

Creaky voice variation across language, gender and age in Canadian English-French bilingual speech

Jeanne Brown¹ and Morgan Sonderegger¹

¹McGill University (Canada)

Introduction: This study addresses how non-contrastive voice quality varies among languages and across speakers (as a function of gender and age). *Creaky voice* is characterized by a bundle of acoustic properties, including low pitch, irregular voicing, and decreased glottal airflow. [1] find more creak in English in English-French bilingual speech (in Paris). Conversely, [2] find that Cantonese-English bilinguals' voice quality remains relatively constant across languages. Voice quality in Canadian English-French bilinguals offers a unique opportunity to investigate whether these close-contact varieties exhibit distinct or similar vocal qualities. In English, creak has often been described as an indexical feature, yet there is no consensus on gender differences. Some find more creak in men's speech [3,4], typically using acoustic data, while others find more creak for women [5,6], typically relying on impressionistic audio-visual cues. These inconsistent gender effects may result from methodological heterogeneity and motivate the use of qualitatively different measures to quantify creak. This study asks: 1) Comparing Canadian English and French, how does creaky voice differ across languages? 2) What gender and age differences, if any, emerge when using different measures of creak, and what does this reveal about creak as a sociolinguistic marker?

Method: Spontaneous speech from 9 English-French bilingual speakers born and raised in Ontario or Québec was collected from publicly available online data sources, amounting to roughly 5 minutes of speech per speaker-language pair and 14k vowels in total. The speaker sample will reach 40 by the conference. Data treatment features an automated pipeline (with extensive manual checks), implementing state-of-the-art open-source tools for speech analysis. Audio files were speaker diarized using PyAnnote [7] and transcribed automatically using OpenAI Whisper [8]. Timestamped transcripts were then force aligned using the Montreal Force Aligner [9]. Acoustic analysis consisted of pitch tracking in Praat, providing a proportion of unreliable f0 tracks for each vowel, as well as acoustic measures extracted using PraatSauce [10]: one spectral slope measure (H1*-H2*) as a correlate of glottal constriction, and two Harmonics-to-Noise Ratios (CPP and HNR from 0-500Hz) as indicators of waveform periodicity. Praat voice analysis measures (e.g., jitter and shimmer) will also be included. Statistical significance of social and linguistic effects on acoustic properties of creak was tested using mixed-effects regression models, with fixed effects of language, gender, age and utterance position, and maximal by-word and by-speaker random effects.

Results: The main results for gender (Figure 1 and 2) show that men's vowels have more unreliable f0 tracks ($p < 0.0001$), lower H1*-H2* ($p = 0.03$), lower HNR05 ($p = 0.04$), and somewhat lower CPP ($p = 0.07$), suggesting more glottal closure and higher levels of noise/aperiodicity for male speakers. Male speakers are creakier overall, aligning with studies using phonetic measures [4] but in contrast to those employing impressionistic coding [5,6]. This supports a physiological and/or socially-constrained difference, but crucially reveals that the acoustics of creaky voice are distinct from its perception. Regarding language, English displays more unreliable pitch tracking compared to French ($p < 0.0001$) (Figure 1), but other acoustic correlates of creak do not suggest any cross-linguistic differences (Figure 3). The bilinguals' speech fails to provide definitive evidence that English is creakier than French. Results on age will be presented as well.

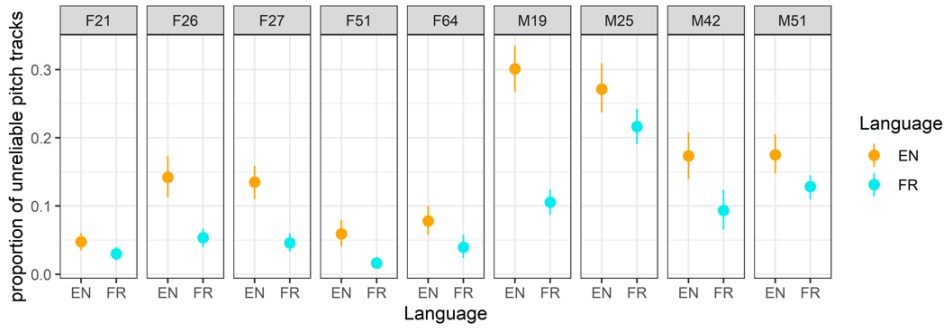


Figure 1: Proportion of unreliable pitch tracks (pitch tracked over < 50% of the vowel) by speaker and language.

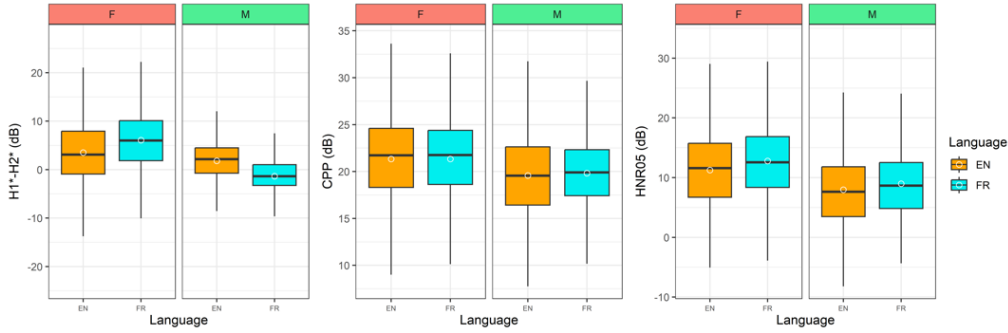


Figure 2: Acoustic correlates of creak plotted by gender (female vs. male) and language.

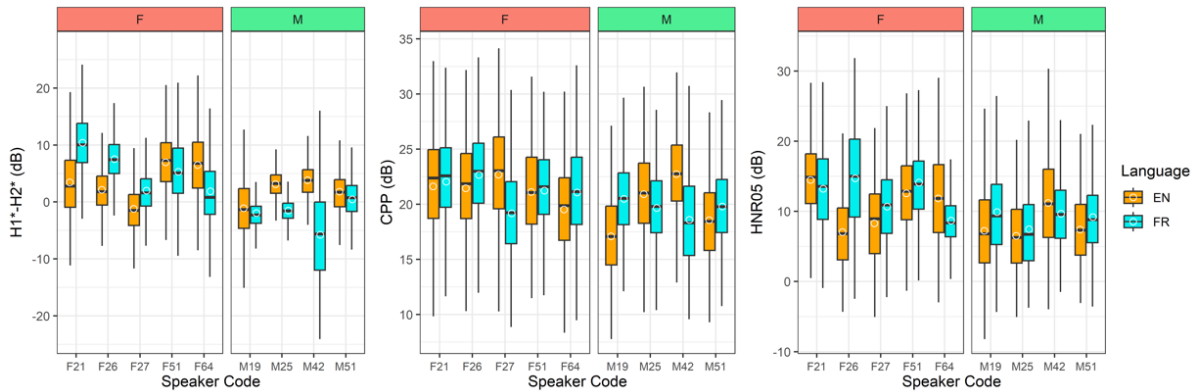


Figure 3: Acoustic correlates of creak plotted by gender, language, and speaker.

References

- [1] Pillot-Loiseau, C., et al. (2019) The evolution of creaky voice use in read speech by native-French and native-English speakers in tandem: a pilot study. *Anglophonia*, 27.
- [2] Johnson, K., & Babel, M. (2023). The structure of acoustic voice variation in bilingual speech. *JASA*, 153(6), 3221-3238.
- [3] Henton, C., & Bladon, A. (1988). Creak as a sociophonetic marker. In L. Hyman & C. Li (Eds.), *Language, speech and mind: Studies in honor of Victoria A. Fromkin* (pp. 3–29).
- [4] Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *JASA*, 87(2), pp. 820–857.
- [5] Podesva, R. J. (2013). Gender and the social meaning of non-modal phonation types. In C. Cathcart et al. (Eds.), *Proc. Berkeley Linguistics Society* 37, pp. 427-448.
- [6] Yuasa, I. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85, pp. 315-337.
- [7] Bredin, H. (2023). “pyannote.audio 2.1 speaker diarization pipeline: principle, benchmark, and recipe”. In *Proc. INTERSPEECH 2023*.
- [8] Radford, A., Kim, J. W., Xu, T., Brockman, G., Mcleavey, C. & Sutskever, I. (2023). Robust Speech Recognition via Large-Scale Weak Supervision. *Proceedings of the 40th International Conference on Machine Learning, in Proceedings of Machine Learning Research*, 202: 28492-28518. Available from <https://proceedings.mlr.press/v202/radford23a.html>
- [9] McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. *Proc. INTERSPEECH 2017*. pp. 498-502. 10.21437/Interspeech.2017-1386
- [10] Kirby, J. (2018). Praatsauce. Available at: <https://github.com/kirbyj/praa sauce>.