BALTISTICA LVIII(2) 2023 195–224
doi: 10.15388/Baltistica.58.2.2517

Evaldas ŠVAGERIS
Vilnius University

THE PHONETIC INTERACTION OF PROSODIC ELEMENTS IN THE BALTIC LANGUAGES: A TENTATIVE THEORETICAL MODEL

Abstract. The aim of this paper is to propose a theoretical model of phonetic interaction between prosodic elements in the Baltic languages. Long syllables, for example, sometimes become the field of the synchronic realization of stress, syllable accent, and intonation. Therefore, the qualitative and quantitative characteristics of certain sounds must represent an entire group of elements in question. A natural question to ask, then, is how the distinctive features of all these elements are realized and combined. A point of reference for the model of this interaction could be a category of sound control/balance, interpreted in terms of F0 dynamics. The data presented in this paper allow us to consider the hypothesis that F0 change, viewed both in terms of paradigmatic and syntagmatic vectors, is the prerogative of phrasal intonation. The other two elements, namely word stress and syllable accents, are to be regarded as factors that regulate the balance of F0 change at the intersyllabic level. Word stress determines the syntagmatic distribution of tonal dynamics of this kind, whereas syllable accents determine the paradigmatic one.
Keywords: Baltic languages; phonetic interaction; prosody; F0 dynamics; jerk; word stress; syllable accent; intonation.

The problem. It is only at first sight that the selection of phonetic categories used for describing the prosodic structure of language can appear self-evident. It has long been argued, for example, that the acoustic structure of word stress in large groups of languages, including Lithuanian, cannot, in principle, be characterized by a single acoustic parameter (Lieberman 1960, 451–454; Lehiste 1970, 142–146; Pakerys 1982, 111–144; Fant, Kruckenberg 1994, 125–144; Girdenis 2014, 265–271; Ladefoged 2003, 90–91; Ortega-Lliebaria, Prieto 2010, 73–97; Plag, Kunter, Schramm 2011, 1–29; Lippus, Asu, Kalvik 2014, 232–235; Gordon, Roettger 2017, 1–11; Zarka, Schuppler, Lozo, Eibler, Wurzwallner

2017, 15–44; van Heuven 2018, 15–59, and others). Moreover, experimental research suggests that the probative value of all qualitative and quantitative acoustic correlates of stress becomes evident only under *ceteris paribus* conditions. This means that acoustic differences can only be found between the same syllable in stressed and unstressed positions. The situation is different when it comes to the syntagmatic relationship between all syllables, where phonetic prominence does not always lie with the stressed syllable. One might suggest that empirical data tend to support paradigmatic identification of stress, but of course, such an assumption is difficult to reconcile with the syntagmatic nature of stress (Lehiste 1970, 106; Girdenis 2014, 267). Another group of prosodic elements, namely syllable accents also show certain complexity. Although their key difference is understood in terms of the dynamics of the fundamental frequency (F0), they are usually accompanied by other acoustic correlates too, especially by duration (cf. experimental research on syllable accents in the Baltic languages: Gerullis 1930, 22; Ekblom 1933; Girdenis 1967, 31–41; Girdenis, Pupkis 1974, 107–125; Girdenis 1974, 160–198; Liepa 1979, 48–151; Pakerys 1982, 156–182; Markus 1991, 57–62; Sarkanis 1993, 62–90; Markus 1993, 39–44; Vaitkevičiūtė 1995, 45–86; Atkočaitytė 2002, 123–143; Bacevičiūtė 2004, 83–113; Kudirka 2004, 233–246; Leskauskaitė 2004, 179–232; Murinienė 2007, 171–186; Bacevičiūtė 2011, 13–26; Švageris 2016, 1–9; Bakšienė 2016; for other languages see: Frings 1934, 110–140; Bruce 1977, 49; Lehiste, Ivić 1986, 39–61; Fisher-Jørgensen 1989, 1–59; Gussenhoven, Peters 2004, 251–285; Fournier, Verhoeven, Swerts, Gussenhoven 2006, 29–48; Jurgec 2007, 195–207; Gussenhoven, van den Beuken 2012, 75–107; Zintchenko 2018, 101–203, and others.) More recent research on syllable accents in the Baltic languages (Švageris 2020, 119–157; 2021, 271–305) is no exception, because the F0 dynamics are described there as a function of time, i.e. distinctive features are determined in terms of the relationship between F0 and the duration. Finally, the phonetic form of the third prosodic element, namely phrasal intonation is probably the most problematic of all. It is usually formalized in terms of two-level tone (high and low) pairs (the so-called ToBI transcription; see Pierrehumbert 1980; Pierrehumbert, Hirschberg 1990, 271–311, Gussenhoven 2004, etc.), but its acoustic expression, of course, is far more complex. Although there has been relatively little research to date on intonation in the Baltic languages, it is regarded as a factor behind high phonetic variability of other elements or, in some cases, even their neutralization (for general overview of intonation research see Fox 2000, 267–329; also Švageris 2020, 121–127).

This phonetic complexity of prosodic elements makes it difficult to understand the mechanism of their interaction. Long syllables, for example, sometimes become the field of synchronic realization of stress, syllable accent, and intonation. Therefore, the qualitative and quantitative characteristics of certain sounds must represent an entire group of elements. The natural question to ask, then, is how, under such conditions, the distinctive features of all these elements are realized and combined. If we fail to resolve (or even address) this issue, the risk of misinterpreting or misattributing acoustic data increases. One should also not rule out the possibility that the commonly invoked concept of phonetic complexity is no more than a convenient compromise that allows us to systematize ambivalent empirical data without either categorically denying or affirming the significance of any acoustic correlate in question.

On the other hand, it seems that this problem could be easily solved by selectively assigning different prosodic elements with different phonetic category. For example, the old practice of assigning a prosodic element with a dynamic, melodic (tonal), or quantitative[1] type is a clear illustration of the logic of this sort. This is the path followed by researchers of Latvian prosody. The Latvian word stress is commonly regarded as a dynamic (Laua 1969, 115; Rudzīte 1993, 87–88; LVG 2013, 120) and the syllable accent a tonal phenomenon (Ekblom 1933; Kariņš 1996, 121; Laua 1997, 105; Grigorjevs, Remerts 2004, 33–50; LVG 2013, 104–106, etc.). However, the relationship between Latvian syllable accents and intonation is much more problematic in this regard (Hualde, Riad 2014, 668–669; Hualde, Riad 2018, 27–31, etc.).[2] To some degree, the same logic is followed in Lithuanian prosody research, too. In early descriptions of Lithuanian syllable accents, for example, different interpretations of syllable accents were proposed, focusing on tonal and the dynamic aspects (Kurschat 1876; Baranowski, Weber 1882, 29; Jaunius 1911, 33).[3] Although intensified experimental research did

[1] On the phonetic precision of all these categories see Girdenis 2003, 252.

[2] Admittedly, no clear dividing line can be drawn by Latvian syllable accents and intonations, either, because the expression of these two elements is mostly related to tonal modulations.

[3] One might observe that this distinction may not have been so clear after all (see Švageris 2018, 21–68).

contribute new empirical data to the analysis of this issue, the sheer volume and diversity of this data made it impossible to fully clarify the phonetic nature of dialectal variations of syllable accents, with the decision to give them neutral, presuppositionless names of "acute" and "circumflex" receiving universal accord (see Girdenis 2003, 273) and becoming another compromise. The genesis of the concept of the acoustic structure of stress was not dissimilar. A seminal work in this field was a study by Pakerys (1982), which established the idea that the phonetic essence of word stress in Standard Lithuanian consists of a complex of acoustic correlates of uneven power. This concept was later adopted and verified by many researchers of Lithuanian dialects (Kazlauskienė 2001, 39–45; Atkočaitytė 2002, 83–103; Bacevičiūtė 2004, 29–44; Leskauskaitė 2004, 124–145; Jaroslavienė 2010, 29–55); although, in the earlier years, research mostly focused on attempting to find one key correlate of stress (Laigonaitė 1958, 71–100; 1978, 78; Kazlauskas 1966, 119–132; Mikalauskaitė 1975, 76; Vaitkevičiūtė 1995, 5). Most recent studies seem to have effectively returned to the issue of what phonetic categories actually are related to word stress (Kazlauskienė, Sabonytė 2018, 55–62; Sabonytė, Goldshtein 2021, 119–128).

Another approach to the solution of the interaction problem is based on the concept of a compensatory mechanism. Although its logic is not clearly articulated, this concept is commonly used to explain the variability in the differential power of acoustic parameters caused by various linguistic factors. These factors include intrinsic prosody, phonological and phonotactic aspects of linguistic systems, the prosodic structure of language, etc. For instance, if there are systems of short and long vowels in a language, the potential of the duration factor in determining the stress under such circumstances may only be minimal (thus, this correlate of stress must be compensated for by another correlate). Conceptually, this is defined by the so-called Functional Load Hypothesis (for its critique see van Heuven 2018, 49–54). Its key idea is appeals to the position of the acoustic correlate within the hierarchy, which is supposed to depend on the functional load of the correlate at other levels of language (see Berinstein 1979, 1–59).[4] However, compensatory

[4] Cf. the idea of Georg Gerullis on the acoustic structure of Lithuanian syllable accents: "Aus der Mischung dieser vier Komponenten, vielmehr je einer Schattierung von ihnen ergibt sich die Resultante, der litauische Akzent. Dabei ist **der Stärkegrad der einzelnen Komponente** bei jedem Akzent verschieden, das gibt ihm seinen eigenen melodischen Charakter" (1930, 22).

modeling of acoustic features requires too high a level of sophistication from the language users, it becomes too difficult to verify the form such a prosodic system experimentally, and finally, the interactional mechanism turns into complicated algorithms (cf. Bruce 1977, 74–92; Garding 1983, 11–25; Pierrehumbert, Hirschberg 1990, 271–311; Gussenhoven, van den Beuken 2012, 75–107, and others). The proposal to rely on the relationships of the ratios of stressed and unstressed syllables in Standard Lithuanian and to take into account the gradation of power of acoustic correlates (Pakerys 1982) could be seen as a clear illustration of the problems associated with this concept.

The aim of this article is to propose an alternative model of phonetic interaction between prosodic elements in the Baltic languages, whose key principle would be based on clearer and simpler logic. It should be emphasized that the goal is not to categorically refute or reject the existing theoretical and experimental achievements in prosodic research, but rather to find a principle that would make it easier to systematize empirical data and do so more clearly.

The basis of phonetic interaction is sound control. As shown above, the characterization of prosodic elements in terms of separate acoustic parameters is often problematic and unproductive. For this reason, there is a need for a category that would reflect both the general principles of linguistic behavior and the basic physics of sound generation (avoiding excessive focus on small acoustic details of the elements in question). One such category could be based on a broader concept of sound control. This choice is motivated by several arguments. First, such a category allows for a more comprehensive assessment of the acoustic structure variability found in the sounds of a language for prosodic purposes; viz. it enables one to discuss the effect of the speaker's primary intention (not) to control the sound on the entire group of parameters. Second, it allows one to avoid high-level sophistication of the issue, which is based on the assumption that the speaker is able to selectively modify acoustic parameters (sometimes even at the micro level) to serve their needs, whatever they may be. Third, this alternative provides a logical framework for understanding the very concept of prosodic interaction, turning possible differences in sound control into various vectors (paradigmatic and syntagmatic).

Although sound control is likely conveyed by all acoustic parameters in one way or another, for the sake of clarity and simplicity we will rely on the dynamics of fundamental frequency (F0). In short, we will focus on F0 variations over time. Each voiced sound[5] is an object of phonation, i.e. the result of the transformation of expiratory airflow energy into acoustic one (for basic principles, see Stevens 2000; Titze 2000; Plant, Younger 2000, 170–177; Hollien 2014, 395–405; Zhang 2016, 2614–2635, etc.). This process is well-defined in the myoelastic-aerodynamic theory of phonation (van den Berg 1958, 227–244; Titze 2006; for recent studies on this topic, see Švec, Schutte, Chen, Titze 2023, 305–313). It suggests that phonation is the result of the interaction between aerodynamic and myoelastic (muscular elasticity and stiffness) forces (for simplicity, we will call them Acting Acoustic Forces or AAF in this paper). The vocal folds, acting like a valve, block the path of the airflow rushing out of the lungs. As they close, subglottal pressure begins to rise, and when it reaches a critical threshold, the lower edges of the vocal folds begin to open, followed by the upper edges. As they separate, the air rushes out with full force, and the subglottal pressure drops instantly. This phase lasts until the restoring force of the vocal folds counterbalances the effect of the pressure drop, and the folds return to their initial position (for criticism of certain aspects of this theory, particularly the impact of the Bernoulli effect on phonation, see Zhang 2016, 2619–2620).

Therefore, following the logic of the myoelastic-aerodynamic theory, we may say that F0 dynamics directly reflect the changes in AAF. We may treat F0 dynamics as a measure of how efficiently expiratory energy is utilized for phonation. Large F0 perturbations show uneven vocal fold vibrations, which are attributable to a certain ratio of subglottal pressure and vocal fold mass/tension. Since these two parameters are non-constant variables, no strict correlation between them is possible, of course.[6] According to the logic of classical mechanics (to be more precise, Newton's second law), all information about the changes in AAF should be provided by the F0 acceleration (i.e. the derivative of changes in F0 over time) parameter. The greater the F0 acceleration of the sound in question (both positive and negative), the greater the change in the AAF. In experimental research, arguably the most popular indicator conveying this sort of information is F0 range. The difference

[5] For the purposes of this paper, it will be held that the main prosodic information is concentrated in this group of sounds.

[6] It is important to keep in mind that the correlation between intensity and F0 is more evident only in the higher pitch registers (Plant, Younger 2000, 170–177; Hollien 2013, 395–405).

between the maximum and minimum pitch points (however, not over time) also indicates changes in the ratio of subglottal pressure and vocal fold mass.

Of course, F0 acceleration cannot be regarded as a parameter of sound control. It is important to assess not only the magnitude of changes between AAFs, but also their dynamics. Tone curves with the same range/acceleration can differ in shape, and it is therefore necessary to find a parameter that would capture the differences in the F0 contours (not just their "height"). In other words, the quantity of force is not the only relevant aspect: we also need to know its distribution over time. Rising-falling pitch contours, for example, indicate variations in the changes between AAFs, and a flat contour illustrates greater stability and a more balanced AAF. This type of information is provided by the derivative of F0 acceleration over time, the so-called jerk (for a physical explanation of it, see Eager 2016, 1–11; Rajpa, Patil 2016, 82–87). This could be considered a measure of the linearity of F0 change. If F0 rises or falls in a linear trajectory, it is reasonable to assume a stronger intention to stabilize and control the activity of vocal folds and to counteract more proportionately the changes in the force acting on them. Of course, the very fact of change indicates changing control, but the type of change is equally informative, as it allows us to discern the speaker's intentions. In terms of physics, elastic bodies like vocal folds are deformed by the forces acting on them, and the deformation itself is a function of the tension and said forces. If the changes in force (acceleration) are small, the jerk values are very small, and the deformation in this case can be considered instantaneous with respect to acceleration/change of force. To better understand the physical meaning of this parameter, we can draw an analogy with the human body's movements. Jerk limits are necessary, for example, to maintain body balance. The position of the body is controlled by the balance of antagonistic muscle forces. Only by controlling these forces can the desired balance be achieved. If one of the forces changes too quickly, the muscles cannot relax or contract fast enough, resulting in a temporary loss of balance. The time of response to changes in force depends on the physiological properties of the muscles and the attentiveness of the brain: an expected change will be stabilized faster than an unforeseen sudden decrease or increase in force (for more details, see Hogan 1984, 2745–2754; Roren, Mazaguil, Vaquero-Ramos, Deloose, Vidal, Nguyen, Rannou, Wang, Oudre, Lefèvre-Colau 2022, 1–14). Since vocal folds also consist of a layer of muscles, these principles, even in the most general sense, can be applied to the AAF interaction, which is represented by F0 dynamics. This may be one argument for including the jerk parameter into the analysis.

Finally, it is important to add that the relationship between these two parameters (i.e. F0 acceleration and F0 jerk) and the concept of sound control proposed in the present study directly depends on the duration of the sound, too. Sounds with the same F0 acceleration and F0 jerk can differ in duration (despite both of them being derivatives of time). For this reason, sounds with identical pitch contours but different in duration should be interpreted differently. It is likely that a stronger intention to balance AAF would be signaled not only by F0 acceleration and F0 Jerk values that are approaching zero, but also by longer sound duration. In this regard, maximum sound control would be understood as an F0 with zero acceleration and with zero Jerk sustained for a longer period of time.

Therefore, based on all the arguments above, we can propose the following working hypothesis: acoustic differences in F0 control could be reflected in the combination of these three parameters:[7]

F0 Range (st)[8]
F0 Jerk (st/s^3)[9]
Duration (s)

The principle of graphical analysis. This section describes the principle of graphic analysis. Given that three acoustic parameters are to be simultaneously analyzed and related to prosodic elements, this requires a three-dimensional graph, where the elements under study are treated

[7] All these parameters were automatically generated using *Praat* script (see Švageris 2020, 119–157; 2021, 271–305).

[8] The F0 range parameter is chosen instead of average acceleration of F0 for purely mathematical reasons. The fact of the matter is that the average value in certain cases (for example, when the contour of the tone is in the shape of a semicircle) is not sensitive enough and does not convey important information about F0 dynamics. Adding a modulus to the formula does not solve the problem either, because in this case information about the general direction of the pitch change - rise or fall - is lost (i.e. it is no longer known whether the acceleration is positive or negative). The F0 range parameter is somewhat of a compromise in this regard, but, of course, a mathematical way to solve this problem must be found in the future.

[9] For mathematical reasons, all F0 jerk values were converted to positive ones (modulus of average differences of F0 acceleration was calculated).

as mathematical functions that would ideally distribute specific phonetic realizations to separate clusters. An absolute zero can be chosen as a reference point that will help to determine possible differences between compared sounds and (according to the conception proposed in this paper) allow to assess the level of sound control. According to our model, any deviation from the zero point indicates a shift in F0 towards greater or lesser control (balance). As already mentioned, the increasing F0 balance should be reflected in a longer sound duration, accompanied by zero-approaching F0 range and F0 jerk values. Obviously, a quantitative parameter of any sound of speech can only take positive values, as there can be no negative or zero duration of sound in principle. In turn, a zero value of the F0 range (and F0 jerk) indicates that the F0 trajectory is a simple horizontal line since not even a minimal change in F0 is present. If a suitable mathematical method could be found, all of F0 changes should be reflected in positive and negative F0 acceleration values.

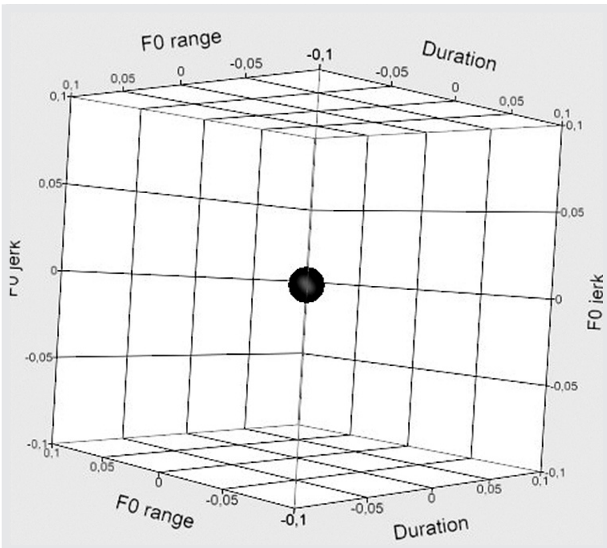


Figure 1. **The reference point in acoustical three-dimensional space**

The change in F0 jerk values, which is the last of the parameters discussed, is dependent on the type of the F0 acceleration. In simple terms, when F0 acceleration is unstable, a greater or lesser deviation in the pitch contour is observed. If both the positive and the negative F0 acceleration stays constant for the entire phase of pronouncing the sound in question (though this never really happens in reality), F0 rises and falls along a linear trajectory, and F0 jerk approaches zero. F0 jerk values change in the opposite direction when two separate parts of the sound differ in F0 acceleration. More precisely, negative F0 jerk values occur in two cases: when the F0 fall suddenly slows down or the F0 rise speeds up; positive F0 jerk values are obtained when the F0 fall accelerates or the F0 rise decelerates. For clarity, let us present the relationship between F0 acceleration and F0 jerk schematically:[10]

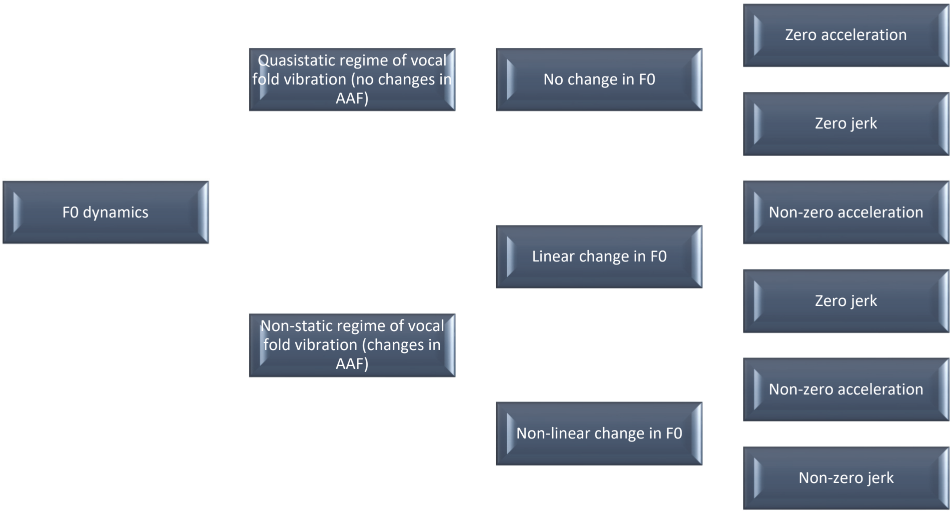


Figure 2. **Relationship between dynamic F0 parameters**

The key takeaway from this diagram is the relationship between F0 acceleration and F0 jerk: constant acceleration (whether it's zero, negative, or positive) results in zero jerk and changing (nonlinear) acceleration results in non-zero jerk. As mentioned above, this dichotomy allows us to think in terms of two categories: the amount of force and its distribution over time. This difference could serve as a starting point for explaining the interaction between prosodic elements: in this case, some elements would be understood as resulting in F0 changes, and others would ensure the stability and balance of those F0 changes over time.

[10] For positive and negative F0 jerk values see Footnote 9.

Analysis of specific cases. **The stress pattern in Standard Lithuanian.**[11] Let us begin the verification of the proposed model by analyzing the stress pattern in Standard Lithuanian. Presented below is a graphical analysis of the minimal pairs [**²ˈlʲeː**lʲeːs; ▲/▲] and [lʲeː**²ˈlʲeːs** ●/●].[12] Although all long stressed syllables in Lithuanian have syllable accents, this factor is neutralized here, as the accent of both syllables (when they receive stress) is identical (in both cases it is circumflex). Our analysis required a word whose phonetic structure would not affect F0 dynamics (at least on lower degree). This happens, for instance, when short vowels are located between plosives (see Gussenhoven 2004, 9). Lateral consonants in the onset and coda of the syllable should reduce this risk. For the same reason, long vowels were chosen for analysis rather than short ones. The only thing to note is that the syllable types differ, with the first syllable being open and the second closed. This could affect the duration of the vowels (open syllable vowels tend to be slightly longer). However, this latter factor should not have a significant impact on model verification. Its significance could increase only if the distribution of the other two parameters shows no clear tendencies.

So, the three acoustic parameters – namely, F0 range, F0 jerk, and duration – transferred to the 3D graph should, in effect, reflect the differences in tonal dynamics (in a broader sense, of sound control or balance) between stressed and unstressed vowels. The words were pronounced with a declarative intonation in sentence-final positions and emphasized with a focus. Vowel positions are marked with different colors: gray dots ● represent the realizations of pretonic vowels [lʲeː], gray triangles ▲ represent posttonic vowels [lʲeːs], black triangles ▲ represent initial stressed [**²ˈlʲeː**] vowels, and black dots ● represent final stressed [**²ˈlʲeːs**] vowels. It can be seen that in the three-dimensional space, the data are grouped into separate clusters, as all vowel groups are distinct from each other (there is a high probability that the ternary acoustic model successfully responds to a stress position change in the word; $p=3.82E-26$).[13] The phonetic realization differs not only between stressed and unstressed vowels but also between unstressed ones too (compare the groups of gray triangles and gray dots). Since the distribution principle

[11] Here and in all other cases acoustic data were taken from previous studies conducted by the present author.

[12] (Lith. *lėlė*; Eng. *a doll*). It is a Nominativus Pluralis in the first form and Genitivus Singularis in the second.

[13] For this purpose, two-way ANOVA with normalized data was performed.

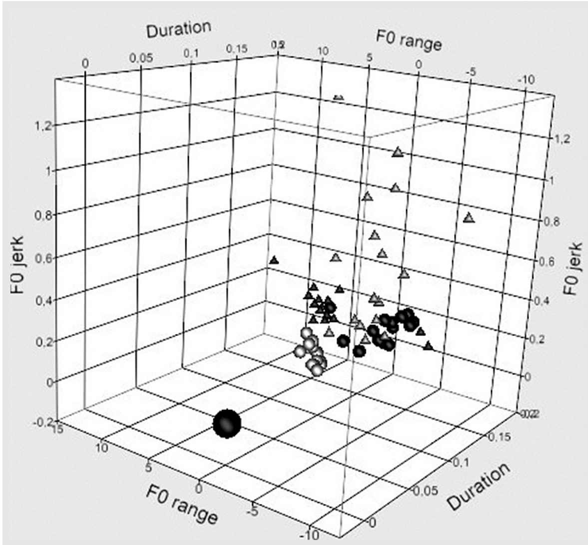


Figure 3. **[2ˈlʲeːlʲeːs; ▲/▲] and [lʲeː2ˈlʲeːs ●/●][14] parameter distribution (declarative intonation, final position, + phrasal stress)**

is the most relevant for our purposes, we will not pay much attention to specific values this time. The realization of posttonic vowels (gray triangles) is distributed most widely in the 3D space. Their large deviation is particularly visible in terms of F0 jerk. This indicates a very low F0 balance, strong deformation, sharp curvature, and "scattering" of F0 points.

This result should not be very surprising, as unstressed syllables in the final position are often subject to reduction and thus to a large F0 perturbation. According to the logic of our model, we could say that the control of the F0 of these vowels is the weakest. In turn, pretonic vowels seem to lose their prosodic weight due to the largest shift towards the center point (the acoustic zero). Their realizations are not only of the shortest duration, but also those closest to zero values of F0 range and F0 jerk. All of this means that, from a phonetic perspective, pretonic and final posttonic syllables are different. This is because the position of the stress determines the distribution of acoustic energy throughout the word: when the stress falls on the final syllable, the

[14] Syllable accents are marked using adopted IPA symbols (Bakšienė, Čepaitienė 2017, 105–135).

pretonic syllables cannot be reduced to the same degree as the posttonic ones, since the energy must be maintained until the stressed syllable. It seems that the prosodic neutrality/passivity and energy transfer functions of these syllables are realized through static F0 dynamics. Their pitch changes are minimized (low F0 range and F0 jerk values) and the duration is reduced, but their remaining greater resistance to reduction allows to distinguish them from posttonic vowels in this case. If the opposite were true, i.e. if the phonetic realizations of unstressed syllables were identical, this could lead to problems in determining the prosodic boundaries of the word. Moreover, this phonetic non-equivalence of unstressed vowels is probably the most problematic one when attempting to determine the prosodic contrast using traditional methods.

As already mentioned above, stressed vowels are not all made equal, either. They occupy an intermediate position between prosodically weak syllables. The greatest difference between them is in terms of the F0 range. It is clear that the stress on the final syllable halted the deformation of the tone to an extent and made it more balanced. This is shown by the graphical difference between the gray triangles and black dots: although the F0 range values of both groups of vowels are more or less the same, the difference is very obvious in terms of F0 jerk (values of the black dots are much farther from zero). In other words, we get the most information about the differentiation of final syllables (stressed and unstressed) from the ratio between F0 range and F0 jerk parameters. Although the range of variation is about same, the dynamics are radically different.

All these tendencies, namely the phonetic differences between unstressed vowels, the intermediate position of stressed syllables in the 3D graph, the noticeable shift of pretonic vowels towards the acoustic zero, the greatest tonal perturbation of final posttonic vowels, and the information provided by the ratio between F0 range and F0 jerk values, all of them suggest that, at least in this specific case, the acoustic effect of stress is created by intersyllablic differences in tonal balance. When stress is placed at the beginning of a word, tonal control only concerns the first syllable, and the final syllable is simply reduced: F0 balance drastically decreases, with its perturbations coming to the fore (we might say that the inertial tonal chain breaks at the initial stressed syllable). It may be presumed that in such cases, the speaker does not attempt to articulate the posttonic final syllable with a lower pitch (or make it shorter and less intense), but rather is more inclined not to articulate (control) it altogether. A different scenario is observed when stress shifts to the final syllable of a word. In that case, the F0 dynamics of the pretonic vowel change into a static state: F0 range and F0 jerk values drastically decrease, and the duration shrinks. This arrangement of tonal features is probably determined by two factors: the need to avoid prosodic competition with the stressed syllable and the necessity to maintain the distribution of acoustic energy, to carry it forward to the final (stressed) vowel (i.e., not to break the tonal chain). Therefore, stress, in this case, functions not as a factor of focusing acoustic energy on a single syllable, but as a means of constructing tonal/prosodic chains and marking their boundaries. The possible acoustic advantage of stressed syllables over unstressed ones is rather a side effect (which is not always detectable), more often observed when comparing stressed initial syllables to unstressed final syllables (due to the latter's significant reduction under certain intonational conditions). In simple terms, the direction of acoustic analysis of word stress should be not vertical (i.e. one should not compare mean or maximum levels of pitch or intensity) but horizontal, i.e. one should focus on the static/dynamic, deformed/non-deformed, linear/non-linear (more generally, controlled / not controlled or balanced/non-balanced) F0 sequences of vowels. It is the extreme element of such sequences, formed by a more prominent change in tonal control, which is to be considered the phonetic expression of the syntagmatic nature of stress.

Of course, all the above arguments would be rendered meaningless if it turned out that the presumed phonetic form of the stress is not resistant to changes in intonational conditions. Figure 4 shows members of the same minimal pair pronounced with an interrogative intonation in the focus. The first thing to note is the increased F0 range values in the final syllables (both stressed and unstressed). This means that the F0 of these syllables, regardless of stress, changes (rises) intensively for the majority of the sound duration. Nevertheless, the data stays divided ($p=1.67E-34$), with the gray triangles and black dots once again forming separate clusters. This differentiation is attributable to the separation of realizations along the F0 range axis. It might be argued that the tone of stressed vowels simply rises to a greater extent. However, one should note that this difference also determines a different relationship between the F0 range and the F0 jerk values. Recall that the increasing values of these parameters weaken tonal control and decreasing values strengthen it. A straightforward logic follows from this:

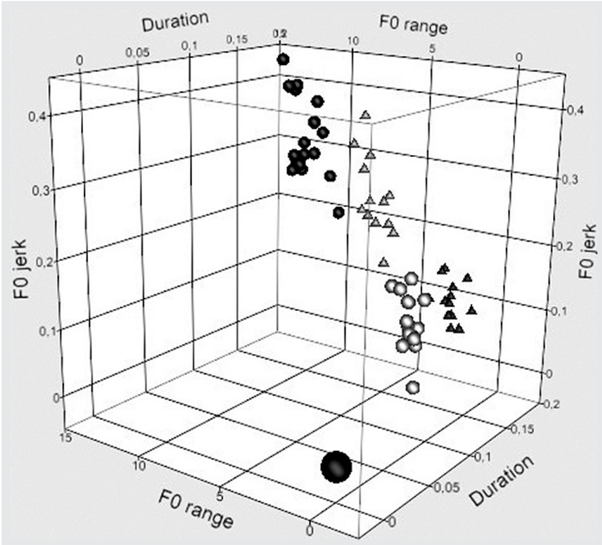


Figure 4. **[^{2}ˈlʲeːlʲeːs; ▲/▲] and [lʲeː^{2}ˈlʲeːs ●/●] parameter distribution (interrogative intonation, final position, + phrasal stress)**[15]

if a tone with the same extent of change has different F0 jerk values, then the control level also differs. The gray triangles, compared with the black dots, represent a vowel group characterized by a smaller pitch range, while their F0 jerk values are almost identical. Therefore, the F0 range / F0 jerk ratio of unstressed final syllables is smaller once again (i.e., the intensity of pitch change is accompanied by an equally intense pitch deformation, resulting in a lower balance of F0) than that of the stressed ones. The only difference is that when the same words were pronounced with statement intonation, the identical ratio was determined by strong posttonic syllable reduction. It is important to understand that the intention to use more acoustic energy does not directly imply its distribution over time. From all these instances, it becomes clear that the decrease or increase in acoustic forces depends on intonation, while the balance of these forces (from a syntagmatic perspective) depends on stress. Of course, when analyzing the differentiation of stress under interrogative intonation, one should not focus exclusively on the final syllables. The difference in the dynamics of stressed and unstressed syllables

15 Data from a previous study is used here (see Švageris 2015).

would receive the greatest functional load only if the initial syllables of the word with different prosodic status would coincide in their tonal structure. Nevertheless, it is evident that a change in intonational conditions has little effect on acoustic structure of pretonic vowels. They are once again closest to the central point of the 3D space. This implies a static, even inertial, change in their pitch. The prosodic weight of stressed initial syllables seems to be increased slightly by more intense F0 dynamics (i.e., higher F0 range and F0 jerk values) and by a larger quantity. Even if we assume that the F0 dynamics of initial syllables, regardless of stress, are very similar, the trend lines connecting the realization of stressed and unstressed vowels would still differ. In other words, the trend line connecting black and gray triangles is steeper than the corresponding trend line connecting gray and black dots. This should also mean a greater change in F0 control between the syllables, an extension of the tonal chain to the end of the word when the stress is on the final syllable, and its interruption when it is on the first. The same tendency was observed in the previous graph (see Figure 3). It should be emphasized that the arguments outline by no means imply strict deviations of the tonal control parameter or its absolute values. The ratio of syllables is more important in this regard. A greater need for making the impression of a more carefully modulated pitch under interrogative intonation, presumably, arises when the stress falls on the first syllable. As previously mentioned, the interrogative intonation raises the level of phonation in the final syllable, thus creating additional competitive conditions for the stressed syllable (we might say that in such cases the distinguishing features of intonation and stress compete with each other). There is less competition when both prosodic elements are in the same zone of activity, i.e. when the final syllable is stressed in a word pronounced with the interrogative intonation and the first syllable with the declarative. In these cases, even a small modulation might be enough to differentiate it from a static (short flat pitch) the pretonic and especially the deformed posttonic final syllables (reduced syllable). Without taking this circumstance into account, the prosodic importance of the differences between stressed and unstressed syllables in one position can be overestimated. In other words, despite established differences between syllables in the same position, the focus should be on the syntagmatic axis: one should compare the dynamics of tonal transfer from one syllable to another. As seen above, declarative intonation resulted in a gradual decrease in acoustic energy between the syllables, while the interrogtive intonation caused an increase. When the stress falls on the first syllable, the acoustic effect of a tonal chains being interrupted at the midpoint of the word in both cases was due to the decreased level of tonal control for both final posttonic vowels (the decreased F0 range and F0 jerk ratio). When the word's final syllable was stressed, the impression of a two-part tone chain (consisting of both syllables of the word) was created. This was likely caused by the acoustic zero-approaching dynamics of the pretonic vowel tone, accompanied by a slightly more balanced F0 range / F0 jerk ratio in the stressed final syllables.[16]

Syllable accents in Lithuanian and Latvian.[17] Since the aim of this article is to explain the principle of prosodic interaction as clearly as possible, rather than providing a detailed acoustic analysis of all prosodic elements (in all possible positions), we will immediately proceed to another key issue. It is also crucial to determine whether the same ternary acoustic model is able to capture the differences between syllable accents in Baltic languages.[18] If our analysis of stress focused on the horizontal (syntagmatic) comparison, this time the vector is rotated to examine the paradigmatic relation between syllables. The only difference, however, is that this time we will look to answer the question of whether the vowels pronounced with different syllable accents differ in tonal control. It should also be noted that although it is sometimes presumed that syllable accents in the Baltic languages used to be realized in non-stressed syllables (according to Saussure's law), currently their differences are more prominent only in stressed syllables.

The following 3D graphs (see Figures 5 and 6) show the phonetic realizations of the syllable accents of two Lithuanian and Latvian dialects.[19] The test words are pronounced with declarative intonation and emphasized with phrasal stress in a central position of the phrase. The phonetic structure

[16] On similar tendencies of the acoustic structure of stress in Lithuanian and Latvian languages see Švageris 2022, 71–95.

[17] It bears mentioning that the functionality of syllable accents, especially of long vowels in Standard Lithuanian and a number of its dialects has been debated for quite some time. The syllable accents of diphthongs differ mainly in the quality of the first vocalic component and seem to be determined by the position of the stress. In other words, one need not invoke the concept of syllable accents in this case (see Kazlauskas 1966, 127; Pakerys 1982, 147).

[18] For syllable accents in the Baltic languages and their interrelation see Endzelīns 1951, 34–48; Rudzīte 1993, 99–115; Girdenis 2014, 287–387.

[19] All data used here comes from the present author's dissertation (see Švageris 2015).

of the Lithuanian examples, where the test vowel is situated between plosives (compared to the Latvian examples, where it is between sonants), may have had some negative impact on the accurate reflection of tonal dynamics. In fact, standard Lithuanian and Latvian syllable accent terms (such as falling, rising, etc.) inherently imply differences in tonal dynamics, which should make them easily detectable using the methodological tools employed in this article. The data distribution illustrated in Figure 5 clearly supports this hypothesis. The realizations of both acute/falling (black dots) and circumflex/rising (gray dots) vowels form separate clusters (p=4.47E-08). A tendency towards maintaining a smoother tonal dynamic (or sound control in a broader sense) is indicated by the longer duration of circumflex vowels and a higher ratio of F0 range to F0 jerk. This means that in such cases, the longer tonal curve is much less deformed (this is particularly evident from the lower F0 jerk values). In turn, acute (falling) vowels are often associated with glottalization (though it is not always regular), which is the main factor causing tonal deformation.

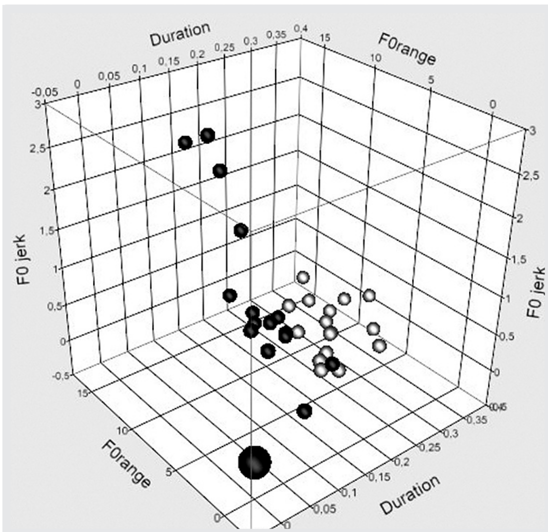


Figure 5. **Lithuanian (dialect of North Žemaitian) syllable accents [^{1}dʲi:ks][20] (acute) and [^{2}dʲi:ks][21] (circumflex)**

[20] Eng. *Will sprout*

[21] Eng. *lazy*

The tone of glottalized vowels most frequently rises sharply in the initial phase of the sound and then falls rapidly or is no longer present in the second phase (when phonation is interrupted in the central part). The result of such articulation is a significant curvature (or even a break) of the tone contour, indicated by sharp increase in F0 jerk values.

The same scenario is repeated in the Latvian examples (see Figure 6). The type of parameter distribution can be considered identical. Vowels of the level tone (gray dots) are characterized by longer duration and a higher F0 range / F0 ratio, while broken tone vowels (black dots) show the opposite characteristics (p=7.37E-15). This recurring pattern of data is especially important for the logical justification of the interaction principle. It shows that intense tonal changes in these cases do not interfere with the realization of syllable accents. In other words, the intonation-related increase in the F0 range (in simple terms, an increase in force) does not prevent syllable accents from achieving different F0 balances (F0 range / F0 ratios + duration). It should also be noted that in the cases discussed, the words contain only one syllable.

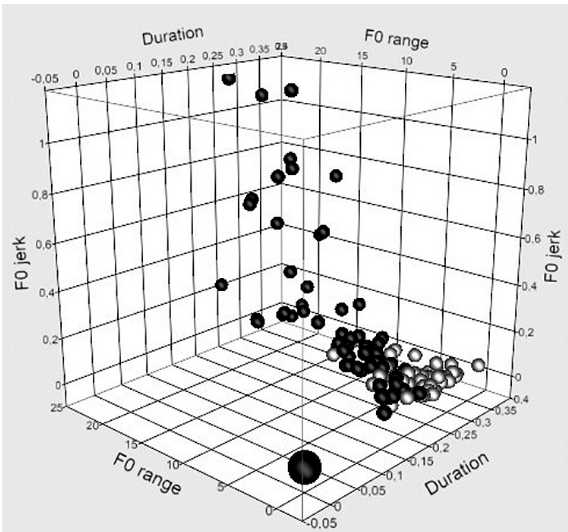


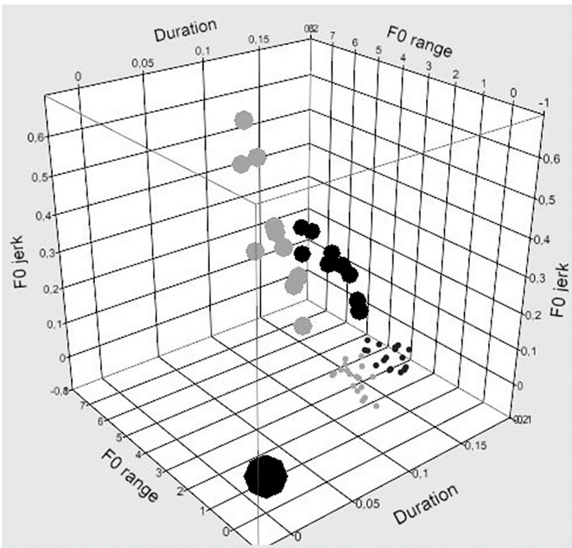
Figure 6. **Latvian syllable accents (Middle dialect) [plâ:ns][22] (broken tone) and [plã:ns][23] (level tone)**

[22] Eng. *thin*

[23] Eng. *A plan*

The situation changes when other syllables in a multisyllabic word start competing with the stressed syllable pronounced with a syllabic accent. The distribution of acoustic energy among all syllables of a word, as well as its inertia, particularly affects the vowels with deformed tones. As already mentioned above, the static tone of pretonic vowels is not reduced to the same extent as that of the unstressed ones in the final syllable of the word. It is probably for the same reason that, under certain prosodic/intonational conditions, acute/broken tones become deglottalized, since the need to maintain the inertia of acoustic energy until the more prominent posttonic syllable (e.g. due to a secondary stress or interrogative intonation) is hardly compatible with strong F0 deactivation, even in the primary stressed syllable (see Kazlauskas 1968, 6; Girdenis 1974, 160–198; 1996, 71–84; Švageris 2020, 119–157). The most favorable conditions for this dynamic type of F0 (as in the cases analyzed) occur when the influence of adjacent (non-stressed) syllables is minimized, i.e. when the glottalized vowel is pronounced before the pause (e.g. at the end of the phrase) and has a strong phrasal stress. All of these symptoms once again confirm that the prosodic structure of a word is the result of a combination of the intensity of tonal change (a relative equivalent of acoustic force) and the its balance (the distribution of acoustic force) over time. The necessity to balance the F0 dynamics of syllabic nuclei and create the effect of a syntagmatic tonal chain is the main factor correcting/modifying/determining the prosodic structure of a word and a syllable.

The influence of phrasal focus on the phonetic realization of word stress. To further develop the ideas of the previous paragraph and verify the interactive model from another perspective, we can examine the effect of phrasal stress (i.e. the intonational factor) on the phonetic realization of word stress. In this case, once again, we use the prosodic data of the Standard Lithuanian. The illustrations (see Figures 7 and 8) show the realization of the same words, both stressed (darker color tones) and unstressed (lighter color tones) with phrasal focus. However, this time, the words were taken from the initial part of the phrase to avoid sentence-final effects (see Berkovitz 1984, 255–256 for more on this). Admittedly, it is not difficult to guess how the intersyllablic tonal dynamics would change when the words are in a prosodically weak position. It has long been argued that the acoustic characteristics of a word tend to fade under such intonational conditions. Now, based on the empirical material presented here, this prosodic phenomenon can be explained in detail. The 3D graphs clearly illustrate that the F0 dynamics of syllables of words that lost phrasal stress "undergo centralization", i.e., regardless of the word stress factor, the tone of the vowels starts approaching the acoustic zero and the F0 change becomes static, losing its prosodic independence ($p = 1.13E-32$). This is particularly evident when the stress is on the first syllable of the word ([^{2}ˈ**lʲeː**lʲeːs ●/●] and [^{2}ˈ**lʲeː**lʲeːs͈ •/•]; see Figure 7). When the same word is emphasized with focus, the previously observed difference becomes apparent: the F0 Jerk values of the unstressed vowels at the end of the word increase (bigger gray dots), indicating weakened tonal control in this position (again, referring to the F0 range / F0 Jerk ratio). When the same final unstressed vowel is pronounced in a weak position (smaller gray dots), the tonal deformation is clearly slowed down, presumably because under such conditions the syllables are accentually slide towards the emphasized word in the phrase. Similarly, the tone of the initial stressed vowel also slides towards the acoustic zero (compare the ratio of bigger and smaller black dots). The acoustic difference between syllables of different prosodic status under different phrasal conditions is maintained only by the duration (the stressed vowels are slightly longer), but it remains an open question how much weight this feature preserves in such cases.



F i g u r e 7. **[^{2}ˈlʲeːlʲeːs ●/●] and [^{2}ˈlʲeːlʲeːs͈ •/•] parameter distribution (declarative intonation, initial position, +/– phrasal stress)**

This tendency remains the same when the location of the stress in words has been changed (p=0.03). When stress shifted to the end of the word, the shift of the tonal dynamics of all syllables of the word towards the acoustic zero remained the same. We can be reasonably sure that the prosodic neutralization of words (or at least a clear weakening) in the weak phrasal position is manifested by the leveling of the tonal dynamics of vowels, which makes them static and fully inert. It may be observed that in this case, the tonal expression of pretonic syllables coincided with the effect of the intonational factor in question on the prosodic structure of the word. For this reason, the tonal characteristics of all vowels in these positions are similar, and only the final syllables of the emphasized words with phrasal stress are separated from the rest in the graph (see the bigger black dots in Figure 8). This proves once again that intersyllabic differences in F0 dynamics are the distinctive feature of stress, detectable using the method applied in this study. Furthermore, the chosen concept of sound control, based on the trinary acoustic model, seems to clarify the prosodic hierarchy. There is good reason to believe that the prosodic structure of the linguistic segments under consideration is primarily determined by the prosodic factors at the phrasal level and only then those at the word (word stress) and syllable (syllable accent) levels.

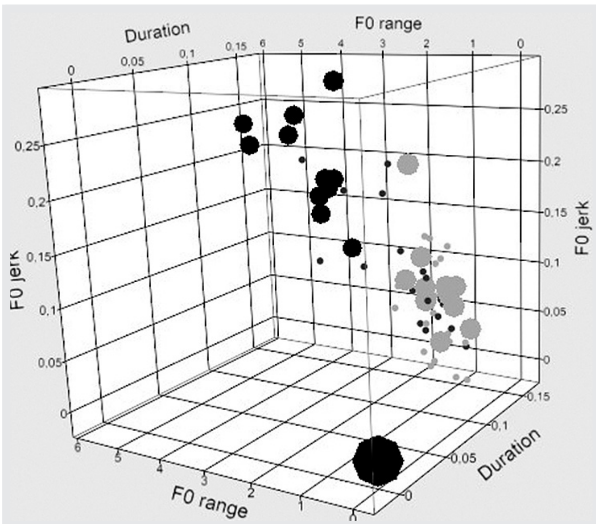


Figure 8. **[lʲeː^{2}ˈlʲeːs ●/●] and [lʲeː^{2}ˈlʲeːs •/•] parameter data distribution (declarative intonation, initial position, +/- phrasal stress)**

The principle of phonetic interaction of prosodic elements. Therefore, based on everything that has been presented in this article, it is possible to clarify and illustrate the principle of phonetic interaction of prosodic elements in the Baltic languages on a graph (see Figure 9). Although the development of this model, understandably, is still in its embryonic stage, the results of a preliminary analysis of empirical data provides at least an initial broad overview that could become the subject of further analysis and critique. The point of reference for the model is the category of sound control/balance, interpreted in terms of F0 dynamics. Any change in vowel tone, in line with the main principles of myoelastic-aerodynamic phonation theory, is determined by the interaction between AAF (aerodynamic expiratory airflow force and the elasticity and tension forces of the vocal folds opposing it). It is important to distinguish two aspects here, the rate of F0 change and its stability. To put it in terms of classical mechanics, we focus on the F0 acceleration and its derivative with respect to time (F0 jerk). Newton's second law, even in a very simplified form, allows us to regard F0 control as expression of changes in AAF and its stability or balance (in other words, of the amount of force and its distribution over time). This very distinction serves as the logical/physical basis for understanding the interaction of prosodic elements: some elements imply the amount of force (the rate of F0 change), and others its distribution over time (stability/instability, continuity/discontinuity of F0 change). For the sake of accuracy, these two parameters must be assessed in the context of the duration parameter, because in very short sounds, the prosodic weight of the F0 dynamics can be strongly restricted.

Figure 9. **The principal scheme of phonetic interaction of prosodic elements**

All data and arguments presented here allow us to consider the hypothesis that F0 change, viewed both in terms of paradigmatic and syntagmatic vectors, is the prerogative of phrasal intonation. As has been seen in the 3D graphs above, especially when comparing the syllable accents in Lithuanian and Latvian languages, differences in the F0 range parameter (which was chosen instead of the acceleration parameter for mathematical reasons) did not affect the type of data differentiation. It was evident that vowels of different prosodic status equally successfully formed separate clusters, simply by shifting up or down along the F0 range axis. Therefore, the rate of F0 change on both the syllabic and intersyllabic levels is determined by intonation. The other two elements, namely word stress and syllable accents, are to be regarded as factors that regulate the balance of F0 change at the intersyllabic level. Word stress determines the syntagmatic distribution of tonal dynamics of this kind, and syllable accents determine the paradigmatic one.

The gray area not yet covered by this interactive model is a clearer understanding of prosodic neutralization phenomena. As has been shown by the distribution of data in the graphs, both words stress and syllable accents lose their phonetic identity in weak positions of the phrase due to the transition of their tonal dynamics to a static-inertial state. However, it is necessary to clarify how this phenomenon depends on the center of intonational emphasis in the phrase. Presumably, description of such prosodic processes may benefit from the concept of inertia. The level of phonation in a syllable, determined by the prominence of phrasal emphasis, probably has a direct effect on the tonal dynamics of adjacent syllables. This scenario is suggested by the blurring boundaries between syllable accents and word stress under certain intonational conditions, but all these aspects require further analysis. This could be a useful direction in developing and improving the interactive model proposed in this study.

FONETINĖ BALTŲ KALBŲ PROZODINIŲ ELEMENTŲ INTERAKCIJA: GALIMAS TEORINIS MODELIS

Santrauka

Šio straipsnio tikslas pasiūlyti fonetinės baltų kalbų prozodinių elementų interakcijos modelį. Teorinis ir eksperimentinis šios krypties tyrimų įdirbis byloja, kad šios problemos

ne(iš)sprendimas didina klaidingo akustinių duomenų interpretavimo ar priskyrimo vienam ar kitam prozodiniam vienetui riziką. Ilgieji skiemenys, pavyzdžiui, kartais tampa kirčio, priegaidės ir intonacijos sinchroninės raiškos lauku, todėl kai kurių garsų kokybiniai ir kiekybiniai rodikliai vienu metu turi atstovauti ištisai elementų grupei. Kyla natūralus klausimas, kaip tokiomis sąlygomis yra realizuojami ir suderinami visų jų skiriamieji požymiai. Kol kas išsamaus atsakymo į šį klausimą neturėta, todėl buvo privalu imtis naujo tyrimo esamai probleminei situacijai spręsti.

Pagrindinis šiame straipsnyje siūlomo modelio atskaitos taškas yra garso kontrolės kategorija, interpretuojama per F0 dinamikos prizmę. Vienoks ar kitoks balsių tono kitimas, atsižvelgiant į svarbiausias mioelastinės-aerodinaminės fonacijos teorijos nuostatas, yra sąlygojamas sąveikos tarp aerodinaminės ekspiracinės oro srauto jėgos ir jai besipriešinančių balso klosčių tamprumo, įtempimo jėgų. Čia svarbu išskirti du aspektus – F0 kitimo intensyvumą ir stabilumą. Ši skirtis ir yra loginis / fizikinis prozodinių elementų interakcijos pagrindas – vieni elementai implikuoja akustinės jėgos kiekį (F0 kitimo spartą), o kiti – tos jėgos distribuciją laike (F0 kitimo pastovumą / nepastovumą, tolydumą / netolydumą). Išanalizuoti duomenys ir išsakyti argumentai leidžia svarstyti hipotezę, kad pats F0 pokytis, žvelgiant į jį tiek pagal paradigminį, tiek pagal sintagminį vektorių, yra frazės intonacijos prerogatyva, o kiti du elementai, kirtis ir priegaidė, laikytini veiksniais, kurie reguliuoja skirtingą F0 pokyčio balansą tarpskiemeniniu lygiu. Kirtis lemia sintagminę tokios tono dinamikos distribuciją, o priegaidė – paradigminę.

REFERENCES

Atkočaitytė, Daiva 2002, *Pietų žemaičių raseiniškių prozodija ir vokalizmas*, Vilnius: Lietuvių kalbos instituto leidykla.

Bacevičiūtė, Rima 2004, Šakių šnektos prozodija ir vokalizmas, Vilnius: Lietuvių kalbos instituto leidykla.

Bacevičiūtė, Rima 2011, Vakarų aukštaičių kauniškių monoftongų priegaidžių fonetiniai požymiai, *Baltistica* 7 priedas, 13–26.

Bakšienė Rima 2016, *Vakarų aukštaičių kauniškių priegaidės: Marijampolės šnekta*, Vilnius: Lietuvių kalbos institutas.

Bakšienė, Rima, Agnė Čepaitienė, 2017, Tarptautinės fonetinės abėcėlės taikymo lietuvių tarmių garsams galimybės, *Baltistica*, 52(1), 105–135.

Baranowski, Anton, Hugo Weber 1882, *Ostlitauische Texte*, Weimar.

van den Berg, Janwillem 1958, Myoelastic-aerodynamic theory of voice production, *Journal of Speech and Hearing Research* 3, 227–244.

Berinstein, Ava E. 1979, A cross-linguistic study on the perception and production of Stress, *UCLA Working Papers in Phonetics* 47, 1–59.

Berkovitz, Rochele 1984, Duration and fundamental frequency in sentence-final intonation, *Journal of Phonetics* 12, 255–265.

Bruce, Gösta 1977, *Swedish Word Accents in Sentence Perspective*, CWK Gleerup.

Eager, David 2016, Beyond velocity and acceleration: jerk, snap and higher derivatives, *European Journal of Physics* 37, 1–11.

Ekblom, Richard 1933, *Die lettischen Akzentarten*, Uppsala.

El Zarka, Dina, Barbara Schuppler, Carina Lozo, Wolfgang Eibler, Patrick Wurzwallner 2017, Acoustic correlates of stress and accent in Standard Austrian German, in Sylvia Moosmüller, Carolin Schmid, Manfred Sellner (Hrsg), *Phonetik in und über Österreich* (= *Veröffentlichungen zur Linguistik und Kommunikationsforschung* 31), Wien: Verlag der Österreichischen Akademie der Wissenschaften, 15–44.

Endzelīns, Jānis 1951, *Latviešu valodas gramatika*, Rīga: Latvijas valsts izdevniecība.

Fant, Gunnar, Anita Kruckenberg 1994, Towards an integrated view of stress correlates, *Working Papers* 41, 42–45.

Fisher-Jørgensen, Eli 1989, Phonetic analysis of the stød in Standard Danish, *Phonetica* 46, 1–59.

Frings, Theodor 1934, Der rheinische und der litauische accent, *Beiträge zur Geschichte der deutschen Sprache und Literatur* (PBB) 58, 110–149, Leipzig.

Fox, Anthony 2000, *Prosodic Features and Prosodic Structure. The Phonology of Suprasegmentals*, Oxford University Press.

Fournier, Rachel, Jo Verhoeven, Marc Swerts, Carlos Gussenhoven 2006, Perceiving word prosodic contrasts as a function of sentence prosody in two Dutch Limburgian dialects, *Journal of Phonetics* 34(1), 29–48.

Gårding, Eva 1983, A generative model of intonation, in Anne Cutler, D. Robert Ladd (eds.), *Prosody: Models and Measurements*, Berlin, Heidelberg: Springer, 11–25.

Gerullis, Georg 1930, *Litauische Dialekstudien*, Leipzig: Markert & Petters Verlag.

Girdenis, Aleksas 1967, Mažeikių tarmės priegaidžių fonetinės ypatybės, *Kalbotyra* 15, 31–41.

Girdenis, Aleksas 1974, Prozodinės priegaidžių ypatybės šiaurės žemaičių žemaičių tarmėje (trukmė, pagrindinis tonas, intensyvumas), *Eksperimentinė ir praktinė fonetika*, 160–198.

Girdenis, Aleksas 2014, *Theoretical Foundations of Lithuanian Phonology*, Vilnius: Eugrimas.

Girdenis, Aleksas, Aldonas Pupkis 1974, Pietinių vakarų aukštaičių priegaidės (prozodiniai požymiai), in *Eksperimentinė ir praktinė fonetika*, 107–125.

Gordon, Matthew, Timo Roettger 2017, Acoustic correlates of word stress. A cross-linguistic survey, *Linguistics Vanguard*, 1–11.

Grigorjevs, Juris, Andris Remerts 2004, Latviešu literārās valodas intonāciju funkcionalitāte mūsdienās*, Baltu filoloģija* 8(2) 33–50.

Gussenhoven, Carlos 2004, *The Phonology of Tone and Intonation*, Cambridge: Cambridge University Press.

Gussenhoven, Carlos, Peters Jörg 2004, A tonal analysis of Cologne Schärfung, *Phonology* 21, 251–285.

Gussenhoven, Carlos, Frank van den Beuken 2012, Contrasting the high rise and the low rise intonations in a dialect with the Central Franconian tone, *The Linguistic Review* 29, 75–107.

Plant Randall L., Ross M. Younger 2000, The interrelationship of subglottic air pressure, fundamental frequency, and vocal intensity during speech, *Journal of Voice* 14(2), 170–177.

Hollien, Harry 2013, Vocal fold dynamics for frequency change, *Journal of Voice* 28(4), 395–405.

Hogan, Neville 1984, An organizing principle for a class of voluntary movements, *The Journal of Neuroscience* 4(11), 2745–2754.

Hualde, José Ignacio, Tomas Riad 2014, Word accent and intonation in Baltic, *Speech Prosody* 7, 668–672.

Hualde, José Ignacio, Tomas Riad 2018, Interaction between word accent and intonational boundaries in Latvian, in *Sitxth International Symposium on Tonal Aspects of Languages* (TAL2018), Berlin, 27–31.

Jaroslavienė, Jurgita 2010, *Rytinių kauniškių Prienų šnektos fonologija*, Daktaro disertacija, Vilnius: Vilniaus universitetas.

Jaunius, Kazimieras 1911, *Lietuvjų kalbos gramatika*, Peterburgas.

Jurgec, Peter 2007, Acoustic analysis of tones in contemporary standard Slovene: preliminary findings, *Slovenski jezik / Slovene Linguistic Studies* 6, 195–207.

Kariņš, Krišjānis A. 1996, *The Prosodic Structure of Latvian*, PhD Thesis, University of Pennsylvania.

Kazlauskas, Jonas 1966, Fonologinė kirčio raidos baltų kalbose interpretacija, *Baltistica* 2(2), 119–132.

Kazlauskienė, Asta 2001, Pietinių vakarų aukštaičių tarmės balsių spektras ir prozodija, *Lituanistica* 47 [2001/3], 39–45.

Kazlauskienė, Asta, Regina Sabonytė 2018, F0 in Lithuanian, the indicator of stress, syllable accent, or intonation?, *Human Language Technologies – The Baltic Perspective*, 55–62.

Kudirka, Robertas 2004, Dabartinės latvių kalbos priegaidės ir jų akustiniai požymiai (trukmė, intensyvumas, pagrindinis tonas, kokybė), *Baltistica* 39(2), 233–246.

Kurschat, Friedrich 1876, *Grammatik der Littauischen Sprache*, Halle.

Ladefoged, Peter 2003, *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques*, Blackwell Publishing.

Laigonaitė, Adelė 1958, Dėl lietuvių kalbos kirčio ir priegaidės supratimo, *Kalbotyra* 1, 71–100.

Laigonaitė, Adelė 1978, *Lietuvių kalbos akcentologija*, Vilnius: Mokslas.

Laua, Alise 1969, *Latviešu literārās valodas fonētika*, Rīga: Zvaigzne.

Laua, Alise 1997, *Latviešu literārās valodas fonētika* (4. pārstrād. un papildin. izd.), Rīga: Zvaigzne ABC.

Lehiste, Ilse 1970, *Suprasegmentals*, Cambridge.

Lehiste, Ilse, Pavel Ivić 1986, *Word and Sentence Prosody in Serbocroatian*, London, Cambridge.

Leskauskaitė, Asta 2004, *Pietų aukštaičių vokalizmo ir prozodijos bruožai*, Vilnius: Lietuvių kalbos institutas.

Lieberman, Philip 1960, Some acoustic correlates of word stress in American English, *The Journal of the Acoustical Society of America* 32(4), 451–454.

Liepa, Elmārs 1979, *Vokālisma un zilbju kvantitāte latviešu literārajā valodā*, Rīga: Zinātne.

Lippus, Pärtel, Eval Liina Asu, Mari-Liis Kalvik 2014, An acoustic study of Estonian word stress, *Speech Prosody* 7, 232–235.

LVG 2013 – Daina Nītiņa, Juris Grigorjevs (red.), *Latviešu valodas gramatika*, Rīga: LU Latviešu valodas institūts, 2013.

Markus, Dace 1991, Bērzgales izloksnes zilbes tonēmas, *Latvijas zinātņu akadēmijas vēstis* 523 [1991/2], 57–62.

Markus, Dace 1993, Krītoša zilbes intonācija dažādās latviešu valodas prosodiskajās sistēmās, *Baltistica* 28(1), 39–44.

Melnik-Leroy, Gerda Ana, Jolita Bernatavičienė, Gražina Korvel, Gediminas Navickas, Gintautas Tamulevičius, Povilas Treigys 2022, An overview of Lithuanian intonation: a linguistic and modelling perspective, *Informatica* 33(4), 795–832.

Mikalauskaitė, Elžbieta 1975, *Lietuvių kalbos fonetikos darbai*, Vilnius: Mokslas.

Murinienė, Lina 2007, *Rytinių šiaurės žemaičių fonologija: vokalizmas ir prozodija*, Vilnius: Lietuvių kalbos institutas, Vilniaus pedagoginis universitetas.

Ortega-Llebaria Marta 2010, Acoustic correlates of stress ins Central Catalan and Castilian Spanish, *Language and Speech* 54 (1), 73–97.

Pakerys Antanas 1982, *Lietuvių bendrinės kalbos prozodija*, Vilnius: Mokslas.

Pierrehumbert, Janet 1980, *The Phonology and Phonetics of English Intonation*, PhD Thesis, Harvard University.

Pierrehumbert, Janet, Julia Hirschberg 1990, The meaning of intonational contours in the interpretation of discourse, in Philip R. Cohen, Jerry Morgan, Martha E. Pollack (eds.), *Intentions in Communication*, Cambridge MA: MIT Press, Bradford Books, 271–311.

Plag, Ingo, Gero Kunter, Mareile Schramm 2011, Acoustic correlates of primary and secondary stress in North American English*, Journal of Phonetics* 39(3), 362–374.

Plant, Randall L., Ross M. Younger 2000, The interrelationship of subglottic air pressure, fundamental frequency, and vocal intensity during speech, *Journal of Voice* 14(2), 170–177.

Rajpa, Avni, Hemant A. Patil 2016, Jerk minimization for acoustic-to-articulatory inversion, in *9th ISCA Workshop on Speech Synthesis Workshop, 13–15 September 2016, Sunnyvale, USA*, 82–87.

Roren, Alexandra, Antoine Mazarguil, Diego Vaquero-Ramos, Jean-Baptiste Deloose, Pierre-Paul Vidal, Christelle Nguyen, François Rannou, Danping Wang, Laurent Oudre, Marie-Martine Lefèvre-Colau 2022, Assesing smoothness of arm movements with jerk: a comparison of laterality, contraction mode and plane of elevation. A pilot study*, Frontiers in Bioengineering and Biotechnology* 9, article 782740.

Rudzīte, Marta 1993, *Latviešu valodas vēsturiskā fonētika*, Rīga: Zvaigzne.

Sabonytė, Regina, Yonatan Goldshtein 2021, F0 and duration changes in unstressed vs. stressed syllables connected to postlexical stress and sentence type in Standard Lithuanian, in *Perspectives on Nordic Phonology: Selected Papers from the Fifth Fonologi i Norden Meeting*, 119–128.

Sarkanis, Alberts 1993, *Latviešu valodas Augšzemes sēlisko izlokšņu prosodija un vokālisms: Eksperimentālie pētijumi*, Disertācija humanitāro zinātņu doktora grāda iegūšanai, Viļņas Universitāte.

Stevens, Keneth N. 2000, *Acoustic Phonetics*, Cambridge, London: The MIT Press.

Švageris, Evaldas 2015, *Lietuvių ir latvių tarmių monoftongų priegaidžių akustiniai požymiai: lyginamoji analizė*, Daktaro disertacija, Vilnius: Vilniaus universitetas

Švageris, Evaldas 2016, Pradedant diskusiją dėl Frydricho Kuršaičio priegaidžių rekonstrukcijos, *Lietuvių kalba* 10, 1–9.

Švageris, Evaldas 2018, Rytų aukštaičių uteniškių ilgųjų balsių priegaidės: istorinė analizė ir eksperimentinis tyrimas, *Baltistica* 53(1), 21–68.

Švageris, Evaldas 2020, Intonacijos įtaka akustinei priegaidžių raiškai: eksperimentinis latvių bendrinės kalbos tyrimas, *Baltistica* 55(1), 119–157.

Švageris, Evaldas 2021, Prozodinių baltų kalbų elementų interakcijos tyrimo tęsinys: audicinė latvių kalbos priegaidžių analizė, *Baltistica* 56(2), 271–305.

Švageris, Evaldas 2022, Akustinė lietuvių ir latvių kalbų kirčio struktūra: lyginamoji analizė, *Baltu Filoloģija* 31(2), 71–95.

Švec, Jan G, Harm K. Schutte, C. Julian Chen, Ingo R. Titze 2023, Integrative insights into the myoelastic-aerodynamic theory and acoustics of phonation. Scientific tribute to Donald G. Miller, *Journal of Voice* 37(3), 305–313.

Vaitkevičiūtė, Valerija 1995, *Lietuvių bendrinės kalbos priegaidės*, Vilnius: Enciklopedija.

van Heuven, Vincent 2018, Acoustic correlates and perceptual cues of word and sentence stress: towards a cross-linguistic perspective, in Rob Goedemans, Jeffrey Heinz, Harry van der Hulst (eds.), *The Study of Word Stress and Accent: Theories, Methods and Data*, Cambridge University Press, 15–59.

Titze, Ingo R. 2000, *Principles of Voice Production*, Iowa City, IA: National Center for Voice and Speech.

Titze, Ingo R. 2006, *The Myoelastic Aerodynamic Theory of Phonation*, Iowa City, IA: National Center for Voice and Speech.

Zintchenko, Jevgenij 2018, *The Production of Lexical Tone in Croatian*, PhD Thesis, Johann Wolfgang Goethe-Universität.

Zhang, Zhaoyan 2016, Mechanics of human voice production and control, *Acoustic Society of America* 140(4), 2614–2635.

Evaldas ŠVAGERIS
Taikomosios kalbotyros institutas
Vilniaus universitetas
Universiteto g. 5
LT-01513 Vilnius
Lithuania
[*evaldas.svageris@flf.vu.lt*]