Individual speaker characteristics of creaky phonation: a comparative study of English and Urdu

Gea de Jong-Lendle[1]*, Sarmad Hussain* [2]*, and Francis Nolan*[3]
[1]*Institut für Germanistische Sprachwissenschaft, Philipps-Universität Marburg,Germany*
gea.dejong@staff.uni-marburg.de
[2]*CLE, Al-Khawarizmi Institute of Computer Science, UET, Lahore, Pakistan*
sarmad.hussain@kics.edu.pk
[3]*Institu Dept. of Theoretical & Applied Linguistics, University of Cambridge, UK*
fjn1@cam.ac.uk

The definition and the role of creak

Creak, also called vocal fry or glottal fry, has been defined by Catford (1964:32) as: ' Low frequency (down to about 40Hz) periodic vibration of a small section of the vocal folds.' Although the phenomenon of creak has received quite some attention in the past (Michel and Hollien 1968, Hollien et al 1966, Moore 1971, Coleman 1963) the exact physiological mechanics of creak are still unknown. The typical auditory quality of creak can be described as a series of separate taps in rapid sequence and it is therefore easily audible. In a spectrogram creak shows up as slightly irregular vertical striations. See Laver (1980) for a detailed overview on creak. In many tone languages syllables with low or falling tones are often accompanied with this type of phonation. In languages like English, creak has often been associated with the paralinguistic feature of turn taking; a speaker can use a falling intonation and creak as a signal that his/her turn has been accomplished, yielding the floor to the listener. When a speaker uses creak throughout an entire utterance or sentence, it is assumed that he/she is bored (Laver 1980:126) or wants to signal indifference.

Creak in a forensic context

Despite the fact that creak can almost be considered as a normal part of communication and is therefore present in a large number of speech recordings, it has received comparatively little attention within the forensic community. In Hudson et al. (2007) an early attempt was made as part of a larger project on the F0 statistics of 100 Standard Southern British English (=SSBE) speakers. Here it was found that, with the minimum threshold of 50Hz, creak showed up in the histogram in 66% of the cases. In 34% of the cases it can be assumed that either 1) creak is absent, or 2) creak was not measured as it was below the lower F0 threshold. The latter possibility being the more likely one of the two. In 3% of the cases the F0-mode was actually found in the creaky part of the H0 histogram instead of the F0- range covering the modal voice. In other words, in the bimodul distribution the creak-peak was higher than the modal peak. Observing the different bimodul histogram patterns, all speakers could be largely divided into 5 different groups.

Objectives

The aim of this investigation is to explore into more detail the speaker identifying potential of creak and creaky phonation.

The following questions are of interest:

1. Per speaker the percentage of creak in a 3 min. segment of normal spontaneous speech

2. F0 histograms when F0 measured with a F0 minimum threshold below 50Hz
3. Is the starting F0-point for creak the same for a falling intonation pattern as for a rising pattern?
4. Could the above patterns for the use of creak be different across languages/ethnological backgrounds? For example, what happens when speakers of a particular language speak at a much higher average speaking F0? Can one expect to find less creak? Here, we would like to compare English with Urdu as it was found that in SSB-English the average F0 mode was 102Hz (Hudson et al. 2007) and in Lahore-Urdu this value was calculated to be 129Hz (deJong et al. 2012).

As creak is difficult to measure due to its irregular behaviour, a 'quality control' is included to compare auditory detection of creak with f0 values produced by PRAAT. In addition, stretches of low f0 are examined to check that these correspond to auditory judgments, and aren't just e.g. halving errors occurring for other reasons.

If the speaker identifying potential is confirmed measuring groups of a small size (n=30), data could then be generated for a larger population. The forensic community could then be provided with probabilistic data concerning the use of creak.

Materials

Speakers: 30 male speakers of SSB-English + 30 speakers of Lahore-Urdu
Speaking style: spontaneous
Software: PRAAT (histograms are produced by a PRAAT script)

The SSB-English speaker recordings come from the DyViS -project in Cambridge. For a detailed description see Nolan et al. 2006a. The Lahore-Urdu speaker recordings come from the URDU-project in Lahore. For a detailed description see Sarfraz et al. 2010.

References

Catford, J.C. (1964) Phonation types: the classification of some laryngeal components of speech production, in Abercrombie, D. et al (19064) pp. 26-37.

Coleman, R.F. (1963) Decay characteristics of vocal fry, *Folia Phoniatrica* **15**: 256-63.

deJong,G., Hussain, S., Irtza, S. and Hudson, T. (2012) F0 statistics for speakers of Urdu, Proceedings, IAFPA 2012, Santander, Spain.

Hollien, H., Moore, P., Wendahl, R.W. and Michel, J.F. (1966) On the nature of vocal fry, *Journal of Speech and Hearing Research*, **9**:245-7.

Hudson, T., de Jong, G., McDougall, K., Harrison, P. and F. Nolan, (2007). F0 statistics for 100 young male speakers of Standard Southern British English. ICPHS conference proceeding, In: J. Trouvain and W. Barry.

Laver, J. (1980) *The phonetic description of voice quality*, Cambridge University Press.

Michel, J.F. and Hollien, H. (1968) Perceptual differentiation in vocal fry and harshness, *Journal of Speech and Hearing Research*, **11**:439-43.

Nolan, F., McDougall, K., de Jong, G. and Hudson, T. (2006a). 'A forensic phonetic study of 'dynamic' sources of variability in speech: the DyViS project.' In: P. Warren and C.I. Watson (eds.), Proceedings of the 11th Australasian International Conference on Speech Science and Technology, 6-8 December 2006, Auckland: Australasian Speech Science and Technology Association, 13-18.

Sarfraz, H., Hussain, S., Bokhair, R., Raza, A.A., Sarfraz, Z., Pervez, S., Mustafa, A., Javed, I. and Parveen, R. (2010). Speech corpus development for a speaker independent spontaneous Urdu Speech recognition system, *Proceedings of O-COCOSDA*, Kathmandu, Nepal.