The prosody of rhetorical vs. information-seeking questions in Icelandic
Nicole Dehé, Bettina Braun, Daniela Wochner

Department of Linguistics, University of Konstanz, Germany
{nicole.dehe, bettina.braun, daniela.wochner}@uni-konstanz.de

Abstract
This paper reports on a production experiment investigating the prosodic realization of rhetorical questions (RQs) as compared to information seeking questions (ISQs) in Icelandic. It looks at two question types: polar questions (Borðar einhver límónur? ‘Does anybody eat limes?’) and wh-questions (Hver borðar límónur? ‘Who eats limes?’).

The main results are as follows. (i) In both question types, the boundary tone fails to contribute to the distinction between ISQs and RQs. It is L% almost across the board. (ii) The semantic difference between ISQs and RQs is reflected in the nuclear accents: In wh-questions, ISQs have more monotonal (H*/!H*/^H*) pitch accents, while RQs have more bitonal ones (mostly L+H*/!L+H*/L+H*). In polar questions, nuclear accents are mostly L+H, but the timing of the rise differs (more L+H in ISQs, more L+H* in RQs). (iii) The first word of the utterance and the nuclear syllable are longer in RQs than in ISQs in both question types. Within the nuclear syllable, both onset and rhyme are lengthened.

Taken together, prosody helps to distinguish between ISQs and RQs, but the terminus of the intonational contour (boundary tone) is not essential.

Index Terms: question intonation, rhetorical questions, Icelandic, pitch timing, duration

1. Introduction
This paper reports on a production experiment, which investigates the prosody of rhetorical wh- and polar questions as compared to string-identical information-seeking questions in Icelandic.

1.1. Information-seeking questions vs. rhetorical questions
Neutral questions (aka information-seeking questions; ISQs) perform the directive speech act of requesting information. Polar questions “request an answer that specifies whether the proposition expressed by their sentence radical holds or does not hold” [1, p. 1747], i.e. the expected answer may be “yes” or “no”. Wh-questions “create an open proposition by leaving parts of the description of the proposition unspecified” [1, p. 1744]; in Icelandic, the open parameter is represented by the h- pronoun (e.g. hver ‘who’, hvar ‘where’, hvernig ‘how’). The expected answer is one that provides information about the open parameter. Rhetorical questions (henceforth RQs) are formally (i.e., surface-syntactically) interrogatives, but differ from ISQs in discourse function. Based on much previous literature, [2] summarize their characteristics as follows: (i) RQs do not expect an answer [3], [4], [5], [6], [7], (ii) RQs have the feel of an assertion [8], [9], [6], [10], [11], and (iii) RQs do not have to but can optionally be answered [12], [10].

1.2. Icelandic question intonation
Icelandic question intonation has not yet been the focus of much experimental research. The default intonational contour of both polar questions and wh-questions is falling to a low boundary tone (L%) [13], [14], [15], [16]. According to [15, p. 323], questions with rising intonation (H%) “have special connotations” such as, for example, impatience or surprise. We therefore expect ISQs to terminate in L% regardless of question type (polar vs. wh).

According to [15], the typical nuclear pitch accent in Icelandic questions is a rise from a low target on the accented syllable (L*+H). Combined with the low boundary tone, the typical nuclear contour is thus a rising-falling one (L*+H L%). Importantly, the intonation of Icelandic polar questions thus differs crucially from those in related languages, whose polar questions typically end in a rise to H% (e.g. English: [17], [18], [19], [20], [21], and German: [22], [23]).

According to [14, p. 477], the typical nuclear contour of a wh-question in Icelandic starts high, has an optional high peak (H*) prenuclear accent associated with the wh-word, followed by a H* nuclear accent, then falls towards L%.

[15, pp. 322-323] maintains that Icelandic statements and polar questions differ in the type of nuclear accent. While the early rise (L*+H) is the typical nuclear accent in statements (see also [24]), the late rise L*+H is typical of polar questions, both followed by a fall to L%. Therefore, in Icelandic, nuclear pitch accents types help to distinguish between illocution types (statement: L*+H; polar question: L*+H; wh-question: H*), while boundary tones do not contribute to this distinction.

1.3. The prosody of ISQs and RQs
It is a common assumption in the semantic and prosodic literature that ISQs and RQs are prosodically different (e.g., [2], [8], [9]), but actual studies on the prosody of RQs are rare, deal almost exclusively with English, and typically do not directly compare string-identical ISQs and RQs (e.g. [6], [18], [3]). A preliminary study for German identifies phonological and phonetic differences between ISQs and RQs, specifically differences in the distribution of boundary tones and nuclear pitch accent types, longer durations in RQs than ISQs, and more breathy voice quality in RQs [25], [26]. Yet more recently, [27] find similar results for English. Specifically they show (i) that edge tones distinguish between ISQs and RQs only in polar questions (wh-questions have L-L% throughout), (ii) that the type and – in polar questions – the position of the nuclear accent is important, and (iii) that phonetically, longer constituent durations are found in RQs than in ISQs. [27] also show in a perception study that listeners are able to identify the intended illocution based on prosodic realization alone.

There is no previous literature on the prosody of RQs in Icelandic.
1.4. Hypotheses

Based on the introduction given in Sections 1.1 through 1.3, we hypothesize that (i) ISQs and RQs differ in their prosodic realization. (ii) Phonologically, the type of nuclear pitch accent does, but the boundary tone does not, contribute to this distinction. (iii) Phonetically, based on previous findings in related experiments for English [27] and German [25], [26], we predict longer durations in RQs as compared to ISQs in the prenuclear and nuclear region of the utterance.

2. The experiment

A production experiment was designed to study the prosodic realization of ISQs vs. RQs in Icelandic. It was carried out in a sound-attenuated room at the University of Iceland in June 2017.

2.1. Materials

Twenty-one pairs of wh- and polar interrogatives were constructed, as well as accompanying contexts (one context triggering an information-seeking, one a rhetorical interpretation, for each question pair). Target interrogatives and contexts were translated from English [27] as closely as possible by a native speaker of Icelandic. Each question was felicitous in both an information-seeking context and a rhetorical one, resulting in 21 quadruples, see Tables 1 and 2 below (contexts in Tables 1 and 2 are translated from Icelandic). Contexts were created such that in contexts triggering an ISQ reading of the target interrogative, the answer was obviously not known to the speaker and would instead have been highly informative. The description of the context situation was therefore followed by a sentence starting You would like to know or similar (Tables 1 and 2, left columns). In contexts triggering an RQ reading of the target interrogative, there was no uncertainty about the answer. On the contrary, the answer to the RQ was obvious from the given context, i.e. common ground in the imaginary situation. This was achieved by the string However, it is well known that ...; see right-hand columns in Tables 1 and 2.

In addition, 34 fillers were created, along with a context (parallel to the experimental items). They were 19 verb-second (V2) exclamatives (e.g. Rosalega leis Nina vel! 'How well Nina can read!'), seven exclamatives starting enn hve (e.g., Enn hve Lena jarf ad lara mikri: 'How much Lena has to learn?') six neutral verb-first (V1) polar questions (e.g. Er hann med doktorgráðu? 'Does he have a doctorate?'), and two V1 sentences functioning as requests for help (Geturðu hjálpðu mér víd ...? 'Can you help me with ...'). Finally, there were three practice items of parallel make-up (one wh-question, one polar question, and one V2 exclamative).

2.2. Procedure

Two basic experimental lists were constructed. Each list contained both polar and wh-questions, and both illocution types. The members of the quadruplets were distributed across the two lists such that one list contained eleven polar and ten wh-questions, the other list contained 10 polar and 11 wh-questions. Illocution type was thus manipulated within-subjects. The same polar or wh-question occurred twice in each list, one in an ISQ context, the other in an RQ context. For example, the items in Tables 1 were members of List 1, the items in Table 2 appeared in List 2. The 34 filler items were added to each list. Participants were randomly assigned to one of the two experimental lists. The experimental lists were randomized for each participant separately with the constraint that two readings of a question were separated by at least four other trials. Each experiment started with three practice trials. After these, there was a short break, which participants were allowed to use for questions, if anything was unclear. The experiment was controlled using the experiment software Presentation (Neurobehavioral-Systems, 2000). Each trial started with the visual display of the context, which the participant had to read silently. After a button press, the target interrogative appeared on the same screen. Participants were instructed to read each context situation carefully and to utter the target and filler sentences as naturally as possible. The recording started simultaneously with the appearance of the target interrogative on screen. Participants pressed a button to proceed to the next trial. The recording was stopped at this point. Participants were allowed to repeat the target in case of mistakes. The whole experiment lasted about 25 to 30 minutes. The contexts were presented in black Calibri 40 font and the target sentences in blue Calibri 40 font, all on white background. The participants were instructed to produce their utterances in such a way that they were suitable in the given context. No feedback was provided during the experiment. Productions were recorded using a headset-microphone (Shure SM10A) and digitized directly onto a PC (44.1kHz, 16Bit).

Table 1: Contexts and target polar interrogatives

<table>
<thead>
<tr>
<th>Context for ISQ</th>
<th>Context for RQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>At a party you offer cake which contains limes. You would like to know which of your guests eat this fruit and whether they would like to try the cake. You say to your guests:</td>
<td>Your aunt offers limes to her guests. However, it is well known that this fruit is too sour for it to be possible to eat it. You say to your aunt:</td>
</tr>
<tr>
<td>Target Q: Borðar einhver límónur? ('Does anybody eat limes?')</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Contexts and target wh-interrogatives

<table>
<thead>
<tr>
<th>Context for ISQ</th>
<th>Context for RQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>At a party you offer cake which contains limes. You would like to know which of your guests eat this fruit and whether they would like to try the cake. You say to your guests:</td>
<td>Your aunt offers limes to her guests. However, it is well known that this fruit is too sour for it to be possible to eat it. You say to your aunt:</td>
</tr>
<tr>
<td>Target Q: Hver borðar límónur? ('Who eats limes?')</td>
<td></td>
</tr>
</tbody>
</table>

2.3. Participants

Thirty-two native speakers of Icelandic (aged 20-65 years, 20 female) participated in the experiment. All participants were unaware of the purpose of the study and were given a present from Konstanz for their participation. Of the 32 participants, all participants aged 20-32 (N=21) entered the analysis. This was done to keep the age range of participants comparable to parallel studies on English [27] and German [26]. Four participants were removed from the analysis due to missing files or because they changed the wording of more than one target interrogative. Accordingly, seventeen participants (aged 22-32; average 26.9; 6 male) were analyzed.
2.4. Data treatment and analysis

Two quadruples contained non-native object nouns (Lambada, Bolognese). They were removed from the analysis due to varying (mostly but not only non-native) placement of word stress (N=68). In addition, one item (polar ISQ) had to be removed due to stammering on the part of the speaker. Accordingly, 645 target interrogatives entered the analysis. They were 313 polar (156 ISQs, 157 RQs) and 332 wh-questions (166 ISQs, 166 RQs). All target interrogatives were annotated in Praat [28] by the first author (see Figure 1 for an example). For the phonological analysis, prenuclear and nuclear pitch accents, as well as boundary tones were annotated following previous intonational analyses of Icelandic in the autosegmental-metrical framework (e.g. [24], [16]; second tier from top in Figure 1). Along with H% and L%, boundaries were labelled M% (N=19), if contours ended in a mid-level. They were mostly falls to mid from a nuclear pitch peak (N=17), comparable to the calling contour, which has been analyzed as a downstepped H phrase accent (H-) (i.e., H*H- in [29]). Since a prosodic hierarchy or the category of phrase accent has yet to be established for Icelandic, the observed fall to mid (instead of low) is analyzed as a fall to the auxiliary category M% here. More rarely, there were slight final rises after a fall from a nuclear pitch peak, labeled LM% (N=2). Bitonal (rising L+H) pitch accent were annotated L*+H, if L was aligned within the stressed syllable and H was aligned in the syllable following the stressed syllable; they were annotated L+H*, if H was aligned within the stressed syllable. Note that Icelandic has word-initial primary stress throughout with very few exceptions in non-native words; e.g. [15]. The duration of the nuclear syllable, its onset and rhyme were automatically extracted from the segmental annotation (third tier from top in Figure 1).

For the statistical analysis of accent types and boundary tones, we ran logistic mixed effect regression models with illocution-type (ISQ vs. RQ) as fixed factor and participants and items as crossed random factors (adjustment of intercepts). For dependent variables with more than two levels, one level was coded 1 and all other levels 0 [30]. For the statistical analysis of duration we ran linear mixed effect regression models with the same specifications and model fitting as described above. P-values were calculated using the Satterthwaite approximation of degrees-of-freedom. To avoid Type I errors, p-values were adjusted by means of the Benjamini-Hochberg correction [31].

2.5. Results

This section focuses on two phonological parameters (boundary tones, nuclear pitch accents) and one phonetic one (duration).

2.5.1. Boundary tones

Contours are overwhelmingly falling across illocution types and question types, with 100% L% in RQs (see Table 3). In wh-questions, there are significantly more L% in RQs than in ISQs (ß=3.6, SE=1.1, z=3.4, p<0.001). There are more M% in wh-IsQs than in wh-RQs, but the effect could not be verified statistically because of the missing instances in RQs.

2.5.2. Nuclear pitch accents

Effects were found for type of nuclear accent for both question types. The main difference for wh-questions is between monotonal and bitonal pitch accents, the main difference for polar questions is the timing of the nuclear rise. Specifically, more H*/!H*/^H* accents were found in wh-IsQs than in wh-IsQs (ß=3.2, SE=0.6, z=5.3, p<0.001); see Table 4 for distribution. In polar questions, more L*+H (late rise) accents were found in ISQs than in RQs (ß=1.9, SE=0.3, z=5.6, p<0.0001). The difference for polar questions is illustrated in Figure 2.

2.5.3. Duration

Longer durations were found for RQs than for ISQs across question types (wh and polar) in both the nuclear and the prenuclear region of the target interrogatives. First, illocution-type had a significant effect on the duration of the initial words.

Figure 1: Example target interrogative and annotation (polar question, RQ, vp28, female)

Figure 2: Late rise (ISQ; left) vs. early rise (RQ; right) in polar questions; same item and speaker as Figure 1; nuclear syllable marked by rectangle; C1f: beginning of stressed syllable, C1: beginning of syllable following the stressed syllable
of the utterances. In polar questions, the verb was on average 86.7ms longer in RQs than in ISQs (ß=86.9, SE=0.01, df=16.7, t=6.4, p<0.001). In wh-questions, the wh-word was on average 58.4ms longer in RQs than in ISQs (ß=60.7 SE=0.01, df=18.8, t=4.6, p<0.0005). Second, there was a significant effect of illocution-type on the duration of the nuclear syllable in both question types. The accented syllable was on average 62.9ms longer in RQs than in polar RQs (ß=30.9, SE=0.003, df=14.8, t=9.9, p<0.0001); RQ: 110.3ms; ISQ: 80.6ms. It is on average 44.9ms longer in RQs than in ISQs (ß=32.9, SE=0.005, df=19.4, t=7.3, p<0.001); RQ: 158.5ms; ISQ: 125.9ms. The rhyme in polar questions is on average 32.6ms longer in RQs than in ISQs (ß=32.9, SE=0.005, df=19.4, t=7.3, p<0.001); RQ: 158.5ms; ISQ: 125.9ms. In wh-questions, the rhyme is on average 50.2ms longer in RQs than in ISQs (ß=50.2, SE=0.007, df=19.4, t=7.0, p<0.0001); RQ: 205.7ms; ISQ: 296.4ms.

For the analysis of the duration of the onset consonant (Vt-C1t in Figure 1) and the rhyme (C1f-Vt in Figure 1) of the nuclear syllable, all items were excluded whose stressed syllable started with a vowel (e.g. innmat’innards’), unless these vowels were preceded by a syllable-initial (onset) glottal stop. 543 items remained for analysis (272 wh: 136 ISQ, 136 RQ; 271 polar: 135 ISQ, 136 RQ). The analysis showed significant effects of illocution-type for both onset and rhyme (see Figure 3). The onset consonant is on average 29.7ms longer in polar RQs than in polar ISQs (ß=30.3, SE=0.003, df=14.8, t=9.9, p<0.0001); RQ: 110.3ms; ISQ: 80.6ms. It is on average 44.9ms longer in wh-RQs than in wh-ISQs (ß=43.3, SE=0.004, df=26.4, t=10.7, p<0.0001); RQ: 124.8ms; ISQ: 83.5ms. The rhyme in polar questions is on average 32.6ms longer in RQs than in ISQs (ß=32.9, SE=0.005, df=19.4, t=7.3, p<0.001); RQ: 158.5ms; ISQ: 125.9ms. In wh-questions, the rhyme is on average 50.2ms longer in RQs than in ISQs (ß=50.2, SE=0.007, df=19.4, t=7.0, p<0.0001); RQ: 205.7ms; ISQ: 296.4ms.

3. Discussion

First, for ISQs the results confirm observations in previous literature based mostly on introspective data (e.g., [14], [15]). In particular, both polar and wh-questions typically end in L%, with around 10% M% in wh-ISQs, but remember that these are also falls, although not to low but to mid. While the typical nuclear accent in wh-ISQs is H*, polar ISQs typically have a nuclear late rise (L+H*). Second, the results confirm the hypotheses put forward in Section 1.4 about the comparison of ISQs and RQs. (i) Despite some overlap in nuclear accent distribution, ISQs and RQs clearly differ in their prosodic realization. (ii) The boundary tone does not contribute to the distinction between illocution types. This is unlike English, where boundary tones distinguish illocution types in polar questions [27], and unlike German, where boundary tones are different between ISQs and RQs in both polar and wh-questions [26]. In Icelandic, nuclear accent type crucially differs between ISQs and RQs. Specifically, we find monotonal H*/H*/H* vs. bitonal L+H*/L+H*/L+H* in wh-questions, and differences in the timing of the rise in L+H accents in polar questions (more later peaks L+H* in ISQs, and more earlier peaks L+H* in RQs). (iii) Longer durations were observed in RQs than in ISQs, both in the prenuclear field (wh-word in wh-questions, verbs in polar questions), and in the nuclear syllable. Longer durations in the prenuclear and nuclear field of RQs were also found for English [27]. For Icelandic the lengthening of the onset consonant of the nuclear syllable is particularly noteworthy. It is well known that under stress and emphasis, the rhyme (the vowel in open syllables and the consonant in closed syllables) is lengthened (e.g. [15]), but lengthening of the syllable onset has not yet been reported. It is conceivable (and will have to be confirmed in future research) that increased lengthening of the syllable onset marks illocution type and/or other aspects of meaning independent of the normal lengthening of the rhyme under stress.

The results are compatible with [15]’s observation that the nuclear pitch accent type used in Icelandic utterances reflects illocution type: [15] maintains that the early rise (L+H*) is typical of statements (see also [24]), while the late rise (L*+H) is typical of polar questions. The present study confirms the late rise for polar ISQs, while polar RQs typically have an early rise (L+H*), similar to statements in [15] and [24], reflecting their interrogative syntactic form but non-question function. In wh-questions, the difference is between monotonal nuclear H* in ISQs (see also [14], [15]), and bitonal nuclear rises in RQs. All nuclear pitch accents are typically followed by L%, i.e. all intonational contours, across question types and illocution types, are typically terminated by a final fall. This also confirms earlier, introspective work on question intonation in Icelandic ([13], [14], [15]), which introduces L% as the default boundary tone in Icelandic interrogatives unless overridden by special connotations such as surprise or impatience. The present results therefore also suggest that these kinds of special connotations must be set apart from illocution type. While in Icelandic, illocution type (assertive, information-seeking, rhetorical) fails to be phonologically signalled by the boundary tone (default L% throughout), but are instead marked by pitch accent type, special connotations, which in principle are possible with all illocution types, are signalled by a non-default boundary tone.

4. Conclusions

We conclude that the boundary tone does not, but nuclear pitch accents do play a role for the distinction between ISQs and RQs in Icelandic, thus adding evidence to previous work observing identical boundary tones (L%) but different pitch accents for different illocution types. In addition, duration contributes to the distinction between ISQs and RQs.

5. Acknowledgements

We thank Sigirður Sæunn Sigurðardóttir for careful translations of the materials, Sigirður Sæunn Sigurðardóttir and Christiane Ulbrich for carrying out the experiment in Iceland, Jörgen Pind and Arn Kristjánsson for giving us access to the sound-attenuated room at the University of Iceland, Tolli Eythorsson for providing equipment, and Margrét Pálsvík and Sigiríður Sigurjónsdóttir for help with finding participants.
6. References


