[5] Rialland, A., 2007. Question prosody: an African perspective. *Tones and tunes*, *1*, pp. 35-64.

[6] Rialland, A., 2009. The African lax question prosody: Its realization and geographical distribution. *Lingua*, *119*(6), pp. 928-949.

[7] Chung, Y. and Rose, S., 2012. Duration and Pitch in Moro Polar Questions. In 43rd Annual Conference on African Linguistics, Tulane University.

[8] Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. Glot International, 5(9/10), 341-347.

[9] Peirce, J., Gray, J., Halchenko, Y., Britton, D., Rokem, A., & Strangman, G. (2011). PsychoPy–A Psychology Software in Python.

[10] Gandour, Jackson T. (1978). The Perception of Tone. Tone: A Linguistic Survey, pp. 41- 76, ed. by Victoria Fromkin. New York: Academic Press.

[11] Connell, B., & Ladd, D. R. (1990). Aspects of pitch realisation in Yoruba. Phonology, 7(1), 1-29.

Prosodic marking in Russian multiple wh-questions: A sentence production study

Pavel Duryagin[1] and Arthur Stepanov[2]
[1]*Ca' Foscari University of Venice, Italy*, [2]*University of Nova Gorica, Slovenia*

Although the prosody of regular constituent/*wh*-questions is increasingly often discussed in the context of theories of prominence, focus and prosodic constituency (e.g [1]), prosodic contours of multiple *wh*-questions (cf. English *Who bought what?*) received surprisingly little attention in the literature so far. Slavic languages offer an additional dimension of interest to the problem as all wh-phrases are typically fronted in the clausal left periphery forming a wh-cluster, cf. Russian (1). Investigation of prosody of multiple wh-questions in these languages is likely to bring new theoretical insights, in particular, regarding realization of units with very similar prosodic properties, limits of prosodic autonomy of wh-phrases and the degree of mapping between prosodic and syntactic boundaries (cf. [2]).

Here we report the results of a sentence production study investigating the prosody of Russian multiple wh-questions. The target stimuli included 12 multiple wh-questions with 4 Russian wh-pronouns (*kogo* 'who.Acc', *komu* 'who.Dat', *kogda* 'when', *kuda* 'where to') in all possible word orders, as well as 8 single wh-questions containing non-interrogative pronouns *jego* 'he.Acc', *jemu* 'he.Dat', *togda* 'then', *tuda* 'there' in preverbal position instead of the second wh-pronoun; see (1)-(2). 20 questions and statements of comparable length were included as fillers, for the total of 40 stimuli. Because producing multiple wh-questions is known to be associated with increased contextual demands, each sentence was preceded by a short preamble story setting up an appropriate contextual situation. Participants were asked to read the preamble to themselves and then naturally pronounce a sentence that followed the story. The stimuli were presented in a pseudo-randomized order. The data from 20 native speakers of Standard Russian (14F, mean age 27.8, $\sigma = 3.5$) were subjected to analysis.

A total number of 327 target stimuli tokens were manually labelled in Praat [3]. Within the initial two-word cluster, the presence/absence of local pitch maxima and the alignment of peaks were measured, as well as mean f0 in semitones for each of the four vowels in the initial four-syllable cluster. The mean pitch and peak alignment were measured within the vowel boundaries (not the entire syllables) to minimize microprosodic effects of non-identical consonants.

The analysis of pitch contours showed that, while the first wh-word is obligatorily assigned a pitch accent (cf. [4, 5]), in multiple questions the second wh-word can be optionally assigned a pitch accent phonetically realized as a downstepped second f0 peak (Fig. 1). In single wh-questions, however, the second word (always a non-interrogative pronoun in our data) is normally not associated with a visible f0 peak (Fig. 2). The presence of the second peak is confirmed by mean pitch measurements. Mixed-effect regression analysis showed a robust main effect of question type (single vs. multiple, $t(301) = -6.505$, $p < 0.001$) on absolute values of mean f0 measured on the stressed vowel of the second constituent. Moreover, as stylized contours in Fig. 3 indicate (see especially the values for vowel 4), even when no visible f0 peak is associated with the second wh-word, wh-pronouns are produced at a higher pitch than their non-interrogative counterparts. The effect of question type remains strong ($t(239.5) = -4.294$, $p < 0.001$) even after 63 contours with double peaks, mostly multiple wh-questions, are excluded from the regression model. In addition, both low and high boundary tones are licensed at the end of the intonational phrase independently of the presence of high target on the second word. The alignment of peaks varies to a high extent, from early to late and delayed, and correlates with the choice of the edge tone.

Our results suggest that in Russian multiple wh-questions the non-initial interrogative pronouns are regularly produced at higher pitch than their non-interrogative counterparts. This is generally in line with the class of prosodic theories that treat wh-phrases on a par with contrastively focused phrases (cf. [6]): in both cases, the increased prosodic prominence is manifested in a heightened pitch. The downstep within a wh-cluster suggests that both wh-phrases belong in a single prosodic domain. It is

yet unclear what factor(s) determine an optional visible second peak in a wh-cluster. One possibility has to do with distribution of prosodic weight across two wh-items in line with individual preferences in assessing relative contextual importance of alternatives introduced by each wh-item.

(1) Kogda kogo   razbudili?
when who-acc. woke
'Who did they wake up when?'

(2) Kogda jego   razbudili?
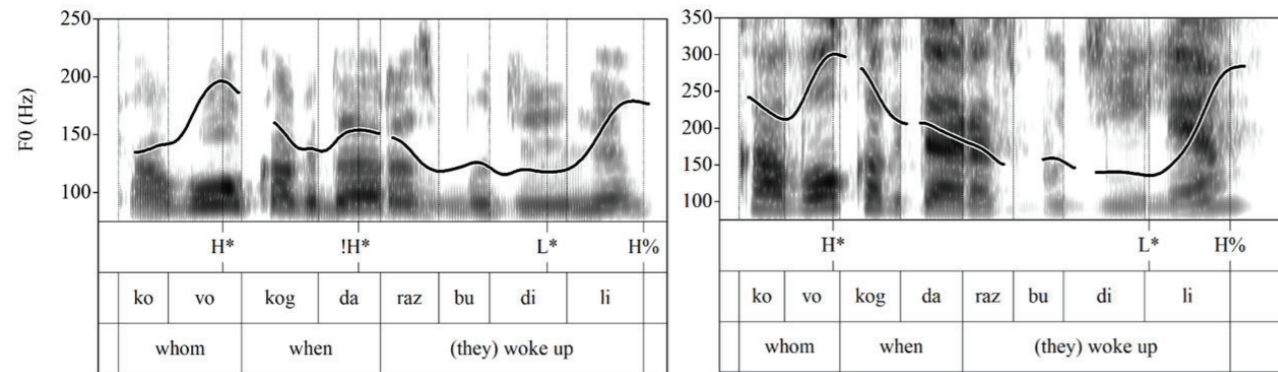*when he-acc. woke*
'When did they wake him up?'

Figure 1. *A multiple question with a visible second f0 peak pronounced by a male speaker (left) and without a visible second f0 peak pronounced by a female speaker (right).*
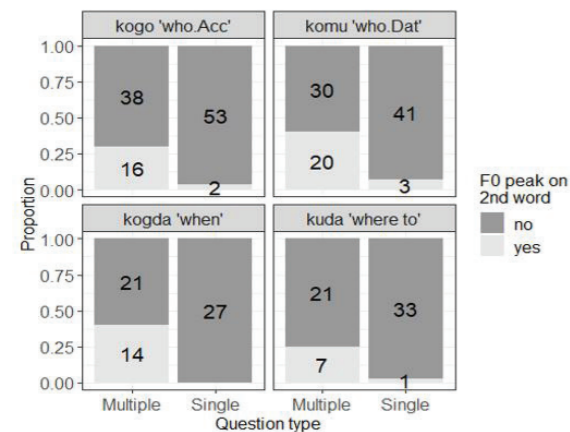

Figure 2. *Distribution of tokens with and without a visible pitch peak on second constituent, each facet of the grid represents a different initial wh-word.*
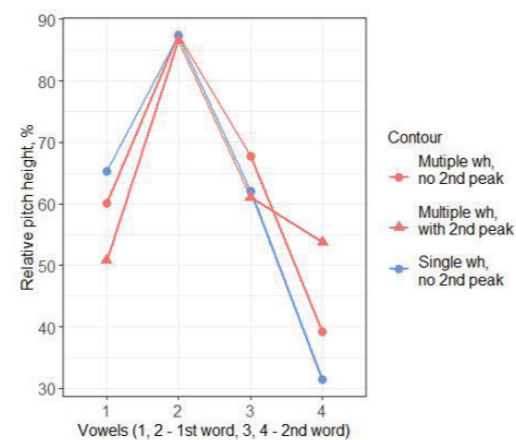

Figure 3. *Stylization of mean f0 (in %, relative to the min and max f0 in the utterance; averaged between speakers and contexts; the results for rare 2-peaked single wh-questions are omitted).*

References

[1] Ozerov, P. 2019. This Is Not an Interrogative: The Prosody of "Wh-Questions" in Hebrew and the Sources of Their Questioning and Rhetorical Interpretations. *Language Sciences* 72, 13-35.

[2] Selkirk, E. 2011. The syntax-phonology interface. In Goldsmith, J., Riggle, J. & Yu, A. (Eds.), The Handbook of Phonological Theory. 2nd edition. Oxford: Blackwell Publishing. 435-485.

[3] Boersma, P., & Weenink, D. 2020. *Praat: doing phonetics by computer*. Computer program.

[4] Bryzgunova, E. 1980. Intonation. In Shvedova, N.Y. (Ed.), *Russian Grammar I*. Moscow: Nauka, 96-122.

[5] Duryagin, P. 2020. On some factors affecting the choice of tune in Russian wh-questions. In *Proc. 10th International Conference on Speech Prosody 2020*, 235-239.

[6] Büring, D. 2016. *Intonation and Meaning*. Oxford Surveys in Semantics and Pragmatics 3. Oxford: Oxford University Press.

The prosody of information-seeking and echo *wh*-questions in English

María Biezma[1], Bettina Braun[2], Angela James[2]
[1]*Linguistics Department and Spanish and Portuguese Studies, UMass Amherst, USA*
[2]*Department of Linguistics, University of Konstanz, Germany*

Inquisitive utterances with *wh*-words may be used to plainly seek information (InfoS) or to ask about what has just been said, echo questions (EcQs), be it because the questioner didn't perceive the previous utterance properly (EcQPer) or because the speaker can't believe it (epistemic: EcQEp). This holds for utterances with a fronted *wh*-word (*and when is your brother coming?*) or final *wh*-word (*and your brother is coming when?*) ([1]). It is often claimed that the semantics of EcQs is different from that of InfoSs. The different semantics is justified by assumed (categorical) prosodic differences ([1-4]). E.g., [1,2] argue that EcQs show a complex pitch accent, L+H*, and a final rising contour H-H%. All parts of the utterance are given (and presumably deaccented) while only the *wh*-word is focused ([1-3]). There is, however, little empirical evidence supporting these claims (but see [5] for German *wh*-final). [6], on the other hand, argue that EcQ-interpretations arise mostly by discourse considerations. We present data from a study on English wh-questions, investigating prosodic differences across conditions (InfoS/EcQPer/EcQEp) in two *wh*-word orders. This is a first step towards deciding between semantic-formal analyses (which rely on cues that justify a different semantics) and discourse-based proposals (in which echoicity is derived discursively).

We manipulated the position of the *wh*-word (fronted vs. final, between-subjects in two experiments) and illocution type (EcQPer, EcQEp, InfoS, within-subjects) on 24 *wh*-utterances and three contexts (one for each condition). For each word order, the sentences had the same sentence structure, number of syllables and stress patterns. We tested 12 native English participants for each experiment (between 18 and 35 years, 12 female, 12 male). In each trial the context and target interrogative were shown on screen for as long as the participants needed. Participants then said the target sentence out loud and proceeded to the next trial. They were asked to utter the sentence in the most natural way given the context. The productions were annotated at the word level by one of the authors. F0-tracking errors were manually corrected, and the f0-values were processed using ProsodyPro ([7]). The continuous analysis of f0-contours revealed differences across conditions (more in fronted wh-words, Fig. 1), but these cannot be linked to pitch accents on content words. Half of the utterances were annotated using Mae-ToBI ([8]). To assess the reliability of the prosodic annotation, another author annotated 20% of the productions (accuracy 76%, Cohen's kappa 72%). Furthermore, word durations and the f0-range of accented words were extracted. ToBI labels were analyzed using a logistic hierarchical regression model, durations and f0-excursions with linear-mixed effects regression models; both with participants and items as crossed random intercepts ([9]); p-values were adjusted using the Benjamini-Hochberg correction, [10]). Results showed that *wh*-words were mostly rising (L* H-H% in final position: 88%, 90%, and 65% in EcQEp, EcQPer, and InfoS, respectively; L*+H in fronted position: 90%, 88%, and 55%), significantly more so in the two echo-conditions than the InfoS condition, see Fig (2a). The noun was typically accented in both word orders, with no differences across conditions (mostly L*+H when the *wh*-word was in final position and H* when the *wh*-word was fronted). The f0-range in the *wh*-word differed between EcQEp and EcQPer and between EcQEp and InfoS, but not between EcQPer and InfoS ($p > 0.5$), see Fig (2b). The same pattern held for the duration of the *wh*-word.

In sum, there are a number of phonological and phonetic differences across conditions: the *wh*-word was more often rising in echo than information-seeking questions, but there do not seem to be prosodic differences that allow us to make categorical differences between