Baltic J. Modern Computing, Vol. 13 (2025), No. 2, 407-418
https://doi.org/10.22364/bjmc.2025.13.2.06

Acoustic Cues of Prosody Phrasing in Lithuanian at Different Speech Rates

Asta KAZLAUSKIENĖ, Gailius RAŠKINIS,
Eidmantė KALAŠINSKAITĖ-ZAVIŠIENĖ

Vytautas Magnus University, V. Putvinskio str. 23-204, LT-44243 Kaunas, Lithuania

asta.kazlauskiene@vdu.lt, gailius.raskinis@vdu.lt,
eidmante.kalasinskaite-zavisiene@vdu.lt

ORCID 0000-0002-6135-1635, 0000-0001-5793-868X, 0009-0003-6007-6621

Abstract. The aim of this research is to determine whether pauses, decreases in intensity, pitch lowering, lengthening of the final sound, and certain voice quality features are equally significant in marking phrase boundaries at different reading rates. The research material consists of 9 recordings of Aesop's fable "The North Wind and the Sun". This text was read by three male actors 3 times: at a natural reading rate, at a fast reading rate, and at a slow reading rate. The presence and duration of pauses, the intensity and F0 of the entire phrase and its final syllable, as well as the duration, intensity, F0, jitter, shimmer, and HNR of the final vowel at the end and in phrase-internal position, were analysed.

The results indicate that all parameters can serve as indicators of phrase boundaries, particularly for intonational phrases. Intonational phrases can be differentiated from intermediate phrases based on the level of intensity decrease, jitter, and to some extent, pauses and F0 lowering. However, among these parameters, only duration-related features, such as pauses and the lengthening of the final sound, show a more prominent sensitivity to changes in reading rate.

Keywords: Lithuanian, prosodic phrasing, phrase, pause, intensity, fundamental frequency, duration, voice quality, speech rate

1. Introduction

Prosodic phrasing involves the grouping of words into larger units, known as intermediate phrases (ip) and intonation phrases (IP). This grouping is influenced by a syntactic and prosodic structure of a language, as well as by semantic and pragmatic factors, and individual speaker characteristics (e.g. speech rate, emotions, and experience in speaking aloud). While phonetic features of phrase boundaries are likely universal, their significance may vary across languages. In many studies, the most important indicators of phrase boundaries include pitch and intensity changes, pausing, segmental lengthening (Peters, 2003; Michelas and D'Imperio, 2010; Petrone et al., 2017; Brugos et al., 2018; Żygis et al., 2019; Harrington Stack and Watson, 2023; Liu et al. 2023; de Souza, 2023; Steffman et al. 2024; Volín et al., 2024), and non-modal phonation

(Kohler, 2000; Crowhurst, 2018). However, as syntactic structure varies between languages, phrasing patterns in one language cannot simply be transferred to another.

Prosodic phrasing affects the perception and production of language phenomena. For appropriate clause and sentence processing, prosodic and syntactic boundaries must align (for research on this relationship, see Watson, Gibson, 2004; Himmelmann, 2022).

Research on prosodic phrasing based on acoustic cues is important for improving the performance of human language technologies such as automatic speech recognition (ASR), speech synthesis, and natural language understanding (NLU) systems. Acoustic cues, including pitch variation, duration, and pauses, are fundamental in signalling phrase boundaries in spoken language. These cues help disambiguate sentence structure, clarify meaning, and improve the flow of communication. By developing algorithms that can accurately detect phrase boundaries using these acoustic signals, ASR systems can produce more accurate transcriptions, particularly in real-time applications where syntactic structure impacts comprehension.

In speech synthesis, accurately predicting pause durations based on speech tempo can lead to more natural and expressive prosody, significantly improving the overall quality of generated speech. For NLU systems, precise identification of phrase boundaries allows for better parsing and understanding of input, leading to improvements in tasks like machine translation and dialog systems. Ultimately, this research is crucial for refining the ability of language technologies to handle the complexities of spoken language, making them more effective in diverse applications such as virtual assistants, automated transcription services, and language learning tools.

Given this context, it is crucial to identify the basic features that signal intonation phrase and intermediate phrase boundaries and to determine which of these features are consistent. Research on the Lithuanian language in this area is still in its early stages. Dereškevičiūtė and Kazlauskienė (2022), Kazlauskienė, Dereškevičiūtė, and Sabonytė (2023) have shown that phrase endings are signalled by a decrease in intensity, F0 lowering, lengthening of the final sound, and an increase in amplitude perturbations. While pauses are a possible indicator of phrase boundaries, they are not mandatory.

2. Aim, Material and Methods

The research presented in this paper represents one stage of a broader study. The aim of this stage is to determine whether pauses, decreases in intensity, pitch lowering, lengthening of the final sound, and certain voice quality features are equally significant in marking phrase boundaries at different reading rates. For this purpose, a controlled-read speech analysis was conducted.

To carry out the experiment, three native male speakers (professional actors) read Aesop's fable *The North Wind and the Sun*. A translation of this fable is usually chosen to illustrate the International Phonetic Alphabet, as these illustrations provide accounts of the sound structure of different languages. The text comprises seven sentences, 68 lexemes, and 405 sounds, and it was read three times: at a natural rate, a fast rate, and a slow rate. Breaks were provided between the recordings. The speakers were instructed to read the text clearly, paying attention to the semantic and grammatical connections between words. They were allowed to choose the phrasing of the sentences freely, and the phrasing did not necessarily have to be identical across all recordings. Recordings were made in a professional sound recording studio.

Until further research is conducted, intonation phrases in read Lithuanian are identified with sentences. In this study, an intonation phrase is defined as a phrase consisting of one intermediate phrase or the final intermediate phrase within a sentence. The number of intonation phrases in each recording was 7.

Intermediate phrases are identified based on acoustic, syntactic, and semantic criteria. The number of intermediate phrases per recording varied across speakers and reading rates. Including the last intermediate phrase of an intonation phrase, as well as intonation phrases consisting of only one intermediate phrase, speaker S1 produced 24 intermediate phrases during fast reading, 22 during natural reading, and 28 during slow reading. Speaker S2 produced 23 intermediate phrases in the fast reading, 27 in the natural reading, and 29 in the slow reading. Speaker S3 produced 23 intermediate phrases in the fast reading, 24 in the natural reading, and 25 in the slow reading.

To investigate whether **intensity** and **fundamental frequency** (F0) signal phrase boundaries, the intensity and F0 of entire intonation and intermediate phrases, as well as the nucleus of the final syllable in each phrase, were compared. F0 values were extracted from audio recordings using an autocorrelation technique with sinc interpolation to enhance the precision of lag and peak height estimates (Boersma, 1993).

To assess whether the final sound in a phrase is lengthened, the **durations** of the vowels [eː] and [ɑː] were compared. These vowels occur both at the end of a word within a phrase and at the end of intonation or intermediate phrases, making them suitable for such comparisons.

Additionally, to explore whether aperiodicity, stability, and noisiness in the voice signal serve as indicators of phrase boundaries, several voice quality features of phrase-final vowels were compared to those in phrase-internal positions. The analysed features included:

- **Jitter:** The period-to-period variability in pitch, estimated using waveform-matching, where the duration of each period is determined via a cross-correlation technique (Boersma, 2009).
- **Shimmer:** The average absolute difference between the amplitudes of consecutive periods, normalized by the average amplitude.
- **Harmonics-to-Noise Ratio (HNR):** The ratio of the relative power of the periodic (harmonic) component to the noise component. HNR values were calculated using the technique described by Boersma (1993)

All the above-mentioned acoustic features were extracted manually using the Praat software tool (Boersma and Weenink, 2018). As this study focuses on acoustic parameter ratios for individual speakers, data normalization was deemed unnecessary.

The ratios in this study are calculated either by dividing one analysed quantity by another or by converting the difference in decibels to a ratio. For instance, if the average fundamental frequency (F0) of the nucleus of the final syllable in a phrase is 83 Hz, and the average F0 of the voiced units in the entire intermediate phrase is 124 Hz, the resulting ratio is 0.67 (83/124 = 0.67). Similarly, if the intensity at the end of the phrase is 55.7 dB and the average intensity of the entire intonation phrase is 61.1 dB, the resulting ratio is calculated as $10^{(55.7-61.1)/10} = 0.288$

3. Results

3.1 Duration of Pauses and Pausing

Pauses can occur for a variety of reasons, including psychological, physiological, rhythmic, and other factors. One of their primary linguistics functions is to divide speech into distinct intonation units. Consequently, pauses (or prosodic breaks) are likely indicators of phrase boundaries.

Table 1 presents the pause duration ratios at different reading rates. The natural reading rate was used as a baseline for calculating ratios, with the pause duration ratio determined by dividing the mean pause duration during slow or fast reading by the mean duration for natural reading. All speakers tended to shorten pauses between intonation phrases during fast reading and lengthen those between intermediate phrases during slow reading.

Table 1. Duration ratios of pauses. S1, S2, and S3 denote the speakers; IP refers to an intonation phrase or the last intermediate phrase of the intonation phrase; ip represents an intermediate phrase; N, F, and S correspond to natural, fast, and slow reading rates, respectively.

Researched unit	**S1**	**S2**	**S3**
ip (F:N:S)	0.7:1:1.2	0.3:1:1.5	0.7:1:2
IP (F:N:S)	0.4:1:1.1	0.2:1:1.3	0.6:1:1.4

When comparing pauses between intermediate phrases in fast and natural reading, it was observed that two speakers shortened these pauses by one-third during fast reading, while one speaker shortened them by nearly two-thirds. These shortenings were statistically significant. In this study, statistical significance was assessed using the *t-test*, with a significance level of 0.05.

Conversely, during slow reading, the speakers extended the pauses from 1.2 to 2 times, with these differences also being statistically significant.

Pauses between intonation phrases were shortened by two to four-fifths during fast reading. During slow reading, pauses after intonation phrases were extended slightly less than those between intermediate phrases, ranging from 1.1 to 1.4 times. Regardless of whether the reading rate increased or decreased, the differences in pause durations remained statistically significant.

Pauses can not only be shortened but may also be entirely omitted when the reading rate is increased. Conversely, additional pauses may be introduced when the rate is slowed. An analysis of the number of pauses revealed that during fast reading, the number of pauses decreased by half to one-tenth compared to natural reading. The difference between natural and slow reading was considerably smaller: on average, slow reading introduced only about 1.5 times more pauses than natural reading. Importantly, only pauses between intermediate phrases were omitted or added. Pauses between intonation phrases were consistently present after each sentence, with one exception in the fast-rate recordings where such a pause was omitted.

It is essential to assess the proportion of total recording time occupied by pauses at different reading rates. For fast reading, pauses accounted for only 4% to 14% of the total duration (S1—13%, S2—4%, S3—14%). In contrast, at a natural reading rate, pauses constituted over 20% of the total duration (S1—22%, S2—28%, S3—26%). During slow reading, pauses occupied even more time, ranging from 27% to 38% (S1—27%, S2—35%, S3—38%).

An analysis of the duration and number of pauses, as well as the total time occupied by pauses in the recordings, indicates that pauses consistently mark the boundaries of intonation phrases, irrespective of the reading rate. However, pauses between intermediate phrases are treated differently. During fast reading, these pauses may be completely omitted, rendering them unreliable markers for identifying intermediate phrases. Conversely, in slow reading, additional pauses may occasionally be inserted within a phrase, creating intermediate phrases not based on syntax or semantics, and thus necessitating caution when identifying phrase boundaries based solely on pauses. Nonetheless, a pause reliably separates the two phrases.

3.2 Intensity

The results of the research reveal certain tendencies. As predicted, intensity decreases at the end of any phrase (see Table 2). However, the range of the decrease varies significantly: for intermediate phrases, it ranges from 0.16 to 0.60 times, and for intonation phrases, it ranges from 0.01 to 0.13 times, with the last syllable exhibiting lower intensity than the entire phrase.

The data in this study do not indicate a strong correlation between intensity decreases and reading rate for intermediate phrases. During natural reading, the intensity decreases quite consistently for all speakers, averaging a decrease of 0.27 times. In fast reading, speaker S1 shows a significantly smaller intensity decrease at the end of the phrase compared to natural reading. Conversely, speakers S2 and S3 demonstrate a slightly greater intensity decrease during fast reading than during natural reading, though the differences are minimal. When reading slowly, speakers S2 and S3 exhibit even greater intensity decreases, whereas speaker S1 shows the opposite trend: although the decrease in intensity is more prominent than during fast reading, it remains less prominent than during natural reading.

Table 2. Intensity ratios of phrases and their final syllables. σ marks the final syllable of a corresponding phrase

Researched unit	**S1**	**S2**	**S3**
ip_N:σ(ip_N)	1:0.29	1:0.27	1:0.24
ip_F:σ(ip_F)	1:0.60	1:0.21	1:0.22
ip_S:σ(ip_S)	1:0.34	1:0.16	1:0.19
IP_N:σ(IP_N)	1:0.03	1:0.13	1:0.03
IP_F:σ(IP_F)	1:0.03	1:0.12	1:0.01
IP_S:σ(IP_S)	1:0.03	1:0.05	1:0.03

At the end of an intonation phrase, the intensity decreases more than at the end of an intermediate phrase. In 5 out of 9 cases, the intensity of the last syllable is 0.03 times lower than that of the entire intonation phrase. It can be argued that the decrease in intensity at the end of an intonation phrase is even less influenced by reading rate than the decrease in intensity at the end of an intermediate phrase. This conclusion is supported by the data from speakers S1 and S3, where the decrease in intensity is equal. Speaker S2 maintains a similar intensity difference during natural and fast reading, while slow reading results in a greater intensity decrease.

To claim that a decrease in intensity at the end of a phrase can reliably serve as an indicator of a phrase boundary, it is necessary to determine whether this decrease is not merely characteristic of word-final positions in general. Only the vowels [ɑ:] and [e:] were used at the end of both phrase types and at the end of a word in a phrase-internal position. Therefore, a comparison was conducted using the data of these vowels.

The results confirm the conclusion drawn from comparing the results for the entire phrase and its last syllable. In all cases, the intensity at the end of a phrase is lower than that of the last syllable of a word not at the end of a phrase (see Table 3). Furthermore, the last vowel of an intonational phrase (V_{IP}) exhibits lower intensity than that of an intermediate phrase (V_{ip}), and even more so than the last vowel of a word within a phrase ($V_{\#}$).

The decrease in intensity at the end of an intonation phrase is not influenced by reading rate, as speakers consistently decrease intensity in this position regardless of the reading rate. However, the data for intermediate phrases shows greater variability. Speaker S1 demonstrates the greatest decrease in intensity during natural reading, a moderate decrease in slow reading, and the least decrease in fast reading. Speaker S2 shows the largest decrease in intensity during fast reading, the smallest during slow reading, and a moderate decrease during natural reading. Speaker S3 exhibits minimal variation in intensity decrease, with the largest decrease during fast reading, slightly less during natural reading, and the least during slow reading.

Table 3. Intensity ratios of vowels in the final syllable. V marks vowel

Researched unit	**S1**	**S2**	**S3**
$(V_{\#}:V_{ip}:V_{IP})_N$	1:0.18:0.02	1:0.24:0.06	1:0.16:0.01
$(V_{\#}:V_{ip}:V_{IP})_F$	1:0.63:0.02	1:0.15:0.03	1:0.12:0.004
$(V_{\#}:V_{ip}:V_{IP})_S$	1:0.43:0.04	1:0.19:0.02	1:0.18:0.02

The difference between the intensity of the entire phrase and that of the final syllable is statistically significant in all cases. This suggests that the decrease in the intensity of the final syllable serves as a reliable indicator of phrase boundaries.

The present study did not identify consistent patterns related to reading rates, likely due to the limited amount of material examined.

3.3 Fundamental Frequency

Changes in F0 at the end of a phrase can either fall or rise. A decline in F0 is typical of utterances that express a complete thought, while an increase in F0 often signals a question or an incomplete thought. The research material did not include intonation phrases ending with rising F0, although 21% of the intermediate phrases ended with rising F0. Only phrases with a falling F0 were selected for this study, as those with rising F0 require semantic and syntactic analysis and were therefore excluded at this stage of the research.

F0 (as well as voice quality features) were undefined at the end of 45% of the intonation phrases due to creaky voice or devoicing. This also indicates the boundary of the phrase, as all three speech production systems—the respiratory, phonatory, and articulatory—are returning to a neutral position. These cases were also excluded from this analysis.

The data obtained from our material indicate that the F0 of the final syllable is, on average, one-fifth lower than the F0 of the entire intermediate and intonation phrase (see Table 4). Speaker S2 lowers F0 consistently at the end of both types of phrases, with minimal dependence on the reading rate. In contrast, the other two speakers show variability in F0 lowering.

At the end of intermediate phrases, speaker S3 lowers F0 equally (i.e. one-fifth) in natural and slow reading but reduces it by nearly two-fifths in fast reading. Speaker S1 halves the F0 in fast reading but lowers it similarly in slow and natural reading (nearly one-fifth).

At the end of intonation phrases, speaker S3 halves the F0, and this is independent of the reading rate. Whereas for speaker S1, no significant decrease is observed. Statistically significant differences are marked with an asterisk in Table 4.

Table 4. F0 ratios of phrases and their final syllables

Researched unit	**S1**	**S2**	**S3**
ip_N:σ(ip_N)	1:0.8*	1:0.8*	1:0.8
ip_F:σ(ip_F)	1:0.5*	1:0.8	1:0.6*
ip_S:σ(ip_S)	1:0.9	1:0.8*	1:0.8
IP_N:σ(IP_N)	1:0.9	1:0.8*	1:0.5*
IP_F:σ(IP_F)	1:1	1:0.8*	1:0.5*
IP_S:σ(IP_S)	1:0.9	1:0.9	1:0.5*

Similar tendencies are also shown by the analysis of the last vowels F0 ratios in the middle and at the end of the phrase (see Table 5). In the S1 speaker recordings, the F0 of the last vowels at the end of the phrases is one tenth lower than in the middle of the phrases, and this lowering is not dependent on the type of phrase or the reading rate. Although, as can be seen from Table 4, analysing F0 for all (not only identical) last syllable centres and the entire intermediate phrase more significant differences were observed, especially in the case of fast reading.

Speaker S2 vowel data shows more variety than the data for the whole phrase and all the last syllables. There is no F0 lowering of the last vowel in slow reading, and much more (almost 30% and 40%) lowering of the last vowels of the phrase in fast reading.

Speaker S3 shows very similar results when comparing the last vowels with each other and with the F0 of the whole phrase only his speech failed to identify F0 at the end of the intonation phrase in the fast reading.

Table 5. F0 ratios of vowels in the final syllable.

Researched unit	S1	S2	S3
$(V_{\#}:V_{ip}:V_{IP})_N$	1:0.9:0.9	1:0.9:0.8	1:0.7:0.6
$(V_{\#}:V_{ip}:V_{IP})_F$	1:0.9:0.9	1:0.7:0.6	1:0.6:–
$(V_{\#}:V_{ip}:V_{IP})_S$	1:0.9:0.9	1:1:1	1:0.7:0.5

Overall, F0 is consistently lowered at the end of a phrase, but there were no clear patterns in the F0 changes at the end of intonation phrases that were related to reading rate. In the intermediate phrases F0 was lowered more in fast reading than in natural and slow reading, while F0 lowering in slow and natural reading was very similar.

3.4 Duration of the Final Sound

The analysed text is relatively short, and there are few same sounds used in the positions under investigation. Nevertheless, this feature was included in the current study to analyse the duration of [eː] and [ɑː] vowels. It should be noted that the duration of long vowels at the end of a word is particularly complex in Lithuanian due to frequent shortening and reduction. However, since the speakers involved in the study were professional actors, their vowels retained the duration and tension characteristic of Standard Lithuanian. Otherwise, vowel duration results would need to be interpreted with caution.

Although it was anticipated that the longest vowels would occur at the end of intonation phrases, the results do not support this hypothesis. The data are ambiguous but certain tendencies can be seen (see Table 6). All speakers lengthen the final vowels of intonation phrases by an average of one-third during natural reading. Speakers S2 and S3 also lengthen the final vowels of intermediate phrases, and these vowels are lengthened more than the final vowels of intonation phrases.

During fast reading, none of the speakers significantly lengthen the final vowels of intermediate phrases, and only speaker S2 exhibits lengthening at the end of intonation phrases.

In slow reading, all speakers, on average, lengthen the final vowels of intermediate phrases by one-third. However, the final vowels of intonation phrases are lengthened by only one-tenth.

Statistically significant differences were observed in three cases, which are marked with an asterisk in Table 6. This can be attributed to the large variance in the length of vowels at the end of phrases, as indicated by the high standard deviation.

Table 6. Duration ratios of vowels in the final syllables

Researched unit	S1	S2	S3
$(V_{\#}:V_{ip}:V_{IP})_N$	1:1:1.4*	1:1.5*:1.3*	1:1.3:1.2
$(V_{\#}:V_{ip}:V_{IP})_F$	1:1.1:1	1:0.9:1.3	1:1:0.9
$(V_{\#}:V_{ip}:V_{IP})_S$	1:1.3:1.1	1:1.2:1.1	1:1.4:1.1

Therefore, while vowel lengthening at the end of a word can reliably signal the end of a phrase, the absence of vowel lengthening does not reliably indicate that it is not the end of a phrase. Furthermore, only one consistent pattern was observed regarding reading rate: speakers do not prolong final vowels during fast reading.

3.5 Voice quality features

Higher jitter values indicate greater frequency perturbations, which are expected to be highest at the end of an intonation phrase, lower at the end of an intermediate phrase, and lowest at the end of a word in a phrase-internal position.

The findings of this research reveal that jitter values at the end of intonation phrases are higher compared to those in the middle of a phrase or at the end of an intermediate phrase, and jitter values at the end of intermediate phrases are higher compared to those in phrase-internal positions. However, due to large variation in the data, only one-third of the differences are statistically significant (9 out of 27; these are indicated with a dash in Table 7). Furthermore, in 11 out of 36 cases, jitter values were undefined at the end of intonation phrases due to devoicing, which is particularly characteristic of speaker S3. These cases were excluded from the calculation of jitter and shimmer ratios and statistical significance. As mentioned earlier, devoicing clearly signals the end of a phrase. No consistent patterns of jitter changes related to reading rate were observed.

Shimmer measurements, which indicate increased amplitude perturbations, are expected to be higher at the end of larger prosodic units. However, our analysis does not fully support this hypothesis. While shimmer values at the end of intonation phrases are higher than those in phrase-internal positions, the intermediate phrase data are contradictory. In only half of the cases (5 out of 9) are shimmer values higher at the end of intermediate phrases than in phrase-internal positions. Moreover, in 7 out of 9 cases, shimmer values at the end of intermediate phrases are lower than those at the end of intonation phrases.

Only three cases (S1 and S3 natural reading, S2 slow reading) align with the general pattern: shimmer values are highest at the end of intonation phrases, lower at the end of intermediate phrases, and lowest in phrase-internal positions. Unfortunately, none of these differences, nor other differences observed, are statistically significant.

As with jitter, no clear correlation was found between shimmer values and reading rate.

The HNR data show that vowels at the end of an intonation phrase tend to be produced with less periodicity, regardless of the reading rate. The HNR of the final vowel in an intonation phrase is statistically significantly lower than that of vowels in phrase-internal positions (except for speaker S1 during natural and fast reading, where this parameter is equal).

The data for intermediate phrases are less consistent. In 4 out of 9 cases, the HNR at the end of an intermediate phrase is lower than in phrase-internal positions, while in 5 cases, it is higher. Thus, it can be argued that HNR cannot reliably differentiate between intonation and intermediate phrases.

Table 7. Jitter, shimmer, and HNR ratios of vowels in the final syllables

Researched unit	S1	S2	S3
jitter			
$(V_{\#}:V_{ip}:V_{IP})_N$	1:2.3*:3.5*	1:1.1:2.5*	1:1.8:2
$(V_{\#}:V_{ip}:V_{IP})_F$	1:1.3:1.9*	1:1.1:4.2*	1:1.4:–
$(V_{\#}:V_{ip}:V_{IP})_S$	1:1.6:1.9*	1:1.8*:2.1*	1:1.8:8*
shimmer			
$(V_{\#}:V_{ip}:V_{IP})_N$	1:1.3:2.2	1:0.9:1.7	1:1.2:2.6
$(V_{\#}:V_{ip}:V_{IP})_F$	1:0.4:3	1:0.7:0.9	1:2.1:–
$(V_{\#}:V_{ip}:V_{IP})_S$	1:1:1.2	1:1.6:2	1:1.4:1.2
HNR			
$(V_{\#}:V_{ip}:V_{IP})_N$	1:0.7:0.3	1:2.1:1	1:0.6:0.2
$(V_{\#}:V_{ip}:V_{IP})_F$	1:1.1:0.2	1:1:1	1:0.5:–
$(V_{\#}:V_{ip}:V_{IP})_S$	1:1.8:0.4	1:0.5:0.3	1:1.5:0.2

To summarise, jitter can be used to identify not only the end of a phrase but also the type of phrase. Shimmer and HNR can indicate the end of an intonation phrase; however, their effectiveness in marking the end of an intermediate phrase should be considered with caution. This study suggests that the reading rate does not have a significant effect on voice quality at the end of a phrase.

4. Conclusions

The results of this stage of the research lead to the following preliminary conclusions:

- **Pauses** consistently mark the boundaries of intonation phrases but are not always necessary for intermediate phrases. An increase in reading rate reliably shortens pauses, while a decrease extends them for both intonation and intermediate phrases. However, pauses between intermediate phrases are less shortened during fast reading and are more extended during slow reading compared to pauses between intonation phrases. There may also be fewer pauses between intermediate phrases in fast reading and more in slow reading.
- **Intensity decreases** are a reliable indicator of phrase boundaries. Based on the level of intensity decrease, it is possible to identify the type of phrase. For intonation phrases, the difference in intensity between the last syllable and the whole phrase is nearly twice as great as for intermediate phrases. Reading rate does not appear to influence the level of intensity decrease at the end of intonation phrases. However, it is unclear whether reading rate affects intensity decrease at the end of intermediate phrases.

- **F0 lowering** consistently occurs at the end of phrases, but its relationship to phrase type and reading rate is not always clear. In fast reading, F0 lowering at the end of intermediate phrases may be more pronounced. Additionally, devoicing of the final syllable reliably signals the end of an intonation phrase.
- **Lengthening of sounds** at the end of a phrase serves as an indicator of phrase boundaries. The data show one clear tendency related to reading rate: speakers accelerate reading by reducing or omitting lengthening of the final vowels of a phrase. This suggests that speakers speed up their reading rate not only by shortening or omitting pauses but also by avoiding the prolongation of final vowels.
- **Jitter** results support the assumption that an increase in jitter is a reliable indicator of phrase boundaries. Jitter can also help distinguish phrase types, as larger intonation units exhibit higher jitter values, indicating a characteristic frequency perturbation.
- **Shimmer** results did not confirm predicted patterns related to phrase type. While shimmer values are higher at the end of intonation phrases compared to phrase-internal positions, intermediate phrase data are inconsistent. This is likely due to the limited scope of the material examined.
- **HNR** data show that vowels at the end of intonation phrases tend to be produced with less periodicity. However, HNR does not consistently signal the end of intermediate phrases.
- The study found **no significant effect of reading rate on voice quality**, as parameters like jitter, shimmer, and HNR remained relatively stable across different reading rates.

All the analysed features signal the end of a phrase to varying degrees. However, according to the data from this study, only the duration-related features—namely, the existence of pauses, their duration, and the lengthening of the final sound—respond to changes in reading rate, while the other features remain relatively constant.

The next stage of the study aims to verify the tendencies identified here using a larger data sample, which will include longer and more varied texts, as well as recordings of spontaneous speech. Additionally, the qualitative characteristics of final sounds should be investigated. Phrase onset data in Lithuanian has not yet been thoroughly analysed, except in perceptual experiments (Kazlauskienė et al., 2023, pp. 87–155). Therefore, these phenomena should be incorporated into future research on prosodic phrasing. The effect of discovered pausing duration patterns on the naturalness of synthetic speech at varying reading rates will be examined using our experimental speech synthesis system.

References

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound, Institute of Phonetic Sciences, University of Amsterdam, *Proceedings* 17, p. 97-110, available at https://www.researchgate.net/publication/2326829_Accurate_Short-Term_Analysis_Of_The_ Fundamental_Frequency_And_The_Harmonics-To-Noise_Ratio_Of_A_Sampled_ Sound

Boersma, P. (2009). Should *jitter* be measured by peak picking or by waveform matching?, *Folia Phoniatrica et Logopaedica* 61: p. 305–308. doi: 10.1159/000245159.

Boersma, P., Weenink, D. (2018). *Praat: doing phonetics by computer, version 6.035*. www.praat.org [computer program], available at https://www.fon.hum.uva.nl/praat/

Brugos, A., Breen, M., Veilleux, N., Barnes, J., Shattuck-Hufnagel, S. (2018). Cue-based annotation and analysis of prosodic boundary events, *Proceedings of the 9th international conference on Speech Prosody 2018*, p. 245-249. doi: 10.21437/SpeechProsody.2018-50.

Crowhurst, M. J. (2018). The joint influence of vowel duration and creak on the perception of internal phrase boundaries, *Journal of Acoustic Society of America*, 143(3): EL147. doi: 10.1121/1.5025325.

Dereškevičiūtė, S., Kazlauskienė, A. (2022). Prosodic phrasing in Lithuanian: preparatory study, *Baltic J. Modern Computing*, 10(3), 317–325. doi:10.22364/bjmc.2022.10.3.05.

Harrington Stack, C., Watson, D. G. (2023). Pauses and Parsing: Testing the Role of Prosodic Chunking in Sentence Processing. *Languages*, 8, no. 3, 157, https://doi.org/10.3390/languages8030157.

Kazlauskienė, A., Dereškevičiūtė, S., Sabonytė, R. (2023). *Standard Lithuanian intonation: phrase centre, boundaries, and marking* (in Lithuanian), https://doi.org/10.7220/9786094675782.

Kohler, K. J. (2000). Linguistic and paralinguistic functions of non-modal voice in connected speech, *Proceedings of the 5th Seminar on Speech Production: Models and Data*, 89-92, available at https://www.ipds.uni-kiel.de/kjk/pub_exx/kk2000_2/51.pdf

Liu, R., Liu, B., Li, H. (2023). Emotion-Aware Prosodic Phrasing for Expressive Text-to-Speech. *arXiv preprint arXiv:2309.11724*, https://doi.org/10.48550/arXiv.2309.11724.

Michelas, A., D'Imperio, M. (2010). Durational cues and prosodic phrasing in French: Evidence for the intermediate phrase. *Speech Prosody*, 2010-196. doi: 10.21437/SpeechProsody.2010-196.

Napoleão de Souza, R. (2023). Segmental cues to IP-Initial boundaries: Data from English, Spanish, and Portuguese. *Prosodic boundary phenomena*, p. 35-86. Language Science Press. doi: 10.5281/zenodo.777752.

Peters, B. (2003). Multiple cues for phonetic phrase boundaries in German spontaneous speech, *Proceedings 15th ICPhS*, Barcelona, p. 1795-1798, available at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_1795.pdf

Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenburger, I., Höhle, B. (2017). Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists, *Journal of Phonetics*, 61, p. 71-92, https://doi.org/10.1016/j.wocn.2017.01.002.

Steffman, J., Kim, S., Cho, T., Jun, S. (2024). Speech rate and prosodic phrasing interact in Korean listeners' perception of temporal cues. *Speech Prosody 2024*, p. 1090–1094, https://doi.org/10.21437/speechprosody.2024-220.

Volin, J., Šturm, P., Skarnitzl, R., Bořil, T. (2024). *Prosodic phrase in spoken Czech*. Prague: Charles University, Karolinum Press.

Żygis, M., Tomlinson, J., Petrone, C., Pfütze, D. (2019). Acoustic cues of prosodic boundaries in German at different speech rate, *Proceedings of ICPhS 2019*, p. 999-1003, Melbourne, Australia. ⟨hal-02097707⟩

Received January 21, 2025, accepted April 16, 2025