

## Perceived naturalness of accents in noun phrases by mono- and bilingual listeners of Russian

Sabine Zerbian<sup>1</sup> & Kristina Barabashova<sup>2</sup>

University of Stuttgart<sup>1</sup>, Saarland University<sup>2</sup>

[sabine.zerbian@ifla.uni-stuttgart.de](mailto:sabine.zerbian@ifla.uni-stuttgart.de)<sup>1</sup>, [kristina.barabashova1@gmail.com](mailto:kristina.barabashova1@gmail.com)<sup>2</sup>

### ABSTRACT

We tested the acceptability of different accent patterns in adjective+noun phrases with contrastive focus on the adjective ( $A_{CF+N}$ ) by mono- and bilingual listeners of Russian in a perception experiment. Both naturally produced as well as lab-recorded stimuli were presented auditorily, embedded in a discourse context which evokes semantic contrastive focus on the adjective in a noun phrase (e.g., *the GREEN<sub>CF</sub> hat*). Listeners had to evaluate how natural the presented audio stimuli sounded to them. The accent patterns varied with respect to accent on the noun, on the adjective or on both. Moreover, the noun phrases occurred as either the first or second noun phrase in the sentence.

Results show that both mono- and bilingual listeners show a sensitivity to prosody in specific information structural contexts in lab-recorded speech. In natural speech and concerning the position of occurrence, only the bilingual listeners show a preference. Bilingual listeners' more stable judgments across phonetically varied stimuli might be attributed to linguistic experience.

**Keywords:** accent, perception, Russian, adjective Heritage Russian, noun phrases, focus.

### 1. INTRODUCTION

In Russian, both word order and pitch accents can be used to indicate information structure in a sentence. Specifically, a pitch accent on a constituent ( $H^*L$ ; indicated by capital letters) can indicate narrow focus on this constituent, as in (1a vs 1b) (Jasinskaja 2014).

- (1) Prosodic expression of focus in Russian
- (a) Marina slušala MUZYKU.
  - (b) Marina SLUŠALA muzyku.  
'Marina listened to music.'

Within a noun phrase, the location of a pitch accent can also indicate narrow focus (Jasinskaja 2014). For example, a pitch accent on the adjective, as in (2a), indicates narrow contrastive focus on this constituent (word order in split constructions as in (2b) is another option) (Sekerina & Trueswell 2011:282).

- (2) Narrow focus in noun phrases in Russian
- a. Položite KRASnuju zvezdočku v Poziciju 4.  
put red.ACC.FEM star.ACC.FEM in Position 4
  - b. KRASnuju položite zvezdočku v Poziciju 4.  
red.ACC.FEM put star.ACC.FEM in Position 4  
'Put the RED star in the Position 4.'

Zerbian et al. (2022) investigated the prosodic realization of contrastive adjective focus in natural speech of mono- and bilingual speakers of Russian, using data from the RUEG corpus (Wiese et al. 2019). The bilingual speakers were heritage speakers of Russian. Heritage languages are "culturally or ethnolinguistically minority languages that develop in a bilingual setting where another sociopolitically majority language is spoken" (Montrul 2016:2).

For monolingual speakers of Russian, a pitch accent on a contrastively focused adjective was tentatively confirmed, at least in informal communicative situations (note though that the number of overall productions was low for this speaker group) (Zerbian et al. 2022).

The data produced by heritage speakers of Russian in the U.S. showed that next to accents on the adjective they more frequently produced double accents (i.e., both on adjective and noun) on  $A_{CF+N}$ , both across formal and informal communicative situations (Zerbian et al. 2022).

The current research therefore sets out to investigate by means of naturalness ratings how mono- and bilingual listeners of Russian perceive different accent patterns in  $A_{CF+N}$ . Both lab-recorded and naturally produced data were used as stimuli.

The study is motivated by a scarcity of perception studies concerning heritage language prosody. Whereas speech production is often tested in heritage language phonology, perception or comprehension is researched less often (Polinsky & Scontras 2019, footnote 6; but see Sekerina & Trueswell 2011, Zerbian 2012; and the short summary in Polinsky 2018: 158-162).

Prior work on the processing of prosody and word order in mono- and bilingual Russian listeners has attested only a weak effect of contrastive accent (Sekerina & Trueswell 2011) and found instead that it is relevant whether the semantic contrast is linguistically explicit or not, i.e., whether the noun phrase in question constitutes the first or the second

in an utterance (Sekerina & Trueswell 2011). Only in the latter case is the contrast linguistically explicit. Therefore, position of occurrence was also included in our perception study as a potentially relevant factor for perceived naturalness next to the accent pattern itself.

## 2. RESEARCH QUESTIONS AND HYPOTHESES

The following research questions were investigated:

- RQ1: Do mono- and bilingual listeners rate different accent patterns in  $A_{CF+N}$  differently with respect to naturalness?
- RQ2: Does the position (i.e., whether the noun phrase is mentioned first or second) lead to a difference in the ratings of accent patterns?

We expect the rating of  $A_{CF+N}$  to depend on ...

- Accent location: Adjective accents are rated high on the natural scale by both mono- and bilingual listeners and clearly different in acceptability from accents on nouns.
- Bilingualism of listeners: Mono- and bilingual listeners differ in their naturalness ratings. The prediction based on our findings in production is that double accents (i.e., on both adjective and noun) are rated as more natural more often by bilingual listeners than by monolingual ones.
- Explicit linguistic contrast: Double accents are rated as more natural if they occur on the first  $A_{CF+N}$  than on the second  $A_{CF+N}$ , because only in the latter is the contrast linguistically explicit.

## 3. METHODOLOGY

### 3.1 Method

The research questions were investigated by means of an online experiment in which listeners were asked to evaluate the naturalness of a spoken noun phrase within a given context.

The study consisted of two experimental parts. Experimental part 1 (RQ1) elicited naturalness ratings for  $A_{CF+N}$  with differing accent patterns, namely an accent on the adjective, on the noun or on both (double accent). In example (3), the  $A_{CF+N}$  in question is *синюю машину* “blue car”.

(3) (speaker code: USbi10FR\_fsR)  
*Èt-a bela-ja mašin-a vresa-l-a-s'*  
 this-F.NOM white-F.NOM car-NOM crash-PST-F-REFL  
 v *sin-ju-ju mašin-u.*  
 into blue-F.ACC car-ACC  
 ‘This white car crashed into a blue car.’

Experimental part 2 (RQ2) explored the role of explicit linguistic contrast for naturalness ratings of

$A_{CF+N}$ s as first and second occurrence in an utterance respectively (*Белая машина* and *синюю машину* in (4), in bold).

(4) (speaker code: USbi33FR\_isR)  
***Bela-ja mašin-a rezko ostanovi-l-a-s'***  
 white-F.NOM car-NOM abruptly stop-PST-F-REFL  
*i vresa-l-a-s'* v ***sin-ju-ju mašin-u.***  
 and crash-PST-F-REFL into blue-F.ACC car-ACC  
 ‘The white car stopped and crashed into the blue car.’

### 3.2 Procedure

The experiment was implemented in Gorilla®. The participants accessed the experiment online via a link and were presented introductory information on the experiment, including informed consent. In a training session the participants got acquainted with the task. Participants saw a written sentence (in Cyrillic) on the screen in which an adjective followed by a noun was marked in bold and italics. A drawing showing the object described in the noun phrase was placed in bottom-middle position of the slide for illustration purposes. On pressing a button, participants could listen to the noun phrase in bold produced by a speaker. They could repeat listening up to 5 times. On the following slide they had to rate how natural the phrase sounded in this context by assigning a number (1 = not at all natural, 7 = very natural). After providing their rating, they were automatically presented with the next item.

In order to make listeners acquainted with the task, the training session progressed from contexts and target nouns being read out, then lab-recorded noun phrases embedded in written contexts, and lastly noun phrases taken from the corpus (thus different speakers) embedded in written contexts. Three examples each were presented during the training session.

After the training session, the experiment started. At the end, participants had to fill in a questionnaire on their linguistic background.

### 3.3 Stimuli

For both experimental parts we used original productions of  $A_{CF+N}$  from the RUEG corpus (see Zerbian et al. 2022 for details), which describe an accident involving two differently coloured cars (which implies a semantic contrastive focus on the adjective).

Only noun phrases were selected in which the adjective was a colour adjective (rather than an inherently deictic adjective such as “other, second”). In addition, only noun phrases were selected which had no hesitation or pronunciation deviation and were of acceptable or good quality (n=30; see below).

Accent on	1 <sup>st</sup> position	2 <sup>nd</sup> position	total
adjective	3	4	7
noun	2	3	5
both	8	7	15

The 27 noun phrases are from 11 different speakers (9 bilingual, 2 monolingual; 2 male, 9 female).

All noun phrases selected have been taken from sentences which contained the two A<sub>CF</sub>+N phrases that were contrasted, thus the contrast was linguistically explicit in the sentence.

The noun phrases were cut out and presented with the transcribed original contexts, leaving out repetitions, hesitations, and grammatical deviations in order not to create any interference from other linguistic features. Corpus samples were treated for noise, and equalized for intensity (loudness) and tempo, using the digital audio editor in Audacity.

Next to samples from naturally occurring speech, listeners were also presented with lab-produced recordings of A<sub>CF</sub>+N phrases, mirroring the three accent patterns on adjective, noun and both, in freely created contexts, as in (5).

(5)  
*Sin-jaja šljap-a mn-e nravit-sja bol'se*  
 blue-F.NOM hat-NOM I-DAT like-PRS.REFL more  
*čem zelën-aja.*  
 than green-F.NOM  
 'I like the blue hat more than the green hat.'

The first part contained 14 items from natural speech. In order to increase and adjust the number of data points, some items were repeated up to three times forming the equivalent number of items across three accent types (9 items in each group). In total, the first experiment comprised 27 stimuli (9 x 3 accent types).

The second part contained 22 items from natural speech. Again, to increase the number of data points, all items were repeated up to three times. In total, the second experiment part included 36 items (9 x 2 accent types x 2 positions).

In addition, 30 lab-recorded stimuli (produced by one single female speaker) were included in the first part (5 x 2 context structure x 3 accent patterns). The second part comprised 20 lab-recorded stimuli (5 x 2 accent patterns x 2 positions).

Altogether, each participant gave 113 naturalness judgements (27 + 36 + 30 + 20).

### 3.3 Participants

Participants were recruited via the platforms Prolific and Mechanical Turk. Requirements specified via the platform included U.S. or Germany as location,

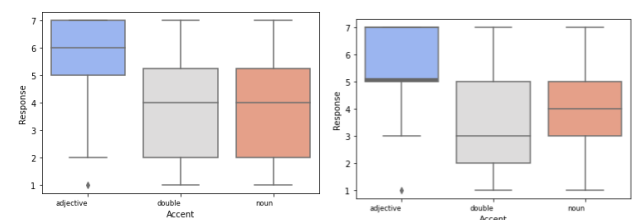
Russian as L1, Russian and English or German as fluent languages. Altogether, we recruited 70 participants. Of those, 25 had to be excluded due to not matching the linguistic requirements needed for the study (e.g., using Russian as a language at home). Of the remaining 45 participants, 16 were raised monolingually, 29 were bilingual. In the following, we concentrate on those 12 participants which we consider heritage speakers as they were either born to Russian speaking parents in the U.S. or Germany or immigrated to the respective countries when they were younger than 5 years (17 participants immigrated at > 20 years of age).

Among heritage Russian listeners, four participants belong to the age group 26-35 years, and eight participants were 18-25 years. Monolingual listeners included seven participants 18-25 years, seven participants 26-35 years, and two participants 51-61 years. Heritage Russian listeners specified the usage of English (6), German (2), and Ukrainian (1) at home. For other purposes, such as work/study and communication with friends, English was selected by 10 participants, German by five participants, and Russian by one participant. Six heritage speakers reported that Russian was a part of their compulsory or additional education. Among monolingual speakers, one participant noted the usage of Chuvash at home. For the purposes of work/study and communication with friends, monolinguals use English (8) and German (2). For 14 monolinguals, Russian was a part of the compulsory or additional education.

## 4. RESULTS

Fig. 1 shows the naturalness ratings for different accent patterns in the context of contrastive adjective focus for the lab-recorded stimuli by bilingual heritage and monolingual listeners.

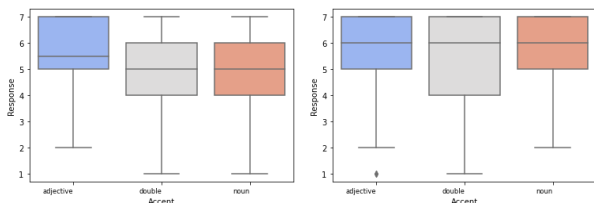
The y-axis shows the assigned values (1= not natural, 7= very natural) to the different accent locations on the x-axis (accent on adjective left/blue, double accent middle/grey, accent on noun right/red). Boxes show the quartiles of the data, whiskers show the rest of the distribution, points show outliers. Horizontal lines in the boxes refer to the median (Waskom 2021).



**Figure 1:** Ratings for lab-recorded stimuli by heritage listeners (left) and monolingual listeners (right)

In the context of contrastive adjective focus, both speaker groups rate an accent on the adjective (heritage: mean=5.4, mono: mean=6.0) as clearly more natural than a double accent (heritage: mean=4.0, mono: mean=3.6) or an accent on the noun (heritage: mean=3.9, mono: mean=4.2).

Figure 2 shows the ratings for the stimuli taken from the corpus.

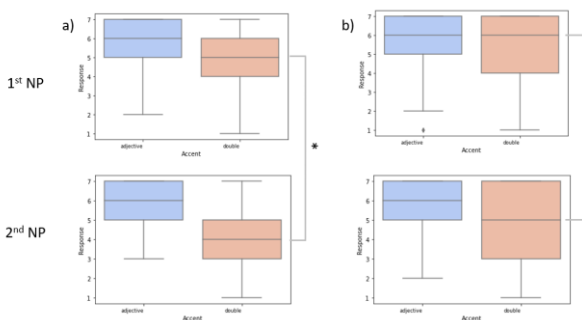


**Figure 2:** Ratings for natural stimuli by heritage listeners (left) and monolingual listeners (right)

For the heritage listeners, accents on adjectives (mean=5.3) are rated higher on average than double accents (mean=4.9) or accents on nouns (mean=4.8), though less clearly so than in the lab-recorded stimuli.

For monolingual listeners, the difference in naturalness ratings depending on the different accent locations disappeared. All accent patterns are rated as similarly highly natural.

Figures 3 and 4 shows the naturalness ratings depending on whether the noun phrase occurred first or second in a sentence. Figure 3 shows the lab-recorded stimuli, Figure 4 the natural stimuli from the corpus.

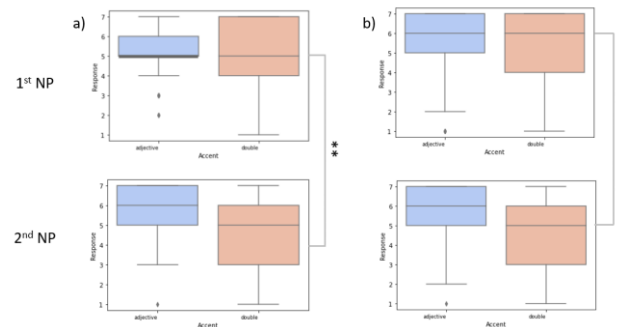


**Figure 3:** Ratings for pre-recorded stimuli by heritage listeners (left) and monolingual listeners (right)

For both monolingual and heritage listeners, accents on adjectives are judged more natural than double accents both in 1<sup>st</sup> and 2<sup>nd</sup> position ( $p < 0.001$ ). However, heritage listeners rate double accents significantly more natural in 1<sup>st</sup> position (mean=4.7) than in 2<sup>nd</sup> (mean=3.9;  $p < 0.05$ ). For monolingual listeners, the position does not matter ( $p > 0.05$ ).

The results for natural stimuli are shown in figure 4. In 1<sup>st</sup> position, double accents (mean=5.3) are as natural as adjective accents (mean=5.3) for

heritage listeners ( $p > 0.05$ ). However, adjective accents (mean=5.8) are still more natural than double accents (mean=5.2) for monolingual listeners ( $p < 0.05$ ). In 2<sup>nd</sup> position, adjective accents are clearly judged more natural than double accents by both groups ( $p < 0.001$ ). Again, double accents are rated higher in 1<sup>st</sup> position (mean=5.3) than in 2<sup>nd</sup> (mean=4.5) by heritage listeners only ( $p < 0.001$ ).



**Figure 4:** Ratings for natural stimuli by heritage listeners (left) and monolingual listeners (right)

## 5. DISCUSSION

The results of the lab-recorded stimuli are in line with the literature and our expectations for both listener groups, namely that adjective accent is judged significantly more natural than an accent on the noun or double accent. Thus, both listener groups have shown to be sensitive to accent and its relation to information structure in Russian.

For the stimuli taken from natural productions, all accent patterns sounded equally (highly) natural, at least to monolingual listeners.

The actual naturalness of the data (including noise, lower audio quality, faster speech, less clear rendition, high speaker variability including age, gender, linguistic biographies possibly including traces of accents) might have indirectly led to convergence of judgments in monolingual speakers. This might be even more so due to the generally weak load of accent in Russian (Sekerina & Trueswell 2010). Bilingual heritage listeners, on the other hand, might benefit from a perceptual advantage that leads to ‘better-than-native perception’ (e.g., Chang 2016). Heritage speakers have more familiarity with variation and accents and might thus adapt more quickly to varied speech.

Last but not least, the majority languages of the heritage listeners were either English or German, two languages which also use variable pitch accent placement for focus in noun phrases. This shared feature in both their languages might have increased the sensitivity to a feature that otherwise has a weaker load in Russian. However, such a confirmatory effect, if it exists, is not equally visible in production.

## 7. REFERENCES

- [1] Jasinskaja, K. 2014. Information structure in Slavic. In: Féry, C., Ishihara, S. (eds), *The Oxford Handbook of Information Structure*, 709-732.
- [2] Wiese, H., Alexiadou, A., Allen, S., Bunk, O., Gagarina, N., Iefremenko, K., Jahns, E., Klotz, M., Krause, T., Labrenz, A., Lüdeling, A., Martynova, M., Neuhaus, K., Pashkova, T., Rizou, V., Tracy, R., Schroeder, C., Szucsich, L., Tsehay, W., Zerbian, S., Zuban, Y. 2019. RUEG Corpus (Version 0.2.0) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.3236069>
- [3] Zerbian, S., Böttcher, M., Zuban, Y. 2022. Prosody of contrastive adjectives in mono- and bilingual speakers of English and Russian: a corpus study. *Proceedings of Speech Prosody 2022*, Lisbon.
- [4] Montrul, S. 2016. *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.
- [5] Polinsky, M., Scontras, G. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition*, 23(1), 4-20. doi:10.1017/S1366728919000245
- [6] Sekerina, I., Trueswell, J. 2011. Processing of contrastiveness by heritage Russian bilinguals. *Bilingualism: Language and Cognition* 14(3): 280 – 300. DOI: 10.1017/S1366728910000337
- [7] Zerbian, S. 2012. Perception and interpretation of intonational prominence in varieties of South African English. In: Braunmüller, K., Gabriel, C. (eds), *Multilingual Individuals and Multilingual Societies*. Benjamins, Amsterdam. 335-348.
- [8] Polinsky, M. 2018. *Heritage Languages and Their Speakers*. Cambridge: Cambridge University Press.
- [9] Waskom, M. 2021. Seaborn: Statistical Data Visualization. *Journal of Open Source Software*, 6(60). <https://doi.org/10.21105/joss.03021>
- [10] Chang, C. 2016. Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition*, 19(4), 791-809. doi:10.1017/S1366728914000261