CREAKY VOICE IN CANADIAN ENGLISH: AN ACOUSTICS-FOCUSED METHOD

Jeanne Brown *1
[1] Department of Linguistics, McGill University, Montreal, QC, Canada

1 Introduction

Creaky voice is a voice quality attributed to vocal fold compression without complete glottal closure. Acoustically, prototypical creaky voice is characterized by three key properties: a low pitch, irregular vocal pulses, and decreased transglottal airflow [1].

Previous work provides competing evidence of gender differences with regard to creaky voice in English: some studies find more creak in men's speech [2, 3] whereas more recent studies frequently report increased creakiness among women [4, 5]. Many of these existing sociophonetic studies of creaky voice rely on impressionistic (auditory and/or visual) coding which are vulnerable to important perceptual biases [6, 7]. The present study aims to investigate the often-cited gender differences in creaky voice use by examining Canadian English, crucially quantifying creak by its acoustic correlates.

2 Method

2.1 Speech sample

ICE-Can: A subset of 29 speakers (three girls, seven women and 19 men) from the International Corpus of English – Canada (ICE-Can) [8] were selected for this study. All speakers were born and raised in Canada, across multiple provinces. Speaker ages range from 12 to 77 years old at the time of recording. The speech samples are collected from broadcast events from the early 1990s.

YT Corpus: 21 Canadian speakers make up the YouTube (YT) Corpus: 11 men and 10 women, all aged between 19 and 76 years old (at time of recording) and born in Ontario or Quebec. The audio data contains spontaneous speech from YouTube videos (podcasts, interviews, radio shows and livestreams) recorded within the last 10 years.

2.2 Acoustic analysis

Several acoustic voice quality measurements were extracted from all vowels at the mid-point using PraatSauce [9], a Praat script for spectral measures. A total of 55 315 vowels were analyzed. The acoustic correlates of interest in this paper are pitch (f0), spectral tilt (e.g., H1*-H2*) and Harmonics-to-Noise Ratios (HNRs), specifically Cepstral Peak Prominence (CPP) and HNR between 0-500Hz (HNR05).

Pitch (f0) tracking: Pitch tracking errors in Praat are often caused by irregular voicing which is typical of creaky voice. Therefore, the percentage of tokens (vowels) that have unreliable f0 tracks is interesting for identifying possible creaky voice. Pitch tracks were considered unreliable if Praat returned "undefined" or 0Hz values for more than 50% of the vowel duration.

Spectral tilt: Spectral tilt measures are acoustic indicators of glottal constriction [1, 3]. H1*-H2* is the difference between the amplitudes of the first harmonic and the second harmonic. With a more open glottis (increased airflow), H1 values are boosted, resulting in a higher spectral tilt value. On the other hand, with a more constricted glottis (reduced airflow), H1 values are decreased and/or H2 values are increased, resulting in a lower spectral tilt value which suggests more creakiness [3, 10].

Harmonic-to-Noise Ratios (HNRs): HNRs such as CPP or HNR05 measure waveform periodicity and symmetry, both indicators of noise [1, 3]. As the name suggests, they are ratios between the levels of harmonics and noise. High HNR values indicate a very periodic vibration with strong harmonics and little noise and conversely low HNR values indicate either weak harmonics or aperiodic vibration (typically caused by noise like aspiration or creakiness) [10]. HNR calculated in the frequencies below 500Hz (HNR05) is a particularly good indicator of noise for creaky voice because aspiration noise will typically occur at higher frequencies, whereas creak noise tends to affect the lower frequency bands [10].

Data cleaning: Due to the direct link between f0 and spectral tilt measures, any H1*-H2* data resulting from unreliable f0 tracks was removed (16% of all H1*-H2* data). As for CPP and HNR05, all data points that fell outside of three standard deviations of the means (by-speaker) were excluded.

2.3 Statistical analysis

Using a script modified from [11], mixed models were conducted in R. Fixed main effects were gender, age, corpus, utterance position, vowel height and stress; only gender effects will be reported here. Random effects included by-speaker random intercepts.

3 Results

3.1 Pitch track

Table 1 shows that the percentage vowels with unreliable pitch tracks is higher for men (19.2%) than for women (12.0%), suggesting more irregular voicing, an acoustic indicator of creak (among others).

3.2 Spectral tilt measures (H1*-H2*)

The spectral tilt measure selected in this paper, H1*-H2*, shows significant differences for the main effect of Gender. Men's vowels have lower H1*-H2* values than women's (GenderM: $\beta = -23.30$dB, $t = -6.01$, $p < 0.05$), which is also

* jeanne.brown@mail.mcgill.ca

observable in Figure 1 below (left panel). These results suggest that men produce their vowels with a more constricted glottis than women, indicating a tendency towards more creakiness.

Table 1: Percentages and number of vowels with unreliable f0 tracks (proportion of vowel f0 tracked < 0.5) by gender.

Gender	% unreliable	n unreliable	n
F	12.0%	2357	19681
M	19.2%	6851	35634
Total	16.6%	9208	55315

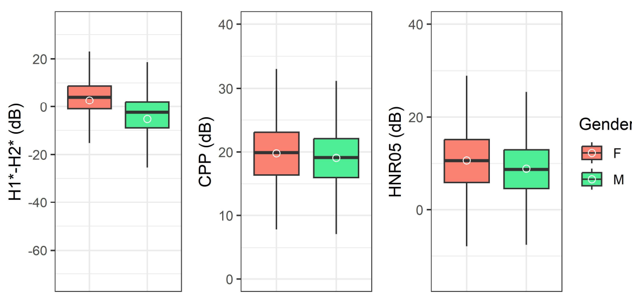

Figure 1: H1*-H2*, CPP and HNR05 values (dB) by gender.

3.3 Harmonics-to Noise-Ratios (HNRs)

Cepstral Peak Prominence (CPP): CPP differences for the main effects of Gender are also significant, men's vowels showing lower CPP values than women's in Figure 1 (middle panel) and in the mixed model summary (GenderM: β = -2.34dB, t = -2.56, p = 0.01). Lower CPP values indicate more noise and/or aperiodicity across all frequency bands, which is indicative of more creakiness.

HNR below 500Hz (HNR05): While Figure 1 (right panel) does seem to show HNR05 differences by Gender, these do not reach statistical significance. The trend does follow previous results for the other acoustic properties of voice quality, with men producing lower HNR05 values than the women (GenderM: β = -1.97dB, t = -1.75, p = 0.09). Low HNR05 values point to more noise in the low frequency band between 0 and 500Hz, which can be attributed to increased creakiness for men compared to women.

4 Discussion

This study presents substantial evidence for more creakiness in men's speech than in women's speech. Men's vowels show less reliable f0 tracks, lower spectral tilt measures (H1*-H2*) and lower HNR measures (CPP and HNR05). Altogether, these acoustic cues indicate more glottal closure and higher levels of noise/aperiodicity, ultimately providing empirical support for more creak among men than women. This result aligns with that of older studies [2, 3] but challenges more recent and well-cited findings [4, 5].

A perceptual explanation for these conflicting findings is that creakiness is much more salient in women's voices because they have higher habitual pitch ranges and often lower their pitch considerably to produce creaky voice. Conversely, men's voices, usually situated at lower pitch ranges, do not require large pitch shifts and as such, creakiness is less perceptually marked. Since most studies examining creaky voice implement impressionistic coding of creaky tokens, a perceptual bias is likely to impact the data. This seems to coincide with work on the interaction of non-modal phonation and pitch [6, 7].

Notable limitations concern the speech sample as well as the data cleaning process and statistical analyses. The current speaker sample is skewed towards men in both quantity of speech and number of speakers. While the precise effect of this gender disbalance is unclear at this time, future work should aim for a more balanced sample. Additionally, the data cleaning and statistical analysis requires more careful consideration to assure that outliers are treated correctly, models are well-calibrated to the data and interactions can be properly interpreted.

Overall, this study highlights the importance of acoustic measures in quantifying creak and provides new insight into the relation between creaky voice and gender.

References

[1] Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. *ICPhS*.

[2] 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). London, England: Routledge.

[3] Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. Journal of the Acoustical Society of America, 87(2), 820–857. https://doi.org/10.1121/1.398894

[4] Podesva, R. J. (2013). Gender and the social meaning of non-modal phonation types. In C. Cathcart, I. Chen, G. Finley, S. Kang, C. S. Sandy, & E. Stickles (Eds.), Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society, pp. 427-448.

[5] Yuasa, I. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? American Speech, 85, 315-337. 10.1215/00031283-2010-018.

[6] Davidson, L. (2020). Contributions of modal and creaky voice to the perception of habitual pitch. Language, 96(1), e22-e37. https://doi.org/10.1353/lan.2020.0013

[7] White, H., Gibson, A., Penney, J., Szakay, A., & Cox, Felicity. (2022). The influence of pitch and speaker sex on the identification of creaky voice by female listeners. In Proceedings of the Eighteenth Australasian International Conference on Speech Science and Technology, ASSTA, 81-85.

[8] Greenbaum, S., & Nelson, G. (1996). The International Corpus of English (ICE project). World Englishes, 15:3–15.

[9] Kirby, J. (2018). Praatsauce: Praat-based tools for spectral analysis. Available at: https://github.com/kirbyj/praatsauce

[10] Keating, P., Kuang, J., Garellek, M., Esposito, C., & Khan, S. (2023). A cross-language acoustic space for vocalic phonation distinctions, Language, 99(2), pp. 351-389. DOI:10.1353/lan.2023.a900607

[11] Brunelle, M., Brown, J. & Thị Thu Hà, P. (2022). Northern Raglai voicing and its relation to Southern Raglai register: evidence for early stages of registrogenesis. Phonetica, 79(2), 151-188. https://doi.org/10.1515/phon-2022-2019