, sustained. The most favourable conditions for it to arise comprise the following combination of parameters: an extended duration + narrow range. The narrow range does not leave any space for any more abrupt pitch rate, while the extended duration (because of the correlation law in play) brings it even closer to the minimum level, i.e. a level and long pitch contour. This model is represented by circumflex tone vowels.[16] A converse model tends to generate a much greater (relative!) pitch change. The long vowels that represent it rarely ever leave the impression of a sustained sound (unless there is strong emphasis), their sound is shorter, more abrupt, and more energetic. The most favourable for this model is the opposite combination of parameters (i.e. short duration + wide pitch range). In these conditions we observe the most intense change of pitch, while the average difference between points on the curve tends to be nearer its maximum value (which is the difference between the highest and the lowest frequencies of the specific sound in question, i.e. its range). This model is to be associated with acute vowels.

It is important to note that this alternative way of data analysis (by technical means) does not provide any new interpretations of acoustic types of tones. The conclusions arrived at through acoustic impressions (which amount to descriptions of the acoustic effect rather than listing the values of specific parameters), by authors who researched tones in northern Samogitian dialects, essentially point to the same two acoustic models discussed here. Take, for example, the tones written in musical notes in the 19th century by Friedrich Kurschat (Kurschat 1876, 59)[17]:



Aleksas Girdenis explains the underlying acoustic differences between these tones written in musical notes as based on the contrast between the rising and the falling pitch, as well as the differences in range (Girdenis 2008: 381–404). Both of these parameters (namely, the contours and ranges), according to our data, are neither regular nor reliable indicators of tones. Kurschat himself, without the means to have technical data at his disposal, made use of intuitive descriptions for the acoustic impressions of vowels, i.e. ‘pushed’ (also, ‘sudden’) pitch (acute) and sustained pitch (circumflex). The duration of vowels illustrated in notes is not identical (quarter note + four thirty second notes [acute] ≠ quarter note + half note [circumflex]). Their exact

[16] The material for these claims comes only from the linguistic data obtained from the one speaker.
[17] Cf. interpretations of Serbo-Croation tones from a similar period (see Lehiste, Ivić 1986: 3–22).

duration is, of course, impossible to determine, and only an approximate calculation can be made. A reference point for measuring duration could be the word's last vowel. Kurschat always marked it with a quarter note. The approximate duration of a short, unstressed, back vowel could be something like ~100–120 ms (cf. Ekblom 1925: 48–57, Pakerys 1982: 45). Based on this relationship (where a quarter note equals 100 ms), we can also calculate the duration of long stressed vowels in the root of the word. This is done by deriving a simple mathematical equation. If we let x stand for the quarter note, we may write down the quantitative value of long syllables as follows: acute → (x + 0.5x), and circumflex → (x + 2x). The second part of the acute vowel comprises four thirty-second notes, the quantitative value of which equals half the quarter note (if the quarter note is x, then half of it will be 0.5x). The respective part of the circumflex vowel is illustrated with the half note (which comprises two quarter notes), so its quantitative value can be said to be 2x. A few basic arithmetic calculations show that the approximate duration of the acute vowel could be 150 ms, while that of the circumflex vowel would be 300 ms (so the circumflex vowel is as much as twice as long!). When attempting a reconstruction of these latter tones, written in musical notes, Girdenis did not take this quantitative difference into account and assumed the same duration for both tones.[18] Had Kurschat observed the differences between the acoustic types of tones only with respect to the horizontal trends of pitch movement (i.e. had he supposed solely the contrast between rising and falling pitch) and never attempted to relate the vowel duration and the rate of pitch change, it would be hard to explain his choice to use notes of different quantitative values and display them in such specific manner. In other words, the dynamics of the acute vowel pitch change could be equated with that of the circumflex vowel and fully captured as follows:

[18] It remains unclear as to why he took this step. He writes: 'The transition from one note to another [i.e. four thirty-second notes of the acute vowel] should not be abrupt (which is shown by the legato sign that connects them)...' This claim can be contested. Legato, as a musical term, indicates a consistent, smooth transition from one sound to the next (but it plays no role in in the quantitative value of notes!).

The width of the range would in this case remain unchanged (5 semitones), the duration and the respective rates of change would be assumed to be equal, and there would be differences only between directions of pitch movement and range widths (which exactly corresponds to Girdenis's conclusion). However, this is *not* how Kurschat chose to depict the acute tone. On the contrary, the illustrative material suggests that his intention was to make a more subtle depiction of differences in pitch change. They are clearly expressed by the contrast between four thirty-second notes (acute) and one half note (circumflex). In spite of the initial leap, the circumflex tone pitch remains levelled in a single range for two thirds of the entire vowel duration.[19] This combination of parameters (extended duration + even pitch) provides the perfect acoustic environment for the effect of a sustained sound (the combination of this type is probably the rationale for the term 'sustained tone'). The pitch of the acute tone, conversely, tends to approach the opposite, combinatory model (shorter duration and an intense, abrupt change of pitch). More favourable conditions for a more energetic acute pitch dynamics are also provided by its wider range (the range of the acute tone is 5 ht, while that of the circumflex tone is 3 ht). Taken together, these characteristics enable one to reconstruct the same two combinations of parameters as those identified above, namely a more intense pitch change (the average rate of change between points on the pitch contour) + a wider range, and shorter duration (acute) on the one hand, and slower main pitch change, a narrower range, and more extended duration (circumflex) on the other. Although one cannot make use of the relative values (because there are no pitch curves divided into small time intervals), the more abrupt pitch change noticed by Kurschat in shorter (or acute) vowels and the slower pitch change in the longer vowels corresponds to the same law of correlation between duration and pitch, which predicts a slowing down in the relative pitch (i.e. a gradual extension of the vowel) as the vowel gets longer[20] (for the relationship between vowel duration and pitch contours, see, e.g., Zhang 2002).

---

[19] The half note remains on the same line in the stave for this whole period of time, so the movement of pitch (change/rate) is very close to zero.

[20] The equality sign can be put not only between tones written down in musical notes by Kurschat and the correlation law proposed in the present work. Experimental research into other languages shows that the rate of the rising pitch change is usually slower than that of the falling one (which was perfectly depicted in musical notes by Kurschat)(see, e.g., Sunberg 1973: 39-47, Ohala, Ewan 1973: 345, Xu and Sun 2002: 1399–1413), and vowels are longer (Pakerys 1982: 156). If one follows this

Similar combinations of parameters are also noted as distinctive features of tones by other scholars who undertook experimental research into tones in Lithuanian dialects (e.g., Girdenis 1967 [2000]: 76–88, 1996 [2001]: 243–256, Remenytė 1990: 60–78, Bacevičiūtė 2001: 144, Urbanavičienė 2005: 152–155, Murinienė 2007: 234, an others). In dialects with, among other things, more or less audible differences between long monophthong tones, the circumflex vowel is usually thought to be longer, to have a narrower range, and to exhibit a less abrupt change of pitch (usually rising), with the acute vowel bearing the opposite characteristics. One problem with this, however, is that their regularity is not always supported by parameter analysis results, and there occurs, therefore, a logical gap between the intuitively felt acoustic difference and the trends displayed by the actual material (unfavourable to this difference). A switch from absolute to relative values of the parameters under analysis could be one of the solutions to this problem and perhaps also help to demonstrate the distinctiveness of tones.[21]

Intensity

The role of the last acoustic feature that we analyse, i.e. intensity, in determining differences between tones should, to stay on the safe side, for now be more objectively regarded as indeterminate (without rejecting any of the above claims). Even though no objective, mathematical, and empirically justified arguments have been found to demonstrate correlations with other prosodic parameters, it does, after all, seem possible to distinguish the vowels in question by means of the dynamics of voice intensity features. One thing is clear: intensity plays absolutely no role in tone

---

assumption (namely that the pitch of the vowel falls slower than it rises), then languages with such a contrast between pitch contours should have the same law in play.

[21] It seems that this method could enable us technically to contrast tones in other languages too (by no means suggesting the type of their acoustic effect). For example, as can clearly be seen from the parameter data in the research into tones of the Franconian dialect of Arzbach, pitches of Class 1 and Class 2 are distinguished by not only the horizontal pitch curve movement (at some points it is totally converse) but also the relationship between prosodic parameters. The Class 1 pitch curve change is always smaller, with the pitch itself always being longer and with a narrower range. It can clearly be seen that in all of the indicated sentence intonation positions this pitch would generate a higher pitch sustainability parameter, while in the Class 2 pitch this value would be lower (the pitch in tones of this type tends to fall more abruptly, is shorter, and has a wider range [Köhlein 2011: 57, 66; also cf., Gussenhoven, Peters 2004: 251–285, Prehn 2011: 59, Markus 1991: 61; for similar parallels, see Jakobson 1931: 247–267, Gandour 1977: 54–65; Jørgensen 1989: 1–59]).

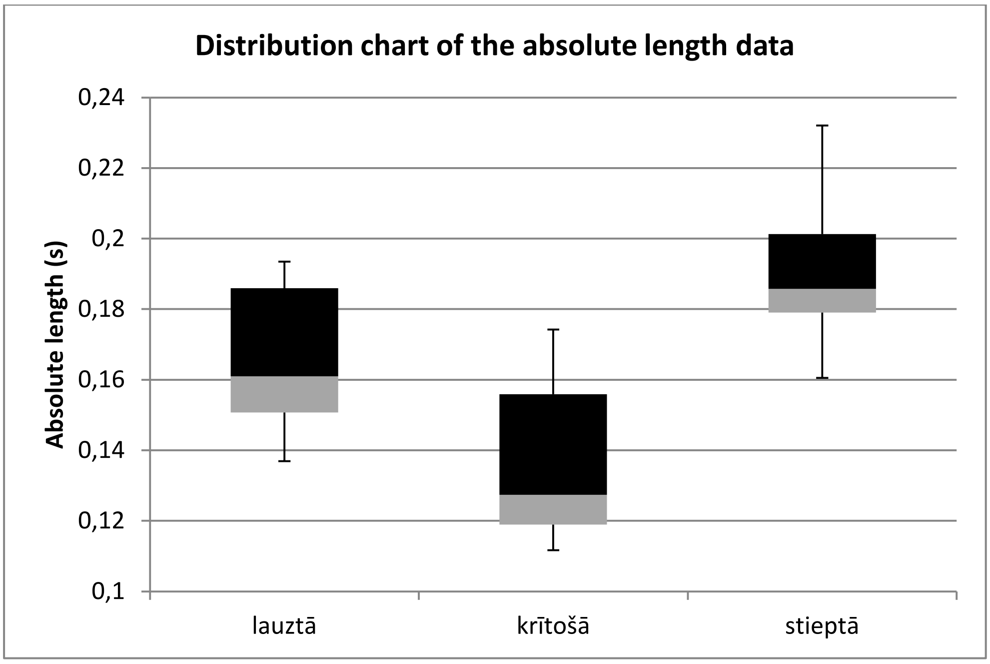
contrastivity in one of the speakers, and we can talk about distinction in terms of intensity characteristics based on linguistic material obtained from only one speaker. If the features observed in the research material obtained from the one of the speakers are not accidental (i.e. if the trends they show are characteristic not only of him but of a considerable part of the dialect community), then they could be regarded as factual grounds for talking about the wave of diminution of differences between long monophthong tones spreading from eastern to western Lithuania. The data gathered necessitate an adjustment of the presupposition about the process of one-way transformation of circumflex vowels to acute vowels. The evidence for this comes from the features of vowel duration and main pitch in the younger speaker (especially the decreasing duration of the circumflex vowels and their reduced sustainability), but the scattering of intensity peaks and the rate parameters (both absolute and relative) do not allow the equality sign to be put between his assimilated tones and some particular tone in the other speaker (the bit older one). The model suggested by the vowel intensity features in the younger speaker's stands in an intermediate position in relation to the features characteristic of the other speaker's tones. The acute of the one of speakers can be recognised through his tendency to intensify the initial part of vowel. Moreover, the rate of intensity change is greater in these vowels (the average rate of the intensity change is greater, and the sustainability coefficient lower). The circumflex tone displays a completely different direction of activity. In vowels that represent it, the central or the final parts are often the more intensive, with lower mean rate of change and a greater sustainability coefficient. As yet, one cannot tell for sure whether these groups of features alone are sufficient to serve as a reliable means to distinguish tones, in spite of the statistical results that seem to speak in favour of difference. It remains possible that the methodology used in this work is simply not powerful enough to grasp the intensity features that distinguish the tones (i.e. it might fail to eliminate the effect of other factors).

Acoustic features of monophthong tones in central Latvian dialect (Valmiera subdialect)

Duration

The analysed material from the Latvian dialect forces one (just as in the case of the northern Samogitians) to talk about two levels of differentiation of vowels in question, which could theoretically be associated with tones. The relationships between features, namely the asymmetric distribution of limit samples in charts, differences between medians (in terms of their expressions in per cent), and the tendency of one of the tones to be ahead of another by a part of its sample values, enables one to note that the degree of disorder of data is not one hundred per cent (see Chart 4). The predominant position, no doubt, is taken by the sustained vowel (that it leads in both oppositions against other vowels is obvious), while the duration differences between the falling and the broken tones are less than negligible (although, according to some parameters, one could make a cautious guess that the duration of the falling tones differs from the sustained tones more than it does from the broken tones). Because there was no considerable divergence between data analyses of the two speakers in the study, we have no grounds to claim that the possibility for the vowels to differ in the quantitatively narrower base (one of the speaker's vowels were shorter) is lower with regard to the prosodic feature at hand.

Chart 4. Distribution chart of the absolute length data of the minimal pair *lùoks : luõks : luôks*.



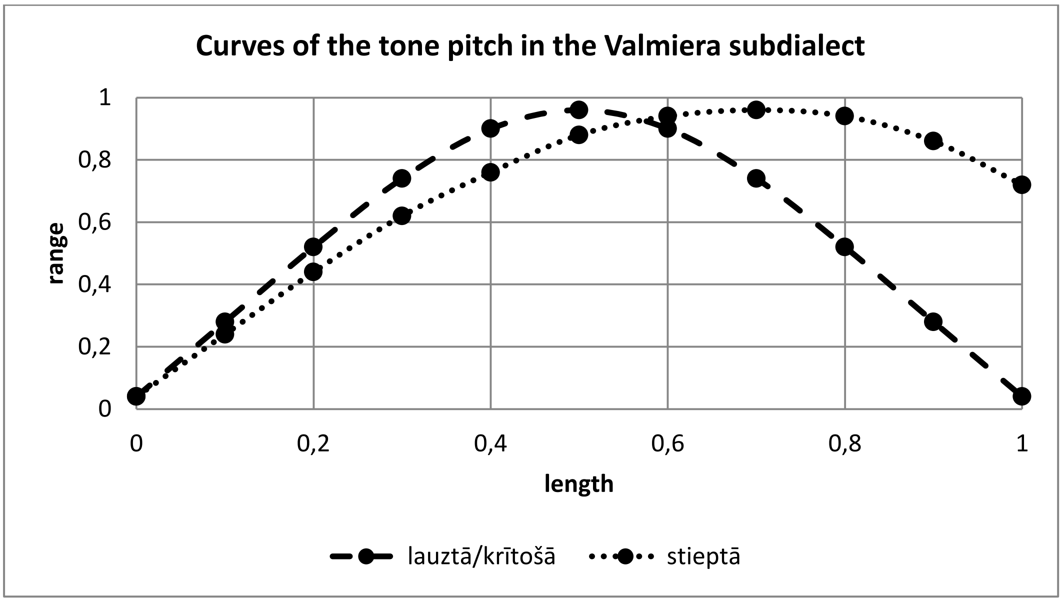
The second level of differentiation, dictated by the results of the Wald–Wolfowitz test, two times out of three forces one to draw an unfavourable conclusion regarding tone contrastivity. Although the tendency of the sustained tone being ahead of other tones is obvious, it is insufficiently large to also be confirmed by the statistical parameters. Descriptions of charts indicate that the peak values of unsustained vowel samples (i.e. the falling and the broken tone vowels) usually reach the medians of the sustained vowel. This means that roughly a half of data under comparison fall within the general interval. The similarity of dispersion data (in both dialects) allows us to make several remarks about the theoretical side of the issue. The claims, considerations, and observations about two levels of distinction, differences between speakers speech rate and intonation, a presumably systemic and non-systemic shortening of vowels, and the relationship between the duration of vowels in question and word structure specify the phonetic environment that directly affects vowel duration. In a living language, the effect of this environment is always there, and it should therefore be predictably impossible to specify the exact intervals of duration of the vowels in question, and, in turn, to identify tones using specific parameter values by technical means.[22] Nor is the problem solved by calculating the relationship between the vowel in question and the word structure, because the duration of the tone syllable and the other phonemes of the syllable of the same word do not seem to be constant or directly affected only by the tone (which means that this relationship itself is not constant). The minimal pair method can show only the fact of the opposition itself, and is not able to determine the exact values of parameters. So the contribution of the factor of duration to tone identification is more appropriately assessed through the correlation aspect, i.e. via the acoustic effect that is created by the relationship between vowel duration and other prosodic features (primarily, the pitch).

Pitch

The parameter analysis performed in this work allows one to note that the speakers of the Latvian subdialect make a more regular use of the features of pitch in distinguishing tones than do the Samogitian speakers (which is demonstrated by statistical calculations); however, they, both audibly and according to the lower sustainability parameters, pronounce them less clearly (than, for example, one of the northern Samogitians). One outstanding feature of Latvian tones is the differentiated pitch contours. Based on them, the sustained tone in the Valmiera subdialect could be reasonably called as a rising tone, whereas the falling and the broken vowels could be described as those of the rising-falling tone (see Diagram 4). The regularity of differences of the horizontal movement of these curves produces a strong visual impression, but it should not mislead us to suppose that Latvian tones are differentiated by means of melodic modulations. One cannot accept such a conclusion because of the narrow pitch ranges of the vowels analysed: oftentimes, they did not exceed the limit of the small tierce. In such phonetic conditions, the identification of pitch (its fall and rise) requires high musical sophistication and is far too complex for language users (it should be borne in mind that many speakers have no musical ear at all). Analysis of pitch abruptness and sustainability has showed that these latter contours reflect the same rate of pitch change differences as those identified in analysing north Samogitian tones.

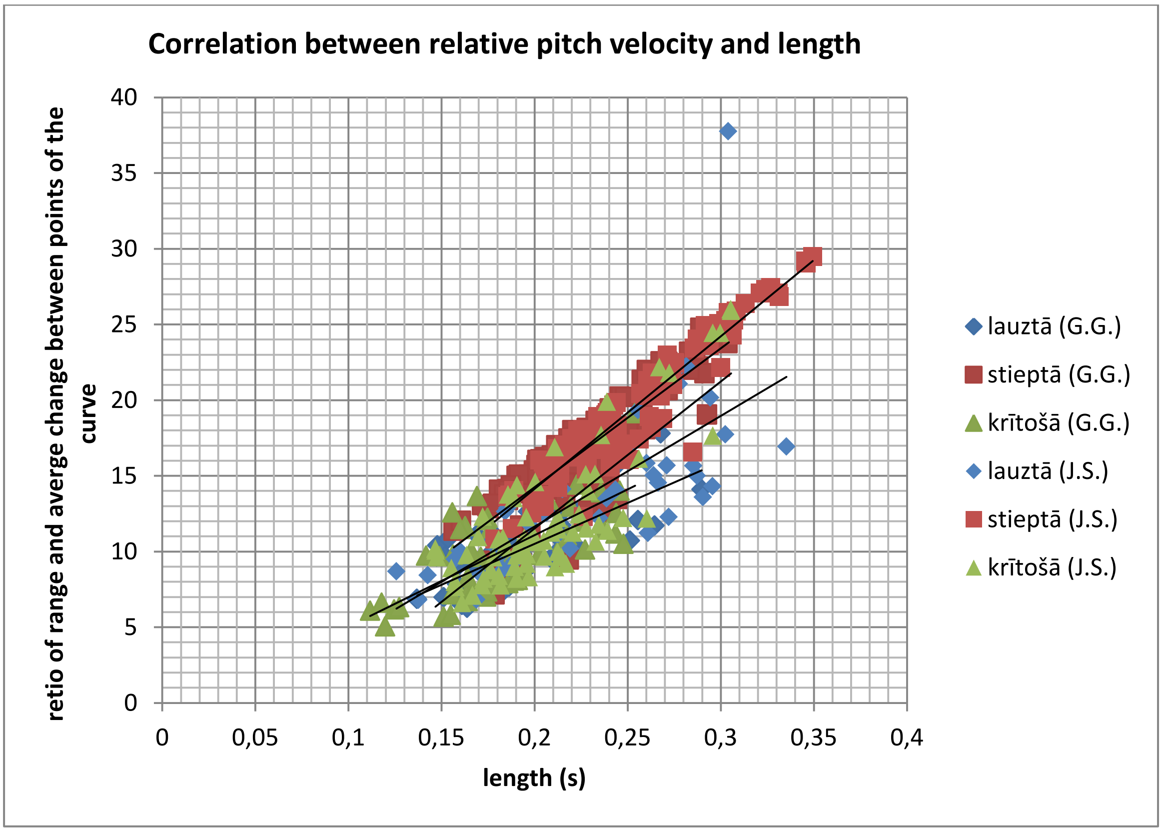
[22] There is no way for the obtained results to be used in, for example, any cutting-edge speech technology. It might be worth considering the following question: what kind of vowel duration *could* be used by future speech recognition devices for identifying tones?

Diagram 4. Curves of the tone pitch in the Valmiera subdialect



The rising pitch, in identical conditions (i.e. when the range and the duration of the vowels under comparison correspond to each other), changes half as fast as the rising-falling pitch (presumably because of the more intense pitch change in the initial and the final parts of the vowel). The functioning of the threefold tonal system in this subdialect could be grounded in the rate of pitch change parameters (admittedly, supported by the data obtained from only one of the speakers); however, the latter needs to be converted to a twofold system after one calculates pitch sustainability coefficients and performs their statistical assessment. The claim about the distinction between two rather than three tones (i.e. that the broken and the falling tone are no longer distinguished) is supported by all other features: the broken and the falling tones are rather made similar than distinguished by similar (or, better even, identical) pitch contours, ranges, dispersions of peaks, and sustainability coefficients. As one clear feature of the broken tone, namely the break in pitch condition by glottalization, has disappeared, there does not seem to be any other acoustic manner for compensating this lack of glottalization. A resurfacing of the very same pitch rate, range, and duration dependence relationships points to a clear phonetic feature, namely the inverse proportionality between relative pitch rate and duration: the longer the vowel, the smaller the relative pitch change rate, and conversely (simply put, increasing duration = stretching pitch) (see Diagram 5). This law-like relationship is a factual argument in favour of a search for tone differences by means of relative values of acoustic features (acoustic effects caused by combinations thereof). Attempts to directly relate the absolute (non-relative) parameter values with tones, as has been shown, is less accurate, despite its pointing to the same two acoustic effects (the sustained and the unsustained).

Diagram 5. Correlation between relative pitch velocity and length



The parallels between the features, which one can draw from the experimental base of the present work, suppose that the acoustic nature of tones in both languages is identical. Key to identifying tones is pitch change rate. This latter conclusion was also reached by Girdenis in his discussion on the distinguishing features of tones in northern Samogitia and, for that matter, in Lithuanian. (Girdenis 1998: 37–38; 2003: 269). Without opting for parameter analysis via their relative values, however, he could not provide a firm support for such a conclusion through factual data. The most optimal technical-mathematical parameter for contrasting tones is the sustainability coefficient that reflects the relative rate of pitch change (see Table 3).

Table 3. Sustainability coefficient according to Student criterion

Participant	G.G.						J.S.					
Tone	˜	ˆ	˜	`	ˆ	`	˜	ˆ	˜	`	ˆ	`
Mean	3.74	2	3.74	1.98	2	1.98	4.9	3.37	4.9	2.68	3.37	2.68
Dispersion	1.94	0.74	1.94	0.83	0.74	0,83	3.84	4.05	3.84	2.63	4.05	2.63
Number of trials	106	52	106	58	52	58	103	42	103	68	42	68
T value	-9.59		9.3		0.134		-4.23		8.13		1.88	
p value (*T*<=*t*) (two-sided alternative)	**0.0000000000000000003**		**0.0000000000000000008**		0.9		**0.000006**		**0.00000000000001**		0.06	
t critical value (two-sided alternative)	1.98		1.98		1.98		1.99		1.98		1.99	

One should bear in mind that the calculation of the value of the latter includes three parameters, namely the range, mean rate of pitch change, and vowel duration, and all of them should therefore be treated with equal attention. We thus have conclusively established that tones are combinations of acoustic parameters rather than isolated modulations of vocal notes (their visual contours) that are determined by strict absolute parametric values (such as strict duration intervals, ranges, tone contours, etc.). Points of intersection between experimental research to date into Latvian tones and the results of the present work are the same as those identified in summarising the tonal features in the northern Samogitian dialect. The law of correlation allows one to say that the pitch of the sustained tone is more sustained and has low rate in cases of extended duration and low range (scholars to date have associated such cases with exactly this, i.e. the sustained, tone), while the pitch of the unsustained tones, namely the broken and the falling tones, changes much more intensely, which is due to their shorter duration, wider range, and glottalization (in the case of the broken tone)(see also Zhang 2002). To put it simply, the choice of the alternative model of data analysis has served to confirm, rather than contradict, the acoustic effects identified by previous scholarly work. The parameters that were identified could serve as grounds for only a cautious guess that tones in Valmiera in the beginning of the previous century may have been much clearer and much more vivid than I heard today.

Intensity

The final prosodic feature, intensity, seems to contribute probably the least to the identification of tones in the Valmiera dialect. To remain cautious, however, one should say that their role in tone contrastivity is unclear. Analysis of peaks, rate of intensity change, and sustainability coefficients, together with the statistical analysis of data, point to the same vague conclusions outlined in the end of the section on the northern Samogitian tones. Despite the resurfacing of several correlative equalities between parameters familiar from the pitch parameter analysis, the most obvious of which is increasing range = greater rate of change, and also the direct proportionality of duration and the relative intensity parameters' values (i.e. the range and the average rate of change between points of the curve), this had no considerable effect on the degree of tone contrastivity. If the distinctive features of the pitch are our established reference point, then we must note that the tone system, once we include in our analysis the parameters of the intensity of voice, falls apart at the very first stage of analysis. The clear distinction between the sustained tone according to the features of pitch from the other two (namely, the broken and the falling tones) was supported by the dispersion of peaks, curve contours, mean rate of pitch change, and sustainability coefficients (thus, a whole complex of features tangled by correlative relations). Identifying similarities between the tone intensity features in both speakers becomes a complex task at the very first step, i.e. the peak dispersion analysis. The data obtained from the one of the speakers from Valmiera allows easily detect distinction of the broken tone (because of its tendency to cluster the peaks closer to the initial part of the vowel), while the chart of data regarding the other speaker allows an easier detection of the distinction of the sustained tone (more often than not, the peaks here were closer to the end of the vowel). The initial expectation was that the correlational parameter analysis (encompassing range, rate of change, their relative values, and duration) would point to the same differences in the dynamics of change (as those found in the case of pitch), namely a slower change of voice intensity will distinguish the sustained tone from the other two, while the broken and the falling tones will allow for a description through the relatively more abrupt and energetic voice intensity. These expectations were modified by the research data (it also helped to rethink any possible correlative relations between pitch and intensity). One might say, albeit only very hypothetically (based exclusively on average sustainability values, the differences between which are often minimal), that the graduated series of tones (the sustained tone > the broken tone > the falling tone) with regard to pitch corresponds to the intensity data. As of now, we are not able to validate this probable and intuitively plausible claim with the methods used in the present work. As John Ohala rightly remarks about such ambivalent cases, 'in order to gain some understanding of the shape of the pattern of speech sounds, including the direction of their change over the years, one must examine how the human articulatory and perceptual systems operate' (Ohala 1975: 289). This issue is likely to be resolved once we find out what external factors affect the intensity indicators under comparison and learn to ignore them.

Conclusions

1. The analysed tones of long vowels in dialects of Lithuanian (Telšiai subdialect, northern Samogitia) and Latvian (Valmiera) dialects are acoustically differentiated by combinations of the pitch and duration. The tones on long vowels in both dialects represent the same set of two acoustic models. The first of these models is identified through a longer, more even, and more sustained change of pitch. It seems to encompass the north Samogitian circumflex and the Valmiera sustained tone. The second acoustic model can be recognised through a more abrupt and more intensely changing pitch, whose the acoustic character is strengthened by a shorter vowel duration (when compared to the opposite tones). The latter seems to encompass the north Samogitian acute vowel and the Valmiera falling and the broken tones.
2. One of the key points of reference in identifying the distinctive features of tones is the correlation of acoustic parameters, namely the direct proportionality between vowel duration and the ratio of the two indicators of the pitch, i.e. range and mean rate of pitch change. As the vowel duration increases, so does the ratio between pitch parameters. If the long vowels are differentiated by duration, they are at the same time also automatically differentiated through the said ratio.
3. With the disappearance of glottalization, there was no acoustic means left to compensate for this feature, which distinguished the broken and the falling tones in the Valmiera subdialect. The tonal system of this subdialect, therefore, can no longer be considered threefold. A clear distinction from the other two (the broken and the falling tone) is only in place in the case of the sustained tone. The diminution of differences between the broken and the falling tone is shown by their shared complex of features: similar peak dispersions, rate of pitch change, contours, and duration that are hardly different nowadays.
4. The difference between tones is better indicated by the relative, rather than the absolute, values of acoustic parameter. The obtained results of the study (the differences between tones confirmed by statistical criteria) served to justify the selected method of analysis (i.e. the inclusion of the relative parameters into the analysis) and bridged the logical gap between the intuitively felt acoustic difference between tones and the contrary trends often suggested by the absolute parameters (i.e. the denial of any difference between the analysed prosodic units).

5. The melodic value of pitch contours (pitch modulation) is a feature of secondary importance only. Under identical phonetic conditions (i.e. when the parameters of vowel duration and range are identical), the rate of change of the rising pitch is slower, while that of the rising-falling one is more intense. This means that the contours illustrate different levels of pitch change (the shape of the curve can be both a factor that slows down the change of pitch as well as a factor that intensifies it). This insight allows one to put the equality sign between the main distinguishing features of tones in both the northern Samogitian subdialect and the Valmiera subdialect: In both it is the relative pitch change (i.e. it enables an interpretation of tones via the opposition between the sustained and the unsustained tones).
6. The most reliable technical-mathematical parameter for identifying tones, which could also be useful in improving speech technologies (e.g. tone recognition) is the vowel pitch sustainability coefficient. It is obtained by multiplying the ratio of pitch range and average change by the entire duration of the vowel.
7. The role of intensity in tone contrastivity is manifestly smaller than that of the main pitch and duration. Even though the statistical analysis sometimes indicates reliable differences between the intensity indicators (i.e. the average rate of change and the sustainability coefficients), they do not allow for a clear, systematic conception of the said role. The voice intensity indicators should rather be regarded as secondary. It is not impossible, however, that research into other dialects using a different methodology, would give a different assessment of its role.

LIETUVIŲ IR LATVIŲ TARMIŲ MONOFTONGŲ PRIEGAIDŽIŲ AKUSTINIAI POŽYMIAI: LYGINAMOJI ANALIZĖ

Santrauka

Darbo objektas, temos aktualumas. Disertacijos **objektas** – lietuvių ir latvių tarmių monoftongų priegaidės, jų akustiniai požymiai. Nors iš pradžių būta minties į disertaciją įtraukti didesnį tarmių skaičių, tačiau eksperimentinės medžiagos gausa privertė pakoreguoti šią pradinę intenciją ir apsiriboti dviem iš jų – centriniais šiaurės žemaičiais telšiškiais ir Latvijos vidurio tarmės vidžemiškuoju arealu (Valmieros šnekta). Tarmių pasirinkimą lėmė keletas priežasčių. Pirma, norėta tirti ir lyginti tas tarmes, kuriose, sprendžiant iš ligšiolinio eksperimentinio kitų mokslininkų įdirbio, priegaidės gerai skiriamos. Tik įvykdžius šią sąlygą, galima tikėtis produktyvios analizės (priešingu atveju nėra objektyvaus pagrindo kalbėti apie lietuvių ir latvių tarmių priegaidžių akustinius panašumus ir skirtumus). Antra, latviškoji Valmieros šnekta yra iš vadinamojo trijų priegaidžių ploto. Trinarė priegaidžių sistema yra tapusi sutartiniu atskaitos tašku kitų Latvijos tarmių priegaidžių sistemų aprašymams (daugelyje tarmių nuo trinarės priegaidžių sistemos yra pereita prie dvinarės). Pagal pastarąją tradiciškai yra nurodomi lietuvių ir latvių kalbų priegaidžių santykiai (pagal vadinamąjį Endzelyno dėsnį; žr. Endzelīns 1951: 35–41)[23].

Eksperimentinių tyrimų, skirtų latvių ir ypač lietuvių tarmių priegaidžių prozodiniams požymiams nustatyti, nestinga[24], tačiau iki šiol neturėta lyginamojo pobūdžio darbų, kuriuose, pasitelkus vieną metodiką, būtų išnagrinėta ir palyginta gyvųjų baltų kalbų tarmių priegaidžių akustinė bazė, įvardyti jos panašumai ir skirtumai. Per nemažą laikotarpį susikaupusi priegaidžių akustinės prigimties interpretacijų gausa yra akivaizdus probleminio disertacijos klausimo, – kokie yra patys universalieji priegaidžių požymiai, – rodiklis. Atlikus šį darbą, būtų galima palyginti baltų kalbų priegaides su atitinkamais kitų kalbų prozodiniais vienetais. Probleminis yra ilgųjų balsių priegaidžių klausimas (dalis mišriųjų dvigarsių ir dvibalsių priegaidžių skyrimo krūvio tenka kokybiniams, t.y. garsų artikuliacijos, požymiams), todėl visas dėmesys šįkart telktas būtent į šią vokalizmo grupę. Dar J. Gerulio praėjusio amžiaus viduryje kalbėta apie priegaidžių (pirmiausia, ilgųjų balsių) niveliacijos bangą,

23 Apie išimtis iš šio dėsnio žr. Stundžia 1985: 142–151.

24 Lietuvių ir latvių priegaidžių eksperimentinių tyrimų apžvalgos pateikiamos vėliau.

besiritančią iš rytinės Lietuvos dalies vakarinės link (Gerullis 1930), todėl svarbu išsiaiškinti, kiek toli ir kaip šio kalbinio reiškinio progresuota. Verifikuoti reikia ir Valmieros trinarės priegaidžių sistemos funkcionavimo šiomis dienomis klausimą (t.y. ar galima teigti, kad vis dar skiriamos visos trys priegaidės). Be to, nuolat tobulėjančių technologijų amžiuje aktualus ir naudingas galėtų būti ir techninis disertacijos įdirbis. Užčiuopus universalius priegaidžių požymius, techninės jų išraiškos galėtų praversti tobulinant kalbos atpažinimo ir sintezavimo technologijas.

Darbo tikslas ir uždaviniai. Darbo **tikslas** – išanalizuoti, palyginti ir aprašyti akustinius latvių (vidurio tarmės) ir lietuvių (šiaurės žemaičių telšiškių) tarmių ilgųjų monoftongų priegaidžių požymius.

Uždaviniai:

1. Pateikti išsamią šiaurės žemaičių ir Latvijos vidurio tarmės priegaidžių tyrimų apžvalgą.
2. Surinkti, apdoroti ir paruošti analizei eksperimentinę medžiagą (surengti dialektologines išvykas, įrašyti medžiagą ir t.t.).
3. Pasiūlyti, motyvuoti ir panaudoti šių dienų eksperimentinės fonetikos lygį atitinkantį techninės duomenų analizės (automatizuoto parametrų sugeneravimo) metodą.
4. Išnagrinėti ilgųjų balsių trukmės, pagrindinio tono ir intensyvumo rodiklių grupes, nustatyti jų tarpusavio koreliacijos laipsnį ir sąsajas su priegaidėmis.
5. Visus, galimai priegaidžių implikuojamus, ilgųjų kirčiuotų balsių diferencinius požymius iliustruoti grafiškai, verifikuoti jų skirtumų tikimybę statistiniais kriterijais.
6. Įvardyti akustinius lietuvių ir latvių tarmių priegaidžių panašumus ir skirtumus.

Ginamieji teiginiai

1. Lietuvių ir latvių tarmių priegaides akustiniu požiūriu diferencijuoja pagrindinio tono ir trukmės kombinacijos.
2. Balsių tono kitimo staigumas yra atvirkščiai proporcingas jų trukmei. Kuo balsis yra trumpesnis, tuo jo tonas (santykiniu požiūriu) kinta staigiau, ir atvirkščiai, – kuo balsis ilgesnės trukmės, tuo jo tono santykinis kitimas yra lėtesnis.
3. Patikimiausias techninis-matematinis priegaidžių identifikavimo parametras – tono tęstumo koeficientas (santykinio tono kitimo rodiklio sandauga iš balsio trukmės).
4. Priegaidžių tono kontūrų forma pirmiausia rodo ne melodines moduliacijas, o skirtingus balsio tono ištęstumo lygius.

Darbo struktūra. Disertaciją sudaro:

1. Įvadas. Jame nurodytas disertacijos objektas, suformuluoti tikslai ir uždaviniai, išryškintas temos aktualumas.

2. Darbo metodikos skyrius. Jame supažindinta su darbo metodika, pasirinktomis techninėmis duomenų analizavimo priemonėmis.

3. Tiriamoji dalis. Ją sudaro tyrimų apžvalgos (kiekvienos tarmės atskirai) ir skyriai, parengti pagal prozodinių požymių grupes – trukmės, pagrindinio tono ir intensyvumo.

4. Išvados.

5. Literatūros sąrašas ir disertacijos priedai.

Išvados

1. Akustiniu požiūriu tirtųjų lietuvių ir latvių tarmių ilgųjų balsių priegaides skiria pagrindinio tono ir trukmės kombinacijos. Šiaurinių žemaičių telšiškių ir Latvijos vidurio tarmės (Valmieros šnektos) ilgųjų balsių priegaidės atstovauja tiems patiems dviem akustiniams modeliams. Pirmąjį jų galima identifikuoti pagal lėtesnį, tolydesnį, labiau ištęstą tono kitimą. Jam priskirtinos šiaurinių žemaičių cirkumfleksinė ir Valmieros šnektos tęstinė priegaidės. Antrąjį akustinį modelį galima pažinti iš staigesnio, intensyviau kintančio tono, kurio akustinį charakterį sustiprina mažesnė (palyginti su opozicinėmis priegaidėmis) balsių trukmė. Prie jo priskirtinos šiaurinių žemaičių akūtinė bei Valmieros šnektos krintančioji ir laužtinė priegaidės.
2. Vienas svarbiausių atramos taškų priegaidžių skiriamiesiems požymiams nustatyti yra akustinių parametrų koreliacija – balsių trukmės tiesioginis proporcingumas dviejų pagrindinio tono rodiklių – diapazono ir vidutinio tono kitimo – santykiui. Didėjant balsio trukmei, didėja ir tono parametrų santykis. Jei ilgieji balsiai yra diferencijuojami trukmės, juos automatiškai skiria ir nurodytasis santykis.
3. Išnykus glotalizacijai, neliko akustinės priemonės, kuri kompensuotų šį požymį, skiriantį Valmieros šnektos laužtinę ir krintančiąją priegaides. Dėl šios priežasties pastarosios šnektos priegaidžių sistemos nebegalima laikyti trinare. Nuo kitų dviejų (laužtinės ir krintančiosios) priegaidžių ryškiai atsiskiria tik tęstinė priegaidė. Laužtinės ir krintančiosios priegaidžių niveliaciją rodo bendras jų požymių kompleksas: panaši pagrindinio tono maksimumo taškų sklaida, tono kitimo intensyvumas, kreivių kontūrai, menkai besiskirianti trukmė.
4. Priegaidžių skirtumą geriau rodo ne absoliučiosios, bet santykinės akustinių parametrų reikšmės. Gautieji tyrimo rezultatai (statistinių kriterijų patvirtinti priegaidžių skirtumai) pateisino pasirinktą analizės metodą (santykinių parametrų įtraukimą į tyrimą) ir panaikino loginį atotrūkį tarp intuityviai jaučiamo akustinio priegaidžių skirtumo ir absoliučiųjų parametrų dažnai suponuojamų priešingų (skirtumo tarp tiriamųjų prozodinių vienetų nebuvimo) tendencijų.
5. Tono kreivių kontūrų melodinė vertė (tono moduliacija) yra antraeilis požymis. Vienodomis fonetinėmis sąlygomis (esant identiškiems lyginamų balsių trukmės ir diapazono parametrams) kylančio tono kitimas yra lėtesnis, o kylančio-krintančio –

intensyvesnis. Tai reiškia, kad kontūrai iliustruoja skirtingus tono kitimo laipsnius (kreivės forma gali būti tiek tono kitimą lėtinantis, tiek intensyvinantis veiksnys). Šis atradimas leidžia dėti lygybės ženklą tarp šiaurės žemaičių ir Valmieros šnektos priegaidžių pagrindinio skiriamojo požymio – santykinio tono kitimo (t.y. interpretuoti priegaides pagal tęstinio ir netęstinio tono priešpriešą).

6. Patikimiausias techninis-matematinis priegaidžių identifikavimo parametras, galintis praversti tobulinant kalbų technologijas (pvz., priegaidžių atpažinimą), yra balsių tono tęstumo koeficientas. Jis gaunamas tono diapazono ir vidutinio kitimo santykį padauginus iš viso balsio trukmės.
7. Intensyvumo vaidmuo skiriant priegaides yra akivaizdžiai mažesnis nei pagrindinio tono ir trukmės. Nors statistinė analizė kartais rodo patikimus skirtumus tarp intensyvumo rodiklių (vidutinio staigumo, tęstumo koeficientų), tačiau iš jų negalima susidaryti aiškesnio sisteminio vaizdo. Balso jėgos rodikliai veikiau laikytini antraeiliais. Neatmestina, kad tiriant kitas tarmes ir taikant kitokią metodiką, intensyvumo vaidmenį būtų galima vertinti ir kitaip.

Darbo aprobavimas

Disertacijos tema paskelbti **straipsniai**:

1. Švageris, Evaldas 2013, Monoftongų priegaidžių skirtumai jaunimo kalboje, *Jaunimo kirčiavimo polinkiai: bendrinės kalbos normos ir vartosenos tendencijos* (straipsnių rinkinys), 77–88: VDU.
2. Švageris, Evaldas 2014, Tone Features in the Latvian Dialect in Lithuania, (*I W o B A V I I I (2 0 1 2) Proceedings of the 8 th international workshop on Balto-Slavic accentology*) (straipsnių rinkinys), 355-366.
3. Švageris, Evaldas 2015, Akustinių požymių koreliacijos svarba lietuvių kalbos priegaidžių skyrimui, *Baltistica* 50 (1), 89–111.

Skaityti **pranešimai** tarptautinėse konferencijose:

1. *Acoustic features of tone in Latvian dialect of Lithuania*, IWoBA VIII (International Workshop on Balto-Slavic Accentology), 2012 m. liepos 6–8 d., Novi Sadas, Serbija.
2. *Eksperimentinių prozodijos tyrimų naujovės: šiaurinių žemaičių telšiškių priegaidžių analizė*, Šiuolaikinės kalbos tyrimai ir problemos (21-oji tarptautinė Jono Jablonskio konferencija), 2014 m. spalio 3 d., Vilnius.
3. *Akustinių požymių koreliacijos svarba lietuvių kalbos priegaidžių skyrimui*, Baltų kalbų fonetika ir fonologija (tarptautinė mokslinė konferencija profesoriui Aleksui Girdeniui (1936–2011) atminti), 2014 m. spalio 30–31 d., Vilnius.
1. *Lietuvių, latvių ir slovėnų kalbų priegaidžių akustiniai panašumai,* XII tarptautinis baltistų kongresas, 2015 m. spalio 28–31 d., Vilnius.

Evaldas Švageris (born 1984) completed his Lithuanian and Latvian language studies at Klaipėda University in 2007 (BA in Linguistics), where he also obtained his Master's degree in 2009 (MA in Linguistics). In 2011–2015, he studied for his PhD at Vilnius University. He has published three articles on the topic of the dissertation, gave three talks at international conferences, and participated in a number of scientific research projects.

Evaldas Švageris (g. 1984 m.) 2007 m. Klaipėdos universitete baigė lietuvių filologijos ir latvių kalbos studijas (įgytas filologijos bakalauro laipsnis), 2009 m. – ten pat ir kalbotyros magistro studijas (įgytas filologijos magistro laipsnis). 2011–2015 m. studijavo Vilniaus universiteto doktorantūros. Disertacijos tema paskelbti trys straipsniai, skaityti trys pranešimai tarptautinėse konferencijose, dalyvauta įvairaus pobūdžio mokslinio pobūdžio projektuose.