

“KARTINA, KORZINA, KARTONKA...” PROSODIC LABELING OF ENUMERATION CONTEXTS IN POETRY READING IN RUSSIAN

Abstract The present exploratory study aimed to test some notions regarding the prosody of enumeration in Russian on a small corpus of readings of a poem *Luggage* by Samuil Marshak performed by six speakers of Standard Russian. The recordings were labeled manually in *Praat*. A high degree of interspeaker variability was attested in the data. In accordance with previous findings, the data demonstrates that the choice between two pitch accents most widely used to mark enumeration in Russian (L*+H and L*) correlates with speech tempo. Within the lines read at a lower speaking rate the speakers produced more frequently the combination of accents H* L* H%, a tune referred to as IK-4 in traditional descriptions of Russian prosody.

Keywords Prosody, Pitch accent, Enumeration, Russian, Speaking tempo.

1. Enumeration prosody in Standard Russian

The variability of prosodic strategies in the contexts of enumeration (listing three or more objects one after another within a sentence) attracted considerable interest in previous studies of Russian intonation. Traditional descriptions

of Russian prosody based on Bryzgunova's framework (1980) distinguish three main *intonational constructions* (IKs) normally marking non-final elements of enumeration lists:

- 1) IK-3 (a rising pitch movement associated with the prominent syllable and followed by a fall on post-stressed syllables whenever they are available);
- 2) IK-6 (a rising pitch movement associated with the prominent syllable and followed by a level high pitch on poststressed sequence);
- 3) IK-4 (a complex fall-rise pitch movement, with a trough loosely associated with the prominent syllable and followed by high pitch on poststressed sequence of syllables).

According to Bryzgunova, the choice between the three constructions is merely a stylistic decision with no semantic consequences. IK-3 is seen as a neutral choice typical for colloquial speech, while the latter two are claimed to be characteristic for reading aloud, with IK-4 being a more “business-like”, “official”, formal variant, and IK-6 a more “solemn” and “emotional” one (Bryzgunova 1977: 102).

Corpus studies conducted by Yanko (e. g., 2008, 2017) outline an even higher degree of variability in Russian enumeration contexts. The researcher adds two patterns with no final rise to the list of tunes typical for enumeration: IK-2 (a steep fall associated with the prominent syllable) and IK-5 (a “hat pattern”). Unlike Bryzgunova, Yanko assumes that the apparent prosodic variability has to do with the information structure, namely, some of the accents are more likely to mark the theme and rheme within the phrase. Another

notion put forward by the researcher is *рассказ по порядку* (a structured narration; a narration “in an orderly fashion”), which is claimed to be a pragmatic meaning typically conveyed by IK-4.

Some experimental data regarding Russian enumeration contexts can be found in Rathcke’s perception study (2013). It should be noted, however, that Rathcke bases her experimental design and analysis on the assumption that the non-final elements of lists in Russian are marked by a combination of tones $L^*+H H\%$, and not by the polar question pattern $L^*+H L\%$, a notion that runs counter to the traditional descriptions of Russian prosody treating IK-6 and IK-3 as equally possible options for enumeration. The study showed that truncation (that is, elimination due to the lack of poststressed voiced segments) of the boundary tone in Russian does not lead to complete neutralization between $L^*+H H\%$ and $L^*+H L\%$. Russian participants used higher pitch scaling as a cue to identify truncated polar questions and to distinguish them from truncated enumeration contexts.

The phonetics of the “fall-rise” contour IK-4 marking enumeration has not been thoroughly discussed in experimental literature. To fill this gap, in our previous production study (Duryagin 2021) we compared the phonetic realization of pitch accents produced in enumeration contexts with pitch accents produced within the phrase type that most typically requires IK-4, an elliptic contrastive question. Based on the measurements of tonal targets’ alignment, we analyze IK-4 as a combination of nuclear accent L^* (phonetically manifested as an F_0 trough associated with the center of the prominent syllable) and a boundary tone $H\%$ (realized as a local maximum of which the alignment

varies to a high degree). This combination of accents is often preceded by an optional high tonal target which we analyze as representing an underlying prenuclear accent H^* .

Our previous study generally confirmed the variability of tonal configurations attested in previous research. We further tested the effects of reading tempo on the phonetic realization of nuclear accents in enumeration contexts (the speakers were instructed to read the text three times: in normal, slow, and fast tempo). A significant effect of slow reading on the relative alignment of tonal targets was found, with both local minima and maxima aligned later in absolute and relative measurements under this condition. These findings can be interpreted as reflecting a tendency to substitute more neutral patterns containing L^*+H nuclear accent with a more salient tune $H^* L^* H\%$ in slow speech without the time pressure.

We conclude this introduction by positing some proposals that regard the prosodic marking of enumeration context and are based on the literature review and our experimental data:

- 1) Non-final elements of an enumeration list in Russian can be marked by three “rising” tonal configurations: $L^*+H H\%$ (IK-6, in Bryzgunova’s framework), $L^*+H L\%$ (IK-3), $H^* L^* H\%$ (IK-4). In addition, occasional use of “falling” configurations $H^*+L L\%$ (IK-2) is attested in corpus data;
- 2) the factors affecting the choice of pitch accents and boundary tones in enumeration contexts are not yet clear; however, we see a correlation between slow speaking tempo and the higher probability of $H^* L^* H\%$ (IK-4) configuration being chosen.

These claims require further testing with experimental and corpus methods. In this paper, we report the results of a pilot study that aimed to test the validity of our previous experimental findings on a small corpus containing readings of Samuil Marshak’s poem *Luggage* («Багаж») by six adult speakers of Standard Russian.

2. The corpus and its prosodic labeling

The poem *Luggage* («Багаж») by Samuil Marshak was first published in 1926. It tells a story of a lady who travels in train with her cumbersome luggage. Most importantly for the present study, the description of the luggage presents a typical enumeration list with seven objects: «... диван, чемодан, саквояж, картина, корзина, картонка и маленькая собачонка» («a sofa, a suitcase, a valise, a painting, a basket, a cardboard box and a small dog»). This refrain is repeated eight times in the poem creating the comic effect along with the final twist of the story.

Besides the abundance of enumeration contexts in Marshak’s poem, the other reason why it was chosen for the present study is its popularity. Since it became a classic often read to preschool and primary school children, multiple authentic readings of this poem are freely available online. We used the video sharing platform YouTube to obtain the data for our study. After excluding readings performed by underage children and speakers of non-Standard Russian, as well as low-quality recordings, we selected for the analysis a sample of six readings (see the links to publicly available YouTube videos in the *Resources* section below). Readings 1 and 2 are performed by professional voice actors (both male); reading

6 is performed by an actress who works in a puppet theatre; reading 5 is performed by a film and theatre actress; readings 3 and 4 are performed by readers who do not identify themselves as professional actors (one male and one female). All speakers are adult native speakers of Standard Russian with no apparent signs of regional or dialectal phonetics.

The recordings were labeled manually in *Praat* (Boersma & Weenink 2021). Out of the eight repetitions in the original text of the poem, two (from the fifth and the seventh verses) were excluded from the analysis because they contain the refrain in a modified form. Within each refrain, only the six non-last elements of the list («...диван, чемодан, саквояж, картина, корзина, картонка...») were prosodically labeled. In total, our corpus included 216 tokens (six words in six different verses, read by six speakers).

Similarly to the procedure adopted in (Duryagin 2021), we analyzed the tokens by labeling local fundamental frequency (F0) minima and maxima using the *Praat* pitch tracing function with manual corrections for the imprecisions of the algorithm (octave jumps, microprosodic effects, etc.). After that, we marked the observed pitch events with ToBI-like labels: H*, L*, L*+H and H*+L for pitch accents, as well as H% and L% for boundary tones.

AM labeling of our data was by no means a trivial task. Most importantly, since half of our tokens («...диван, чемодан, саквояж...») were the words with ultimate stress, it was not always easy to distinguish between the non-truncated L* H% and the truncated L*+H (T%), where the symbol T% in brackets refers to a truncated tone L% or H%. To distinguish between the two configurations, we used the following assumptions based on the previous findings (see the examples of labeled tokens with final stress in Figure 1 below).

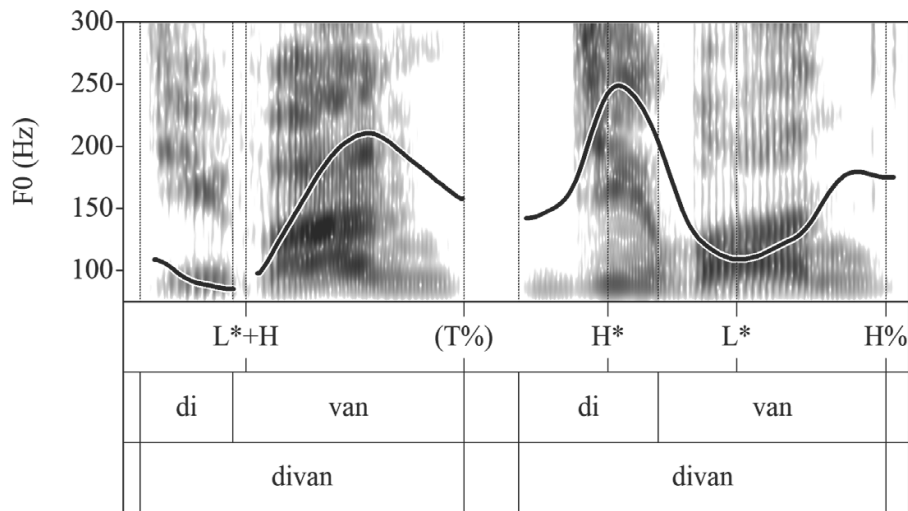


Figure 1. Spectrograms and pitch contours with examples of labeling for two pronunciations of the word *divan*, by speakers 2 (left) and 1 (right).

The recordings were concatenated in *Praat*, the script (Elvira García 2017) was used to produce the illustration.

- 1) L* pitch accent is phonetically manifested with a trough associated with the prominent syllable. Low level pitch is attested in the initial part of the stressed vowel. On the contrary, in case of the nuclear L*+H, the rising transitional pitch movement starts early in the stressed vowel, and no signs of level low Fo are found within the prominent syllable;
- 2) nuclear L* is normally preceded by an optional pre-nuclear accent H* associated with one of the prestressed syllables (usually the first syllable of the word; however, some speakers occasionally associate this accent with the first prestressed syllable). Nuclear L*+H is normally not preceded by pre-nuclear accents;
- 3) in H* L* H% configuration (Bryzgunova’s IK-4), the boundary tone is not truncated. As various descrip-

tions of Russian prosody indicate (e. g., Bryzgunova 1973: 79; Kodzasov 2009: 18), to compensate for the lack of segmental material, Standard Russian speakers utilize the tune-text negotiation strategy of vowel lengthening. On the contrary, the boundary tone following the L*+H nuclear pitch accent is normally truncated, which might lead to (incomplete) neutralization between L*+H L% and L*+H H% (Rathcke 2013). No vowel lengthening was expected in this case (though occasional cases of vowel and coda consonant lengthening were found in the data).

In tokens with final stress, it was often not possible to distinguish between L*+H (L%) and L*+H (H%), since the truncated boundary tone was not represented phonetically (except for the rare cases of coda lengthening). We will further refer to these truncated configurations as L*+H (T%).

It should be added that Marshak's poem previously attracted attention of studies in Russian prosody. Most importantly, Bryzgunova's learning book (1977) offers listening and reading this poem to the students of L2 Russian and contains its complete labeling in IK terms (IK-3, IK-4 and IK-6). Comparing our labeling to Bryzgunova's lies outside the scope of this paper, though it might be of interest for the future studies as a means for testing the degree of inter-labeler agreement.

3. Results. Interspeaker variability

The results of labeling are presented separately for each speaker in Table 1.

Table 1. Number of tonal configurations attested in corpus.

Speaker (gender)	H* L* H% (IK-4)	L*+H T%			H*+L L% (IK-2)	De-ac- cented
		L*+H H% (IK-6)	L*+H L% (IK-3)	L*+H (T%) (truncated)		
1 (M)	17	13	1	5	0	0
2 (M)	2	12	5	15	1	1
3 (M)	24	4	1	6	1	0
4 (F)	28	1	1	0	6	0
5 (F)	25	4	4	0	3	0
6 (F)	15	7	6	6	0	2

A high degree of interspeaker variability was attested in our data. As Table 1 shows, on the one hand, speakers 3, 4 and 5 clearly demonstrated preference for the use of H* L* L% tune. However, all of them also occasionally chose the configurations with the L*+H nuclear accent, followed by high or low boundary tones with boundary tones absent or truncated. An opposite result is attested for the speaker 2 who rarely used the “fall-rise” IK-4 tune and chose the L*+H nuclear accent in most tokens. Finally, speakers 1 and 6 utilized both tunes with comparable frequency.

Rare occurrences of a falling tune H*+L L%, a steep pitch fall associated with the prominent syllable, were present in our data. Interestingly, closer analysis of these contexts showed that 6 of the total 11 belong to the same verse of the poem produced by speaker 4. In this verse, the railroad workers become aware of the fact that the dog is missing and count the luggage in fright («*В испуге считают багаж...*»). We might speculate that the reader used this tune to illustrate the characters’ emotional state.

Besides that, we labeled three tokens (one produced by speaker 2 and two produced by speaker 6) as deaccented.

In these cases, no pitch movements are associated with the stressed syllables, the words are not perceived as prominent, and we treat them as constituting a common prosodic unit with the next prominent word; see a similar interpretation in (Bryzgunova 1977: 232).

4. Results. Speaking rate

Since our previous study (Duryagin 2021) showed that the choice of tune in enumeration contexts correlates with the speech tempo, we used our labeled data to test these findings. For each line of the refrain, (1) «диван, чемодан, саквояж», and (2) «картина, корзина, картонка», we measured the speaking rate in syllables per second. We chose the lines (and not separate words) as the domain for the measurements to account for the length of pauses between the words. As a result, speaking rate for different lines produced by different speakers ranged from 2.04 to 8.01 syll./s (with mean 3.88, $\sigma = 1.14$).

Based on the previous findings, our hypothesis was that in lines pronounced at a lower speaking rate the speakers are more likely to produce the H* L* H% tune (IK-4), while in faster reading they tend to produce more frequently the group of tunes with L*+H nuclear accent (IK-3 and IK-6).

To test the hypothesis, we used a subset of our data from which 14 tokens labeled as deaccented or as containing the H*+L L% tune were excluded. For the remaining 202 observations, we fitted a binary mixed-effects logistic regression model with the dependent variable *tune* (H* L* H% or L*+H T%) and independent variables *speaking rate* and *verse* (repetition of the refrain), as well as random intercepts

for *speaker* and *word*, using R package *lme4* (Bates *et al.* 2015). The model outcome is presented in Table 2 created with *sjPlot* package (Lüdtke 2021).

Table 2. Logistic regression model output for the dependent variable *tune* (with H* L* H% as the reference category).

Predictor	Odds ratios	Confidence intervals	<i>p</i>
(Intercept)	0.00	0.00 – 0.02	< 0.001
verse [2]	0.91	0.24 – 3.46	0.892
verse [3]	0.44	0.11 – 1.72	0.237
verse [4]	0.54	0.12 – 2.44	0.428
verse [6]	0.63	0.15 – 2.64	0.528
verse [8]	0.08	0.01 – 0.49	0.006
speaking rate	6.93	3.23 – 14.87	< 0.001
Observations	202		
Marginal R2 / Conditional R2	0.462 / 0.746		

As Table 2 shows, in our model the predictor *speaking rate* has a significant effect on the choice of tune ($p = .006$). In accordance our previous findings, the speakers produced H* L* H% tune significantly more often in lines read at a slower tempo. It should be noted that this finding does not necessarily imply causation. It might be the case that both prosodic features (the “fall-rise” tune and the slow speaking tempo) are utilized by speakers to mark a common piece of pragmatic or stylistic information.

In this regard, another intriguing finding captured by the model is the effect of *verse*. While no significant differences between verses 1-4 and 6 were found, the last repeti-

tion of the refrain appears to be different. The eighth verse, in which the railway worker controls the receipt and thoroughly indicates to the lady each part of the luggage claiming that none of the objects is missing, is characterized by the most frequent use of the H* L* H% tune ($p = .006$). This finding supports the claims that the choice of IK-4 is typical for formal contexts (Bryzgunova 1977: 114), or that it specifically marks the speakers' intention to perform the enumeration in a structured, ordered fashion (Yanko 2008: 167).

5. Conclusions

The present pilot study allowed us to test some existing notions about the prosody of enumeration in Russian and to provide further evidence for our previous findings regarding the distribution of pitch accents in these contexts and its correlation with speech tempo. It also illustrates the high degree of speaker variability in the choice of tune for enumeration.

We are aware that the current results are to be treated with caution. First, larger speaker samples are required to validate the conclusions concerning the factors affecting the distribution of pitch accents. Since our data suggests that using poetic reading is a viable method of elicitation of prosodic material, we suggest that more data can be collected in controlled experimental conditions. Second, the process of labeling can be optimized by using more than one labeler. Though in most cases distinguishing between nuclear L*+H and L* is straightforward, testing the degree of inter-labeler agreement could validate our notions that the two tunes are perceived categorically by Russian speakers. These adjustments in methodology are reserved for future work on the topic.

List of abbreviations

F0 – fundamental frequency

IK – intonational construction within Bryzgunova’s framework

ToBI – Tones and Break Indices; a prosodic labelling system adopted within the autosegmental metrical framework

H – ToBI label for a high tonal target and the underlying high tone

L – ToBI label for a low tonal target and the underlying low tone

T – undefined (high or low) tone, e.g., a truncated tone

H*, L*, L*+H, etc. – pitch accents; the accents associated with prosodic heads

H%, L% – edge tones; the accents associated with prosodic boundaries

Bibliographical references

BATES D. *et al.*, *Fitting Linear Mixed-Effects Models Using lme4*, in «Journal of Statistical Software», 67(1), 2015, pp. 1-48.

BOERSMA P., WEENINK D., *Praat: doing phonetics by computer* [Computer program]. Version 6.1.38, retrieved 2 January 2021 from <http://www.praat.org/>.

BRYZGUNOVA E.A., *Osnovnye tipy intonacionnykh konstrukcij i ich upotreblenie v russkom jazyke*, in «Russkij jazyk za rubežom», 1, 1973, pp. 75-81.

—, *Zvuki i intonacija russoj reči*, Russkij jazyk, Moscow 1977.

—, *Intonacija*, in ŠVEDOVA N.JU., *Russkaja grammatika*, Nauka, Moscow 1980.

DURYAGIN P., *Fonetičeskaja realizacija russkich tonal'nych akcentov v kontekstach sopostavitel'nogo voprosa i perečislenija*, in «Vestnik Sankt-Peterburgskogo universiteta. Jazyk i literatura», 2021 [In print].

- ELVIRA GARCÍA W., *Create pictures with tiers* [Praat script]. Version 4.4, retrieved 2 January 2021 from https://github.com/wendyelviragarcia/create_pictures.
- KODZASOV S.V., *Issledovanija v oblasti russkoj prosodii*, Jazyki slavjanskikh kultur, Moscow 2009.
- LÜDECKE D., *sjPlot: Data Visualization for Statistics in Social Science*. [R package]. Version 2.8.7, retrieved 2 January 2021 from <https://CRAN.R-project.org/package=sjPlot>.
- RATHCKE T., *On the neutralizing status of truncation in intonation: A perception study of boundary tones in German and Russian*, in «Journal of Phonetics», 3-4 (41), 2013, pp. 172-185.
- YANKO T.E., *Intonacionnye strategii russkoj reči v sopostavitel'nom aspekte*, Jazyki slavjanskikh kul'tur, Moscow 2008.
- , *Intonacija perečislenija v russkom jazyke*, in SELEGEJ V.P., *Komp'juternaja lingvistika i intellektual'nye tehnologii: Po materialam konferencii Dialog 2017*, Izd-vo RGGU, Moscow 2017.

Resources (last access 2 January 2021)

- Speaker 1. <https://www.youtube.com/watch?v=DkPeJpj8RoU>
- Speaker 2. https://www.youtube.com/watch?v=4JBzhbh_85I
- Speaker 3. <https://www.youtube.com/watch?v=Kz41wp7dwss>
- Speaker 4. <https://www.youtube.com/watch?v=BOWXrgQiaVU>
- Speaker 5. https://www.youtube.com/watch?v=ZE_vqIF7ooc
- Speaker 6. <https://www.youtube.com/watch?v=52mrcmZ66Is&t>

PAVEL DURYAGIN
Ca' Foscari University of Venice

PARTE II

LESSICO E MORFOLOGIA

