INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

Prosodic Factors Influencing Vowel Reduction in Russian

Daniil Kocharov, Tatiana Kachkovskaia, Pavel Skrelin

Department of Phonetics, Saint Petersburg University, Russia
[kocharov, kachkovskaia, skrelin]@phonetics.pu.ru

Abstract

Unstressed vowels in Russian are reduced in both duration and quality, but these two manifestations of vowel reduction do not have to be observed simultaneously. In order to investigate this question, we analysed the reduction pattern of words in such contexts where lengthening is induced by prosodic factors: prominence and pre-boundary lengthening. The study is based on a large corpus of read speech. The following results were obtained: (1) as expected, both contexts increase vowel duration; (2) under prosodic prominence vowels undergo less qualitative reduction, while pre-boundary lengthening has no effect on qualitative reduction; (3) additionally, it was shown that prominence mainly affects the pretonic part of the word, while pre-boundary lengthening—the post-tonic part. Thus, an increase in vowel duration does not always cause a decrease in qualitative reduction, which may serve as evidence against the idea that qualitative reduction is caused by quantitative reduction. Additionally, these results may serve as an argument for the idea that the two processes—vowel reduction and temporal organization of utterance—are autonomous.

Index Terms: vowel reduction, segmental duration, prosodic prominence, pre-boundary lengthening, Russian

1. Introduction

Reduction, or weakening, of vowels in lexically unstressed syllables is a universal phenomenon observed in typologically different languages including Dutch, English, Finnish, French, German, Russian and other languages [1], [2], [3], [4], [5]. Phonetically, vowel reduction implies a change in the formant values causing a shrinkage of the vowel space.

Reduction of unstressed vowels is especially noticeable in Standard Russian [6], [7], [8]. The position of lexical stress in Russian is not fixed, and the main correlate of lexical stress is quantity [9].

For Russian, the two key factors influencing the extent of vowel reduction are position relative to word boundary (absolute-initial or absolute-final vs. non-absolute-initial or non-absolute-final) and position relative to the stressed syllable (pretonic vs. post-tonic, immediately-pretonic vs. non-immediately-pretonic). Traditionally, there are two degrees of reduction: (1) immediately-pretonic or absolute-initial; (2) non-immediately-pretonic, but not absolute-initial, and post-tonic [6, 10].

In continuous speech the traditional vowel reduction pattern might change under the influence of global prosodic factors which affect segmental duration. Thus those words in a sentence which are most important for conveying information often carry nuclear stress or logical stress, i. e. they are prosodically prominent. It is possible that such words are pronounced more carefully, leading to weaker vowel reduction. This claim is supported for Dutch in [3] where acoustical analysis revealed the influence of sentence accent on both steady-state formant frequencies and duration of vowels.

It is often stated [11], [7] that vowel quality change is a result of decreased vowel duration: the shorter the vowel, the less time the articulatory system has to perform the articulatory program—thus, the more reduced the vowel. The relation between quantitative (durational) and qualitative (formant) reduction remains an open question. In order to answer this question, it is reasonable to consider vowels which are lengthened due to factors other than position relative to the stressed syllable or word boundaries—such as pre-boundary (phrase-final) lengthening. Being a universal phenomenon [12], pre-boundary lengthening might be governed by other mechanisms than those governing vowel reduction. It is of great interest whether the reduction pattern of the word changes in phrase-final position or remains the same. A similar hypothesis could be formulated concerning prosodic prominence, which also causes lengthening in many languages [13][p. 32].

The present study is aimed at answering the following major questions.

1. How does vowel reduction depend on prosodic prominence and pre-boundary lengthening?
2. What is the relation between qualitative reduction and quantitative reduction?

2. Experimental material

The corpus used in this study is CORPRES (Corpus of Professionally Read Speech) developed at the Department of Phonetics, St. Petersburg State University [14]. The annotated part of the corpus was used in this research. It contains recordings of several large texts read by 8 professional speakers; the total duration is over 30 hours with over 1.1 million sounds. Along with the orthographic tier it contains a prosodic tier with annotation of boundaries of tone units, accented and prominent words. Additionally the corpus contains two phonetic tiers:

- manual transcription produced by expert phoneticians and based on perceptual and acoustic analysis;
- automatic transcription produced by the grapheme-to-phoneme transcriber following the orthoepic rules of Standard Russian, including the rules describing the processes observed in connected speech (such as assimilation at word boundaries, lack of stress on clitics etc.)

On both tiers broad (phonemic) transcription is used. The system used in CORPRES for segmental transcription contains six vowels: /i/, /e/, /ɑ/, /o/, /u/, /ɨ/ (following the Leningrad (St. Petersburg) Phonological School [15], [16], [7]). Additionally, the annotation tiers contain information on each vowel's position relative to stress within the clitic group. The annotation system is described in detail in [14]. As vowel quality was assessed by trained phoneticians, we assume that these judgements objectively reflect the formant structure of the analysed vowels.

When manual and automatic transcription differ we observe a significant change in vowel quality which results in assigning

http://dx.doi.org/10.21437/Interspeech.2019-2918

it a phoneme type different from the phoneme type assigned by the orthoepic rules. Thus, the first tier reflects both systematic and unsystematic changes, while the second tier—only systematic. The difference between these tiers, if present, is a result of unsystematic changes. Therefore, the percentage of cases where the two transcriptions differ is a measure of unsystematic change. Given a large corpus of data, the key assumption of our study is that both the amount of unsystematic vowel alternations and quantitative vowel reduction are influenced by the same factors, and the effect of this influence is the same. In other words, we assume that the amount of unsystematic vowel alternation may be used to estimate the influence of the factors in question on qualitative vowel reduction.

3. Experimental method

A comparison of the phonetic tiers of manual and rule-based transcription shows that even in professionally-read speech it is not uncommon to observe vowel and consonant omissions or replacements: e.g. 'обстоятельство' /ɑpstɑˈjɑtʲjɪstvɨ/ ('circumstances') instead of /ɑpstɑˈjɑtʲjɪlʲstvɑ/—an omission of /ɪ/ between /tʲ/ and /lʲ/ and a replacement of absolute-final /ɑ/ by /ɨ/.

In order to calculate the number of vowel omissions or replacements in different positions relative to the stressed syllables, the corpus data were analyzed automatically. The prosodic unit we are focused on is an accentual phrase (clitic group), i. e. content word and its surrounding clitics if present, e. g. 'а до этого' ('and before that') pronounced as /ɑdɑˈetɑvɑ/. Those words whose status is arguable (e. g. pronouns) were considered content words if they contained a stressed vowel, and a clitic otherwise.

For each accentual phrase in the corpus the two types of transcription were aligned using a modified version of Levenshtein algorithm [17]. Then for each vowel the following parameters were calculated:

- numerical parameters: duration (z-score-normalized duration calculated over each phoneme and each speaker), omissions/replacements ('1' or '0' depending on whether the transcriptions for the vowel differ or not);
- linguistic parameters
  - position of the vowel within the word: for pretonic vowels: '1' for the 1-st syllable, '2' for th 2-nd syllable etc.; for post-tonic vowels: '1' for the 1-st post-tonic syllable, '2' for th 2-nd post-tonic syllable etc.;
  - for unstressed vowels: the length of the pre-tonic or post-tonic part of the word where the vowel occurs.

Additionally, we subdivided corpus data by the features needed for our analysis, namely: position of the word within the intonational phrase: non-final vs. final; prosodic prominence of the word: non-prominent vs. prominent.

Z-score-normalized duration. Since the absolute vowel duration differs for different types of vowels—e.g., /ɑ/ is usually longer than /ɪ/—a reliable comparison of vowel duration requires calculation of z-score, i. e. duration value normalized by the mean and standard deviation [18].

Positive normalized duration values indicate that the duration of the vowel is higher than the mean for this phoneme; negative normalized duration values correspond to duration values below the mean. The value itself is measured in standard deviations. For example, in CORPRES the mean duration of /ɑ/ ranges between 75 and 90 ms for different speakers, while the mean duration of /ɪ/—between 55 and 70 ms. The standard deviation for the vowel /ɑ/ ranges between 15 and 35 ms depending on the speaker, for the vowel /ɪ/—between 20 and 23 ms.

For the calculation of the vowel duration only that part of the vowel was taken where formant structure could be observed. Voiceless parts of vowels and parts with tone but no formant structure were therefore excluded from the analysis (see [19]).

Omissions/replacements percentage. The percentage of omissions and replacements was calculated for each position within the word. Similar calculations were performed for the duration values.

Grouping. The values for unstressed vowels were grouped according to the length (the number of syllables) of the pretonic or post-tonic part of the word where it occurs. (Compound words containing more than one stressed syllable were excluded from the analysis.) We assume that the length of the post-tonic part of the word has no crucial influence on the vowel reduction processes within the pretonic part, and vice versa—the length of the pretonic part of the word has no crucial influence on the vowel reduction processes within the post-tonic part. Thus, for example, the words 'кора' /kɑˈrɑ/ ('bark'), 'палата' /pɑˈlɑtɑ/ ('chamber') and 'количество' /kɑˈlʲitʃʲistvɑ/ ('quantity') can be grouped together to analyse the vowel reduction processes in one-syllable pre-tonic groups, despite the fact that they have a different number of post-tonic syllables.

Then for each part (pretonic and post-tonic) N vowel positions were analyzed, where N is the length of the pre-tonic or post-tonic part of the word.

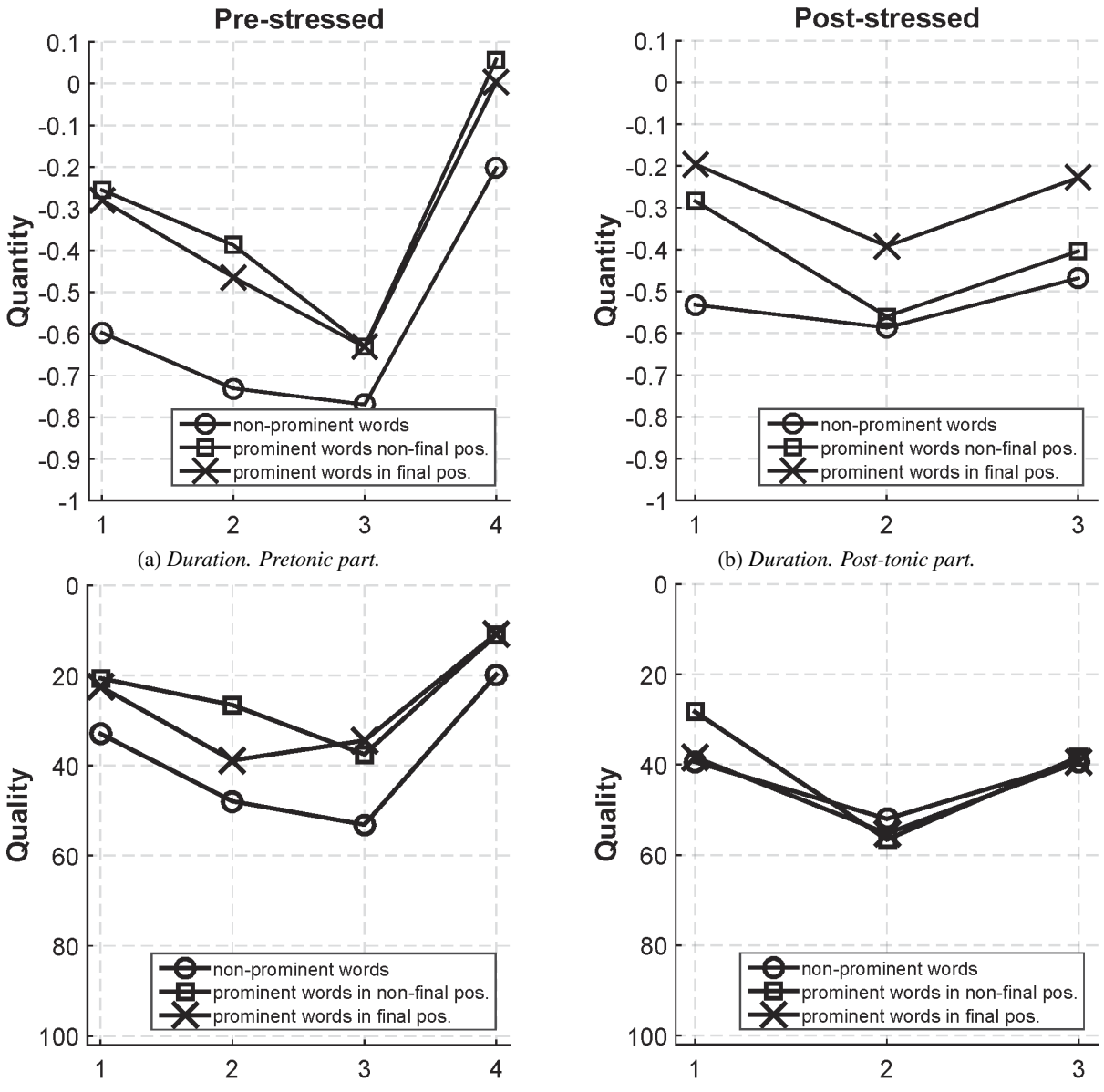
In order to estimate the significance of our observations and speculations about vowel reduction in Russian, statistical analysis was performed using R. First, each sample was tested for normality by means of Shapiro-Wilk test [20]. For normally distributed data Student's t-test was used, otherwise—Wilcoxon signed-rank test [21].

4. Results and Discussion

A series of experiments were carried out to test how prosodic prominence and pre-boundary lengthening influence vowel reduction in Russian. There are two major prosodic phenomena which might influence vowel duration and accuracy of pronunciation. On the one hand, those words in a sentence which are prominent—i. e. carry a nuclear stress or a logical stress—tend to be pronounced more carefully than all other, prosodically neutral words. On the other hand, words at the end of a prosodic unit tend to be longer than other words due to the final lengthening effect.

In order to estimate the influence of these two phenomena on vowel reduction, it was reasonable to compare duration and omissions/replacements percentage for the following groups of words—according to their position relative to the end of the intonational phrase and the presence or absence of prosodic prominence:

1. A. non-prominent words in non-final position: 51785 words;
2. B. prominent words in non-final position: 10359 words;
3. C. non-prominent words in final position: 4989 words;
4. D. prominent words in final position: 58122 words.



(a) *Duration. Pretonic part.*

(b) *Duration. Post-tonic part.*

(c) *Omissions/replacements percentage. Pretonic part.*

(d) *Omissions/replacements percentage. Post-tonic part.*

Figure 1: *Omissions/replacements percentage (%) and normalized duration (z-scores) of vowels in three positions within the phrase: non-prominent and non-final words (marked with circles; group A), prominent and non-final words (marked with squares; group B), and prominent and final words (marked with crosses; group D). The pretonic and post-tonic parts of the word are presented separately; the X axis shows the index of the syllable counted from the beginning of the word (for the pretonic parts) or from the stressed syllable (for the post-tonic parts).*

Our data show that both of these factors positively influence vowel duration, i. e. the presence of at least one of them leads to vowel duration increase ($p < 0.001$). More accurate analysis of the data reveals that these factors affect vowel lengthening in different ways (see table 1). In case of word prominence, pre-tonic vowels are lengthened more than post-tonic vowels (see comparison of groups A and B in table 1: 0.26 vs. 0.13, $p<0.001$). In the phrase-final position, post-tonic vowels are lengthened more than pre-tonic vowels (see comparison of A and C in table 1: 0.29 vs. 0.18, $p<0.001$). In other words, prosodic prominence affects mostly the pretonic part of the word, whereas pre-boundary lengthening affects mostly the post-tonic part of the word. The asymmetry observed here for prosodic prominence is opposite to what was found for English and Dutch [22]: for those languages the lengthening effect was stronger on the post-tonic part of the word. Thus, the distribution of lengthening among the word seems a language-specific phenomenon.

In theory, words in group D might benefit from *both* of these factors. A comparison of these words with non-prominent words in non-final position (i. e. prosodically neutral words) shows that pretonic vowels are lengthened in the same way as in

Table 1: *Vowel duration. Results of statistical analysis of the variation between different word sets: A—non-prominent words in non-final position; B—prominent words in non-final position; C—non-prominent words in final position; D—prominent words in final position.*

Word sets	Vowel position	Mean difference	p-value
B vs A	pretonic	0.26	**<0.001**
	post-tonic	0.13	**<0.001**
C vs A	pretonic	0.18	**<0.001**
	post-tonic	0.29	**0.005**
D vs A	pretonic	0.24	**<0.001**
	post-tonic	0.29	**<0.001**
D vs B	pretonic	–0.02	0.34
	post-tonic	0.16	**<0.001**
C vs D	pretonic	0.06	**0.07**
	post-tonic	0.00	0.15
C vs B	pretonic	–0.08	**0.06**
	post-tonic	0.16	**0.03**

Table 2: *Omissions/replacements percentage. Results of statistical analysis of the variation between different word sets: A—non-prominent words in non-final position; B—prominent words in non-final position; C—non-prominent words in final position; D—prominent words in final position.*

Word sets	Vowel position	Mean difference	p-value
A vs B	pretonic	9.41	**<0.001**
	post-tonic	2.68	0.20
C vs D	pretonic	7.23	0.013
	post-tonic	-0.03	0.82
A vs D	pretonic	5.93	**0.004**
	post-tonic	1.93	0.64
C vs B	pretonic	10.46	**0.001**
	post-tonic	0.72	0.84
D vs B	pretonic	3.48	0.017
	post-tonic	0.75	0.85
A vs C	pretonic	1.97	0.51
	post-tonic	-1.97	0.65

case of prominent words, and post-tonic vowels are lengthened in the same way as in case of pre-boundary position (see comparison of groups A and D in Table 1: 0.24 vs. 0.29, p<0.001). Thus we conclude that, firstly, these factors influence vowel lengthening relatively independently and, secondly, the impact of these factors is not additive: the overall durational increment is equal to the largest of the two.

Another proof to support this claim is shown by the comparison of post-tonic vowel duration in prominent words: they are longer in final position than in non-final ($p < 0.001$), whereas the duration of pretonic vowels does not change. A secondary argument for this is a statistically insignificant, but observed, tendency towards the lengthening of pretonic vowels in words in final position in case of prominence (p=0.07). Our data on pre-boundary lengthening agree with a generally accepted view that the more distant a vowel is from the boundary the less the pre-boundary lengthening effect. The data on prominence impact on vowel reduction may be explained by the fact that the speaker pronounces the pretonic part more carefully than the post-tonic part.

Another crucial result is that the qualitative reduction of post-tonic vowels does not depend on these factors (see Table 2). This may indicate that syllables after the stress are prosodically weaker than the syllables before the stress. Similar phenomenon was described by M. Baltazani for Greek [23]. This is supported by the high syntactic and semantic redundancy of Russian: formally, inflections and suffixes play a major role in linking words together within a sentence but lexical meaning of words help the listener predict the links between them. These morphemes are highly probable to appear in the post-tonic part of the word, and their recognition does not depend on their pronunciation [24]. In this case the major phonetic role is played by rhythmic and morphonological factors, i. e. what morphemes occur in the word and what their phonetic form is.

The omissions/replacements percentage of pretonic vowels is correlated with prosodic prominence just as we hypothesized, which is illustrated by comparison of prominent and non-prominent words in all positions (see Table 2). Vowels in prominent words have lower omissions/replacements percentage than vowels in non-prominent words. At the same time, final position has no impact on qualitative vowel reduction—see comparison of omissions/replacements percentage for non-prominent words in final and non-final positions.

Figures 1a–1d illustrate the patterns of vowel duration and omissions/replacements percentage depending on word prominence and position within the tone unit. Here we only present data for 4-syllable pretonic parts and 3-syllable post-tonic parts, but the cases of shorter parts show similar tendencies. The vowels in prominent words are longer: duration values of the post-tonic vowels in non-prominent words are lower than those in prominent words; the pretonic vowels in non-prominent words also tend to be shorter than in prominent words. The case of omissions/replacements percentage is different: the vowels within pretonic groups in prominent words only change slightly less than in non-prominent words, while the vowels in post-tonic positions do not show any difference between words in the two positions. The correlation between vowel duration and omissions/replacements percentage holds true only for pretonic vowels.

5. Conclusions

Our studies have shown that pretonic and post-tonic parts of the word are not symmetrical with respect to the influence of prosodic factors. Prosodic prominence and phrase-final position do not influence qualitative vowel reduction in post-tonic syllables. Prosodic prominence decreases the omissions/replacements percentage for pretonic vowels, while pre-boundary lengthening does not reveal any influence on the quality of pretonic vowels. Both prosodic prominence and phrase-final position increase vowel duration, the former influencing mostly the pretonic part while the latter—the post-tonic part. However, the effect of these factors is not additive. Thus, an increase in vowel duration does not always cause a decrease in the omissions/replacements percentage, which may serve as evidence against the idea that qualitative reduction is caused by quantitative reduction. Additionally, these results may serve as an argument for the idea that the two processes—vowel reduction and temporal organization of utterance—are autonomous.

6. References

[1] C. Mooshammer and C. Geng, "Acoustic and articulatory manifestations of vowel reduction in German," *Journal of the International Phonetic Association*, vol. 38, no. 2, pp. 117–136, 2008.

[2] R. J. J. H. van Son, O. Bolotova, M. Lennes, and L. C. W. Pols, "Frequency Effects on Vowel Reduction in Three Typologically Different Languages (Dutch, Finnish, Russian)," in *Proceedings of INTERSPEECH 2004*, 2004, pp. 1277–1280.

[3] D. R. van Bergem, "Acoustic vowel reduction as a function of sentence accent, word stress, and word class," *Speech Communication*, vol. 12, no. 1, pp. 1–23, 1993.

[4] P. Delattre, "An Acoustic and Articulatory Study of Vowel Reduction in Four Languages," *International Review of Applied Linguistics in Language Teaching*, vol. 7, no. 4, 1969.

[5] K. M. Crosswhite, "Vowel reduction," in *Phonetically Based Phonology*, B. Hayes, R. Kirchner, and D. Steriade, Eds. Cambridge University Press, 2004, pp. 191–231.

[6] R. I. Avanesov, *Russian Standard Pronunciation [Russkoe literaturnoe proiznoshenije]*. Prosveschenije, 1984.

[7] L. Bondarko, *Phonetics of Contemporary Russian [Fonetika sovremennogo ruskogo jazyka]*. SPbSU, 1998.

[8] K. M. Crosswhite, "Vowel reduction in Russian: A unified accountof standard, dialectal, and "dissimilative" patterns," *University of Rochester Working Papers in the Language Sciences*, pp. 107–172, 2000.

[9] L. V. Bondarko, L. A. Verbitskaya, and M. V. Gordina, *Fundamentals of General Phonetics [Osnovy obschej fonetiki]*. Academia, SPbSU, 2004.

[10] J. Barnes, *Strength And Weakness at the Interface: Positional Neutralization in Phonetics And Phonology*. Mouton de Gruyter, 2006.

[11] B. E. F. Lindblom, "Spectrographic Study of Vowel Reduction," *Journal of the Acoustical Society of America*, vol. 35, no. 11, pp. 1773–1781, 1963.

[12] J. Vaissière, "Language-Independent Prosodic Features," in *Prosody: Models and Measurements*, ser. Springer Series in Language and Communication, A. Cutler and D. R. Ladd, Eds. Springer Berlin Heidelberg, Jan. 1983, no. 14, pp. 53–66.

[13] D. Hirst and A. Di Cristo, "A survey of intonation systems," in *Intonation Systems. A Survey of Twenty Languages*, D. Hirst and A. Di Cristo, Eds. Cambridge University Press, 1998, pp. 1–44.

[14] P. A. Skrelin, N. B. Volskaya, D. Kocharov, K. Evgrafova, O. Glotova, and V. Evdokimova, "A fully annotated corpus of russian speech," in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA), 2010, pp. 109 – 112.

[15] L. V. Scherba, "New problems of linguistics [ocherednye problemy jazykovedenija]," *Izvestia of The Academy of Sciences of the USSR*, vol. IV, no. 5, pp. 173–186, 1945.

[16] L. R. Zinder, *General phonetics [Obschaja fonetika]*. Vysshaja shkola, 1979.

[17] D. Kocharov, "Automatic alignment of phonetic transcriptions for russian," in *Speech and Computer*, A. Ronzhin, R. Potapova, and V. Delic, Eds. Springer International Publishing, 2014, pp. 123–128.

[18] C. W. Wightman, S. Shattuck-Hufnagel, M. Ostendorf, and P. J. Price, "Segmental durations in the vicinity of prosodic phrase boundaries," *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1707–1717, Mar. 1992.

[19] A. Turk, S. Nakai, and M. Sugahara, "Acoustic segment durations in prosodic research: A practical guide," in *Methods in Empirical Prosody Research*, S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augursky, I. Mleinek, N. Richter, and J. Schliesser, Eds. Mouton de Gruyter, 2006, pp. 1–26.

[20] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality," *Biometrica*, vol. 52, pp. 591–611, 12 1965.

[21] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[22] T. Cambier-Langeveld and A. E. Turk, "A cross-linguistic study of accentual lengthening: Dutch vs. English," *Journal of Phonetics*, no. 27, pp. 255–280, 1999.

[23] M. Baltazani, "Prosodic rhythm and the status of vowel reduction in Greek," in *Selected papers on theoretical and applied linguistics from the 17th international symposium on theoretical and applied linguistics*, vol. 1, 2007, pp. 31–43.

[24] L. Bondarko and L. Verbitskaya, "On phonetic features of poststressed flexions in modern Russian [O foneticheskih harakteristikah zaudarnyh fleksij v sovremennom russkom jazyke]," *Voprosy jazykoznanija*, no. 1, pp. 37–49, 1973.