Prosodic characteristics of canonical and non-canonical questions in Estonian

Heete Sahkai¹, Eva Liina Asu², Pärtel Lippus²

¹Institute of the Estonian Language, Tallinn, Estonia
²Institute of Estonian and General Linguistics, University of Tartu, Estonia

Abstract

This paper presents a comparison of the prosodic characteristics of canonical questions with two types of non-canonical interrogative utterances in Estonian. The data consisted of string-identical interrogative sentences with the question-word kuidas (‘how’) elicited in three readings: information-seeking question (ISQ), rhetorical question (RQ) and surprise question (SQ). A three-way distinction between the three utterance types emerged. First, there was a binary distinction between canonical and non-canonical questions in mean pitch, utterance duration and voice quality: non-canonical questions were characterised by lower mean pitch, longer duration and a larger proportion of non-modal (creaky) voice quality. Second, there was a three-way distinction in pitch range: ISQs had the narrowest and SQs the widest pitch range while RQs were in-between the two. Third, SQs were further distinguished from ISQs and RQs by a different placement of focal accent and the accentuation of pronouns. There were, however, no differences in intonational pitch accent types and boundary tones between the three utterance types.

The results imply that the lower mean pitch signals the indirect illocutionary force of the non-canonical questions while the longer duration, non-modal voice quality and larger pitch range indicate their affective nature. SQs are additionally associated with a specific information structure.

Index Terms: pitch, duration, voice quality, speech act, information-seeking questions, rhetorical questions, surprise questions, information structure

1. Introduction

The goal of this study is to investigate the role of prosody in disambiguating indirect speech acts by comparing string-identical canonical and non-canonical questions in Estonian. By canonical or information-seeking questions (ISQs) we mean syntactically canonical interrogative sentences that are uttered by the speaker in order to request information from an addressee who is presumed to possess that information, cf. [1]. By non-canonical questions we mean syntactically canonical interrogative sentences that are uttered for other purposes than to seek information. These include rhetorical and surprise questions. Rhetorical questions (RQs) imply that the speaker and the addressee are already committed to a similar and obvious answer [2]. Surprise questions (SQs) express the surprise of the speaker at an unexpected state of affairs or piece of information that s/he has witnessed or learned.

Cross-linguistically, several prosodic differences have been identified between string-identical RQs and ISQs in various Germanic languages (see [3] for an overview), and between SQs and ISQs in French [4]. In both cases the utterance types were shown to differ in their intonation. Additionally, in both studies ISQs were found to be shorter than RQs and SQs.

Earlier research on Estonian compared the prosody of string-identical RQs and ISQs [5], and string-identical SQs and ISQs [6]. This work showed that both SQs and RQs differed from ISQs in that they had a longer duration, lower mean pitch and a larger proportion of creaky voice. There were no significant differences in the occurrence of intonational pitch accent types or boundary tones between these question types. However, both RQs and SQs differed from ISQs in terms of the distribution of pitch accents: RQs were characterised by a denser pitch accent distribution while SQs had the same number of pitch accents but a potentially different placement of the nuclear pitch accent and more frequently accented pronouns. This implies that SQs may have a characteristic information structure. As a further difference, SQs displayed a significantly wider pitch range than ISQs while RQs had a narrower pitch range than ISQs. These results suggest that RQs and SQs share several prosodic characteristics but may differ from each other with respect to pitch accent distribution and pitch range.

In the above-mentioned studies on Estonian, RQs were examined using polar questions and wh-questions with the interrogative pronoun kes (‘who’), and SQs using wh-questions with the interrogative pronoun or determiner mida (‘what’). The present study focuses on interrogative sentences containing the interrogative form kuidas (‘how’). In some languages, the interrogative form with the meaning ‘how’ has a second reading in addition to the manner and means reading: a reason reading [7]. The Estonian interrogative word kuidas can likewise have both the manner/means reading ‘how’ and the reason reading ‘why, how come’. How-questions with the reason reading have been shown to be triggered by expectancy disconfirmation and, as a result, to express surprise [7, 8, 9]. Additionally, kuidas-interrogatives can be RQs that express the speaker’s and addressee’s common assessment of the proposition conveyed by the interrogative.

The prosody of how-questions with manner/means vs. reason reading has been studied in French comment-questions by [9]. It was found that reason-comment questions were marked either by a characteristic final tone or by a different focal accent placement. Phonetically, reason-comment questions were characterised by slower speech rate and breathy voice. Additionally, the occurrence of laughter was noted.

An open topic regarding non-canonical questions is their illocutionary force. Existing studies on the prosody of Estonian ISQs in comparison with statements [10, 11, 12] suggest that their main prosodic correlate is a higher mean pitch, cf. [13] for other languages. Consequently, the significantly lower pitch found in SQs and RQs [5, 6] can imply that these questions are not information-seeking. On the other hand, SQs and RQs share a property of ISQs that distinguishes the latter from statements, namely the lack of H+L pitch accent which is relatively common in the nucleus position of Estonian statements. This suggests that non-canonical questions differ prosodically not only
from canonical questions but also from statements, constituting a speech act of their own.

The present study aims to elaborate on the previous findings by carrying out a direct comparison of the prosody of three types of string-identical Estonian kuidas-interrogatives: ISQs with kuidas in the manner/means reading, SQs with kuidas in the reason reading, and RQs which express the speaker’s and addressee’s common assessment of the proposition conveyed by the interrogative. The following predictions are made:

- RQs and SQs have a similarly lower F0, longer duration and a larger proportion of creaky voice quality than ISQs;
- RQs have a narrower and SQs a wider pitch range than ISQs;
- RQs have a denser intonational pitch accent distribution while SQs may have a different (but not denser) pitch accent distribution, incl. nuclear accent placement and accentuation of pronouns;
- ISQs, RQs and SQs do not differ in terms of pitch accent types and boundary tones.

2. Materials and method

2.1. Materials

The materials included ten interrogative sentences which each consisted of four words and had the structure [WH-WORD SUBJECT COMPLEMENT VERB]. The subject was always the short form of the second person singular pronoun sa (‘you’). The complement and the verb varied depending on the sentence. The complement was either lexical (e.g. Kuidas sa tööle lähed? ‘how 2SG work.ALL go.2SG!’ “How are you going to work?”) or pronominal (e.g. Kuidas sa seda müüt? ‘how 2SG it.PAR sell.2SG’ “How are you selling it?”). For each interrogative sentence three different contexts were created to elicit the three different readings. In order to facilitate the elicitation the contexts for ISQs explicitly stated “You want to know. . .”; the contexts for SQs included “You are surprised that. . .”; and the contexts for RQs contained “It is obvious to both of you that. . .”.

In addition to the 30 test items the materials contained 60 fillers. The fillers were various declarative sentences that were likewise elicited by a context description.

2.2. Informants

The analysis is based on the data from eight informants: all female speakers of Standard Estonian between 26 and 45 years old (mean age 32). The informants were remunerated for their participation.

2.3. Procedure

The recordings were made in the sound-treated recording booth of the phonetics laboratory of the University of Tartu using a Praat [14] demo script. The informants were asked to first silently read the context description that appeared on the computer screen, and when ready to proceed to the next slide where the test sentence was displayed for 5 seconds. Each recording session was preceded by three trial contexts. All the materials were presented to each participant in randomised order at one sitting. Each context and test item appeared only once.

The recordings were segmented using an ASR force-aligner [15] and the segmental boundaries were manually corrected. The data was annotated for intonational pitch accents and

Figure 1: Average utterance duration for ISQs, RQs and SQs.

boundary tones following [12] by the first two authors, and for the occurrence of creaky voice by the third author. The utterance duration and pitch from 100 equidistant points were extracted using a Praat script. The following F0 measures were calculated: utterance mean F0, pitch range between the 5% and 95% quantile within the utterance, utterance-initial and final F0 as the mean of the first and the last vowel, accordingly. The results were tested in R [16] using the packages lme4 [17] and lmerTest [18]. All F0 measures were converted from Hertz to semitones (st) with reference to the speaker’s mean. The acoustic measures of duration and F0 were fitted with linear mixed models with condition (levels ISQ/SQ/RQ) as fixed effect, and random intercepts for speaker and utterance. The durations were log-normalised in order to approach a normal distribution. The occurrence of creaky voice as a binary factor was tested with a logistic mixed model with the same fixed and random factors.

3. Results

3.1. Duration

There was a significant difference in utterance duration ($F(2, 221)=83.4, p<0.001$). The average duration of ISQs was 1.03 seconds (s) and a post-hoc test showed that ISQs were significantly shorter than RQs and SQs (both $p<0.001$). The average duration of SQs was 1.21 s which is slightly longer than that of RQs (1.18 s), the difference being not significant (see Figure 1). The solid line in Figures 1-5 shows the LMM estimate.

3.2. Mean F0

There was a significant effect of question type to the mean F0 of the utterance ($F(2, 227.05)=4.06, p=0.019$) which is presented in Figure 2. Post-hoc testing showed that ISQs were significantly different from RQs and SQs ($p<0.05$) with there being no significant difference between RQs and SQs. The overall differences between the question types were small: mean F0 for ISQs was 0.21 s, for SQs 0.24 s, and for RQs 0.12 s.

3.3. Pitch range

There was a significant difference between question types in pitch range ($F(2, 221)=46.17, p<0.001$) which is presented in Figure 3. ISQs had the narrowest (6.4 st) and SQs the widest pitch range (9.8 st) while RQs were in-between the two (7.9 st). Post-hoc testing confirmed all differences at $p<0.001$.
3.4. Initial and final F0

There was a significant effect of question type on the initial F0 of the utterance (F(2, 220.2)=7.27, p<0.001). Post-hoc testing showed that the initial F0 of SQs (4.2 st) was significantly higher (p<0.01) than that of ISQs (3.4 st) and RQs (3.2 st) whereas the difference between ISQs and RQs was not significant. There was no difference between the three question types in their final F0; all utterances ended at about the same level approximately -2.5 st lower than the speakers’ mean F0.

3.5. Voice quality

There was a significant difference between question types in the occurrence of creaky voice (χ²(df=2)=14.9, p<0.001) which was present in 38% of ISQs, 55% of RQs and 61% of SQs (Figure 4). Post-hoc testing showed that ISQs had a significantly lower percentage of creaky voice as compared to SQs and RQs (p<0.05) while the difference between RQs and SQs was not significant. Figure 5 shows the proportion of creaky voice in the utterances. There was a significant difference between the three types of question (F(2, 109.6)=4.1, p=0.019). The proportion of creaky voice formed approximately 10% of the utterance duration in ISQs, 12% in RQs and 15% in SQs. The difference was significant only between ISQs and SQs (p<0.05).

3.6. Intonation and accentuation patterns

The question types did not differ in terms of intonational pitch accents and boundary tones. The predominant pitch accent was H*+L (563 of the total of 573). Only two utterances of the total of 240 ended with a high boundary tone (H%); in all other cases the final boundary was low and marked by default as 0%.

The average intonation contours are shown in Figure 6 separately for utterances with a pronominal (left panel) and with a lexical complement (right panel). Most of the utterances were produced with one of the following three accentuation patterns. The pattern where the accented constituents were the wh-word and the verb (WH-subj-compl-VERB) was typical of the utterances with a pronominal complement for all question types (in total 89 instances). The pattern where the only unaccented constituent was the subject (WH-subj-COMPL-VERB) was particularly typical of ISQs and RQs with a lexical complement (91 instances). The pattern where the accented constituents were the wh-word and the complement (WH-subj-COMPL-verb) was characteristic of SQs (34 of the total of 52 instances). The distribution of pitch accents in ISQs and RQs was very similar while that of SQs differed in two respects. First, in SQs the nuclear accent was equally often on the complement as on the verb while in ISQs and RQs it was predominantly on the verb. Second, in SQs the pronominal complement was more likely to be accented: 21 SQs vs. 3 ISQs and 6 RQs.

4. Discussion

The goal of this study was to compare the prosody of string-identical canonical and non-canonical questions in Estonian testing various predictions based on earlier work. As predicted, ISQs had a significantly higher mean F0 than RQs and SQs while RQs and SQs did not differ from each other. Initial F0 was significantly higher in SQs, which could be attributed to their higher level of expressivity and emotional load. As also predicted, the duration of ISQs was significantly shorter than
that of RQs and SQs that did not differ significantly from each other, although SQs were somewhat longer than RQs. It was also confirmed that ISQs contained significantly less creaky voice than RQs and SQs. Although the difference between RQs and SQs was not significant, the proportion of creaky utterances was somewhat larger for SQs where the duration of creaky voice was proportionally the longest. It is worth noting, however, that there was a large inter-speaker variation as to the occurrence of creaky voice, cf. [19].

The prediction about pitch range being narrower in RQs and wider in SQs than in ISQs was not borne out with respect to RQs: both RQs and SQs had a significantly wider pitch range than ISQs. This finding does not, therefore, confirm the results of an earlier comparison of RQs and ISQs in Estonian [5]. Assuming that a wider pitch range signals a higher level of emotion and expressivity, this could be explained by the fact that the contexts used for eliciting RQs in the present study triggered a higher degree of emotion than those in [5]. The presence of affectivity in RQs is supported by the observation that some of the utterances were accompanied by a sigh. While the pitch range of RQs was significantly wider than that of ISQs, it was significantly narrower than that of SQs. This three-way distinction may imply that SQs involve the highest level of emotion and ISQs the lowest, with RQs in the middle. The fact that RQs were in-between ISQs and SQs also in terms of duration, the proportion of creaky utterances as well as the proportion of creaky voice in these utterances suggests that a longer duration and non-modal voice quality may likewise be associated with affectivity.

The prediction about RQs being characterised by a denser pitch accent distribution and SQs by a different nuclear accent placement was not borne out with respect to RQs: the total number of pitch accents was very similar in all question types. The results of the present study differ, thus, again from [5]. This difference could be explained with a different word order of the test sentences that in [5] included a sentence-medial finite verb, which tended to be deaccented in ISQs and accented in RQs, and in the present study a sentence-final verb, which was accented both in ISQs and RQs. As predicted, SQs differed from ISQs and RQs in nuclear accent placement. The placement as well as the shape of the nuclear accent in SQs indicated a relatively frequent narrow focus on the complement, while in ISQs and RQs the nuclear accent was nearly always on the sentence-final verb. This result confirms earlier findings for Estonian SQs [6] as well as French reason-comment questions [9] in that surprise and counter-expectation are expressed by a characteristic information structure. The results of the present study also conform to the results of the previous study on the prosody of Estonian SQs in the frequent accentuation of pronouns absent in ISQs and RQs.

While cross-linguistic studies comparing the prosody of canonical and non-canonical questions have usually revealed differences in intonational phonology (e.g. [3, 4, 9]), this is not the case in Estonian. As predicted, the three question types did not differ in terms of intonational pitch accent types and boundary tones. The H+L* pitch accent that has only been observed in statements and can therefore be associated with the signalling of declarative force was equally absent from all three question types. This implies that while RQs and SQs are different from ISQs, they are also different from statements, constituting a speech act of their own.

5. Conclusions

This study compared the prosodic characteristics of string-identical canonical and non-canonical questions in Estonian. It appeared that these two types differ prosodically with respect to several characteristics: non-canonical questions have a lower pitch and longer duration and contain a larger proportion of creaky voice quality. Information-seeking, rhetorical and surprise questions were differentiated on the basis of pitch range which was narrowest for ISQs and widest for SQs. As pitch range can be associated with the degree of affectivity, this suggests that SQs involve the highest and ISQs the lowest level of affectivity. Additionally, SQs had a significantly higher initial pitch than ISQs and RQs. As shown in earlier studies, SQs were different from RQs and ISQs also in terms of nuclear accent placement and, consequently, information structure, as well as in terms of the accentuation of pronouns. The three question types did not differ in terms of the number and type of intonational pitch accents or boundary tones.

6. Acknowledgements

This work was conducted within the framework of the project EKKD10 “The prosody and information structure of surprise questions in Estonian in comparison with other languages”. The first author was supported by the Centre of Excellence in Estonian Studies (CEES, European Regional Development Fund).
7. References


