Creaky voice in spontaneous spoken Estonian

Kätlin Aare[1], *Pärtel Lippus*[2,3], *Juraj Šimko*[3]

[1] Department of Linguistics, Stockholm University, [2] Institute of Estonian and General Linguistics, University of Tartu, [3] Institute of Behavioural Sciences, University of Helsinki

k2tlin.a@hotmail.com, partel.lippus@ut.ee, juraj.simko@helsinki.fi

Creaky voice has many functions in spoken communication. In a number of languages (e.g. in Finnish, Swedish, English) it has been observed to function as a turn-yielding discourse marker. The use of voice quality has been associated with sociolinguistic variation, e.g, it is more frequent in informal settings. The use of different voice qualities has been ascribed to specific social groups, e.g., in American English creak is a characteristic of young educated women. Here we observe the use of creaky voice in Estonian. The data consist of spontaneous monologues and dialogues from the Phonetic Corpus of Estonian Spontaneous Speech. The results show that creak appears mainly in informal situations. They reveal an interesting sociolinguistic pattern where creakiness depends on the age and gender of dialogue partners. While creak tends to occur more likely towards the end of a turn, it does not seem to serve directly as a turn taking cue; our data suggest that it in fact suppresses overlaps by interlocutors.

Keywords: creaky voice, sociolinguistic variation, discourse patterns

1. Introduction

Creaky voice is a phonation type typically associated with tightly adducted vocal folds open enough along a portion of their length to allow for voicing. Acoustically, it is a series of irregularly spaced vocal pulses with a decreased acoustic intensity and a lower fundamental frequency compared to modal phonation, although the lowering effect for f0 is not a universal feature (Gordon & Ladefoged 2001).

Cross-linguistically, creaky voice has many functions. A common example is turn-yielding. In Finnish, creaky voice is used systematically in places where it is relevant for turn transition to occur (Ogden 2001). A comparable feature has been reported from Swedish – utterance fragments with final creak signal boundaries more strongly (Carlson et al. 2005). When *yeah* is said in a creaky voice in American English, it signals recipiency and requests for a change in topic, thus occurring speech-finally (Grivičić & Nilep 2004). Creaky voice is also a turn ending marker in English (RP) used regularly with a lowering fundamental frequency to cue the turn (Laver 1994).

Different situations demand a different language use and phonation types are used accordingly. Creaky voice is most prominent in informal situations (Podesva 2007). In Hungarian, creakiness can refer to emotions with a lower emotional activity, such as boredom or relaxation (Kane et al. 2011). In Estonian, creak is also used most often in informal spontaneous dialogues (Aare 2013). If creaky voice has been used throughout the whole utterance in English (RP) it is perceived as boredom, but could be habitual and a part of the speaker's identity (Laver 1994).

Voice quality can distinguish between social groups. For example, young female speakers of American English in California use creakiness significantly more than men and Japanese women with comparable backgrounds (Yuasa 2010). Another study shows that college-aged young women in Virginia have very creaky voices (Lefkowitz & Sicoli 2007). Generally, creakiness has been interpreted as a sign of masculinity or authority and is associated with high social status. Young educated females might thus use creaky voice with a very low fundamental frequency to project an authoritative image (Yuasa 2010).

Speakers exhibit significant variation in voice quality across situations. Podesva (2007) examined intraspeaker variation in the speech of a gay student who uses phonation types depending on the situation: modal voice quality for professional purposes and when communicating with his family, but falsetto to construct a diva persona. He suggests that this speaker uses phonation types to expand his f0 range to the physical extreme, with falsetto at the high end and creaky voice at the low. Another example are the girls belonging to the Chicana gang in Northern California who use creaky voice when they need to construct a hardcore persona (Mendoza-Denton 2011).

Previous studies have shown that voice quality, especially creaky voice, is exploited differently across languages and among various social groups and communicative situations. While there is evidence of certain groups using it more, it is unclear whether it is the same for Estonian or whether creakiness is used only when speaking with somebody from a specific social group. In addition, so far there is no plausible proof of creaky voice being used as a turn-transition marker in Estonian.

2. Material and Methods

The data comes from the University of Tartu Phonetic Corpus of Estonian Spontaneous Speech (http://www.keel.ut.ee/foneetikakorpus). This corpus consists of recordings of monologues and dialogues, hand labelled on phoneme level. The dialogues are between two speakers and carried out in a sound-detected room, the monologues are academic lectures recorded in a lecture hall with a

head-mounted microphone. All speakers were aware of being recorded for the corpus. The recordings have been made 2006–2013.

Table 1. Number of speakers in dialogues and monologues grouped by gender and age.

	Dialogue		Monologue	
	F	M	F	M
older	11	14	5	5
younger	19	13	1	1

For this study 47 hours of speech was used: 39 hours of dialogues and 8 hours of monologues from 64 speakers (Table 1). Five speakers participated in both a dialogue and a monologue, 7 speakers participated in more than one dialogue. The duration of each recording is approximately 30 minutes, varying from 15–50 minutes. The age ranges from 20 to 85 years, the mean being 36.5 years. The speakers were divided into two age groups based on the year of birth 1973.

The creaky voice in the recordings has been assessed in three steps. Firstly, the labellers have noted the voice quality during the phoneme level segmentation in Praat. Then, creaky voice has been detected with the Voice Analysis Toolkit (https://github.com/jckane/Voice_Analysis_Toolkit). Finally, the manual and automatic labelling was manually checked and corrected. There were misses and false alarms in both manual and automatic creak detection, agreement between the two methods was approximately 70%.

3. Results

For each speaker, the amount of creakiness was calculated as a percentage of being creaky over the time being in voice. Figure 1 compares the percentage of creakiness between speaking style (dialogue vs. monologue) and speaker gender and age group.

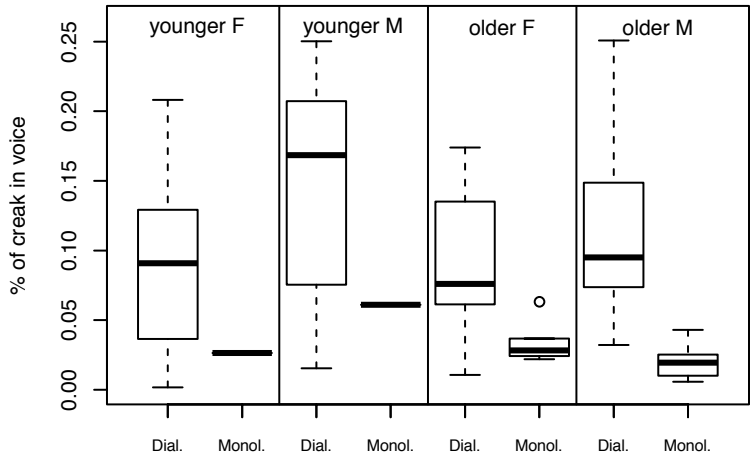

Figure 1. The amount of being creaky when being in voice as a function of speaking style, gender and age group.

Table 2. The minimal linear mixed effect model of creakiness in monologues and dialogues with significant interactions.

	Estimate	*t* value
Intercept	0.095	9.686
gender *M*	0.033	2.295
Style *monologue*	-0.056	-2.689
gender *M*:style *monologue*	-0.058	-2.003

The amount of creakiness varied greatly from speaker to speaker from 0.1% to 25%, but there was no speaker who did not creak at all. The amount of creakiness was tested with a mixed model with *age*, *gender* and *style* as fixed effects and the speaker as a random effect. Table 2 contains the significant main effects and interactions revealed in a minimal model. The speaker *age* did not have a significant effect. The intercept yields the estimate of average percentage of creak for female speakers in dialogues (about 10%). There was a significant effect of *gender*: in dialogues, the male speakers creak about 3% more than females. The significant main effect of *style* shows less amount of creakiness in monologues than in dialogues (decrease by 5.6% for females). As shown by the interaction the male speakers decrease the percentage of creakiness in monologues significantly more than the female speakers (by approximately further 6%).

Looking at dialogues we asked to what extent was the amount of creakiness influenced by characteristics of the interlocutors. Figure 2 shows the percentage of creaky voice as a function of the speaker's (SP1) age and gender depending on the speaking partner's (SP2) age and gender.

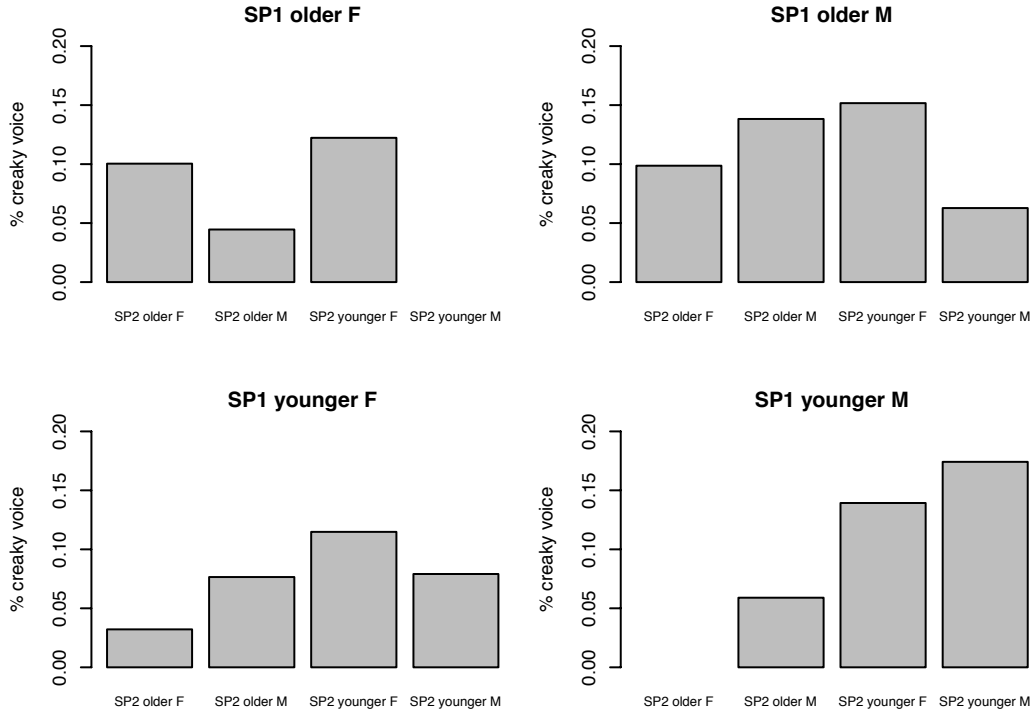

Figure 2. The percentage of creakiness depending of the speaking partner's age and gender.

The rather complicated patterns shown in Figure 2 are illuminated by a linear mixed model with *SP1 gender* and *age* and *SP2 gender* and *age* as fixed effects and the speaker as a random effect, see Table 3. Consistently with the above analysis, there was a strong main effect of *SP1 gender*, male speakers creaking considerably more than females. The main effects of speakers' *age* were not significant, although the results suggest that creakiness tends to decrease when speaking to older partners. There were two significant interactions. The *SP1 age*:*SP2 gender M* interaction shows that when speaking to males the percentage of creaky voice significantly decreased with speakers age, older speakers creaked significantly less. On the other hand, *SP1 gender M*:*SP2 age* indicates that male subjects were creaking significantly less when speaking to older dialogue partners.

Table 3. The minimal linear mixed effect model including both dialogue partners with significant interactions.

	Estimate	*t* value
(Intercept)	0.1196	5.104
SP1 gender *M*	0.1099	3.442
SP1 age	0.0002	0.307
SP2 age	-0.0006	-1.597
SP1 age:SP2 gender *M*	-0.0006	-2.708
SP1 gender *M*:SP2 age	-0.0021	-2.786

Next we ask how is creakiness distributed within a dialogue turn, and how is this distribution linked to overlaps between the partners. We look at the relative occurrence of overlaps and its dependence on creakiness. The speech was divided into Inter-Pausal Units (IPU) containing speech between two pauses longer than 15 ms. A typical IPU in our data ranged between 0.5 and 2 seconds and contained 2–8 words.

For each IPU, we extracted 100 points equally distributed in time which were checked for creakiness. Also for each of these points we registered whether the second speaker was in voice. Similarly, we logged whether SP2 was speaking during the 200 ms following the SP1's IPU. For both the SP1's IPU and the SP2 action we also checked, whether it contained lexical words or backchannels (fillers, words such as *ahah, mhmh, ja*). These measurements were used to estimate the likelihood of creakiness and overlaps within the normalized IPUs by dividing the number of cases with creaky voice and overlap, respectively, by the total number of all relevant IPUs.

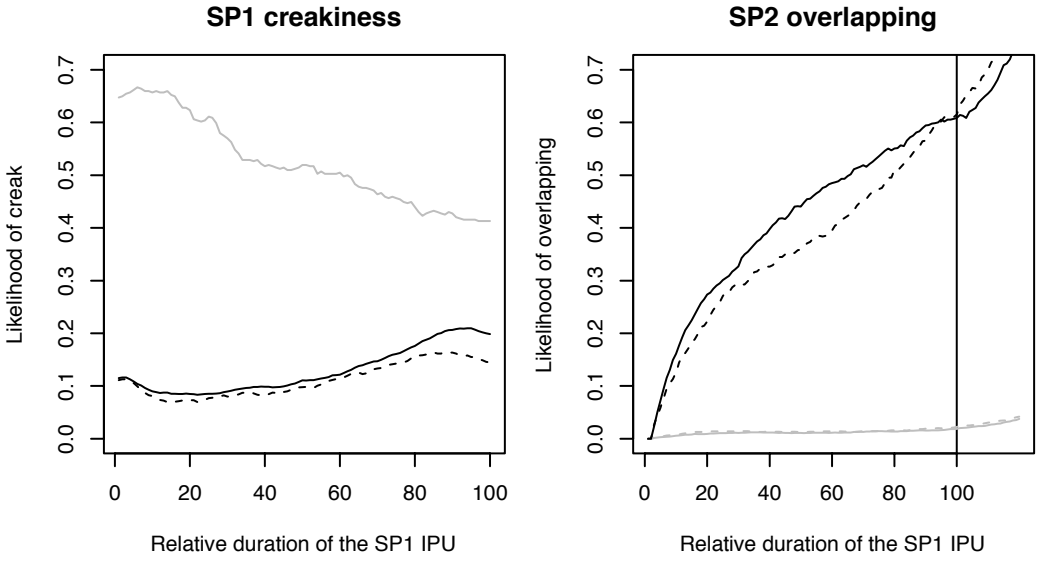

Figure 3. Left: Probability of SP1 being creaky during lexical speech when SP2 is not overlapping (black solid line) and when SP2 is overlapping (black dashed line), and when SP1 is backchanneling (grey). Right: Probability of SP2 overlapping during SP1's IPU with lexical speech when SP1 is modal (black solid line) and creaky (black dashed line), and with backchanneling (grey).

Figure 3 depicts the results, the left panel shows the likelihood of creakiness within speaker's IPU and the right panel the probability of overlapping speech by SP2. For lexical speech (non-backchannels), the creakiness increases towards the end of IPU. The likelihood of speech being creaky also seems to be lowered when there is an overlap by the second speaker within the IPU. In contrast with lexical speech, there is more creakiness at the beginning than at the end of SP1's own backchannels.

The likelihood that the second speaker overlaps SP1's lexical IPU with lexical speech is greater when the SP1 speaks with modal voice than when there is a creak within his or her IPU. The presence and distribution of backchannels does not seem to be influenced by creak.

This analysis was also performed separately for SP1 gender and age groups, but the results were qualitatively the same as those pictured in Figure 3, with the exception of older male speakers for which the percentage of creakiness was greater when there was an overlap within their IPU.

4. Discussion and Conclusions

Our data show that creakiness occurred far more often in dialogues than in monologues. It should be noted that the monologues were recorded in a lecture hall before audience, while the dialogues were recorded in a studio with speakers sitting close to each other. Consequently, the settings are

not equal; the monologues can be less creaky just because the speaker must speak up to be heard. When speaking louder and with a higher pitch the voice is less likely to creak.

Our results do not reveal a straightforward link between creakiness and turn-taking. The tendency to creak towards the end of turns might be associated with the declining pitch and intensity within a phrase. Our findings suggest that creakiness in fact somewhat inhibited the overlapped speech and the likelihood of overlap was greater during turns containing only modal voice. A possible explanation is that creaky speech is more difficult to parse and therefore requires greater attention from the interlocutor and impedes her propensity for taking the floor.

The age and gender influence the creakiness in a manner that is open to sociolinguistic interpretations. Creak might be associated with informal attitude and reflect the traditional social hierarchy. Indeed, at least male subjects creaked significantly less when speaking to older dialogue partners. Also, the tendency of older speakers creaking more to females can be seen as supporting this interpretation. However, it must be noted that everybody does creak to some extent and before reaching general conclusions, more thorough analysis needs to be performed. For example, the current version of corpus is not perfectly balanced (missing some age/gender combinations of partners).

Acknowledgements

The study was supported by the Estonian Research Council grant IUT2-37, and the National Program for the Estonian Language Technology Project EKT4.

References

Aare, K., 2013. *Kärin eesti keele spontaanses kõnes*. BA Thesis. Tartu: University of Tartu. Available at: http://hdl.handle.net/10062/30862.

Carlson, R., Hirschberg, J. & Swerts, M., 2005. Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. *Speech Communication*, 46(3-4), pp. 326–333.

Gordon, M. & Ladefoged, P., 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29(4), pp.383–406.

Grivičić, T. & Nilep, C., 2004. When phonation matters: The use and function of yeah and creaky voice. *Colorado Research in Linguistics*, 17(1). Available at: http://www.colorado.edu/ling/CRIL/Volume17_Issue1/paper_GRIVICIC_NILEP.htm.

Kane, J. et al., 2011. On the use of creak in Hungarian spontaneous speech. In W.-S. Lee & E. Zee, eds. *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.

Laver, J., 1994. *Principles of phonetics*, Cambridge ; New York, NY: Cambridge University Press.

Lefkowitz, D. & Sicoli, M., 2007. Creaky voice: Constructions of gender and authority in American English conversation. In 106th annual meeting of the American Anthropological Association. Washington, DC.

Mendoza-Denton, N., 2011. The semiotic hitchhiker's guide to creaky voice: Circulation and gendered hardcore in a Chicana/o gang persona. *Journal of Linguistic Anthropology*, 21(2), pp. 261–280.

Ogden, R., 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association*, 31(1), pp. 139–152.

Podesva, R.J., 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, 11(4), pp. 478–504.

Yuasa, I.P., 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85(3), pp. 315–337.