

Gender effects in creaky voice production in English and Mandarin: Men are creakier

Meixian Li¹, Yao Yao¹, Charles. B. Chang²

¹The Hong Kong Polytechnic University, ²City University of Hong Kong

In addition to semantic meanings, creaky voice can carry socio-indexical meanings about speakers' affect, personality, and social identity [1-2] that often interact with their gender. For example, North American English listeners tend to associate creakiness with young women [3-4]. But are women actually creakier than men? Production studies report mixed findings, which are often difficult to compare due to methodological disparities [5-9].

The present study addressed these limitations by examining creaky voice production by American English and Mandarin Chinese speakers using parallel speech corpora. Each corpus consisted of recordings from 40 native speakers (Mandarin: 20F, aged 20-35; English: 20F, aged 18-44), each reading 120 emotion-neutral sentences in the target language. The sentences were controlled for length (8-12 syllables) and designed to be representative of the entire phonological inventory (segments, tones) of the language. Sentence recordings were manually segmented into voiced portions (consecutive vowels and voiced consonants), which were measured for four acoustic features: fundamental frequency (F0), amplitude difference between the first and second harmonics (H1-H2) corrected for F1, harmonic-to-noise ratio (HNR) up to 3500 Hz, and cepstral peak prominence (CPP). For all of these measures, creaky voice shows lower values than modal voice [10], although F0 is more heavily influenced by speaker gender. Separate linear mixed-effects models were fitted for each acoustic measure of the voiced portions ($n=66,941$ over both languages), with Language (English, Mandarin; reference level = English), Gender (female, male; reference level = female), and their interaction as dummy-coded fixed effects and by-sentence and by-speaker random intercepts.

Overall, we found that **men tended to be creakier than women**. In English, there was a significant Gender effect on F0 ($\beta=-83.58$, $p<.001$) and a marginal Gender effect on H1-H2 ($\beta=-2.04$, $p=.055$). There was also a significant Language effect on F0 in females: Mandarin females showed higher F0 than English females ($\beta=34.01$, $p<.001$). Moreover, three out of the four acoustic measures showed a significant Language $\times$ Gender interaction: F0 ($\beta=-18.29$, $p=.024$), H1-H2 ($\beta=-4.77$, $p=.002$), and HNR ($\beta=-4.80$, $p=.047$). No significant effects were found for CPP. As illustrated in Figure 1, the Language $\times$ Gender interactions were mainly due to Mandarin males having particularly low acoustic values, suggesting that this group shows the strongest degree of creakiness of all. Post hoc analysis of estimated marginal means confirmed that for English speakers, the Gender effect was significant only for F0; however, both H1-H2 and HNR showed the same trend of males having lower values. In Mandarin, the Gender effect was significant for F0, H1-H2, and HNR, all showing lower values in males than females; a similar trend was seen for CPP.

Taken together, our findings run counter to the perception that women are creakier than men. The lack of evidence in this study for the alleged female bias in creaky voice production supports previous production studies that challenge the notion of creaky voice as a young female speech style in American English. If anything, our results on English suggest a male bias in creaky voice production, consistent with [5-6]. Furthermore, our results on Mandarin provide even clearer evidence of a male bias in creakiness. This pattern dovetails with recent perception findings showing that creaky male voices are negatively evaluated by Mandarin listeners, especially male listeners [9]. The coexistence of robust creakiness in male speakers and its negative evaluation by peers strongly suggests that, as in English, creaky voice carries gendered socio-indexical meanings in Mandarin (e.g., an affected mature-sounding masculine persona, which listeners dislike). These results contribute new evidence of crosslinguistic variations in the socially patterned use of voice quality and further attest to the importance of linguistic and cultural diversity in sociophonetic research.

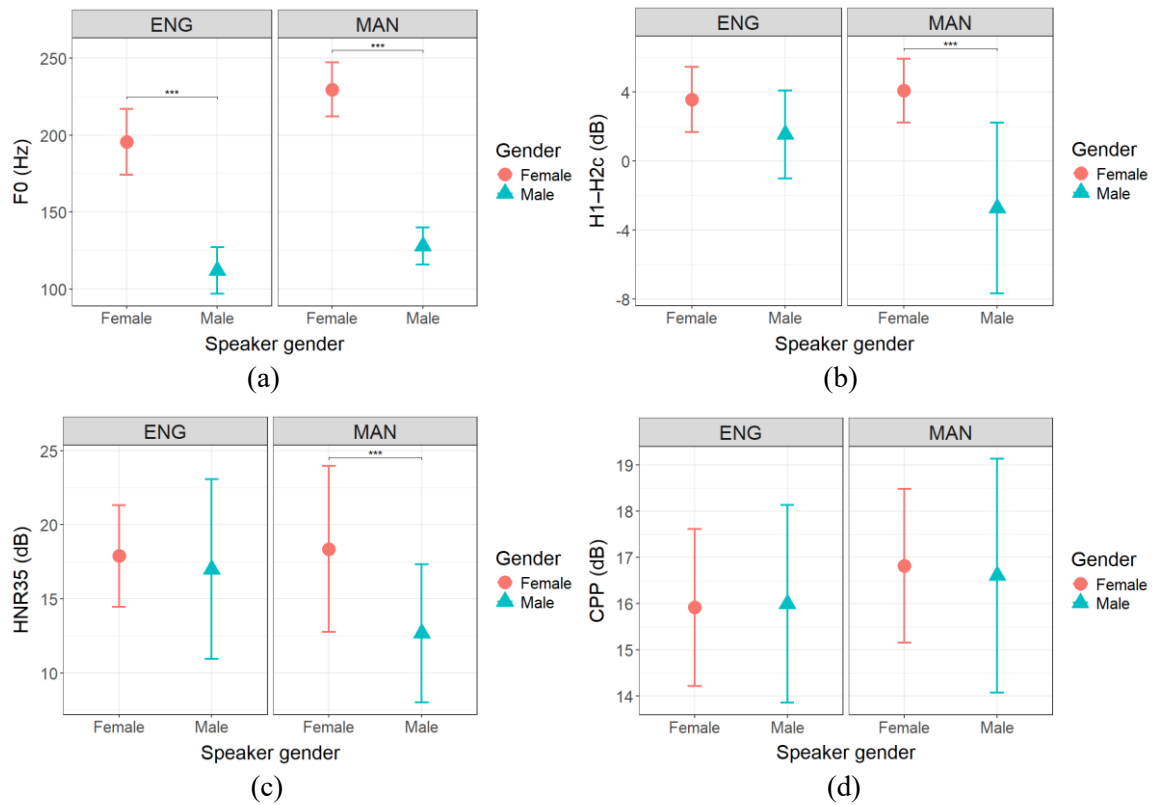


Figure 1. Average values of acoustic measures separated by language (ENG = English, MAN = Mandarin) and speaker gender: (a) F0; (b) H1-H2; (c) HNR; (d) CPP. Error bars indicate ± 1 standard error of the mean. Significance code: *** $p < .001$.

References

- [1] Dallaston & Docherty (2020). The quantitative prevalence of creaky voice (vocal fry) in varieties of English: A systematic review of the literature. *PLoS One*.
- [2] Podesva & Callier (2015). Voice quality and identity. *Annual Review of Applied Linguistics*.
- [3] Wolk et al. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*.
- [4] Yuasa (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*.
- [5] Irons & Alexander (2016). Vocal fry in realistic speech: Acoustic characteristics and perceptions of vocal fry in spontaneously produced and read speech. *Journal of the Acoustical Society of America*.
- [6] Brown & Sonderegger (2025). A sociophonetic study of creaky voice across language, gender and age in Canadian English-French bilinguals. *Journal of Phonetics*.
- [7] Kuang (2017). Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *Journal of the Acoustical Society of America*.
- [8] Li & Mok (2023). A perception study on voice quality and stance in Mandarin Chinese. *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2023)*.
- [9] Yao et al. (2024). Perceiving the social meanings of creaky voice in Mandarin Chinese. *Proceedings of the 12th International Conference on Speech Prosody (Speech Prosody 2024)*.
- [10] Keating et al. (2015) Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*.