



On some factors affecting the choice of tune in Russian wh-questions

Pavel Duryagin

Ca' Foscari University of Venice, Italy

pavel.duryagin@unive.it

Abstract

The results of a production experiment that investigated prosodic variability in Russian information-seeking wh-questions are reported. Three wh-questions in four focus conditions, with and without initial particle *a*, were elicited in a reading task from 20 native speakers of Russian. The data generally corroborate prior descriptions and demonstrate that a large inventory of tunes is used by Russian speakers in wh-questions. Namely, several patterns with one or two “falling” pitch accents (downstepped and non-downstepped) can be recognized in the data, as well as one “rising” pattern containing a high edge tone. Preliminary phonological analysis is proposed for these tunes.

The effects of two factors on the choice of the “nuclear pitch accent + edge tone” configuration (“falling” H*+L L-% vs. “rising” L* H-%) were tested statistically. The results demonstrate that contrastive focus condition restricts the use of the “rising” pattern while the presence of phrase-initial particle *a* has an opposite, but weaker effect on the choice of tune.

Index Terms: intonation, wh-questions, contrastive focus, Russian

1. Introduction

While Russian yes/no questions have received considerable attention in experimental research within an autosegmental metrical framework (AM) over the past two decades [1, 2, 3], Russian wh-questions remain largely understudied. However, the variability of tunes in information-seeking wh-questions has been attested since the earliest analyses of Russian intonation. Namely, the most widespread traditional description, Bryzgunova’s set of holistic intonational constructions, generally characterized a typical unmarked pattern for this phrase type in Russian as a high fall located on a question word followed by a low deaccented post-wh-word sequence [4]. However, Bryzgunova identified other possible prosodic realizations of Russian wh-questions, namely, a fall-rise and a flat hat pattern. Rather vaguely, the scope of use of the fall-rise pattern was characterized as “questions... in contrastive relationship” with the preceding phrase, usually an element in a series of consecutive questions. It was further stated that wh-questions with a fall-rise tune when taken out of context are perceived as containing “a shade of discontent or edification”. Similar nuances of meaning (“impatience, annoyance”) were postulated for “flat hat” patterns in wh-questions [4, pp. 114–115].

Several Russian researchers proposed different accounts on this topic. Svetozarova characterized the high fall associated with a wh-word as a “rarer type of intonation” [5]. In her description wh-questions are typically characterized by a “flat hat” contour with a rise on the initial question word and

a fall on the final word of the phrase. Similarly, Kodzasov defined the flat-hat contour with medial declination as unmarked [6]. As an alternative, a prenuclear fall followed by a nuclear downstepped fall (“a pair of falling tones connected by a level tone”, in his terms) was listed as an unmarked option. Instead, only two typical contexts were proposed for the use of a nuclear fall on a wh-word with rest of the phrase deaccented: 1) when the post-wh-word sequence has been actualized in the preceding phrase; 2) in an “exam question” (apparently, formal insistent question). In his later work [7], Kodzasov included the relations between interlocutors as a factor affecting the choice of tune: according to his data, a speaker in a position of domination tends to produce a falling F0 movement on a question word; otherwise, a flat hat pattern is chosen. Similarly to Bryzgunova, he reserved the use of “fall-rise” pattern only for a sequence of questions.

Considerable controversy in impressionistic descriptions of Russian wh-question intonation indicates the necessity of experimental research on the topic. However, prior laboratory studies did not pay special attention to the apparent variation in Russian wh-question tunes. Odé’s IPO-style analysis of Russian intonation [8] does not deal with pragmatic interpretations of the tunes and therefore does not discuss wh-questions in detail. However, more recent revision of her model states that the high fall located on a question word (HL* accent, in Odé’s notation) “sounds a bit impolite and impatient” [9]. In accordance with preceding studies, Odé states that the same pitch accent normally marks narrow focus on content words in declaratives.

Finally, Igarashi’s study of Russian wh-questions [10] recognizes the distinction between a “peak pattern” and a “hat pattern” but chooses to concentrate only on the patterns with a fall on the accented wh-word and deaccented post-wh-word sequence. The researcher lists the following distinctive features of this wh-question pattern (as opposed to yes/no questions): 1) non-late peak alignment; 2) lower peak scaling; 3) “absence of the low turning point at the onset of the accented syllable”. Two possible AM interpretations of wh-question nuclear accent are therefore proposed: H* or H*+L.

Despite the apparent discrepancies in various descriptions and the lack of experimental evidence, the variability of tunes in Russian wh-question is generally acknowledged. Such a variability is not uncommon cross-linguistically. Since the interrogative function is marked lexically by the wh-word, prosodic instruments tend to be utilized by speakers to convey various types of information. It has been shown for other languages that the choice of tune in wh-questions can be affected by information structure [11, 12], degrees of politeness [13], expected answer [14], speaking style (read vs. spontaneous speech) [15]. To my knowledge, no comparable experimental studies on Russian were conducted.

Within the scope of the present study the effects of two linguistic factors were tested. First, the role of information

structure was examined. Namely, in the present paper the results for two focus conditions are reported: unmarked broad focus and contrastive focus on a phrase final verb. Previous studies in Russian prosody indicate that the difference between these two focus conditions can be potentially neutralized prosodically since the same nuclear accent can be associated with phrase-final words in both contexts.

Second, analysis of wh-questions extracted from the largest spoken corpus of Russian available, the Multimedia Russian corpus [16], indicates a potential effect of phrase initial word *a* on the choice of tune. Recent extensive semantic analysis of Russian phrase initial *a* [17] defines it as a polysemic “discursive connector” that has three main functions (contrast, inconsistency, addition). The use of this particle is obligatory in Russian elliptical questions and optional in all other question types. However, corpus data demonstrate that *a* is commonly used in wh-questions where it is frequently accompanied by the “fall-rise” tune. I therefore intend to verify this observation experimentally.

The aim of the study is thus twofold. First, it aims to collect a set of data that will make it possible to confirm experimentally the variability of tunes in Russian wh-questions and describe the tunes in terms of accent placement, pitch scaling and alignment. Second, the present paper aims to extend current knowledge of the factors that can affect the choice of tune by testing the effects of information structure (broad focus vs. contrastive focus on the predicate) and the effect of phrase-initial particle *a*.

2. Methods

2.1. Participants

20 native Russian speakers took part in the experiment (14 F, 6 M, mean age 26.5, $\sigma = 5.7$). All participants speak Standard Russian (as judged by the experimenter) and reside in Moscow (except for one speaker who has recently moved from Moscow but continues using Russian on a daily basis). None of them reported having any speaking or hearing impairments.

2.2. Materials and procedure

In order to elicit the pronunciation of wh-questions, a set of dialogues was composed. Each dialogue was read aloud by a participant together with the experimenter. During the reading task the participants had the opportunity to familiarize themselves with each dialogue before reenacting it.

Each scripted dialogue/stimulus contained four or five utterances: 1) one phrase or two phrases uttered by the participant (an “out of nowhere” question; a statement + a question); 2) a reaction of the experimenter; 3) a statement of the participant + a wh-question; 4) a second reaction of the experimenter (one or two declaratives); 5) optional concluding statement of the participant.

Three Russian question words were selected for the study: [ˈgdʲe] ‘where’, [kɐgˈda] ‘when’ and [pətʃiˈmu] ‘why’. These words have 1, 2 and 3 syllables respectively and ultimate syllable stress. For each question word, comparable dialogues were composed with wh-questions in four different focus conditions: broad focus (BF), contrastive focus on the predicate (verb) in phrase-final position (CFP), contrastive focus on the subject (noun or pronoun) in phrase-medial position (CFS) and a short question containing only the wh-word (SQ). Only the results for BF and CFP are reported here.

Also, each dialogue was elicited with and without a phrase-initial particle *a*. The participants had been warned before the experiment that some of the dialogues would occur more than once. Consequently, the speakers thought that they were reading some dialogues twice, though in fact the stimuli never repeated and actually varied in the presence / absence of the phrase-initial particle *a*. As a result, during the reading task none of the participants became aware of the manipulations with the particle.

In total, 24 dialogues containing experimental stimuli were composed, with 3 wh-words (‘where’, ‘when’, ‘why’) in 4 focus conditions (BF, CFP, CFS, SQ) and 2 particle conditions (presence and absence of *a*). All stimuli together with 26 fillers (dialogues similar in structure to the experimental ones) were presented to the participants in 4 different orders. Similar dialogues were consistently separated by several unrelated contexts.

Only the results for the first two focus conditions (BF and CFP) with and without *a* are reported in the present paper, which makes a total of 240 tokens elicited. See below an example of a broad focus dialogue translated in English.

[Participant]: *Where is your brother? Isn't he going to the cinema with us?* [Experimenter]: *I don't know. He's probably at home, busy with his computer as always.* [P.]: *It will be boring without him. Why didn't you call him?* Russian: *‘(A) почему вы ему не позвонили?’* [E.]: *If you want you can call him yourself. I can give you his phone number.* [P.]: *Sure, I'll write it down.*

2.3. Annotation

The elicited wh-questions were segmented and acoustically annotated in Praat [18]. Some tokens (8 out of 240) were discarded due to disfluencies and reading mistakes that were not corrected by the experimenter during the experimental session. For each phrase, pitch-related and segmental landmarks were labelled: relevant F0 maxima and minima, as well as word and syllable boundaries.

3. Results

3.1. Acoustic analysis and preliminary AM interpretation of the tunes

In accordance with prior descriptions, analysis of the wh-questions elicited during the reading task demonstrates the variability of tunes. According to my analysis, at least five distinct tunes can be recognized in the data. In this section I will introduce each pattern with a brief description of its phonetic implementation and a discussion of its possible AM interpretation. However, it should be outlined that since no conventional AM description of Russian intonation exists at the moment and the data of the present experiment are limited, only tentative phonological analysis can be proposed here.

3.1.1. Accented wh-word followed by a deaccented post-wh-word sequence

The data do not support existing claims that the pattern characterized by a high fall on the wh-word and post-wh-word deaccentuation is stylistically marked in Russian. This tune was commonly used by most of the speakers in broad focus contexts. Phonetically, the pattern is characterized by a high F0 peak achieved early on the wh-word stressed syllable followed by a gradual falling F0 movement located at the

stressed syllable and several unstressed syllables (if present). The exact timing of the falling movement is presumably influenced by the segmental material and may be indicative of the edge tone association in Russian. However, the phonetic details of this association are yet to be studied.

This tune has been analyzed in [10] as containing a H*+L nuclear accent followed by L- phrase tone and L% boundary tone. I adopt Igarashi's analysis with a minor adjustment. Since there is no experimental data demonstrating that Russian has phonological pitch movements specifically aligned at the phrase-final syllable, Rathcke in [19] proposes using a sole L% label for low edge tones in Russian. The lack of L- (and H-) labels in her analysis aims to demonstrate the lack of empirical evidence for the necessity to introduce phrase tone in AM description of Russian. However, I tend to adopt a different option outlined in Rathcke's analysis and use the combined L-% and H-% labels, similar to the ones adopted in GToBI [20]. In my view, the use of T-% label for edge tones better reflects the observation that Russian edge tones are associated not with the last syllable of the phrase, but rather with the sequence of syllables that follow the last pitch accent (see, for example, the low stretch in Russian polar question tune that follows the nuclear L*+H).

Therefore, H*+L L-% is used as a preliminary transcription for the wh-question tune introduced in the present section (see Fig. 1; all illustrations were created with Praat script [21]).

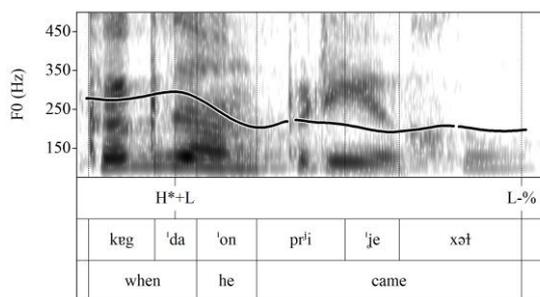


Figure 1: *F0 contour, spectrogram and annotation of a phrase 'Когда он приехал?' ('When did he come?'), BF condition, as produced by speaker ES.*

3.1.2. Accented wh-word followed by a downstepped fall

While the pattern with deaccented post-wh-word sequence is not uncommon in the data, patterns with nuclear prominence assigned to the final word (always a verb in our set of stimuli) were more frequent under both focus conditions.

First, a pattern with two falling F0 movements can be recognized in the data: a fall on the wh-word (phonetically similar to the one described in the previous section) followed by a second fall which starts early in the stressed syllable of the phrase-final word. It is worth noting that the F0 maximum is regularly located on the first pre-stressed syllable; this typical "set-up" movement had been previously reported for pitch falls in Russian, see, e. g., [8, pp. 101–102]. The second peak has lower scaling compared to the first one (see Fig. 2) which makes it necessary to introduce the notion of downstep in the analysis of Russian wh-questions. I use the commonly adopted downstep label "!" and analyze this tune as H*+L !H*+L L-%. However, precise phonological interpretation of downstep in Russian requires a separate investigation and falls out of the scope of the present study.

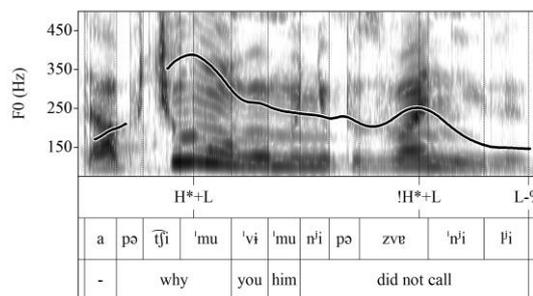


Figure 2: *F0, spectrogram and annotation of a phrase 'А почему вы ему не позвонили?' ('Why didn't you call him?'), BF condition, as produced by speaker EK.*

3.1.3. Accented wh-word followed by a non-downstepped fall

In accordance with previous descriptions, the "hat pattern" was present in the data in both focus conditions. In this tune, the F0 maximum is achieved late in the stressed syllable of the wh-word or after it and is followed by a declining high plateau. The plateau ends with a fall that starts at the pre-stressed syllable of the final word (see Fig. 3). I analyze this pattern as a sequence of a prenuclear H* accent followed by a non-downstepped H*+L accent and a L-% edge tone.

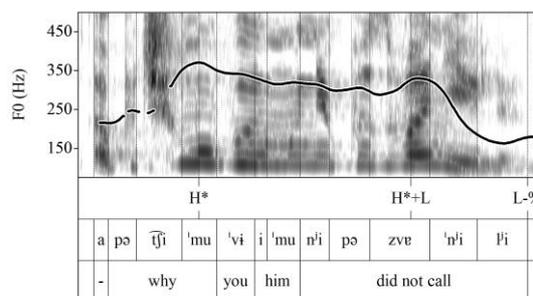


Figure 3: *F0, spectrogram and annotation of a phrase 'А почему вы ему не позвонили?' ('Why didn't you call him?'), BF condition, as produced by speaker NC.*

3.1.4. Deaccented wh-word followed by a fall on the phrase-final verb

According to the data, a contrastive focus condition enables optional wh-word deaccentuation. F0 starts at the default medium level and gradually rises to the stressed syllable of the word in focus (phrase-final verb). I analyze this pattern similarly to the one described in section 3.1.1, as H*+L L-%, but with nuclear accent located differently (see Fig.4).

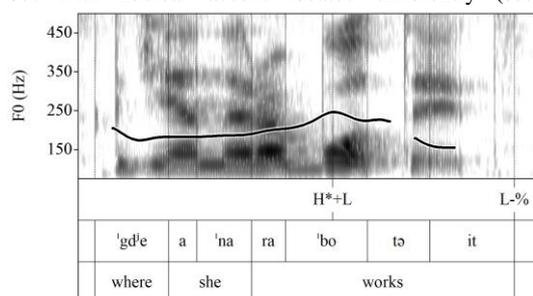


Figure 4: *F0, spectrogram and annotation of a phrase 'Где она [работает]?' ('Where does she [work]?'), CFP condition, as produced by speaker IK.*

3.1.5. Accented wh-word followed by a low plateau and an edge rise

Finally, the “fall-rise” pattern was common in the data. To my knowledge, this tune has not yet been interpreted in AM analyses of Russian. However, the present data and existing traditional descriptions (see abundant literature on “fourth intonational construction”, or IK-4, within Bryzgunova’s model) allow us to propose a preliminary AM interpretation.

Phonetically, in this tune F0 starts at the default medium or low level, then a high target is reached late in the stressed syllable of the wh-word. The second prominent syllable is associated with a low plateau. The plateau ends with a rise that starts late in the stressed syllable or immediately after it. As soon as the high target is reached, level high pitch is maintained. It is important to note again that the edge tone is not associated with the phrase-final syllable, but with the whole postnuclear sequence. This pattern will be analyzed here as H* L* H-% (or L+H* L* H-%); see Fig. 5.

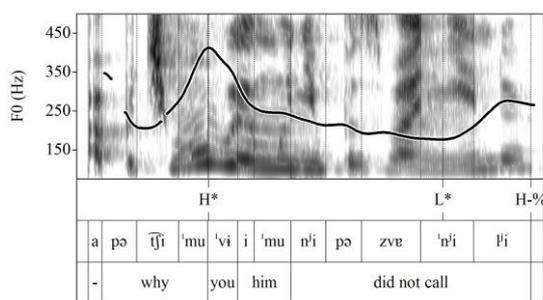


Figure 5: F0, spectrogram and annotation of a phrase ‘А почему вы ему не позвонили?’ (‘Why didn’t you call him?’), BF condition, as produced by speaker KV.

3.2. Factors affecting the choice of tune: statistical analysis

It should be noted that categorization of tunes listed in sections 3.1.1–3.1.4 is far from being straightforward. Several known problems arise when classifying these tunes, e. g., the criteria for the distinction between downstepped and non-downstepped tunes are not clear, as well as the uniform strategy for labelling prominent words. As a result, it is not always possible to assign a tune to a particular phonological category. The phonetic implementation is clearly affected by segmental material and several other factors not manipulated in the experiment. However, in my analysis these patterns differ only in prenuclear accent type and in nuclear accent localization, but not in the configuration of nuclear accent and edge tone which is always H*+L L-%. I therefore assume that these sequences of accents can be treated as constituting a single category of “falling” tunes.

On the contrary, the H* L* H-% tune (see section 3.1.5), is clearly distinguishable from the “falling” group as it contains a distinctive rising F0 movement associated with the postnuclear syllable sequence. It is therefore possible to model the choice of tune at least at the level of binary decision. To test the effect of two experimental conditions, a mixed-effects logistic regression model was fitted by means of *lme4* R package [22, 23] with tune category as a dependent variable (“rising” vs. “falling”); focus condition, particle presence, wh-word as fixed effects and a random intercept for the participant.

The results of regression modelling are presented in Table 1. Strong effects of focus condition and the presence of particle *a* were found: the stimuli were significantly more likely to be produced with “rising” tunes in broad focus condition and in the presence of phrase-initial particle. Also, the effect of wh-word was significant, but only for when-questions. The latter effect was unexpected and can be explained by lexical givenness of the verb attested only in broad focus dialogues with “when”. A possible effect of this feature was not considered in the preparation of stimuli.

Table 1: The model output for binary regression analysis of the tune choice.

Predictors	Odds Ratios	CI	P
(Intercept)	0.00	0.00 – 0.03	<0.001
particle: yes	5.95	2.22 – 15.91	<0.001
focus: broad	167.41	30.18 – 928.51	<0.001
word: when	0.21	0.06 – 0.65	0.007
word: why	0.42	0.14 – 1.25	0.119
N of speakers	20	AIC	165.503
σ^2	3.29	Marginal R ²	0.556
Observations	232	Conditional R ²	0.768

4. Discussion and conclusions

The present experiment aimed to examine variability of Russian wh-question tunes that is reported in literature and test the effect of two linguistic factors on the choice of tune. The study broadly supports earlier observations that wh-question intonation in Russian varies to a high degree.

On one hand, a group of tunes can be recognized with the F0 fall located on the last prominent stressed syllable (which can be the stressed syllable of the phrase-initial wh-word). Following [10], I analyze this tune as containing a H*+L nuclear accent followed by a L-% edge tone. The data shows that the choice of nuclear accent placement, as well as the presence and configuration of prenuclear accents in these tunes is affected by several factors that are yet to be studied.

On the other hand, the pattern with a low plateau associated with the last prominent syllable that is followed by final rising was common in the data. I preliminarily analyze this pattern as H* L* H-%. Contrary to several existing descriptions, no apparent evidence was found for pragmatic markedness of this tune in wh-questions.

Statistical analysis of the effect of two factors that were manipulated in experimental design on the choice of nuclear accent showed that the choice of tune is affected both by information structure and by the presence of phrase-initial particle *a*. It can be inferred that while in broad focus contexts the whole inventory of tunes is available for the speaker (except for wh-word deaccentuation), contrastive focus imposes restrictions on the choice of tune in favor of the falling nuclear accent associated with the word in focus.

The analysis also indicates that in broad focus contexts the presence of particle *a* in phrase-initial position increases the likelihood of production of H* L* H-% tune. It is hardly a coincidence that the same tonal sequence combined with particle *a* is obligatory in Russian elliptical questions. Whether this effect is due to some sort of diachronic expansion of this tune or it has to do with common pragmatics, remains an open question. Understanding the exact mechanisms of this interplay requires further investigation.

5. References

- [1] T. V. Rathcke, "A perceptual study on Russian questions and statements," *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, vol. 37, pp. 51–62, 2006.
- [2] Y. Igarashi, "Intonational Patterns in Russian Interrogatives — Phonetic Analyses and Phonological Interpretations," in *Prosody and Syntax: Cross-linguistic Perspectives*, Y. Kawaguchi, I. Fonagy, and T. Moriguchi, Eds. Amsterdam: John Benjamins, 2006, pp. 175–196.
- [3] V. Makarova, "The Effect of Pitch Peak Alignment on Sentence Type Identification in Russian," *Language and Speech*, vol. 50, no. 3, pp. 385–422, 2007.
- [4] E. A. Bryzgunova, "Intonacija," in *Russkaja grammatika*, vol. 1, Moscow: Nauka, 1980, pp. 96–122.
- [5] N. Svetozarova, "Intonation in Russian," in *Intonation Systems: A Survey of Twenty Languages*, D. Hirst and A. Di Cristo, Eds. Cambridge: Cambridge University Press, 1998, pp. 264–277.
- [6] S. V. Kodzasov, "Kombinatornaja model' frazovoj prosodii," in *Prosodičeskij stroj russkoj reči*, T. M. Nikolayeva, Ed. Moscow: Institut russkogo jazyka RAN, 1994, pp. 85–123.
- [7] S. V. Kodzasov, "Fazovaja simvolika tona," in *Logičeskij analiz jazyka. Semantika načala i konca*, N. D. Arutyunova, Ed. Moscow: Indrik, 2002, pp. 310–320.
- [8] C. Odé, "Russian Intonation: A Perceptual Description, [Volume 1]," *Studies in Slavic and General Linguistics*, vol. 13, pp. 1–180, 1989.
- [9] C. Odé, Transcription of Russian Intonation: a free interactive research tool and learning module. Accessed on: Dec. 10, 2019. [Online]. Available: <http://www.fon.hum.uva.nl/tori/>
- [10] Y. Igarashi, "Russian Interrogatives and Intonational Categories," in *The Discourse Potential of Underspecified Structures*, A. Steube, Ed. Berlin, New York: Walter de Gruyter, 2008, pp. 227–270.
- [11] N. Hedberg and J. M. Sosa, "The prosody of questions in natural discourse," *Proceedings of Speech Prosody 2002*, 2002, pp. 275–278.
- [12] A. Chen, "Shaping the intonation of WH-questions: Information structure and beyond," in *Questions*, J. P. de Ruiter, Ed. Cambridge: Cambridge University Press, 2012, pp. 146–164.
- [13] M. Baltazani, S. Gryllia, and A. Arvaniti, "The intonation and pragmatics of Greek wh-questions," *Language and Speech*, pp. 1–39, 2019.
- [14] J. M. Sosa, "Wh-Questions in Spanish: Meanings and Configuration Variability," *Catalan Journal of Linguistics*, vol. 2, pp. 229–247, 2003.
- [15] J. Hirshberg "A corpus-based approach to the study of speaking style," in *Prosody: Theory and Experiment*, M. Horne, Ed. Dordrecht: Kluwer, 2000, pp. 271–311.
- [16] Mul'timedijnyj russkij korpus. Accessed on: Dec. 10, 2019. [Online]. Available: <http://ruscorpora.ru/new/search-murco.html>
- [17] A. A. Zaliznjak and I. Mikaelian, "The Russian *a*: toward an integral description," *Russian Linguistics*, vol. 42, pp. 321–344, 2018.
- [18] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 6.1.08, retrieved 5 December 2019 from <http://www.praat.org/>
- [19] T. V. Rathcke, "How Truncating Are 'Truncating Languages? Evidence from Russian and German," *Phonetica*, vol. 73, no. 3–4, pp. 194–228, 2016.
- [20] M. Grice, S. Baumann, and R. Benz Müller, "German Intonation in Autosegmental-Metrical Phonology," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, Sun-Ah Jun, Ed. Oxford: Oxford University Press, 2005, pp. 55–83.
- [21] W. Elvira García, Create pictures with tiers v.4.5. Praat script, retrieved 10 December 2019 from <http://stel.ub.edu/labfon/en/praat-scripts>
- [22] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- [23] D. Bates, M. Maechler, B. Bolker, S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.