



## Creak as a feature of lexical stress in Estonian

Kätlin Aare<sup>1,3</sup>, Pärtel Lippus<sup>1</sup>, Juraj Šimko<sup>2</sup>

<sup>1</sup>University of Tartu, Estonia

<sup>2</sup>University of Helsinki, Finland

<sup>3</sup>Stockholm University, Sweden

{katlin.aare, partel.lippus}@ut.ee, juraj.simko@helsinki.fi

### Abstract

In addition to typological, turn-taking or sociolinguistic factors, presence of creaky voice in spontaneous interaction is also influenced by the syntactic and phonological properties of speech. For example, creaky voice is reportedly more frequent in function words than content words, has been observed to accompany unstressed syllables and ends of phrases, and is associated with relaxation and reduced speech.

In Estonian, creaky voice is frequently used by all speakers. In this paper, we observe the use of creaky voice in spontaneous Estonian in connection to syllabic properties of words, lexical stress, word class, lengthening, and timing in phrases.

The results indicate that creak occurs less in syllables with primary stress than in unstressed syllables. However, syllables with secondary stress are most frequently creaky. In content words, the primary stressed syllables creak less frequently and unstressed syllables more frequently compared to function words. The stress-related pattern is similar in both function and content words, but more contrastive in content words. The probability of creakiness increases considerably with non-final lengthening within words, and for all syllables towards the end of the intonational phrase.

**Index Terms:** lexical stress, creaky voice, spontaneous Estonian, conversation

### 1. Introduction

Creak is a voice quality associated with tightly adducted vocal folds open along a portion of their length to allow for voicing. Acoustically, it is characterized by a series of irregularly spaced vocal pulses with a decreased acoustic intensity and typically a lower fundamental frequency compared to modal voice [1].

The use of creaky voice has mostly been studied from typological, turn-taking or sociolinguistic perspectives. For example, as creak is regularly located at phrase boundaries in several languages (e.g. in Finnish [2], Swedish [3], English RP [4], etc.), it has been investigated as a potential turn-yielding cue. Creak is also reportedly more frequent in informal and relaxed situations (see e.g. [5], [6]). In Estonian, creaky voice occurs more often during informal spontaneous dialogues (on average about 10 % of speaking time for all speakers, but more frequently by younger males) compared to lecture monologues [7].

By nature, spontaneous speech is subject to various phenomena characterizing running unplanned speech, such as reduction and devoicing. For instance, according to [8], devoicing in English occurs more likely in unstressed

syllables, thus often resulting in a breathy voice, and in German spontaneous speech, devoicing is frequently realized as creaky phonation (see also [9]). Creaky voice is characterized by less acoustic intensity than modal voice, usually resulting in a decrease in loudness, and therefore creak has been connected to reduction. Unstressed syllables often have lower pitch – in Estonian, primary stressed syllables have higher pitch, while stressed and unstressed vowels have greater pitch variation [10].

As summarized in [10], in Estonian, primary stress falls on the first syllable, with the exception of some loan words, interjections and names, where it can fall elsewhere. The position of stresses follows the trochaic rhythm structure, where disyllabic or trisyllabic words usually consist of a single foot, beginning with a primary stressed syllable which is followed by one or two unstressed syllables. A tetrasyllabic word consists of two disyllabic left-headed feet. Secondary stressed syllables are usually successive odd-numbered syllables. In terms of its acoustic correlates, the Estonian stress has been studied somewhat sporadically, mostly in relation to the quantity system. Acoustically, stress is mostly manifested by the temporal patterns and pitch variability [11, 10], less consistently by intensity [12, 13], and secondarily also by vowel quality [14, 15]. However, the acoustic variability related to stress strongly interacts with the complex three-way quantity system.

In Estonian, creak has mostly been studied from a sociolinguistic point of view [7, 16], and studies on Estonian stress correlates have excluded analyzing creaky speech spurts, although it has been noted that creak is most likely a manifestation of the lack of prominence [10]. In this paper, we focus on prosodic features (lexical stress) and their interplay with grammatical word categories, phrasal structure as well as some social aspects of spontaneous interaction to gain a more thorough understanding of the occurrence and role of creaky voice in Estonian. An important part of our exploration is to investigate the potential communicative function of creak. In other words, is creaky voice simply a manifestation of low-effort voicing, i.e., is it predominantly present in less acoustically prominent portions of speech? Alternatively, does this type of voicing contribute to marking and, possibly, further enhancing contrast between acoustically more and less prominent syllables (stressed/unstressed) and/or words (content/grammar words)?

### 2. Material & methods

The data under investigation is from the collection of University of Tartu Phonetic Corpus of Estonian Spontaneous Speech [17]. The corpus contains spontaneous dialogues recorded in a sound-treated room. The recordings are

manually segmented using Praat [18] TextGrids on word and phoneme level. The subset of the corpus used in this paper has additionally been manually annotated for syllable boundaries and lexical stress. The subset consists of 14 spontaneous dialogues between 20 unique speakers (7 female and 13 male; age range 20–64, mean 32 years). The duration range of the recordings is 25–50 minutes (mean 39 minutes).

Creak has been identified both manually and automatically in the corpus. Manual identification is carried out during the phoneme level annotation of the corpus. Automatic detection is done with the Voice Analysis Toolkit (Matlab, [https://github.com/jckane/Voice\\_Analysis\\_Toolkit](https://github.com/jckane/Voice_Analysis_Toolkit)), specially designed to carry out voice quality analysis [19]. As the last step, all labels marked as creaky by manual and/or automatic detection were checked by a trained phonetician.

The word class information has been added using automatic morphological analyser Vabamorf (online: <https://github.com/Filosoft/vabamorf>). The complex morphological analysis from the analyser’s output has been simplified into three categories: content words (nouns and adjectives), verbs and grammatical words (adverbs, conjunctions, numbers, pronouns, pre- and postpositions).

The final data set consists of 105 766 syllables. The mean number of syllables per speaker is 5288. The distribution of the data between stress and word class is illustrated in Table 1. The following measures were extracted from the TextGrids for each syllable: creak percentage; stress (levels: primary/secondary/ unstressed); position in the word: duration; timing within a phrase; word class (levels: content word/verb/grammatical word); foot quantity; phonological vowel quality. In order to reduce variability, we excluded cases where: the labeler was unsure about the stress conditions (4983 cases); highly reduced syllables (2142); syllables containing diphthongs (7275); syllables from more than 5-syllable words (54).

Table 1: *The number of syllables by stress and word class*

	Content	Verb	Grammatical
Primary	15750	8071	23913
Secondary	2131	762	903
Unstressed	19003	12089	23144

For the analysis, we checked each syllable for creakiness from 100 points. The presence of any portion of creaky voice within a syllable was treated as a binary factor. A logistic mixed effect regression model was fitted using the lme4 [20] package in R [21]. An optimal model was reached by testing the combinations of the linguistic and social factors. Random intercepts were calculated for the speakers. Additional random intercepts were included for the vowel quality as this has been shown to have an effect on creak (e.g. in German [22]), and for the quantity which in Estonian strongly interacts with the stress (e.g. [10]).

### 3. Results

Out of all analyzed syllables, 18 % were creaky. Table 2 shows the distribution of syllables in terms of creakiness and stress.

Table 2: *The number and percentages of creaky/non-creaky syllables in stressed/unstressed syllables.*

	Primary	Secondary	Unstressed
Non-creaky	40177 (84 %)	2942 (78 %)	43585 (80%)
Creaky	7557 (16 %)	854 (22 %)	10651 (20 %)

As Table 2 demonstrates, the absolute number of non-creaky syllables is relatively even in unstressed and primary stressed syllables, while creaky syllables are predominantly unstressed or have primary stress. In all syllable types, the number of creaky syllables is considerably smaller than the number of syllables lacking creak. The following statistical model describes the variation of creak in more detail (Table 3).

Table 3: *The fixed effects of the optimal logistic regression model estimating the likelihood of the syllables to be produced with creaky voice.*

	Est.	Std. Error	z value	Pr(> z )
(Intercept)	-1.831	0.221	-8.29	<0.001
Str <i>Primary</i>	-0.401	0.040	-9.98	<0.001
Str <i>Secondary</i>	0.160	0.062	2.58	0.010
Word <i>Verb</i>	-0.057	0.031	-1.84	0.065
Word <i>Gram</i>	-0.069	0.027	-2.49	0.013
Dur To Phr End	-0.474	0.013	-37.79	<0.001
Lengthening	1.417	0.056	25.52	<0.001
Syl Rank 2	0.039	0.029	1.34	0.180
Syl Rank 3	0.322	0.037	8.64	<0.001
Syl Rank 4	0.455	0.057	7.92	<0.001
Syl Rank 5	0.394	0.129	3.05	0.002
SP2 Gnd <i>Male</i>	-0.254	0.059	-4.29	<0.001
Age Dif	-0.038	0.005	-7.2	<0.001
Str <i>Pri</i> : Word <i>Verb</i>	0.289	0.051	5.69	<0.001
Str <i>Sec</i> : Word <i>Verb</i>	0.079	0.109	0.73	0.468
Str <i>Pri</i> : Word <i>Gram</i>	0.460	0.041	11.33	<0.001
Str <i>Sec</i> : Word <i>Gram</i>	0.002	0.104	0.02	0.985
Str <i>Pri</i> : Dur To Phr End	0.084	0.021	4.07	<0.001
Str <i>Sec</i> : Dur To Phr End	0.042	0.051	0.81	0.417
Str <i>Pri</i> : Lengthening	-0.157	0.087	-1.8	0.071
Str <i>Sec</i> : Lengthening	-0.844	0.337	-2.5	0.012
Dur To Phr End : Lengthening	0.314	0.048	6.52	<0.001
SP2 Gnd <i>M</i> : Age Dif	0.035	0.005	6.41	<0.001

The base value of the intercept is taken to be an unstressed non-lengthened first rank syllable in a Content word, uttered by an interlocutor whose speaking partner is female.

### 3.1. Creaky voice, stress & word class

The results showing the presence of creak in syllables with different stress level and occurring in different word categories are summarized in Figure 1 (although only rough percentages are plotted, as the mixed effect model used for evaluation takes into account interactions and random effect of other analyzed phenomena). Syllables with primary stress have significantly less creak and syllables with secondary stress have significantly more creak than unstressed syllables, and (as revealed by post hoc analyses), this difference is also significant for verbs (at  $p < 0.05$ ) but not for grammatical words (for this word category the percentage of creaky syllables does not significantly differ based on stress level). Primarily stressed syllables are significantly creakier in grammatical words than in verbs, and in verbs than in content words. Interestingly, for unstressed syllables, this trend is reversed, with minutely but significantly greater percentage of creak in syllables from content words than in grammatical ones (the remaining differences are not significant). Finally, the likelihood of creak in syllables with secondary stress does not significantly differ between word classes (at  $p > 0.05$ ).

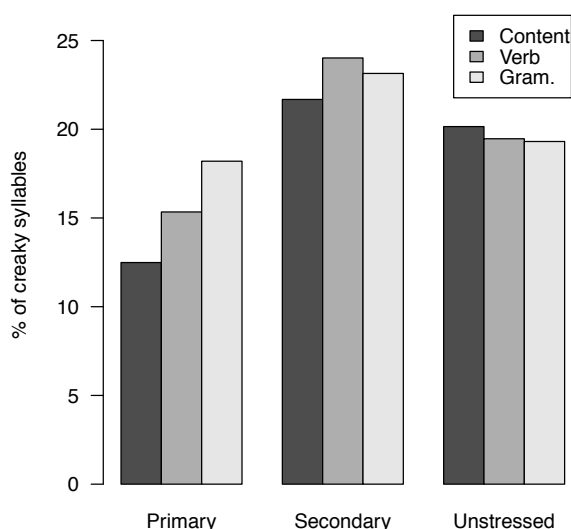


Figure 1: The percentage of creaky syllables in different stress conditions in content words, verbs and grammatical words.

The interactions between the verb and grammatical word classes and primary and secondary stress categories all increase probability of creakiness. More specifically, syllables with primary stress in verbs and grammatical words are significantly creakier than unstressed syllables in the same word class categories. Word class interactions with secondary stress are not significant, but verbs and grammatical words also show tendencies for more creakiness than content words

### 3.2. Creaky voice & lengthening

Lengthening in general significantly increases the occurrence of creakiness, and the interaction between lengthening and duration to the end of the phrase shows that non-final lengthening has a strong effect on the probability of creak. The interaction between lengthening and stress indicates that

lengthened syllables with primary or secondary stress are less creaky than lengthened unstressed syllables, while the difference is greater for lengthened syllables with secondary stress.

### 3.3. Creaky voice & syllable rank

To see whether syllable rank plays a role in which syllables creak more and which less, we investigated words containing up to five syllables. According to the model, there is an overall tendency for the later syllables to be creakier than the syllables in the beginning of a word. More specifically, there is a very significant increase in the occurrence of creak in syllables nr three and four, while the increase in syllable nr five is evident, but less significant.

### 3.4. The timing of creaky voice within the phrase

The timing of creaky voice was calculated by measuring the duration from the beginning of the syllable to the end of the corresponding phrase. Results indicate that the probability of creakiness increases significantly towards the end of the phrase. Figure 2 shows the percentage of creaky syllables during a phrase.

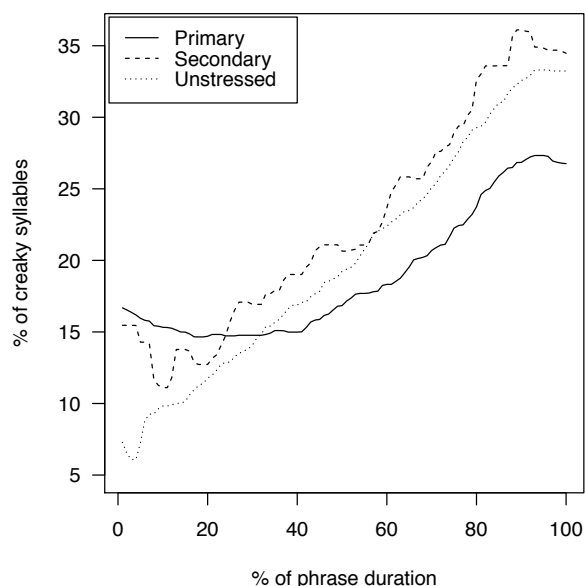


Figure 2: The percentage of creaky voice in syllables with primary and secondary stress and unstressed syllables during a phrase.

The interaction between stress and distance to the end of the phrase is an indication of the sensitivity of creak towards timing in the phrase, depending on the type of stress a syllable is assigned. Relative to unstressed syllables, which become creakier towards the end of the phrase, syllables with primary stress are significantly less creaky near the end of the phrase. Although this result is not significant for syllables with secondary stress, they behave similarly.

### 3.5. Creaky voice, gender & age

While the analysis in this study is not focused on sociolinguistic aspects, we considered some variables of the speakers' background since it has previously been shown that gender and age have a significant impact on the use of creak in spontaneous Estonian. In earlier studies, we have found a

significant effect of speaker and interlocutor gender: male speakers use considerably more creak towards females than towards males, and they use creaky voice significantly more than female speakers. Female speakers use less creak towards male interlocutors. In addition, creak is used less when speaking to an older interlocutor [7, 16]. Although the results presented in this study are based on the same corpus as in [7, 16], the subset used here is considerably smaller and all the gender-age combinations observed in [7] are not represented in this subset.

The results of the present study show that while the speaker's gender is not significant, there is a significant effect of the interlocutor's gender (see Figure 3) – when speaking with a male partner, the speaker exhibits considerably less creak. Additionally, the larger the age difference between speakers, the less the younger interlocutor uses creak. There is a significant interaction between the interlocutor's gender and the speaker's age difference, indicating that there is more creak in a speaker's voice when speaking to a younger female.

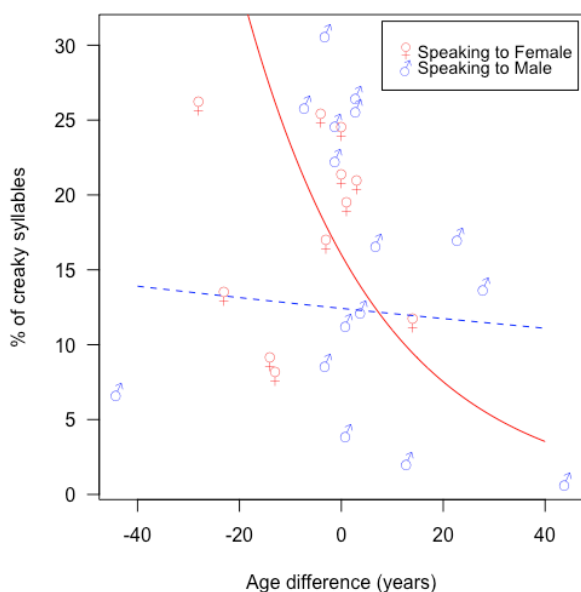


Figure 3: *The percentage of creaky syllables as a function of the interlocutor's gender and the age difference between the speaker and the interlocutor. The regression lines (red for female and blue for male interlocutors) show the probability of the syllables being creaky, estimated by the logistic mixed effects regression model (Table 3).*

#### 4. Discussion

The results presented in this study suggest that the use of creaky voice in spontaneous Estonian can be related to three types of factors: (the lack of) prominence, timing-related properties, and social factors.

Firstly, in relation to prominence we can see that the grammatical words and verbs are generally more creaky than content words that tend to be more acoustically prominent (louder, with higher pitch). Within content words, there is a pattern of primary stressed syllables to be less creaky and unstressed syllables more creaky, again mirroring acoustical prominence of individual syllable. All these findings support the hypothesis by [10] suggesting creak is a manifestation of the lack of prominence.

Interestingly, however, this prominence-creak relationship does not seem to hold universally. Secondary stressed syllables, presumably more acoustically prominent than the unstressed ones, tend to be the most creaky. This pattern can, in part, be related to the pitch movement, where creak is related to low pitch. Also, despite the general trend for the content words to be the least creaky, this only holds for their stressed syllables: the unstressed syllables are actually more or at least equally creaky in content words than in other two word types. This finding of the creak suppression in the primarily stressed syllables and relative increase in the unstressed ones in the content words suggest that creak might play, to some extent, a communicative function of enhancing the contrast between the prominent and less prominent constituents in the content words.

Secondly, creaky voice is influenced by a number of time-related properties, such as lengthening and timing in words and phrases. In particular, creak is favoured by (non-final) lengthening and later timing among the syllables in words and well as within phrases. The phrase final creak is partly again related to the breathing pattern and the pitch movement: the more towards the end of the phrase, the less air we have in our lungs and the lower the pitch falls, due to subglottal pressure becoming weaker. Listeners can interpret this as a signal of turn end, triggering a turn change between interlocutors. Moreover, the speaker can manipulate the use of creak in order to hold the floor by not creaking or for signalling the interlocutor of a possible turn change by creaking.

Finally, the gender and age of the interlocutors also significantly affect the amount of creak. The general pattern of creaking less when speaking with a male interlocutor and more when speaking to a younger female is open to many possible interpretations connected to the nature of the speakers' relationship, attitudes in the society, domination strategies in conversations, etc. – and might be of interest to both linguists and sociologists.

#### 5. Conclusions

The purpose of this study was to observe the use of creaky voice in spontaneous Estonian in connection with syllabic properties of words, lexical stress, lengthening, word class, and timing in phrases. In Estonian, creaky voice is frequently used by all speakers and the feature has therefore been under observation before with a sociolinguistic focus, but not from a detailed prosodic perspective. As creak has so far been excluded from studies looking into stress patterns of Estonian as well, and since creakiness could potentially be connected to those patterns, we focused on a number of prosodic parameters to achieve a better understanding of the occurrence of creaky voice and its realization with regard to lexical stress in Estonian. Our results show that the occurrence of creak is motivated by a number of prosodic features, in addition to sociolinguistic factors favouring it. Most importantly, creakiness occurs during unstressed syllables, especially with lengthening and at the ends of words and phrases.

#### 6. Acknowledgements

The study was supported by the Estonian Research Council grant IUT2-37, the National Program for the Estonian Language Technology project EKT71, and by the Academy of Finland Digital Language Typology project (No. 129XXX).

## 7. References

- [1] M. Gordon and P. Ladefoged, "Phonation types: a cross-linguistic overview," *Journal of Phonetics*, 29, no. 4, pp. 383–406, 2001.
- [2] R. Ogden, "Turn transition, creak and glottal stop in Finnish talk-in-interaction," *Journal of the International Phonetic Association*, 31, no. 1, pp. 139–152, 2001.
- [3] R. Carlson, J. Hirschberg, and M. Swerts, "Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates," *Speech Communication*, 46, no. 3–4, pp. 326–333, 2005.
- [4] J. Laver, *Principles of phonetics*. Cambridge; New York, NY: Cambridge University Press, 1994.
- [5] R. J. Podesva, "Phonation type as a stylistic variable: The use of falsetto in constructing a persona," *Journal of Sociolinguistics*, 11, no. 4, pp. 478–504, 2007.
- [6] J. Kane, K. Pápay, L. Hunyadi, and C. Gobl, "On the use of creak in Hungarian spontaneous speech," in *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong, pp. 1014–1017, 2011.
- [7] K. Aare, P. Lippus, and J. Šimko, "Creaky voice in spontaneous spoken Estonian," in *XXVII Fonetikan Päivät. Turku 25.–26. Lokakuuta 2013. Konferenssijulkaisu*, K. Jähi and L. Taimi, Eds. Turku: Turun Yliopisto, pp. 27–35, 2014.
- [8] J. E. Rodgers, "Vowel devoicing/deletion in English and German," *AIPUK*, 32, pp. 177–196, 1997.
- [9] K. J. Kohler, "Articulatory reduction in German spontaneous speech," in *Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech Production Modelling & 4th Speech Production Seminar. Atrants*, pp. 1–4, 1996.
- [10] P. Lippus, E. L. Asu, and M. L. Kalvik, "An acoustic study of Estonian word stress," *Proc. Speech Prosody 2014*, pp. 232–235, 2014.
- [11] M. Gordon, "Phonetic correlates of stress and the prosodic hierarchy in Estonian," in *Estonian prosody: Papers from a symposium*, I. Lehiste and J. Ross, Eds. Tallinn: Institute of Estonian Language, pp. 100–124, 1997.
- [12] A. Eek and E. Meister, "Simple perception experiments on Estonian word prosody: Foot structure vs. segmental quantity," in *Estonian prosody: Papers from a symposium*, pp. 71–99, 1997.
- [13] M. L. Kalvik and M. Mihkla, "Modelling the Temporal Structure of Estonian Speech," in *Baltic HLT*, pp. 53–60, 2010.
- [14] A. Eek and E. Meister, "Quality of standard Estonian vowels in stressed and unstressed syllables of the feet in three distinctive quantity degrees," *Linguistica Uralica*, 34(3), pp. 226–233, 1998.
- [15] P. Lippus, E. L. Asu, P. Teras, and T. Tuisk, "Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian," *Journal of Phonetics*, 41, no. 1, pp. 17–28, 2013.
- [16] K. Aare and P. Lippus, "Some gender patterns in Estonian dyadic conversations," in *Nordic prosody. Proceedings of the XIIIth conference, Trondheim*, to appear.
- [17] P. Lippus, "Phonetic Corpus of Estonian Spontaneous Speech v.1.0.3," [Online]. Available at: <https://doi.org/10.1515/1-00-0000-0000-0000-0012BL>, 2016.
- [18] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program], version 5.4.18, retrieved from <http://www.praat.org/>, 2016.
- [19] T. Drugman, J. Kane, and C. Gobl, "Data-driven detection and analysis of the patterns of creaky voice," *Computer Speech & Language*, 28, no. 5, pp. 1233–1253, 2014.
- [20] D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, doi:10.18637/jss.v067.i01, 67, no. 1, 1–48, 2015.
- [21] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>, 2015.
- [22] J. Brunner and M. Žygis, "Why do glottal stops and low vowels like each other," in *Proceedings of the 17th international congress of phonetic sciences*, pp. 376–379, 2011.