Prosodic characteristics of exclamations and 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
heete.sahkai@eki.ee, eva-liina.asu@ut.ee, partel.lippus@ut.ee

Abstract

This study investigates the prosody of Estonian exclamations as compared to questions. The materials consisted of string-identical wh-exclamatives and wh-interrogatives that were elicited using contexts prompting the two different readings. The data was recorded from 21 female speakers of Standard Estonian.

The analysis revealed various important prosodic differences between the two speech acts. First of all, exclamations were characterised by the presence of a prominent pitch accent that has also been found to signal exclamatives in other languages. Exclamations also showed a significantly longer duration of the whole utterance which could be associated with their expressive nature. Exclamations and questions did not, however, differ with respect to pitch range. Other characteristics of exclamations were a significantly lower mean f0 including both lower initial and final pitch as well as more frequent occurrence and a larger proportion of creaky voice towards the end of the utterance. Additionally, differences between the two speech acts in the placement of the nuclear pitch accent were found.

Index Terms: intonation, pitch, duration, voice quality, speech act, exclamations, questions, Estonian

1. Introduction

This study compares the prosody of string-identical information-seeking wh-questions (ISQs) and wh-exclamations (EXCs) in Estonian. It forms part of a larger project aiming to identify prosodic features that serve to signal different speech acts and disambiguate indirect speech acts in Estonian, in particular, canonical information-seeking questions as opposed to different types of non-canonical interrogatives. By non-canonical interrogatives we mean syntactically canonical interrogative utterances that are not questions such as rhetorical questions, surprise questions, exclamatory questions and interrogative requests.

The main prosodic features that have been shown to distinguish between Estonian canonical ISQs and non-canonical interrogatives are pitch level, pitch range, duration and voice quality as well as intonational pitch accent distribution, but not pitch accent type or boundary tones [1, 2, 3]. More specifically, both rhetorical questions (RQs) and surprise questions (SQs) were distinguished from canonical ISQs by lower mean pitch, longer duration and a larger proportion of creaky voice quality. Lower mean pitch along with steeper declination has also been found to distinguish Estonian assertions from ISQs [4]. SQs were further distinguished from ISQs and RQs by a larger pitch range, while the pitch range of RQs seems to vary depending on the level of emotional expressivity that they carry. Additionally, the three speech acts were found to differ with respect to the distribution of intonational pitch accents: SQs had a characteristic nuclear accent placement signalling narrow focus, while RQs may have more pre-nuclear pitch accents per utterance. A larger number of upstepped pitch accents is a feature that has been found to characterise Estonian ISQs as opposed to assertions [4], but also SQs as opposed to ISQs.

Prosodic characteristics of exclamations have so far not been studied in depth. In Estonian, a perception study using synthesized monosyllabic stimuli [5] showed that questions and exclamations were similar in that both were characterised by a high pitch peak while the final f0 was higher in exclamations.

Cross-linguistic research on exclamatory vs. canonical questions (e.g. [6, 7] for German, [8] for English) has shown that exclamatory questions are distinguished from canonical questions by the following properties: longer duration; the presence of a highly prominent (usually pre-nuclear) pitch accent, often attracted by pronominal elements [6, 7] or the wh-phrase [8], in German polar exclamatory questions also by the utterance-initial finite verb [7]; the type and distribution of pitch accents as well as the type of nuclear contour; a larger number of pitch accents [7] and prosodic phrases [8]; and a smaller sensitivity to information structure [6, 7].

Based on the previous findings on Estonian exclamations [5] and non-canonical interrogatives [1, 2, 3] and findings on exclamatory questions in other languages [6, 7, 8], we expect Estonian EXCs to differ prosodically from ISQs in terms of duration, pitch characteristics, voice quality as well as intonational pitch accent realisation and distribution. No differences in pitch accent types or boundary tones are expected, however, as these have not been found to signal speech acts in Estonian. In particular, we will examine if Estonian EXCs are distinguished form ISQs by the following characteristics:

1. a longer duration, which has been found in non-canonical questions in Estonian and EXCs cross-linguistically;
2. a different mean pitch, which in the light of the findings on Estonian non-canonical questions could be lower than in ISQs, while following the findings of [5] the utterance-final pitch of EXCs may be higher;
3. an overall wider pitch range, which has been found in expressive non-canonical questions in Estonian (but not in exclamatory questions in German);
4. presence of a highly prominent (pre-nuclear) pitch accent, which may fall on the wh-phrase or a pronoun, as found in German and English exclamatory questions;
5. a different distribution and number of pitch accents, as found in Estonian non-canonical questions and exclamatory questions in English and German;
6. more frequent occurrence of creaky voice quality, which has been shown to characterise expressive non-canonical questions in Estonian.
2. Materials and procedure

The materials of the study consisted of string-identical ISQs and EXCs that were elicited using contexts prompting the two different readings. In total 12 different target sentences and 24 contexts were compiled. Each target sentence had three constituents: a wh-phrase consisting of the adverb kui ‘how’ and a scalar adjective or adverb, a third-person subject (six pronominal and six proper name subjects), and a present tense third person singular finite verb (there were seven lexical and five copular verbs). The context descriptions prompted the speaker to express astonishment with or without an additional positive or negative judgment (admiration or disapproval). Table 1 gives an example of a target sentence along with two context descriptions: one eliciting an ISQ reading (left column) and another an EXC reading (right column). The informants were asked to first silently read the context description that appeared on the computer screen, and when ready they could proceed to read aloud the target sentence that appeared upon the next mouse click.

The data was recorded from 21 female speakers of Standard Estonian in the sound-treated recording booth of the phonetics laboratory of the University of Tartu using a Praat [9] demo script. Each speaker produced 12 questions and 12 exclamations. The final data-set comprised 503 tokens (one item had to be excluded for technical reasons).

The recordings were segmented using an ASR force-aligner [10] and the segmental boundaries were manually corrected. The recordings were manually annotated for intonational pitch accents and boundary tones, and the occurrence of creaky voice. 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 [11] using the packages LME4 [12] and InerTest [13]. The acoustic measures of duration and f0 were fitted with linear mixed models with condition (levels ISQ/EXC) as fixed effect, and random intercepts and random slopes of condition for speaker and item. 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.

Table 1: Example of a test sentence with two context descriptions eliciting the two different readings.

<table>
<thead>
<tr>
<th>ISQ</th>
<th>EXC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your friend explains that their sister is travelling in Uruguay. You want to know whether their sister often makes such trips. You say to your friend:</td>
<td>You hear from your friend that Leena is travelling in Uruguay. You are surprised because only last week Leena had arrived back from a long trip to Cambodia. You say to your friend:</td>
</tr>
</tbody>
</table>

Figure 1 presents the average duration of the utterance for ISQs and EXCs. The duration of ISQs was 1.04 seconds and that of EXCs 1.29 seconds. The difference of 250 ms was significant $F(1, 25)=8.31, p<0.001$.

Figure 2 shows the mean f0 of the utterance. On average the f0 of ISQs was 232 Hz which was significantly higher than that of EXCs (217 Hz) ($F(1, 17.86)=37.88, p<0.001$). Figure 3 presents the pitch range of the utterance, which was 103 Hz in ISQs and 116 Hz in EXCs. The difference is not significant $F(1, 24.99)=4, p=0.056$.

Utterance-initial and utterance-final f0 can be observed from the panels of Figure 6. Utterance-initial f0 was 261 Hz in ISQs, which is significantly higher than the 225 Hz of EXCs ($F(1, 24.5)=47.23, p<0.001$). Utterance-final f0 was 199 Hz in ISQs and 183 Hz in EXCs, and this difference is significant ($F(1, 18.38)=7.79, p=0.012$).

The percentage of utterances containing creaky voice is shown in Figure 4 and the proportion of creaky voice in the utterances in Figure 5. Creaky voice occurred in 50% of EXCs but only in 34% of ISQs ($\chi^2(1)=6.59, p=0.01$). Of all the utterances that contained creaky voice the relative duration of creak was significantly longer in EXCs than in ISQs ($F(1, 16.5)=9.88, p=0.006$). Creaky voice mostly occurred within the final 20% of the utterance duration in both ISQs and EXCs (see bottom panels in Figure 6).

The most common pitch accent in the data was H*+4L (93%). The two accentuation patterns that were predominant in both ISQs and EXCs were the one where the wh-phrase and verb were accentuated (WH-PHRASE - subject - VERB) in the test sentences with the proper subject (45% of ISQs and 35% of EXCs) and where all three constituents received a pitch accent (WH-PHRASE - SUBJECT - VERB) in the sentences with the proper name subject (48% of ISQs and 36% of EXCs).

The nuclear accent was mainly on the verb (94% of ISQs vs. 72% of EXCs) but in EXCs, often also on the wh-phrase (in 23% of EXCs vs. 4% of ISQs).

3. Results

4. Discussion

The aim of this study was to compare the prosody of string-identical wh-ISQs and wh-EXCs in Estonian. Several differences were revealed.
As expected, EXCs were significantly longer than ISQs, which has been shown for EXCs cross-linguistically and for non-canonical questions in Estonian. This can be associated with the expressive nature of EXCs, as has also been suggested by [6] for German. Another characteristic that could be associated with expressivity is a wider overall pitch range of an utterance that was found, for instance, in Estonian SQs. In the current study, however, the difference between ISQs and EXCs in terms of pitch range was narrowly below the level of significance. Similarly, a wider pitch range was not found to be characteristic of German exclamatory questions [6, 7].

EXCs were further distinguished from ISQs by their significantly lower mean, initial and final pitch. This is in accordance with our earlier findings [1, 2, 3] showing high pitch level to be the main prosodic correlate of ISQs, distinguishing them from both assertions and non-canonical interrogatives. The finding about the utterance-final f0 being lower in EXCs differs, however, from the results of a perception study [5] where EXCs were associated with a higher final pitch than questions.

In the current study, EXCs were additionally characterised by a larger proportion of creaky voice quality which has also been found for Estonian RQs and SQs. This could be explained, on the one hand, by their lower mean pitch and, on the other hand, by their expressive nature.

The EXCs also contained a highly prominent pitch accent, which has been found to characterise exclamations cross-linguistically. It was located on the wh-phrase, similarly to what has been observed in English wh-exclamatives [8]. In the German wh-exclamatives studied by [6], the wh-phrase tended to be unaccented; instead, the prominent accent was usually on a pronoun or auxiliary in the utterance-initial region. This difference in the accentuation of wh-phrases could be explained by the fact that in the study of [6] the wh-phrase did not convey a degree. As noted by [7], exclamations are characterised by the presence of a prominent accent, but the location of this accent is variable; if there is an element that independently attracts prominence for discourse-semantic reasons, the prominent accent is likely to fall on this element. It is possible that a wh-phrase consisting of a degree word and a scalar element attracts the prominent accent for semantic reasons. As defined by [14], exclamatives presuppose a proposition that involves an open scalar degree and assert that this degree is higher than the speaker expects. The wh-phrase could thus attract the prominent accent because it denotes the relevant scale and contains the degree word ‘how’.

The most common intonational pitch accent in our data was H*+L. Figure 6 suggests that on the wh-phrase in EXCs it is characterised by a longer and steeper rise to the high tone, a
somewhat later peak, and a longer and steeper fall to the low tone as compared to the H*L accent on the wh-phrase in ISQs. A separate study is needed to investigate the peak-alignment in these two speech acts in more detail.

The nuclear accent placement was on the whole similar in ISQs and EXCs: it was mostly on the utterance-final verb. Nevertheless, in 23% of the EXCs the nuclear accent occurred on the wh-phrase while in ISQs only 4% of nuclear accents were on the wh-phrase. Apart from this difference, the number and distribution of pitch accents was relatively similar in ISQs and EXCs.

As expected, pitch accent type and boundary tones did not distinguish EXCs from ISQs. This is different from German and English where EXCs are distinguished by the nuclear contour or pitch accent types.

5. Conclusion

Estonian wh-exclamations were distinguished from wh-information-seeking questions by characteristics that can be associated with their expressive nature such as longer duration, and a larger proportion of creaky voice quality. Additionally, EXCs were characterised by a highly prominent pre-nuclear or nuclear pitch accent on the wh-phrase, and a significantly lower mean as well as utterance-initial and utterance-final pitch level. The number and distribution of intonational pitch accents was relatively similar in the two speech acts with one notable difference: the nuclear accent was on the wh-phrase in 23% of the EXCs while only in 4% of the ISQs. EXCs and ISQs did not differ with respect to pitch range or intonational pitch accent type and boundary tones.

6. Acknowledgements

We are grateful to Katrin Leppik for carrying out the recording of the data and manually checking the segment boundaries. 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


