# THE PERCEPTION OF SURPRISE QUESTIONS IN ESTONIAN

Pärtel Lippus[1], Eva Liina Asu[1], Katrin Leppik[1], Heete Sahkai[2]

[1]Institute of Estonian and General Linguistics, University of Tartu, Estonia

[2]Institute of the Estonian Language, Tallinn, Estonia

partel.lippus@ut.ee, eva.liina.asu-garcia@ut.ee, katrin.leppik@ut.ee, Heete.Sahkai@eki.ee

## ABSTRACT

The study examines the perception of surprise questions as compared to information-seeking questions. An online perception experiment was carried out in the PCIbex environment using natural stimuli from 6 different speakers from an earlier production experiment. The task of the participants was to evaluate whether the speaker intended to ask for information or to express surprise. The overall correct identification score for information-seeking questions was 84% and for surprise questions 55%, but there was large variability between speakers and items. Higher perception rate of surprise questions is above all related with a longer duration, a wider pitch range, and a lower mean pitch of the utterance.

**Keywords:** surprise, non-canonical questions, information-seeking questions, prosody, perception, Estonian

## 1. INTRODUCTION

This study examines the perception of Estonian surprise questions (SQs) as compared to information-seeking questions (ISQs). The SQs examined in the study are syntactically and lexically canonical wh-interrogatives that have been uttered by the speaker to express surprise. They are compared to string-identical interrogatives that have been uttered in order to request information. SQs differ from ISQs in terms of speech act, as they are not (canonical) questions, as well as in terms of added emotional expressivity, as they express surprise.

Prosodic expression of speech acts is usually associated with the use of phonological pitch accents and edge tones while the expression of emotions has been associated with the gradient use of phonetic realisation, see e.g. [1]. A previous production study [2] showed, however, that Estonian SQs differ from string-identical ISQs by their longer duration, lower pitch, wider pitch range, and a more frequent use of creaky voice quality, but not in terms of pitch accent types or boundary tones. A similar production study examining string-identical SQs and ISQs in French [3] identified both phonological and phonetic differences between the two categories: ISQs were shorter and produced with a faster tempo, and contained a greater proportion of rising final contours. A following perception study [4] confirmed that prosodic cues were sufficient for distinguishing the two conditions without any context provided, but the correct perception rate of ISQs was significantly higher than that of SQs (82% vs. 69%).

The present study addresses two research questions: (1) Which acoustic features correlate most with the perception of surprise? (2) Does the context of surprise influence the recognition rate of SQs? The broader aim of the study is to investigate which factors contribute to the perception of surprise and speech acts.

## 2. MATERIALS AND METHOD

### 2.1. Stimuli

The stimuli of the perception test were recorded as part of a production study [2] and included 24 interrogative sentences: 12 with the question word *mis* 'what' (e.g. *Mis loom see on?* 'What animal is this?') and 12 with the question word *mida*, partitive case of *mis* 'what' (e.g. *Mida sa teed?* 'What are you doing?'). Two contexts were created for each sentence in order to elicit an information-seeking reading and a surprise reading of the sentence. The contexts for eliciting ISQs prompted the participant to imagine a situation that would require asking for information from a knowledgeable addressee. The contexts for eliciting SQs, on the other hand, prompted the participant to act surprised at an unexpected situation (surprise caused by the unprepared mind of the speaker) or a counterexpectational situation (surprise caused by a contrary expectation of the speaker). A subset of the contexts included incongruity (untypicality of the

object of surprise) as an additional cause of surprise, and contexts of another subset were intended to elicit disapproval in addition to surprise. The purpose of the different context types was to test the effect of the context of surprise on the expression and perception of surprise (for a full list of test sentences and elicitation contexts see https://osf.io/b4fek).

The design of the perception test followed largely that of a similar test by [4]. The stimuli consisted of 240 unique utterances produced by six randomly chosen female participants of the production study. An equal number of utterances was randomly chosen from each speaker (40 per speaker). The stimuli were distributed between four lists of 72 so that each list contained 36 ISQs and 36 SQs. Twelve interrogative sentences in each list were uttered once as a SQ and twice as an ISQ by different speakers, and 12 sentences were uttered once as an ISQ and twice as a SQ. There were eight stimuli from each speaker (in total 48 stimuli) that appeared in two lists but no repetitions by the same speaker of the same sentence were included in one list.

### 2.2. Perception test

The perception test was carried out online in the PCIbex environment [5]. 52 native speakers of Estonian (44 female, 7 male, 1 non-binary; 24 to 69 years of age, mean age 42) participated in the test. All but one participant had self-reported normal hearing but as her responses did not differ from the rest of the group they were not excluded. Two participants reported no knowledge of languages other than Estonian (on B2 level or above); 26 reported the knowledge of one, 13 of two, and 10 of more than two other languages.

During the test, upon hearing each stimulus the participants had to decide whether the speaker's intention was to ask for information or express surprise. The test was preceded by a short practice session consisting of four stimuli where the intention of the speaker was indicated (two requests for information and two expressions of surprise).

Due to technical reasons the lists did not receive an equal number of listeners. The total number of listeners per each list was as follows: A 12, B 14, C 11, D 15. Considering the design of the experiment each stimulus received, thus, at least 11 responses and some stimuli up to 29 responses.

### 2.3. Analysis

The results of the perception test were analysed in *R* [6, 7]. Logistic regressions from the *lme4* package [8] were used to assess the effects of the context type and the acoustic factors on the perception of surprise. The effects of the acoustic features on the perception test ratings were further evaluated with a Random Forest model using the package *randomForest* [9].

## 3. RESULTS

The overall average results are shown in Figure 1. The information-seeking questions were recognised (correctly) as questions asking for information in 84% of the cases but the surprise questions were recognised as expressing surprise only in 55% of the cases, which is just above the chance level. In the following subsections, the results are presented in relation to the effect of the context type and the acoustic features of the surprise condition stimuli.

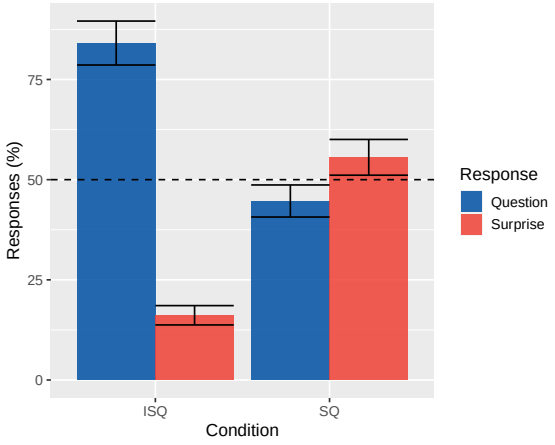


**Figure 1:** Recognition of ISQs and SQs as asking for information (blue) or expressing surprise (red).

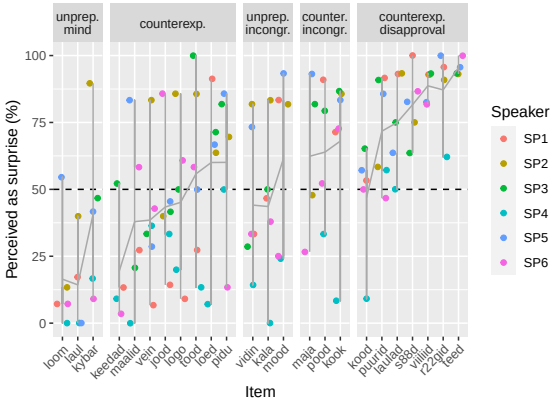


**Figure 2:** The frequency of surprise condition items perceived as expressing surprise grouped by the context type.

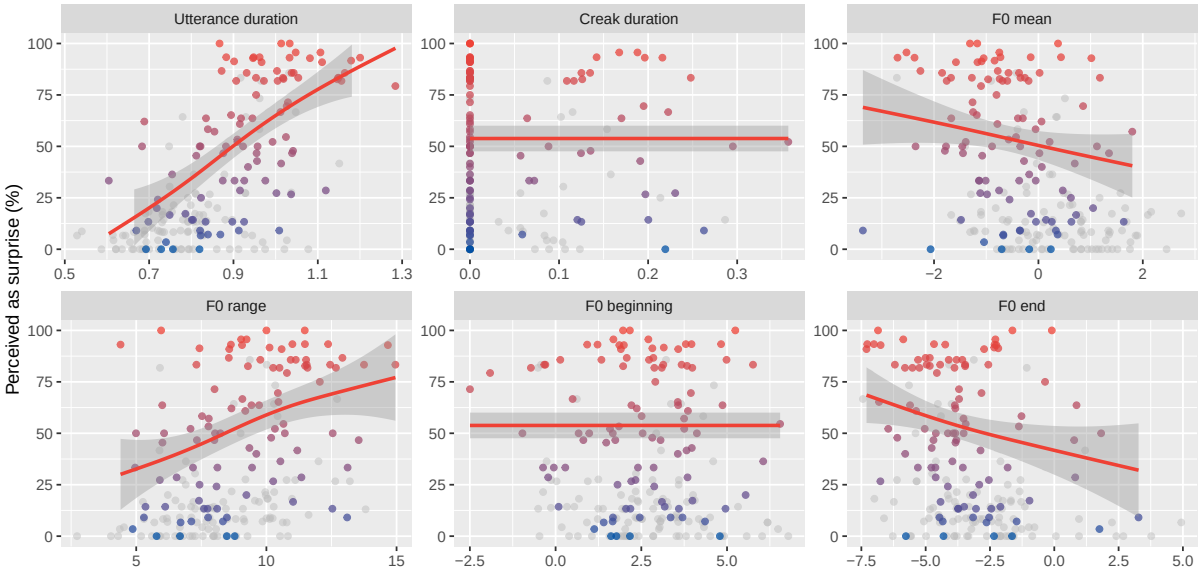


**Figure 3:** Perception of surprise as a function of acoustic features of the surprise condition stimuli (ranging from blue to red according to the rate of perceived surprise; duration in seconds, pitch in semitones re speaker's mean). The question condition stimuli are shown in grey in the background.

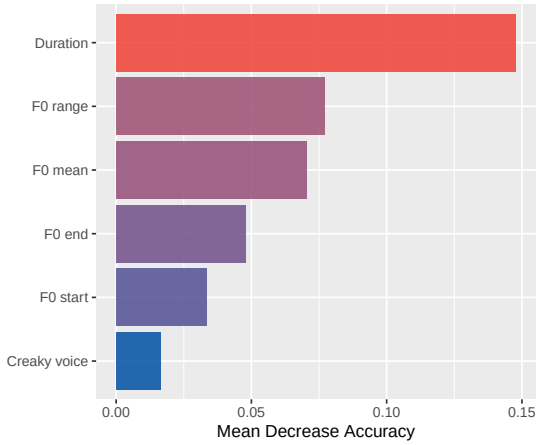


**Figure 4:** Mean decrease of accuracy of the acoustic factors in the random forest model.

### 3.1. Context of surprise

There was a large speaker- and item-dependent variation in the perception of SQs. Figure 2 shows the frequency of surprise condition items perceived as expressing surprise grouped by the context type. It can be seen that some stimuli were rarely identified as expressing surprise (e.g. items *loom*, *laul*, *keedad*) while there were stimuli that were in most of the cases recognised as expressing surprise (e.g. items *teed*, *viilid*). It can also been seen that the stimuli by some speakers were less well recognised as expressing surprise, e.g. in particular those of SP4.

Figure 2 also shows that the recognition of surprise is affected by the context type that was coded in the context eliciting the utterances used for the stimuli. The perception of surprise was the lowest in the case of *unprepared mind* (28%), followed by *counterexpectation* (45%), *unprepared mind combined with incongruity* (47%), and *counterexpectation combined with incongruity* (66%). The perception of surprise was the highest and significantly above chance level in the case of *counterexpectation combined with disapproval* (75%).

A post-hoc test of a logistic regression model (with stimulus and participant as random intercepts) showed that the difference was significant only between the two categories with lower recognition rate (*unprepared mind* and *counterexpectation*) on the one hand and the category with the highest recognition rate (*counterexpectation combined with disapproval*) on the other hand ($p < 0.001$).

### 3.2. Acoustic features

Figure 3 shows the perception of surprise questions as a function of different acoustic features of the stimuli such as utterance duration, F0 range, F0 mean, utterance-initial and utterance-final F0 and the duration of creaky voice quality. There is a

strong positive correlation between the perception rate of surprise and the duration of the utterance as well as the F0 range of the stimuli, and a negative correlation between the perception rate of surprise and the F0 mean and the utterance-final F0. There was no correlation between the perception rate and utterance-initial F0 or the duration of creaky voice. A logistic model showed only a significant effect of utterance duration ($\beta = 8.34$, $z = 7.85$, $p < 0.001$) and F0 range ($\beta = 0.09$, $z = 2.03$, $p = 0.042$) indicating a strong intercorrelation between the different measures of F0.

The effect of the acoustic features on predicting the perception of surprise was tested with a random forest model. The data was split into two subsets: 80% of the data was used for training the model and 20% for testing. The model was trained with the six acoustic factors with 500 trees and 2 variables at each split. The OOB estimation rate of the model was 19.5%. The variable importance is presented in Figure 4. The most important factor is the total duration of the utterance, which is followed by the F0 range and F0 mean. Utterance-final F0 had a stronger effect than utterance-initial F0. The duration of creaky voice was the weakest of the six factors. On the testing sample the prediction accuracy of the model was 81% (error rate 15% for question and 28% for surprise).

## 4. DISCUSSION

This study investigated the perception of information-seeking (ISQs) and surprise questions (SQs) in Estonian using natural stimuli produced by six speakers. ISQs were more reliably recognised than SQs (at 84%), and nearly half of the SQs were perceived as ISQs. A similar outcome (albeit a higher recognition rate of SQs) was obtained for French by [4], who suggest that the interpretation of the stimuli as questions may have been supported by their interrogative syntax and absence of context.

By far the strongest acoustic cue correlating with the perception of surprise was the longer duration of the utterance. Longer duration was also identified as the main phonetic correlate of French SQs [3] and has been found to correlate with non-canonical questions more generally, in particular rhetorical questions [10, 11] and exclamatory questions [12, 13, 14, 15]. The next strongest correlates of perceived surprise were a wider F0 range and lower mean F0. A wider F0 range also distinguishes Estonian exclamations from questions [12] and has been associated with the expression of surprise [16, 17] as well as other emotions [18, 19]. The lower mean pitch of Estonian SQs as compared to ISQs signals a non-questioning speech act [2] and is characteristic of Estonian statements [20, 21, 22], rhetorical questions [10] and exclamations [12] as compared to questions.

Yet another characteristic acoustic feature of SQs that was identified in the production study [2] - a longer duration of creaky voice - did not correlate with the perception of surprise. This can be explained by the fact that there is a lot of individual variation and the present study used the stimuli from only six randomly chosen speakers.

Regarding the effect of different possible causes of surprise (the unprepared mind of the speaker, a contrary expectation by the speaker, and the untypicality of the object of surprise), and the combination of surprise with another emotion/attitude (disapproval), the results suggest that surprise alone is perceptually relatively similar to a request for information. This is especially obvious when caused only by the unprepared mind or counterexpectation on the part of the speaker, while surprise in combination with disapproval is perceptually distinct from a request for information. Such a result might imply that when the expression of surprise is accompanied by the expression of a judgment (for example, that a situation is untypical or disapprovable), the resulting speech act is more clearly perceptually non-questioning. In future work, we intend to examine in more detail the correlation of the acoustic cues of SQs with the different contexts of surprise.

## 5. CONCLUSION

It can be concluded that the acoustic feature that is most robustly related to the perception of surprise questions in Estonian is the longer duration of the utterance. Additionally, such pitch-related features as a wider F0 range and a lower mean F0 contribute to the accurate perception of surprise questions. Regarding the effect of the context of surprise, surprise questions seem to be perceptually more distinct from information-seeking questions when they also express a judgment.

## 6. ACKNOWLEDGEMENTS

We are very grateful to Maud Pélissier for her valuable advice on experimental design. The study was supported by the projects EKKD10 and EKKD114 (the Estonian Ministry of Education and Research) and the basic governmental financing of the Institute of the Estonian Language by the Estonian Ministry of Education and Research.

## 7. REFERENCES


[1] P. Prieto, "Intonational meaning: Intonational meaning," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 6, no. 4, pp. 371–381, Jul. 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/wcs.1352

[2] E. L. Asu, H. Sahkai, and P. Lippus, "The prosody of surprise questions in Estonian," *Journal of Linguistics*, pp. 1–21, Mar. 2023. [Online]. Available: https://doi.org/10.1017/S0022226723000014

[3] A. Celle and M. Pélissier, "Surprise questions in spoken French," *Linguistics Vanguard*, vol. 8, no. s2, pp. 287–302, Jan. 2022. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/lingvan-2020-0109/html

[4] ——, "C'est quoi and qu'est-ce que as surprise questions: a perception study," in *Presentation given at the PHC Parrot final workshop*, 2022-10-28.

[5] J. Zehr and F. Schwarz, "PennController for Internet Based Experiments (IBEX)," Aug. 2022, publisher: Open Science Framework. [Online]. Available: https://osf.io/md832/

[6] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: https://www.R-project.org/

[7] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. [Online]. Available: https://ggplot2.tidyverse.org

[8] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[9] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002. [Online]. Available: https://CRAN.R-project.org/doc/Rnews/

[10] H. Sahkai, E. L. Asu, and P. Lippus, "Prosodic characteristics of canonical and non-canonical questions in Estonian," in *Proc. Speech Prosody 2022*, 2022, pp. 135–139. [Online]. Available: https://www.doi.org/10.21437/SpeechProsody.2022-28

[11] B. Braun, N. Dehé, J. Neitsch, D. Wochner, and K. Zahner, "The prosody of rhetorical and information-seeking questions in German," *Language and Speech*, vol. 62, no. 4, pp. 779–807, Dec. 2019. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0023830918816351

[12] H. Sahkai, E. L. Asu, and P. Lippus, "Prosodic characteristics of exclamatives and questions in Estonian," in *Proc. 1st International Conference on Tone and Intonation (TAI)*, 2021, pp. 41–45. [Online]. Available: https://www.doi.org/10.21437/TAI.2021-9

[13] S. Repp, "The Prosody of *Wh* -exclamatives and *Wh* -questions in German: Speech Act Differences, Information Structure, and Sex of Speaker," *Language and Speech*, vol. 63, no. 2, pp. 306–361, Jun. 2020. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0023830919846147

[14] S. Repp and H. Seeliger, "Prosodic Prominence in Polar Questions and Exclamatives," *Frontiers in Communication*, vol. 5, p. 53, Aug. 2020. [Online]. Available: https://www.frontiersin.org/article/10.3389/fcomm.2020.00053/full

[15] J. Rett and B. Sturman, "Prosodically marked mirativity," in *Proceedings of the 37th West Coast Conference on Formal Linguistics*, D. K. E. Reisinger and M. Huijsmans, Eds. Somerville, MA, USA: Cascadilla Proceedings Project, 2021, publisher: The University of British Columbia Version Number: 1. [Online]. Available: http://www.lingref.com/cpp/wccfl/37/paper3511.pdf

[16] C. Gunlogson, *True to form: rising and falling declaratives as questions in English*, ser. Outstanding dissertations in linguistics. New York: Routledge, 2003.

[17] C. Lai, "Perceiving surprise on cue words: prosody and semantics interact on right and really," in *Proc. Interspeech 2009*, 2009, pp. 1963–1966.

[18] S. Repp and L. Rosin, "The intonation of echo wh-questions," in *Proc. Interspeech 2015*, 2015, pp. 938–942.

[19] T. Bänziger and K. R. Scherer, "The role of intonation in emotional expressions," *Speech Communication*, vol. 46, no. 3-4, pp. 252–267, Jul. 2005. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167639305000890

[20] E. L. Asu, "Downtrends in different types of questions in Estonian," in *Proceedings of the Speech Prosody 2002 conference*, B. Bel and I. Marlien, Eds. Aix-en-Provence: Laboratoire Parole et Langage, 2002, pp. 143–146.

[21] ——, "The phonetics and phonology of Estonian intonation," Ph.D. thesis, University of Cambridge, 2004.

[22] K. Vende, "Phonetic conditioning factors of pitch in Estonian vowels," in *Estonian Papers in Phonetics*, Tallinn, 1973, pp. 46–85.