BALTISTICA LX(2) 2025 323–354
doi: 10.15388/Baltistica.60.2.2576

Evaldas ŠVAGERIS
Vilnius University

A MULTIFACTORIAL ANALYSIS OF PHONETIC WORD STRESS CORRELATES IN STANDARD LITHUANIAN: VERIFICATION OF THE PROSODIC ELEMENT INTERACTION MODEL*

Abstract. This article aims to verify the empirical background of the phonetic interaction of prosodic elements in Standard Lithuanian – specifically word stress and phrase intonation. A multifactorial (MANOVA) analysis of F0 dynamics was performed to achieve this, revealing that the ratio of three acoustic parameters (F0 range, F0 jerk, and duration) depends on word stress. The speaker‘s intention to distinguish a syllable from its environment prosodically can be associated with a relatively wider F0 range, its more linear modulation, and a longer duration of the sound. The overall picture of F0 dynamics correlates with the position of the stress within the word. The primary phonetic cue for achieving the effect of a stressed final syllable is the static phonation of adjacent pre-stressed syllables (F0 acceleration →0 resulting in F0 jerk →0 and reduced syllable duration →0). The need to highlight intersyllabic differences in dynamic F0 profiles diminishes when the word stress is on the first syllable, leading to the assumption that the absolute beginning of the word, under phonetic *ceteris paribus*, acquires greater relative weight from a prosodic point of view. Bearing this in mind, one can conclude that the same F0 dynamics reflect the interaction of prosodic elements. The only distinction is that the intonational factor determines the overall change of F0, while the word stress controls its (intra)syllabic dynamics (variation of F0 over time).
Keywords: prosodic element interaction; F0 dynamics; MANOVA; word stress; phrase intonation; Standard Lithuanian.

* This research was funded by a grant (No. S-MIP-24-90) from the Research Council of Lithuania (Project „Intonation of the Lithuanian language: Functional Types and Interference with Prosodic Units“).

Introduction

If we consider word stress as a correlate of speech rhythm (Dauer 1983, 51–62; 1987, 447–450; Tilsen, Arvaniti 2013, 628–629; Gussenhoven, Riad 2025), its phonetic interpretation must be based on the syntagmatic syllable contrast (for phonological arguments, see Fox 2000; Girdenis 2003). Paradoxically, this restriction complicates efforts to define the distinctive correlates of stress, as direct comparison of syllables within a word poses methodological challenges. Researchers follow the logic that the influence of a linguistic factor on the distribution of acoustic data can only be identified under *ceteris paribus* (all else being equal) conditions. Otherwise, we cannot determine whether the distribution of parameters is influenced solely by one or more potential factors (Pakerys 1982; van Heuven 2018). For this reason, phonetic word stress correlates are often compared paradigmatically. In such cases, the acoustic characteristics of the syllable in stressed and unstressed positions are taken into account.

However, even the paradigmatic analysis does not fully resolve the problem. Studies (including those of the Lithuanian prosody) suggest that the phonetic realisation of stress is complex, simultaneously involving a group of acoustic parameters (duration, intensity, F0, and spectral shape). The type of differentiation is essentially a matter of the prosodic system in the language (Lieberman 1960, 451–454; Lehiste 1970, 142–146; Pakerys 1982, 111–144; Fant, Kruckenberg 1994, 125–144; Girdenis 2003, 252–254; 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; Ericksson, Niebuhr 2023, 1–45, etc.). Such a broad and compromise-based conclusion naturally gives rise to complex algorithms and compensatory mechanisms in order to define stress in spectrograms as easily as possible. Although one cannot rule out such a possibility, it likely indicates that the physiological-physical nature of word stress is not yet fully understood. Too much attention seems to be focused on some stress correlates that only occur under specific conditions. A deeper understanding of speech sound dynamics and the relationship between aerodynamic and biomechanical factors in this process would contribute to solving this problem. Moreover, it seems more helpful to concentrate on acoustic parameter ratios (instead of their absolute values), which, like in mathematics, could be differentiated by prosodic elements, leading to specific sound effects (Švageris 2023, 195–224).

It is important to emphasise that, in this context, the issue of the phonetic interaction of prosodic elements is highly relevant. Suppose an analysis is not limited to verifying phonological opposition alone and instead aims to determine the correlates of stress that would remain invariant across all its positions. In that case, clarifying the essential principle of the interaction in question becomes necessary. We cannot ignore that, as there are many instances when the same phonetic domain, such as a long syllable, becomes a field for the synchronous realisation of all prosodic units (word stress, pitch accents, and phrase intonation), and their phonetic correlates overlap. Although the acoustic complexity of stress can be easily explained by linking different acoustic parameters to various prosodic units, empirical data do not fully support such assumptions. Additionally, the dynamic, melodic, and quantitative types of stress are sometimes considered obsolete (Girdenis 2003).

Thus, it must be acknowledged that the previously mentioned prosodic problem is two-layered. To identify the distinctive correlates of stress, we must first address the issue of the phonetic interaction mechanism among prosodic elements. Considering that acoustic stress complexity has a vague empirical basis, it is essential to establish an interaction principle relevant to the syntagmatic interpretation of word stress. In this context, a potential solution could be to derive higher-grade tone parameters over time (Švageris 2023, 195–224). Various types of dynamic F0 profiles may serve as alternative sources of prosodic information. In simpler terms, this methodological approach distinguishes between the amount of sound energy (F0 level) and its distribution over time (F0 acceleration and F0 jerk). This dichotomy is crucial for understanding the interactive mechanisms of prosodic elements. One can assume that the intonational factor incites changes in F0 itself, while word stress and pitch accents control its balance. Stress determines the syntagmatic distribution of these F0 dynamics, while pitch accents determine the paradigmatic. Therefore, when analysing prosodic data, we must rely on sequences of (non)static, (non)deformed, (non)linear, (non)controlled, and (non)stable F0. The most extreme element, reflecting an intentional change in F0 (phonation activation), should be considered a syntagmatic realisation of word stress (Švageris 2023, 195–224). It is important to note that these correlates of word stress are indifferent to the phrase's intonation and do not directly depend on the inherent properties of speech sounds.

Since the theoretical basis of the interaction model has already been established (Švageris 2023, 195–224), it is now necessary to verify it using more extensive empirical data. Therefore, the most important task here is determining whether F0 dynamics can be considered the primary means of synchronously expressing prosodic elements in Standard Lithuanian, namely, word stress and sentence intonation. The research objectives are as follows:

1. To describe the dynamics of phonation, including its aerodynamic, biomechanical, and physiological aspects, and to provide a broader understanding of the phonetic phenomena analysed by the proposed interaction model;
2. To select minimal pairs for the study that would encompass the larger number of positional stress variants (considering the length and number of syllables, the placement of stress within the word, etc.);
3. To conduct a multifactorial analysis of acoustic data (MANOVA) and assess the influence of the following factors on F0 variation associated with stress: focus, phrase type, the position of the stress, and individual pronunciation (the human factor); to illustrate the most prominent trends graphically;
4. To verify the model of prosodic element interaction based on the empirical findings and to assess its potential for identifying the distinctive correlates of word stress in Standard Lithuanian across various phrasal conditions.

Dynamics of phonation: its physical and physiological aspects

Since phonetic studies on word stress have been examined in greater detail in a previous article (Švageris 2023, 195–224), at this stage, it is beneficial to explore the dynamics of phonation, reflected in F0 variation over time. In the broadest sense, phonation is the conversion of the energy of respiratory airflow into the vibration of the vocal folds; in other words, it results from the interaction of aerodynamic and myoelastic (muscle resistance to deformation) forces (van den Berg 1958, 227–244; Titze 2000; Švec et al. 2023, 305–313; for a more general review, see Zhang 2016, 2614–2635). It is worth noting that additional forms of energy ultimately transform into acoustic energy during phonation (Titze et al. 2016, 398–403). Metabolic energy, for instance, is used to induce muscle contraction;

the respiratory system produces aerodynamic energy to sustain airflow in the vocal tract; elastic energy pertains to the elastic properties of tissues; and kinetic (motion) energy is conveyed by the airflow and the vocal folds (during their vibration). One can generalise that the phonation of a speech sound is a measure of the efficiency of all these types of energy (ibid.). From a physiological perspective, the primary types of phonation are typically described with reference to the following factors: the degree of opening of the vocal folds (abduction and adduction); their length, effective mass, and tension (adjustments in the length, stiffness, and thickness of the vocal folds); the configuration of the vocal tract in the oral and pharyngeal cavities (constriction of supraglottal structures); and the elevation and lowering of the larynx (Hirose 1996, 127). The most commonly distinguished phonation types include null phonation and modal phonation, breathy and creaky voices, whisper phonation (Gordon, Ladefoged 2001, 383–406; for the context of the Lithuanian language, see Kazlauskienė et al. 2023, 87–89). The noise energy level in the acoustic signal fundamentally determines non-modal phonations. While modal phonation sounds also contain noise, they are typically overlapped by strong harmonics (quasi-periodic sound frequencies). The changing ratio of harmonic and noise energies should be regarded as an informative indicator of phonation dynamics. The typical reduction of unstressed final syllables often accompanies a weakening of phonation (a slower or faster transition of sound into noise, followed by highly perturbed F0). Clinical studies assess the extent to which sustained phonation is maintained to determine various voice pathologies, vocal fold damage, dysfunction (Brockmann et al. 2009, 44–53; 2018, 162–168), hormonal system changes, or simply fatigue (Laukkanen et al. 2008, 283–289; Titze et al. 2016, 398–403). In addition to the aforementioned HNR parameter, Jitter and Shimmer, as well as Cepstral Peak Prominence (see Teixera et al. 2013, 1112–1122; 2014, 1228–1237; Murton et al. 2020, 1596–1607), are measured in such cases.

However, the F0 changes not only in the sense of weaker phonation. Since sound frequency is a function of time (specifically, the speed of vocal fold vibration), the laws of classical mechanics enable us to interpret tone modulations in terms of changes and control of acoustic forces. Based on the first (F0 slope) and second (F0 jerk; for a more detailed explanation of the physical meaning of this parameter, see Eager 2016, 1–11) time derivatives, we can assess the degree of tone change and, more importantly, the level of control over it. Relatively, the closer the F0 jerk values to zero, the more linear the tone changes, indicating a greater balance of acoustic forces acting in phonation. It is important to emphasise that these parameters capture the differences, for instance, between a gradual lowering or raising of tone and its uncontrolled fall. Although the main trajectory of F0 in such cases coincides, its phonational nature differs fundamentally. The wider the range (higher F0 acceleration/force) and the smoother the tone changes (lower F0 jerk/greater control), the more pronounced the speaker's intent to emphasise and distinguish the spoken sound, as well as to maintain control over it (and, of course, vice versa). It is worth recalling that, in general, the amount of energy and its distribution over time do not correlate. In elementary terms, we can spend the same amount of time doing nothing or working intensively. For this reason, the value of the time parameter increases only when the F0 dynamics of the syllables coincide. When the dynamic F0 profiles differ, it is better to interpret the duration factor from a relative point of view. Additionally, within the framework of the proposed concept, the auditory impression of emphasis is obtained when tone changes more linearly, over a longer time, and in a broader range. Moreover, the logical approach to evaluating prosodic elements as mathematical functions also has a perceptual basis, as it has been observed that insufficient speech sound quantity (less than 90 ms sound duration) becomes an obstacle to identifying the dynamics of F0 (see Greenberg, Zee 1979, 150–164; Zhang 2002).

Physiological arguments naturally complement this physical perspective. Maintaining a body position requires a constant balance of forces between antagonistic muscles. Suppose an external force (such as a strong gust of wind) suddenly and unexpectedly affects a body. In that case, the muscles cannot contract or relax quickly enough, causing the body to lose its balance temporarily. The reaction time in such cases depends on physiological factors (muscle structure) and brain attentiveness. The predicted effect is balanced more quickly than the sudden and unexpected one. This type of F0 control (differential control) is also conceptually discussed in classical works on human voice physiology (Titze 2000, 211–238).

Furthermore, one can strengthen this argument by the fact that the muscles exerting the most influence on F0 regulation—the cricothyroid muscle and the thyroarytenoid muscle—are innervated by two separate branches of the vagus nerve: the superior and recurrent laryngeal nerves. The ability to activate the interacting antagonistic muscles of the vocal tract at different levels is a crucial condition for balancing their strength (phonation control). For instance, if one muscle is gradually deactivated while another is simultaneously activated, the effective mass of the vocal folds, and thus their vibration frequency, can change over a wide range and at varying rates (ibid.). Although the anisotropy of the vocal folds prevents an oversimplified understanding of the phonation process, it must be kept in mind that changes in phonation can be determined from the F0 trajectories too, allowing us to infer different speaker intentions.

Research material, graphical and statistical data analysis

The material for this study consists of 12 minimal pairs. The paradigms of the verbs *mìnti* [^{1}ˈ**mʲɪn**ʲtʲɪ][1] and *miñti*[2] [^{2}ˈ**mʲɪnˑ**ʲtʲɪ] were selected due to their large number of oppositions regarding phonological syllable quantity, pitch accent, syllable number in the word, and stress position. This choice has the methodological advantage of avoding excessive variation in the phonemic structure of the words. Additionally, it allows us to examine how the interaction mechanism operates when, for example, the same stressed syllable [ˈ**mʲɪ**] appears in different prosodic environments.

The following pairs are included in the list:

mìni [ˈ**mʲɪ**nʲɪ] – *minì* [mʲɪˈ**nʲɪ**][3];
mìna [ˈ**mʲɪ**nɐ] – *minà* [mʲɪˈ**nɐ**];
mìnima [ˈ**mʲɪ**nʲɪmɐ] – *minimà* [ˈmʲɪnʲɪˈ**mɐ**];
mìnimos [ˈ**mʲɪ**nʲɪ^{2}ˈmoːs] – *minimõs* [mʲɪnʲɪˈ^{2}ˈ**moːs**];
mìnti [^{1}ˈ**mʲɪn**ʲtʲɪ] – *mintì* [^{1}mʲɪnʲˈ**tʲɪ**];
miñti [^{2}ˈ**mʲɪnˑ**ʲtʲɪ] – *mintì* [^{2}mʲɪnˑʲˈ**tʲɪ**];
mìntų [^{1}ˈ**mʲɪn**ʲ^{2}ˈtuː] – *mintų̃* [^{1}mʲɪnʲˈ^{2}ˈ**tuː**];
miñtų [^{2}ˈ**mʲɪnˑ**ʲ^{2}ˈtuː] – *mintų̃* [^{2}mʲɪnˑʲˈ^{2}ˈ**tuː**];
užmìnti [ʊʒʲ^{1}ˈ**mʲɪn**ʲtʲɪ] – *užmintì* [ʊʒʲ^{1}mʲɪnʲˈ**tʲɪ**];
užmiñti [ʊʒʲ^{2}ˈ**mʲɪnˑ**ʲtʲɪ] – *užmintì* [ʊʒʲ^{2}mʲɪnˑʲˈ**tʲɪ**];
užmìntų [ʊʒ^{1}ˈ**mʲɪn**ʲ^{2}ˈtuː] – *užmintų̃* [ʊʒ^{1}mʲɪnʲˈ^{2}ˈ**tuː**];
užmiñtų [ʊʒ^{2}ˈ**mʲɪnˑ**ʲ^{2}ˈtuː] – *užmintų̃* [ʊʒ^{2}mʲɪnˑʲˈ^{2}ˈ**tuː**].

---

[1] „To step“
[2] „To ask“
[3] Stressed syllables are highlighted.

As can be seen, the opposition to stressed short vowels, circymflexed long vowels, and both circumflex and acute diphthongs consists of syllables of varying phonological lengths. Two- and three-syllable words were examined, where the word stress occupied all possible positions. Intonationally, three types of phrases (affirmative, interrogative, and imperative) and phrasal focus positions (focused/unfocused) were chosen. The test words were presented both in isolation and within phrases (always in the middle position). When creating sentences, uniform stress placement was maintained in peripheral words. The scheme [(x) –́ ⌣ / test word / –́ ⌣ (x)] was followed to ensure that the rhythm of the phrase would equally affect both members of the minimal pair. One hundred eighty-eight slides were prepared (see Fig. 1 for an example). To eliminate doubt about the placement of word stress during the recordings, all stressed syllables were highlighted in red on the slides, and the word to be stressed within the phrase was underlined. In fact, it was challenging for some pairs to create logical and natural-sounding imperative phrases (in such cases, those phrases were excluded from the material). Of course, this intonation type did not occur when the test words had to be pronounced in isolation.

Figure 1. **Example of a slide with a research phrase**[4]

Two professional male actors, aged 33 (D1) and 32 (D2), were asked to read the material required for the study to convey the intonational aspects of the phrases more clearly and expressively. The first actor is from Alytus (Southern Aukštaitian dialect), while the second is from Zarasai (Eastern Aukštaitian

[4] Sentence in slide: „Is it your brother‘s birthday today?“

subdialect of Utena). Both actors have been living and working in Vilnius for over ten years. With their extensive experience in stage speaking and theatre, they can be regarded as good representatives of Standard Lithuanian. Furthermore, no signs of dialectal speech, potential influence from other languages, or unusual articulation were observed during the recordings. The speakers were instructed to read each slide ten times at a consistent pace, resulting in 3,760 recorded utterances (188 slides * 2 speakers * 10 times). The recordings were conducted with professional equipment at the Vilnius University Phonetics Laboratory, following all methodological requirements for this type of research.

All audio recordings were initially divided into separate files and then annotated using Praat (Boersma, Wennink 2018). Acoustic parameters were automatically calculated from the annotation files with a script (subprogram) developed by Gintautas Tamulevičius, Associate Professor of the Faculty of Mathematics and Informatics at Vilnius University. The obtained data were transferred to standard Excel files and subsequently to the statistical data analysis software JMP (a free annual academic license was utilised). This software serves as a convenient tool not only for illustrating the relationships between the distribution of acoustic data and prosodic elements (Švageris 2023, 195–224) in three-dimensional graphs but also for conducting a range of statistical calculations. After combining and grouping the data into a single working file, a multivariate analysis (MANOVA) was performed. As previously mentioned, from an interactional perspective, prosodic elements can be considered mathematical functions regulating the relationship between acoustic parameters. Since this analysis of acoustic stress correlates aims to verify the model of phonetic interaction, it consistently adheres to previously established assumptions, specifically the methodological provision that the dynamic F0 profiles can be assessed from the ratio of F0 range, F0 jerk, and duration (for details, see Švageris 2023, 195–224). Higher values of F0 range, and especially F0 jerk for a speech sound of shorter duration, should indicate decreasing phonatory efforts (prosodical weakening of the syllable). At the same time, the corresponding opposing distribution of parameters should suggest the speaker's intention to distinguish the syllable from its environment. Schematically, this dichotomy could be represented as follows:

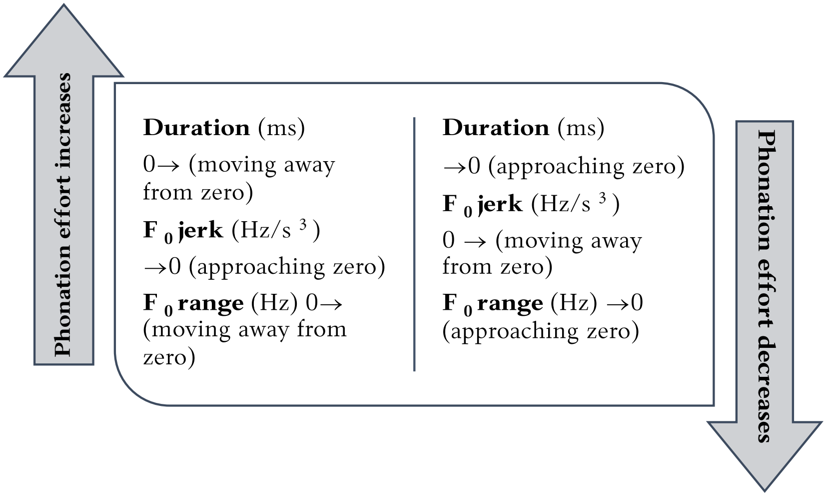


Figure 2. **Relationship between dynamic F0 profiles and acoustic parameter distribution**

The logic of statistical calculations compels us to regard the factor uniting these three parameters as an independent variable. For simplicity, as previously noted, this study refers to it as the dynamic F0 profile. What matters is whether this factor has an interactive relationship with word stress (i.e., whether stress selectively regulates the ratio of these acoustic parameters within a word), and, more importantly, whether this interaction remains resistant to the influence of other independent variables (intonation types, speakers, phrasal focus, etc.). If it does remain resistant, a solid empirical and statistical foundation would emerge to validate the proposed concept of the phonetic interaction of prosodic elements. In this case, two dynamic aspects of phonation would be regarded: the change in F0 over time and the absolute level of F0. The former should be associated with the phonetic expression of word stress (according to the syntagmatic approach) and pitch accent (according to the paradigmatic one). At the same time, the latter relates to the intonational type of the phrase (according to both vectors).

Multivariate analysis of variance (MANOVA) was chosen for this study because of the many factors that may influence the phonetic realisation of word stress. In simple terms, it extends two-way analysis (TWO-WAY ANOVA) to assess relationships among more than two independent variables.

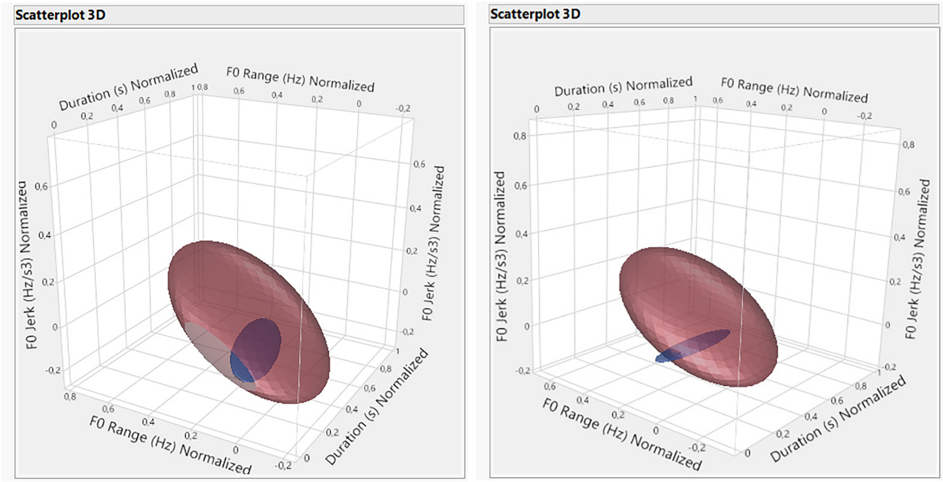
While many factors can often result in ambiguous interpretations of the overall results, it remains necessary for a deeper understanding of the nature and possible limits of phonetic variation in prosodic elements. Furthermore, as noted, it serves as the primary criterion for confirming or denying the working hypothesis—the interaction principle.

Multifactorial analysis of phonetic accent features of Standard Lithuanian

On the one hand, the discussion of statistical analysis results could begin—and seemingly end—with the finding that stressed syllables are clearly distinguishable from unstressed ones from the ratio of three acoustic parameters: duration, F0 range, and F0 jerk (with $p<.0001$ across all minimal pairs). However, this result alone is insufficient to establish an interaction model, as these calculations merely reveal the difference but not its specifics. Understanding phonation changes within a word and how the stress factor influences them requires a much more thorough analysis. It is crucial to clarify how the same groups of acoustic parameters are affected by additional factors: phrasal focus, the position of the stressed syllable within the word, the speaker, and the intonational type of the phrase. Identifying the interaction relations that link all these factors with word stress is important. Once it becomes clear that these relationships have a statistical basis, it is essential to specify how the correlates of stress are modified under varying conditions and, most importantly, which ones remain invariant. Only with such a methodological approach can a more comprehensive picture of the dynamics of F0 and its dependence on stress be developed. The statistical weight and phonetic details regarding the influence of all the aforementioned factors on stress are discussed separately below for consistency and clarity.

Phrasal focus. If we associate the phonetic realisation of stress with a more actively changing and controlled F0 (see Fig. 2), particular attention must be paid to those changes in tone that have an explicit intentional nature. In principle, from this perspective, all phonation dynamics can be divided into three F0 profiles: phonation deactivation (sudden relaxation of the vocal folds, transformation of sound into noise, F0 perturbation), constant or static phonation (where the F0 jerk and F0 range are equal to zero, and the duration approaches zero) and active phonation (where the F0 jerk approaches zero, and the F0 range and duration move away from zero). It is helpful to keep these dynamic differences in mind when analysing the effect of phrasal focus on word stress. Some researchers indicate that determining the word stress in the weak position of a phrase is quite problematic (Girdenis 2003, 253–254; van Heuven 2018, 47). It is interesting that under these intonational conditions, final word stress is particularly weakened. The perception of stress in such words frequently shifts to the first syllable (Pakerys 1982, 108–111; Huss 1987, 86–105). Based on the material of this study, we can try to examine the phonetic aspects of this phenomenon in more detail.

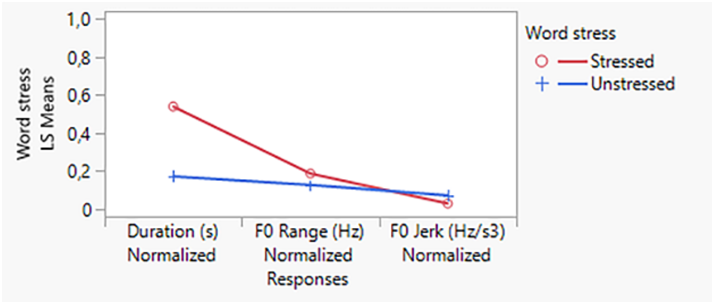
Statistical results indicate that there is not a single case (across all intonational types and speakers) when there is no strong interaction between phrasal focus and word stress (on average, p<.0001). This interaction can be described in more detail in the illustration below (see Fig. 3). It reflects that the distribution of acoustic parameters of the syllables in [^{1}ˈ**mʲɪn**ʲ^{2}ˈtuː] – [^{2}ˈ**mʲɪn·**ʲ^{2}ˈtuː] and in [^{1}ˈmʲɪnʲ^{2}**ˈtuː**] – [^{2}ˈmʲɪn·ʲ^{2}**ˈtuː**] depends on the phrasal focus. In both cases, the probability of this relationship is far from the chance limit (p<.0001 and p = 0.009, respectively). Indeed, the F0 range, F0 jerk, and duration values are normalised here (converted to a scale from 0 to 1) to ensure that the parametric differences do not overshadow each other due to the different types of scales. The red ellipses represent the realisations of the words in the strong position of the phrase, while the blue ones represent those in the weak position. It is clear that the F0 profiles of the words pronounced without phrasal focus, regardless of the location of word stress, acquire a static character (static phonation). It indicates that the change in tone of both syllables under such conditions is minimal, and the values of the F0 range and F0 jerk are significantly close to zero. It explains why the blue ellipses in the graphic space occupy an incomparably smaller area than the red ones, and the statistical probability of interaction remains unquestionable. Since the analysed words always occupied the central position, the end-of-phrase effect was avoided, which could have led to more frequent instances of F0 perturbation (increased F0 jerk values). Therefore, it is pretty apparent that words in weak central positions of phrases are pronounced with inertia. Their tone, changing with zero acceleration and jerk, is one of the strongest physical arguments for this assumption. The probability of completely inarticulate pronunciation is lower, likely due to the need to maintain a certain level of acoustic energy until the stressed word at the end of the phrase.



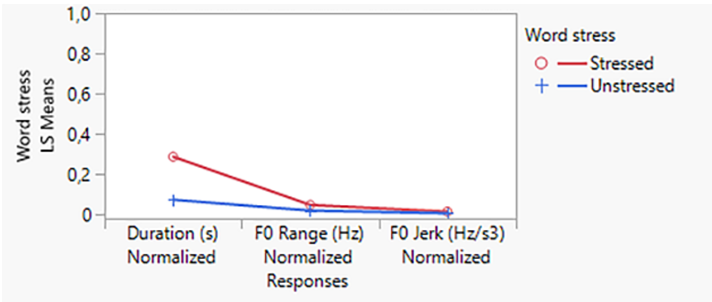
F i g u r e 3. **Dependence of the distribution of acoustic parameters for [^{1}ˈmʲɪn^{j2}ˈtuː] – [^{2}ˈmʲɪn·^{j2}ˈtuː] (left graph; p<.0001) and [^{1}ˈmʲɪn^{j2}ˈtuː] – [^{2}ˈmʲɪn·^{j2}ˈtuː] (right graph; p=0.009) on phrasal focus**

All the ideas discussed above are confirmed even more clearly by the average values of the parameters (see Figs. 4–7). Attention should be paid to how the phonation of stressed syllables is activated in the strong position of the phrase compared to the weak position. A general increase in values along the scale is evident. Relatively, their increased duration is accompanied by an expanded F0 range and relatively stable F0 jerk values. This pattern indicates a more intense tone change in these cases. It should be noted that the F0 dynamics of pre-stressed (longer!) syllables are inertial. The transition from a static tone phase to an active one likely creates the impression of stress in the final syllables. When the exact words are pronounced without phrasal focus, both syllables essentially do not differ in F0 parameters (the tone of both syllables is static). The limitation of emphasis in the phrase to the activation of the stressed syllable F0 is also demonstrated by the transition of unstressed final syllables to static phonation.

Considering the previous instances of problematic word stress identification (P a k e r y s 1982, 108–111), one can assume that stress should be associated with the first syllable in the absence of more significant F0 changes within a word. We suggest that under phonetic *ceteris paribus* conditions, the prosodic weight of the initial syllable is more substantial. For a syllable to

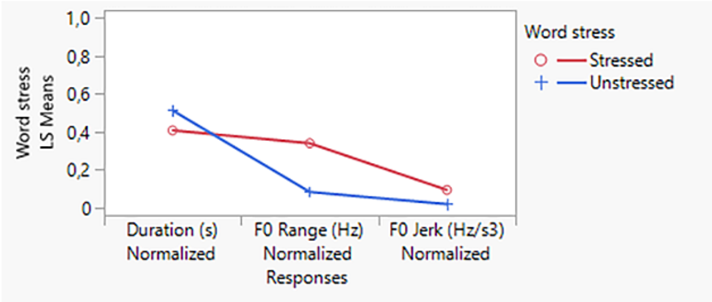


F i g u r e 4. **Distribution of average acoustic parameters of [^{1}ˈmʲɪn^{j2}ˈtuː] – [^{2}ˈmʲɪnˑ^{j2}ˈtuː] (with phrasal focus)**

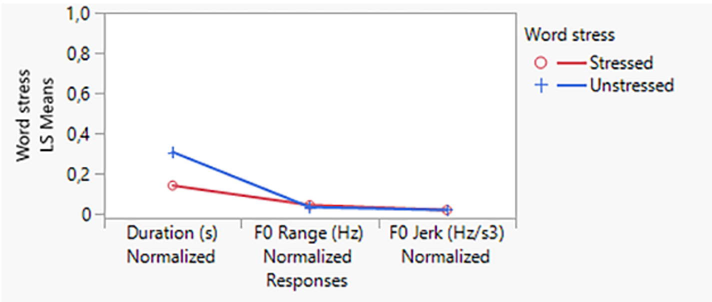


F i g u r e 5. **Distribution of average acoustic parameters of [^{1}ˈmʲɪn^{j2}ˈtuː] – [^{2}ˈmʲɪnˑ^{j2}ˈtuː] (without phrasal focus)**

be acoustically emphasised, the tone change must be intensified, moving away from its static phase. If such a change is lacking, the entire syllable chain is interpreted phonetically as uniform, leading to no intention to distinguish other syllables within the word. In this context, F0 control (the distribution of F0 jerk values) is highly significant, as the tone trajectory can shift even due to simple sound reduction. Basic logic suggests that phonation deactivation (relaxation of the folds) cannot serve as a physiological means for the speaker's intention to emphasise a sound, even though a visual change in tone might be observed. Finally, it should be noted that an almost identical distribution of acoustic parameters has also been observed, making the assumptions applied in this study relevant to all minimal pairs included in the analysis.



F i g u r e 6. **Distribution of average acoustic parameters of [**^{1}ˈmʲɪnʲ^{2}ˈ**tuː] – [**^{2}ˈmʲɪn·ʲ^{2}ˈ**tuː] (with phrasal focus)**



F i g u r e 7. **Distribution of average acoustic parameters of [**^{1}ˈmʲɪnʲ^{2}ˈ**tuː] – [**^{2}ˈmʲɪn·ʲ^{2}ˈ**tuː] (without phrasal focus)**

Intonational type of phrase. It is customary to associate the phonetic correlates of phrase intonation with modulations of F0 levels, namely, high and low tone sequences. Currently, the so-called ToBi transcription is most often used to formalise them (for its theoretical background and application to the Lithuanian language, see P i e r r e h u m b e r t 1980; K a z l a u s k i e n ė, D e r e š k e v i č i ū t ė, S a b o n y t ė 2013). However, in the case of this study, a more relevant question is whether the intonational factor can distort those dynamic features of phonation associated with the phonetic expression of stress in this study. Let us recall that three types of sentences were included in the analysis: affirmative, imperative, and interrogative. Although statistical data show that the intonation type significantly affects the F0 distribution implied by word stress (on average, $p<.0001$), it is notable that this effect is greater when the final syllable of the word is stressed. The dynamics of F0 in stressed final syllables may be somewhat more flexible than in the initial syllables. This is suggested by the [**ˈmʲɪ**nɐ] – [mʲɪ**ˈnɐ**] pair chosen for illustrative purposes. Attention should be drawn to the smaller graphical area occupied by ellipses in the left graph [**ˈmʲɪ**nɐ]. Based on the arguments presented in the previous section, we could assume that the graphical distribution of the data reflects the F0 dynamics of stressed syllables more accurately (especially in the cases of [mʲɪ**ˈnɐ**]). It can be observed that, for example, the tendency to emphasise the stressed final syllable (to expand the F0 range) is more characteristic of words pronounced with affirmative intonation (red ellipse, right graph). According to the ratio of both F0 parameters, the imperative intonation type (green ellipse) is somewhat distinct. The slightly higher F0 jerk value for a similar F0 range indicates that the tone of the final syllables of words stressed with such intonation changes more abruptly and with a lower degree of control. In contrast, when the same word is pronounced with interrogative intonation (blue ellipse), the same F0 values have a longer duration on the stressed vowel. This implies that, at least in relative terms, their level of tone control is higher (the tone change is more linear).

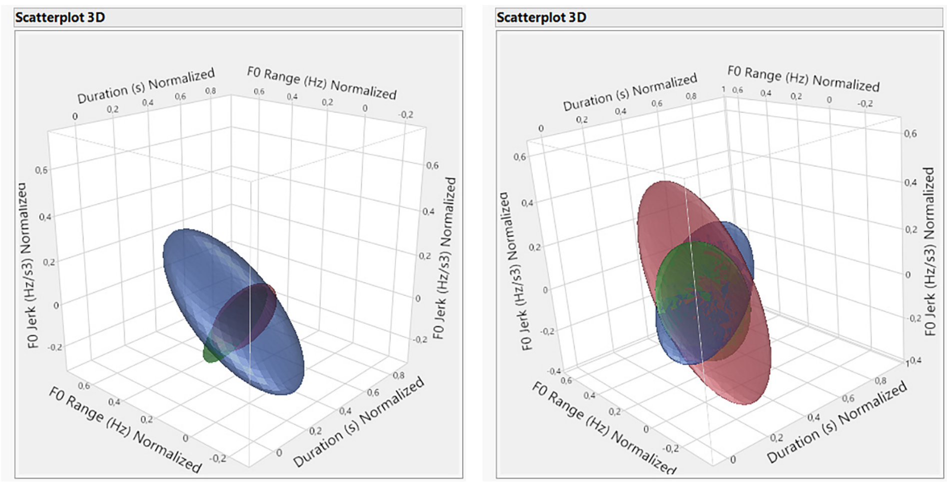


Figure 8. **Dependence of the distribution of acoustic parameters of [ˈmʲɪnɐ] (left graph; p<.0001) and [mʲɪˈnɐ] (right graph; p<.0001) on the intonational phrase type**

These relationships regroup when the stress shifts to the word's first syllable ([ˈ**mʲɪ**nɐ] cases, left graph). The data scattered in the left three-dimensional graph shows that imperative (green ellipse) and affirmative (red) intonations cause the narrowing of the F0 dynamics, while interrogative (blue) intonations tend to expand them. Although the graphical differences are apparent, this inconsistency in the relationship between prosodic salience and intonation type, compared to the [mʲɪˈ**nɐ**] cases, suggests that acoustic stress on a syllable under different intonational conditions depends more on other factors (for example, on the speakers).

When this type of acoustic data distribution is viewed from the perspective of phonetic stress realisation (which is of primary concern in this study), we return to previously stated assumptions. General [mʲɪˈ**nɐ**] averages (see Figs. 9–11) indicate that the statistical interaction between word stress and intonation type of the phrase is, in principle, determined by the dynamics of the phonation of stressed final syllables: a more extended and linear F0 (interrogative intonation), a somewhat sharper and less controlled F0 (affirmative), and the sharpest and least variable F0 (imperative). However, from the standpoint of phonetic interaction between prosodic elements, the most important point is that, despite these minor differences, the nature of the phonetic differentiation of word stress remains consistent. It is clear that the intonation type factor does not distort the core trend in any way: the static F0 pattern of pre-stressed syllables (see the values of the F0 range and the average of the F0 jerk marked by the blue lines in Figs. 9–11). There is growing confidence that this is the primary rhythmic-acoustic condition for producing the impression of a stressed final syllable. The intensity with which the F0 of the stressed syllable changes is less important (as it may vary depending on intonation type); what matters more is that this marks a departure from the zero-acceleration F0 pattern typical of the pre-stressed syllables.

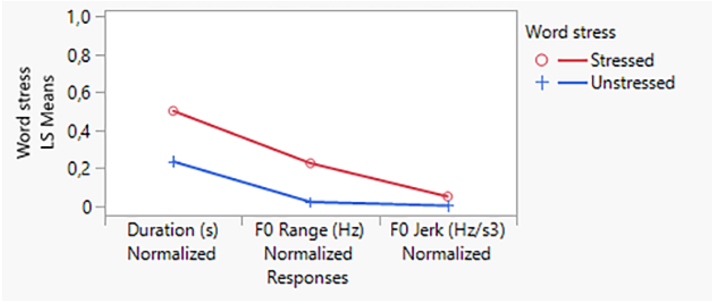


Figure 9. **Distribution of average acoustic parameters of [mʲɪˈnɐ] (interrogative intonation)**

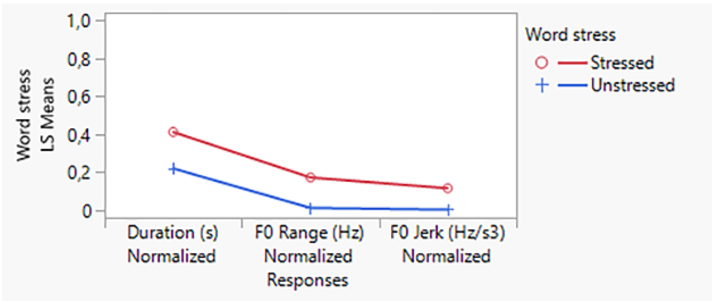


Figure 10. **Distribution of average acoustic parameters of [mʲɪˈnɐ] (affirmative intonation)**

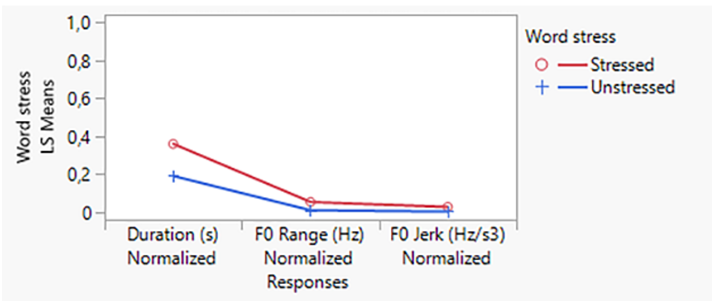


Figure 11. **Distribution of average acoustic parameters of [mʲɪˈnɐ] (imperative intonation)**

When stress shifts to the beginning of a word ([ˈ**mʲ**ɪnɐ] cases), the phonation dynamics of both syllables become similar, and the intonational factor appears to start influencing the acoustic form of the entire word. There seems to be no tangible basis left to connect the phonetic realisation of stress with the chosen group of three acoustic parameters. However, this dilemma is easily resolved if we revert to interpreting stress as a factor that activates phonation. It has already been noted that the similarity of the F0 profiles of syllables is a favourable condition for forming the acoustic stress effect in the first syllable of a word. It is important to emphasise that the F0 of stressed initial syllables in these cases is not static or uncontrolled (perturbed). Instead, it varies somewhat, creating more challenging conditions for other syllables to counterbalance its prosodic weight. If a small tone activation following the static phonation of pre-stressed syllables was sufficient to establish the final word stress impression, then in this case, there is a new point of reference: a noticeably intense change in the F0 of the initial syllable. To surpass this, a greater range and higher tone control are needed (in simple terms, "pushing the acceleration pedal"). If a similar tone trajectory extends across the entire word, then there is no longer an apparent change in phonation (F0 acceleration and increase in control), which is the essential condition for producing the stress effect in the non-initial syllable of the word. It should be noted that word-initial stress becomes even more pronounced if the F0 of subsequent syllables is merely perturbed (due to reduction) or transitions into a static phase (due to possible accentual adhesion). Additionally, a logical conclusion follows: the more the phonation of post-stressed syllables is deactivated (the vowel is reduced, F0 is perturbed), the relatively lower the level of activation of the F0 of the stressed first syllable; the only crucial point is that it is not phonetically weaker. From all this, the principle of the phonetic interaction of prosodic elements becomes clear. The intonational factor apparently determines the overall tone direction, while stress at the syllabic level balances its dynamics.

The position of the stressed syllable in the word. These assumptions can be substantiated even more firmly when the position of stress within the word appears as an independent factor in statistical calculations. It should be admitted that the specificity of the MANOVA model dictates that the relationship between phonetic stress correlates and the stress position is decided according to the paradigmatic relationship of syllables. It is easy to predict that the statistical interaction weight will be most influenced by the

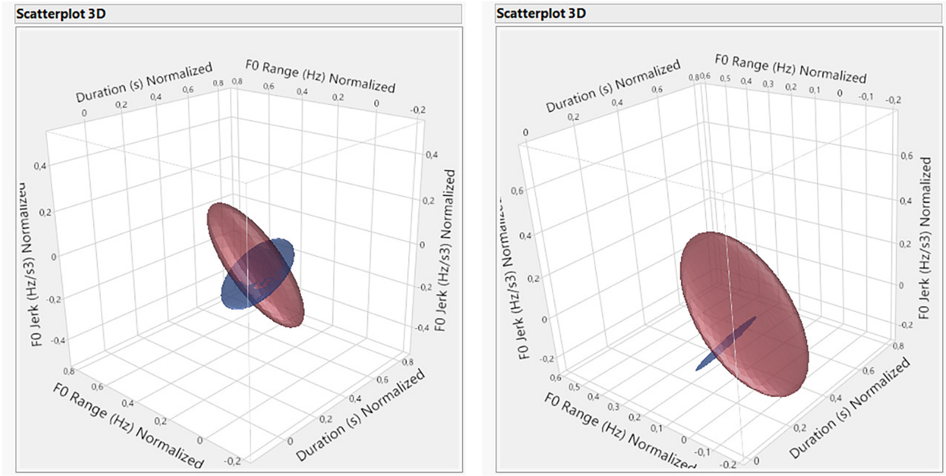


Figure 12. **Dependence of the distribution of acoustic parameters of ['mʲɪnɐ] (left graph; p<.0001) and [mʲɪ'nɐ] (right graph; p<.0001) on the stress position in the word**

often-mentioned F0 statics of pre-stressed syllables, differing from the much freer F0 dynamics when the same syllable is stressed. This assumption is supported by a significant statistical probability (p<.0001). There are no cases where the phonetic realisation of stress does not correlate with its position in the word, and the tone of pre-stressed syllables at the beginning of the word is inertial (F0 range, F0 jerk, and duration →0).

All this can be easily seen from the overall distribution of data in the three-dimensional graphs (see Fig. 12). In these graphs, the blue ellipses represent the distribution of the acoustic parameters of the first syllables of the word. In contrast, the red ones correspond to the final ones. In the right graph ([mʲɪ'**nɐ**] cases), it can be seen that the F0 dynamics of pre-stressed syllables (blue ellipse) are extremely "compressed." It is worth emphasising that this pattern is not negatively affected by either the speaker or the intonation type of the phrase (since these factors do not group the data), further reinforcing the notion that these phonetic parameters are differentiated by word stress. The shift of stress to the word's first syllable (['**mʲɪ**nɐ] graph) paves the way for freer F0 dynamics of post-stressed syllables. However, the essential condition remains: from a phonological point of view, the post-stressed syllables must coincide with the stressed one or be weaker so as to not cause the effect of a pronounced and controlled phonation change (stress!). All these statements are confirmed by the general average values of the parameters presented below.

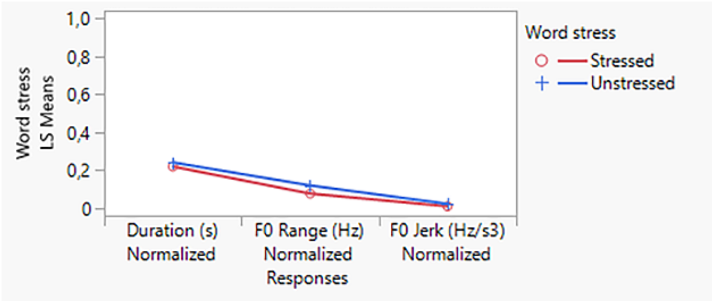


Figure 13. **Distribution of average acoustic parameters of [ˈmʲ**ɪnɐ**]**

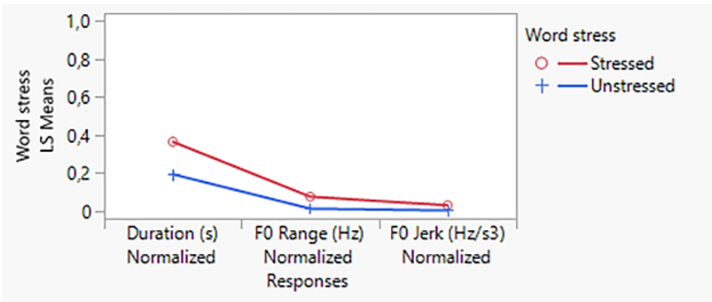


Figure 14. **Distribution of average acoustic parameters of [**mʲɪˈ**nɐ]**

To provide a visual representation of how F0 trajectories in a word reflect the differences in phonation implied by stress, four Praat spectrograms of minimal pairs [ˈ**mʲɪ**nʲɪ] – [mʲɪˈ**nʲɪ**][5] and [ʊʒʲ¹ˈ**mʲɪn**ʲtʲɪ] – [ʊʒʲ¹mʲɪnʲˈ**tʲɪ**] are presented below (see Figs. 15–18). The selected examples differ in their constituent syllables' phonemic structure, number, and phonological length. In addition, all the presented words are pronounced with interrogative intonation to dissociate them from possible stress connections with the absolute F0 level. The trajectories of the blue lines illustrate the nature of the tone change in the spectrograms. As previously noted, the prosodic

[5] Stressed syllables are highlighted.

stress of a syllable is associated with the activation of phonation ("pushing the acceleration pedal"), the dynamic shift of the tone in the direction of **F0 range/duration 0→** (moving away from zero), **F0 jerk →0** (approaching zero). In each illustrated case, the stress is the moment that corresponds to the onset of this tonal change. In the case of the form [**ˈmʲɪ**nʲɪ], the tone rises at the beginning of the word and continues rising until the end of the word, while in [mʲɪˈ**nʲɪ**] the tone rises only from the final stressed syllable. As repeatedly observed, the F0 of pre-stressed syllables is static and inertial (see the vertical arrows in the pre-stressed syllables of each example).

In contrast, the F0 of post-stressed syllables repeats the trajectory of stressed F0 (the example of [ˈ**mʲɪ**nʲɪ]) or transitions to a static phonation (the case of [ʊʒʲ[1]ˈ**mʲɪn**ʲtʲɪ]). This differentiation of F0 dynamics allows us to look at the prosodic structure of words from a more general perspective and to question the human ability to differentiate linguistic units at the acoustic micro-level (differences measured in tens of milliseconds or semitones), especially under the conditions of intense speech tempo. The prosodic weight of a syllable is a matter of human phonational effort. However, it should not be interpreted in terms of F0 levels alone. Instead, it should be viewed through the lens of F0 dynamics (F0 profiles), based on the elementary laws of classical mechanics.

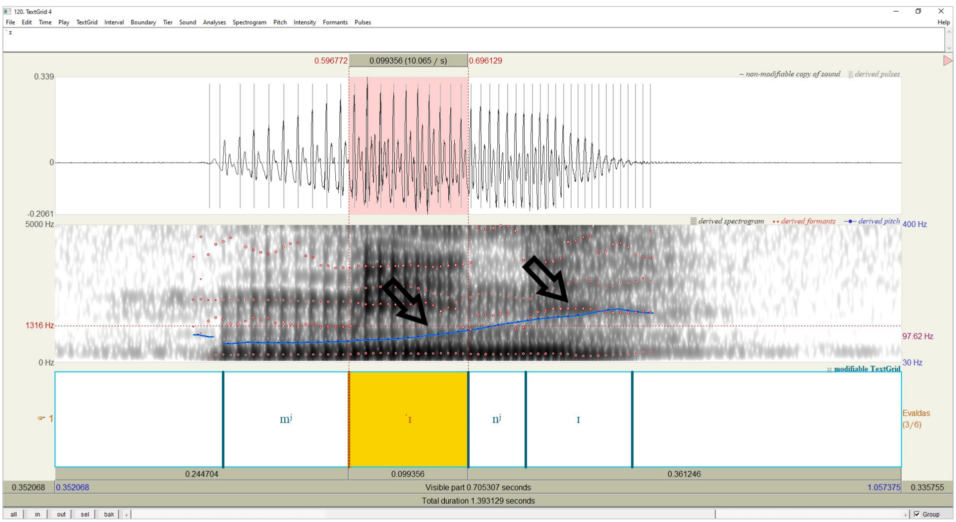


Figure 15. **[ˈmʲɪnʲɪ] F_0 modulation (interrogative intonation)**

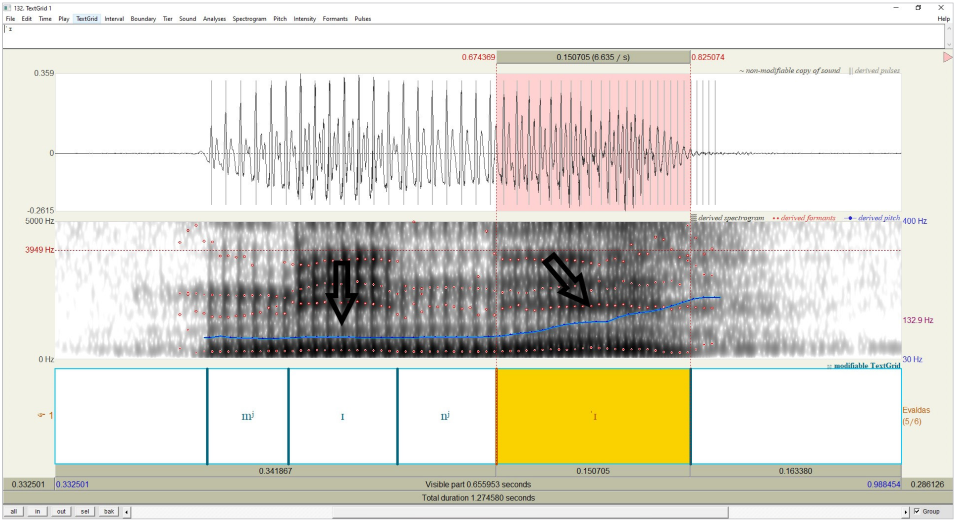


Figure 16. [mʲɪˈ**nʲɪ**] F_0 **modulation (interrogative intonation)**

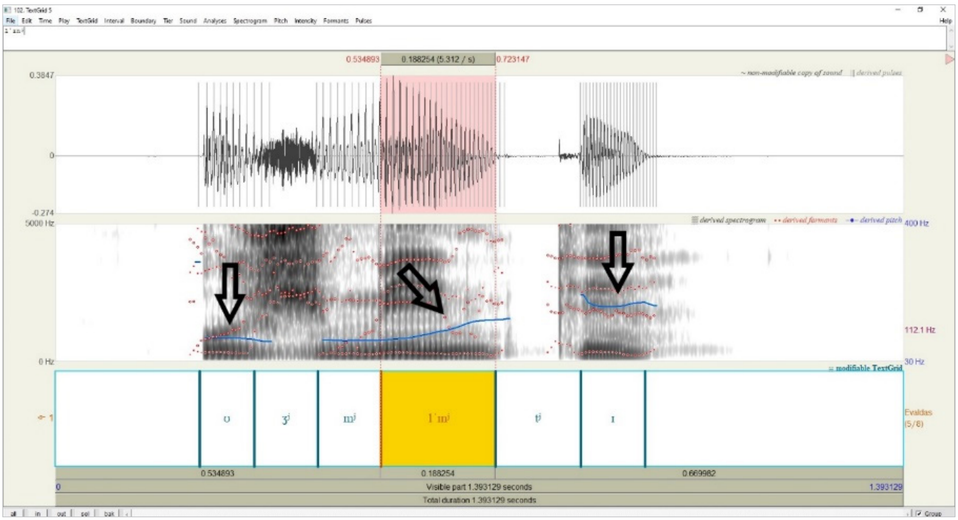


Figure 17. [ʊʒʲ¹ˈ**mʲɪn**ʲtʲɪ] F_0 **modulation (interrogative intonation)**

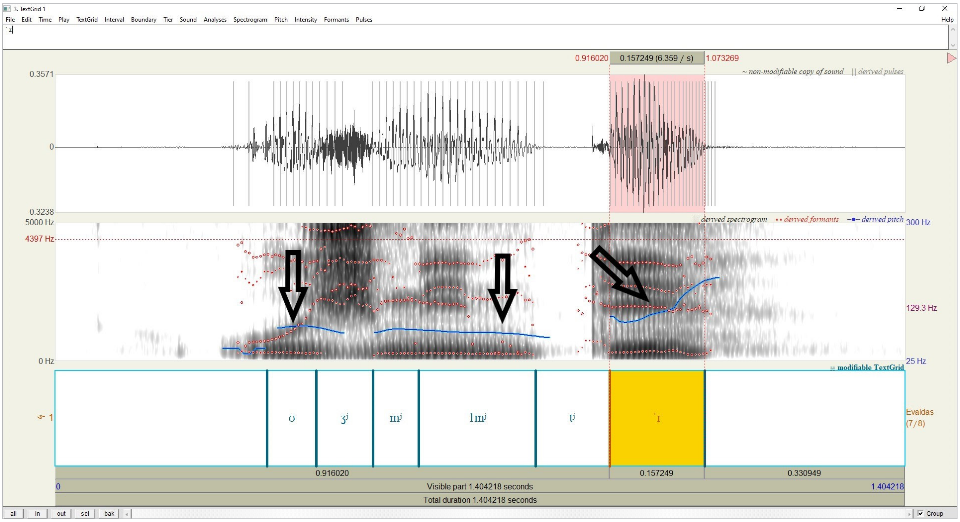


Figure 18. [ʊʒ^{j1}m^{j}ɪn^{j}ˈ**t^{j}ɪ**] **F$_0$ modulation (interrogative intonation)**

The speaker factor. Finally, the effect of speakers on the examined word stress correlates should also be assessed. Although there is a relatively small probability that the previously established trends would be significantly affected by this factor, it is nevertheless important to substantiate this claim with objective arguments for methodological consistency. MANOVA calculations suggest that, on the whole, the connection between phonetic stress realisation and speakers is the weakest. Although in most cases favourable probability values are still recorded (p<0.05), there are also negative values (***mì****nima* Wilks's lambda p=0.726) or parameters approaching the critical limit (***mì****na* p=0.02, minimum p=0.03, etc.). As far as can be judged from the illustrative material (see Fig. 19), the vowel and diphthong syllable nuclei of D1 (red ellipses) that are pronounced relatively longer have the most significant influence on this type of interaction. However, this quantitative advantage is accompanied by a greater degree of its variation (the variance of the blue D2 ellipses concerning the duration scale is smaller). The graph of the minimal pair *mìn****tų̃****–miñ****tų̃*** (right) shows a somewhat steeper angle of inclination of the red ellipse relative to the F0 range and F0 jerk scales (i.e., lower values of the ratio of these parameters), which implies a rather logical assumption that a higher sound quantity is more favourable for achieving a higher degree of phonation control. However, such a statement cannot be absolute because, as

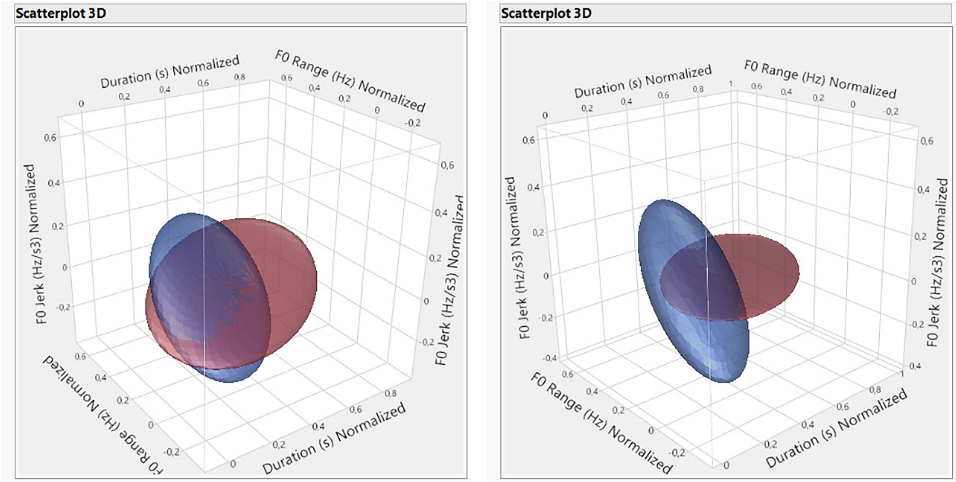


Figure 19. **Dependence of the distribution of acoustic parameters of [**^{1}ˈ**mʲɪn**ʲ^{2}ˈtuː**] – [**^{2}ˈ**mʲɪn·**ʲ^{2}ˈtuː**] (left graph; p<.0001) and [**^{1}ˈmʲɪnʲ**^{2}ˈtuː] – [**^{2}ˈmʲɪn·ʲ**^{2}ˈtuː] (right graph; p<.0001) on the speakers**

the arrangement of the relevant data in the left ***mìntų̃–miñtų̃*** graph shows, a longer total duration of sounds can also be accompanied by quite similar F0 dynamics. The variation in the speech tempo of the speakers (after all, each of the speakers read over 1,000 phrases), possible differences in the level of emphasis in the strong phrase position, etc., could have contributed to such a distribution of parameters. It is probably impossible to completely control these factors. On the other hand, this is another way to verify the relationship between the proposed principle of interaction of prosodic elements and individual pronunciation patterns.

The average values of [^{1}ˈ**mʲɪn**ʲ^{2}ˈtuː] – [^{2}ˈ**mʲɪn·**ʲ^{2}ˈtuː] and [^{1}ˈmʲɪnʲ^{2}ˈ**tuː**] – [^{2}ˈmʲɪn·ʲ^{2}ˈ**tuː**] (see Figs. 20–23) confirm the assumptions made at the beginning of the subsection: the speaker factor does not change or distort the dynamic F0 differences, which in this study are associated with distinctive features of stress. The tendency toward phonation statics in pre-stressed syllables remains unchanged (there is a significant shift in the F0 range and F0 jerk values toward zero; pay attention to the blue lines at these parameters in Figs. 22 and 23). Although the diphthongal structure of the initial syllables of this minimal pair naturally gives them a quantitative advantage over the final syllables (which, regardless of stress, are shorter in all cases), from a differential perspective, it is more important to assess the extent to which this difference affects the phonation types of syllables and their F0 dynamics. Remember that the value of all three acoustic parameters is relative; their ratio is important, not the absolute values. Therefore, we can estimate the prosodic weight of a syllable much more accurately by noting that the static nature of the F0 in pre-stressed syllables (phonation type) becomes a prosodic factor, reducing quantitative differences between syllables in a word. Naturally, the diminished pre-stressed F0 dynamics (zero F0 acceleration → F0 jerk) are also accompanied by a relatively shorter duration of the syllable nucleus.

In contrast, the tone activation of the final stressed syllables that replaces it (F0 acceleration 0→) requires a larger quantitative base. In other words, all these quantitative relations between syllables result from (are a side effect of) the differentiation of their dynamic F0 profiles, not vice versa; therefore, a separate comparison of duration parameters is insignificant from a prosodic

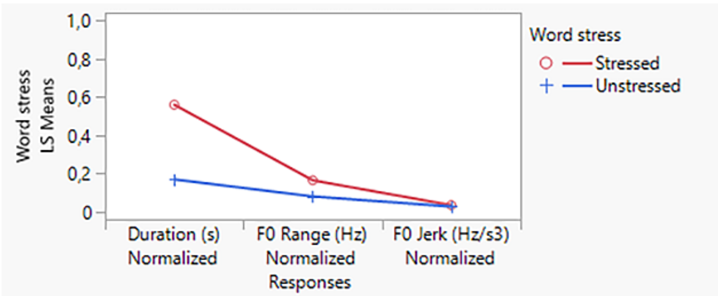


Figure 20. **Distribution of average acoustic parameters (D1) of [^{1}ˈmʲɪn^{j2}ˈtu:] – [^{2}ˈmʲɪn·^{j2}ˈtu:]**

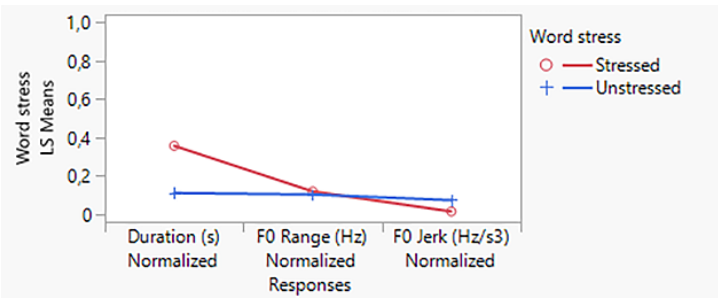
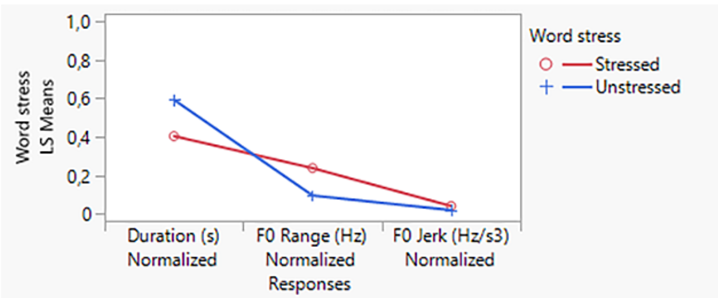
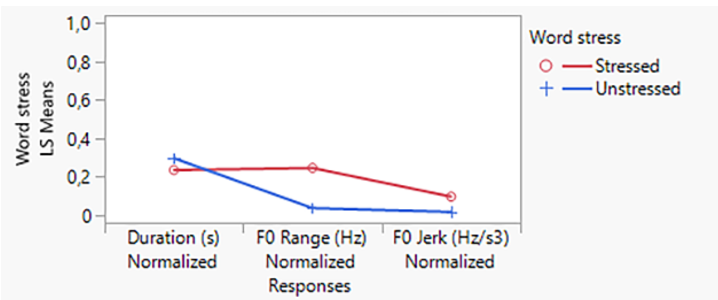


Figure 21. **Distribution of average acoustic parameters (D2) of [^{1}ˈmʲɪn^{j2}ˈtu:] – [^{2}ˈmʲɪn·^{j2}ˈtu:]**

point of view. After all, knowing the differences in the physical duration of the sounds being compared, we cannot say what acoustic effect each of them creates. Simply put, the duration of a phenomenon does not imply its nature (as previously mentioned, during the same amount of time we can be both lazy and hardworking); it can be a side effect only. The situation is different when the prosodic profile of a syllable is viewed from a phonational perspective. If the prosodic weight of a syllable correlates with sound activation and control (quantified by the time derivatives of F0), then differences in the vocal fold tension and their differentiated control (T i t z e 2000) become the elementary physiological cause of the established differences in phonation. The formation of a more complex F0 profile of a syllable (without deactivating phonation or transitioning to its static phase) apparently requires additional physiological efforts, more active control of the muscular fold layer, and a longer time frame (after all, it takes longer to do more thorough and complex work).



F i g u r e 22. **Distribution of average acoustic parameters (D1) of [**$^{1\text{ˈ}}$mʲɪnʲ$^{\mathbf{2ˈ}}$**tuː] – [**$^{2\text{ˈ}}$mʲɪn·ʲ$^{\mathbf{2ˈ}}$**tuː]**



F i g u r e 23. **Distribution of average acoustic parameters (D2) of [**$^{1\text{ˈ}}$mʲɪnʲ$^{\mathbf{2ˈ}}$**tuː] – [**$^{2\text{ˈ}}$mʲɪn·ʲ$^{\mathbf{2ˈ}}$**tuː]**

The need to highlight differences between phonational syllable types decreases when the word stress is on the first syllable of a word. This trend is confirmed by comparing the data grouped by the speaker factor (see Figs. 20–21). The first syllable gains greater prosodic weight, although the F0 profiles of both syllables become somewhat similar. Naturally, the absolute beginning of the word enhances the prosodic power of syllables in this position, since, from a syntagmatic perspective, non-initial syllables are accentually coordinated with the word's first syllable, rather than vice versa. This also means that the F0 dynamics of monosyllabic words should, in principle, be free. It appears that the phonetic expression of stress occurring at the absolute beginning of the word is also freer for the same reasons. The only condition for post-stressed syllables is that they do not exceed the level of activation and control of the F0 of stressed initial syllables. All this suggests that, from a phonational standpoint, the difference between pre-stressed and stressed final syllables is generally greater than that between stressed initial and post-stressed syllables. Of course, this statement cannot be made absolute either, as the prosodic salience (phonation activation and control) of any stressed syllable, as we have seen, depends both on the phrasal focus and the degree to which the speaker emphasises the word itself (compare the F0 ranges of stressed and unstressed [1ˈ**mʲɪn**ʲ2ˈtuː] – [2ˈ**mʲɪn·**ʲ2ˈtuː] syllables in Figs. 20–21). The key point is that, despite all these factors, the phonetic nature of stress differentiation remains unchanged.

Conclusions

Through a multifactorial analysis of F0 parameters, this study provides a solid empirical background for the proposed principle governing the interaction of prosodic elements in Standard Lithuanian. This principle centers on the methodological distinction and interaction between two dynamic F0 fields: the absolute F0 level and (intra)syllabic F0 variation over time.

Prosodic weight correlation with phonation type. Distinctive features of word stress are traceable through the time derivatives of F0, which reflect phonational efforts by the speaker. The prosodic weight of a syllable is correlated with the activation and control of phonation within the vowel or diphthong (i.e., sonorous nuclei of the syllable). MANOVA results demonstrated a clear dependence of word stress on the ratio among three acoustic parameters: F0 range, F0 jerk, and duration. A speaker's intention to emphasise a syllable is fulfilled through a tone that exhibits a broader range, a more linear manner, and a longer duration (lower F0 jerk). This suggests that obtaining a more complex F0 profile necessitates more muscular activation and control of the vocal folds.

Phonetic stress correlates depend on stress position. The overall F0 dynamics in a word significantly depend on the position of the stress. To create the effect of a stressed word-final syllable, the primary prosodic mechanism is the static phonation of the preceding, pre-stressed syllables (characterised by near-zero F0 acceleration and F0 jerk alongside reduced vowel duration). For word-initial stress, the need to highlight intersyllabic differences decreases. The prosodic weight of the initial syllable may be inherently greater because the entire syllable sequence can be interpreted without a significant, pronounced change in F0 activation/control for the stress effect to occur in a non-initial unstressed syllable.

General statement. The interaction of prosodic elements in Lithuanian is reflected in the F0 dynamics. Sentence intonation determines the overall, macro-level direction of F0 variation, while word stress governs the micro-level, syllabic dynamics—the specific variation of F0 over time.

DAUGIAFAKTORINĖ FONETINIŲ BENDRINĖS LIETUVIŲ KALBOS KIRČIO POŽYMIŲ ANALIZĖ: PROZODINIŲ ELEMENTŲ INTERAKCIJOS MODELIO VERIFIKAVIMAS

Santrauka

Šiame straipsnyje siekta verifikuoti empirinį prozodinių lietuvių kalbos elementų – žodžio kirčio ir frazės intonacijos – fonetinės interakcijos pagrindą. Tuo tikslu atlikta multifaktorinė (MANOVA) dinaminių F0 požymių analizė, kurios rezultatai atskleidė, kad trijų akustinių parametrų (F0 diapazono, F0 džerko ir trukmės) santykis priklauso nuo žodžio kirčio, o kalbančiojo intenciją prozodiškai išskirti skiemenį iš jo aplinkos galima sieti su santykiniu požiūriu platesniu diapazonu, tiesiškiau ir ilgesnį laiko momentą kintančiu tonu. Matyti, kad bendrasis F0 dinamikos žodyje paveikslas koreliuoja su kirčio pozicija žodyje. Pagrindinė fonetinė sąlyga kirčiuoto galinio skiemens efektui išgauti yra statinė prieškirtinių skiemenų fonacija (nulinis F0 pagreitis → nulinis F0 džerkas; mažėjanti skiemens trukmė). Poreikis išryškinti tarpskiemeninius dinaminių F0 profilių skirtumus sumažėja, kai žodžio kirtis yra pirmajame skiemenyje, todėl laikytina, kad absoliuti žodžio pradžia fonetinėmis *ceteris paribus* sąlygomis prozodiniu požiūriu turi didesnį lyginamąjį svorį. Turint visa tai omenyje, darytina išvada, kad prozodinių elementų sąveiką atliepia ta

pati F0 dinamika. Skirtumas tik tas, kad frazės intonacijos veiksnys lemia pačią F0 kitimo kryptį, o žodžio kirtis – skiemeninę jo dinamiką (F0 varijavimą laiko atžvilgiu).

REFERENCES

Boersma, Paul, David Weenink 2018, Praat: doing phonetics by computer, https://www.fon.hum.uva.nl/praat/.

Brockamnn, Meike, Michael J. Drinnan, Claudio Storck, Paul N. Carding 2009, Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task, *Journal of Voice* 25, 44–53.

Brockmann-Bauser, Meike, Jörg Edgar Bohlender, Daryush Mehta 2018, Acoustic perturbation measures improve with increasing vocal intensity in individuals with and without voice disorders, *Journal of Voice* 32(2), 162–168.

Dauer, Rebecca M. 1983, Stress-timing and syllable-timing reanalysed, *Journal of Phonetics* 11(1), 51–62.

Dauer, Rebecca M. 1987, Phonetic and phonological components of language rhythm, in Tamaz Gamkrelidze (ed.), *Proceedings of the 11th International Congress of Phonetic Sciences*, Tallinn: Academy of Sciences of the Estonian S.S.R., 447–449.

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

Erickson, Donna, Oliver Niebuhr 2013, Articulation of prosody and rhythm: Some possible applications to language teaching, in Oliver Niebuhr & Malin Svensson Lundmark (eds.), *Proceedings of the 13th Nordic Prosody Conference: Applied and Multimodal Prosody Research,* Sonderborg: Sciendo, 1–45, DOI: 10.2478/9788366675728-001.

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

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

Gick, Bryan, Ian Wilson, Donald Derrick 2013, *Articulatory Phonetics*, Malden, MA & Oxford: Wiley-Blackwell.

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

Gordon, Matthew, Peter Ladefoged 2001, Phonation types: a cross-linguistic overview, *Journal of Phonetics* 29(4), 383–406.

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

Greenberg, Steven, Eric Zee 1979, On the perception of contour tones, *UCLA Working Papers in Phonetics* 45, 150–164.

Gussenhoven, Carlos, Thomas Riad 2025 (expected), *Suprasegmental phenomena in Germanic: Structural and functional interactions with word stress*, in Mark Aronoff (ed.), The Oxford Research Encyclopedia of Linguistics, Oxford University Press.

Hirose, Hajime 1997, Investigating the physiology of laryngeal structures, in William J. Hardcastle, John Laver (eds.), *The Handbook of Phonetic Sciences*, Oxford: Blackwell Publishers Ltd., 116–136.

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

Huss, Volker 1987, English word stress in the post-nuclear position, *Phonetica* 35(2), 86–105.

Kazlauskienė, Asta, Sigita Dereškevičiūtė, Regina Sabonytė 2023, *Bendrinės lietuvių kalbos intonacija: frazės centras, ribos ir žymėjimas*, Kaunas: VMU.

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

Laukkanen, Anne-Marie, Irma Ilomäki, Kirsti Leppänen, Erkki Vilkman 2008, Acoustic measures and self-reports of vocal fatigue by female teachers, *Journal of Voice* 22(3), 283–289.

Lehiste, Ilse 1970, *Suprasegmentals*, Cambridge.

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

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

Murton, Olivia, Robert Hillman, Daryush Mehta 2020, Cepstral peak prominence values for clinical voice evaluation, *American Journal of Speech-Language Pathology* 29, 1596–1607.

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

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

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

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.

Sluijter, Agatha Martha Cornelia 1995, *Phonetic Correlates of Stress and Accent*, Leiden: ICG Printing.

Švageris, Evaldas 2023, The phonetic interaction of prosodic elements in the Baltic languages: a tentative theoretical model, *Baltistica* 58(2), 195–224.

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.

Teixera, João Paulo, Paula Odete Fernandes 2014, Jitter, Shimmer and HNR classification within gender, tones and vowels in healthy voices, *Procedia Technology* 16, 1228–1237.

Teixera, João Paulo, Carla Oliveira, Carla Lopes 2013, Vocal acoustic analysis – jitter, shimmer and HNR parameters, *Procedia Technology* 9, 1112–1122.

Tilsen, Sam, Amalia Arvaniti 2013, Speech rhythm analysis with decomposition of the amplitude envelope: Characterising rhythmic patterns within and across languages, *The Journal of the Acoustical Society of America* 134(1), 628–639.

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

Tizte, Ingo R., Lynn Maxfield, Anil Palaparthi 2016, An oral pressure conversion ratio as a predictor of vocal efficiency, *Journal of Voice* 30(4), 398-406.

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

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

Zarka, Dina El, Barbara Schuppler, Carina Lozo, Wolfgang Eibler, Patrick Wurzwallner 2017, Acoustic correlates of stress and accent in Standard Austrian German, in *Phonetik in und über Österreich: Veröpflichungen zur Linguistk und Kommunikationsforschung* 31, 15–44.

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

Zhang, Jie 2002, *The Effects of Duration and Sonority on Contour Tone Distribution: A Typological Survey and Formal Analysis*, New York: Routledge.

Evaldas ŠVAGERIS
Baltistikos katedra
Vilniaus universitetas
Universiteto g. 5
LT-01513 Vilnius, Lithuania
[*evaldas.svageris@flf.vu.lt*]