Interspeech 2026
27 September - 1 October 2026, Sydney, Australia

A Corpus-Based Study of Creaky Voice Production in English and Mandarin

Meixian Li[1], Yao Yao[1], Charles Chang[2]

[1] The Hong Kong Polytechnic University
[2] City University of Hong Kong
meixian.li@connect.polyu.hk, y.yao@polyu.edu.hk, cbchang@cityu.edu.hk

Abstract

Creaky voice has been extensively examined across varieties of English, but cross-linguistic sociophonetic research on it remains limited. This study presents an acoustic investigation of creaky voice production in Mandarin Chinese and US English, using read-speech corpora comprising 66 Mandarin speakers (34F) and 61 English speakers (28F). Audios were processed via a pipeline of segmentation, annotation, cleaning, and acoustic analysis. 5,874 vowel tokens were measured for F0, H1–H2c, HNR35, and CPP over the 70% mid-portion. Linear mixed-effects analyses revealed men produced creakier speech than women in both languages. Findings contradict the notion that creaky voice is predominantly a feature of young women's speech in US English but align with findings of more creakiness in men's speech. We discuss the Mandarin results in relation to recent work on Mandarin listeners' social perception of creaky voice.

Index Terms: creaky voice, corpus, Mandarin, English, gender, acoustic analysis

1. Introduction

1.1. Creaky voice and gender-specific social meanings

Creaky voice, characterized by low fundamental frequency (F0), irregular vocal fold vibration, and reduced transglottal airflow [1], has been extensively examined in sociophonetic research, particularly within English varieties. Acoustically, creaky voice is commonly identified by a flatter spectral tilt as in depressed H1–H2 values [2], decreased signal-to-noise ratios such as harmonics-to-noise ratios (HNR) [3-5], lowered cepstral peak prominence (CPP) [2, 6-7], and lower and less reliable F0 [3], all of which reflect increased glottal constriction and aperiodicity. These acoustic measures form the basis for much of the empirical work on creaky phonation.

Beyond its phonetic definition, creaky voice has been shown to carry rich socio-indexical meanings, contributing to listeners' judgments about a speaker's affect and mood, personality traits, and social status [8]. Most previous research along this line focused on English, especially American English, with gender being one of the most researched factors in the perception and production of creaky voice. A substantial body of research (e.g., [9-11]) reports negative evaluations of creaky voice in the speech of young women, with creakiness associated with incompetence, disinterest, or unprofessionalism when adopted by female talkers (however, see [12-13] for evidence of less negative or even positive perception of creaky voice). In fact, many people consider creaky phonation a rising speech style that is largely negatively perceived and primarily attributed to females [14].

Despite the female bias observed in perceptual findings, production studies report mixed results in relation to the relative frequency of creaky phonation by female vs. male talkers (see [9] for a review). For example, a recent corpus study of Canadian English–French bilinguals [15] showed that men produce more creakiness (with lower values of H1–H2, CPP, and HNR) than women in both languages, contradicting the widespread notion that young women lead in the production of this feature.

1.2. Crosslinguistic investigation of creaky voice

Outside English, the study of creaky voice, or "creak", remains comparatively limited. While creaky phonation is attested across many languages, much of the existing work on creak focuses on its phonological or prosodic functions rather than its socio-indexical properties. For example, it has been shown that in tonal languages such as Mandarin Chinese and Cantonese, creak frequently marks low-pitch targets, enhances tonal contrasts, and signals tone sandhi processes [16-17], whereas in non-tonal languages, creak often aligns with prosodic boundaries.

Among existing studies examining creaky voice outside of English, gender differences in creak production and perception are also debatable. As mentioned above, the findings of [15] for Canadian French contrast with previous results on English. If we take Mandarin as an example, some studies reported minimal or no gender differences in both the perception and production of creaky voice [17-18], whereas a more recent study [19] found gendered patterns in the perception of creaky voice, with **more negative evaluations for male talkers' creak production** than for female talkers', suggesting major differences from the perceptual results on English.

These findings suggest that the use of creaky voice and its associated social meanings likely vary across languages and cultures. More crosslinguistic research is therefore needed to establish the relationship between gender and perception/production of creaky voice in different languages. Importantly, crosslinguistic comparison would benefit from the use of shared methods, so that any variations observed in the results between different languages cannot be attributed to methodological discrepancies.

1.3. Current study

The goal of this study was to examine creak production in two typologically distinct languages, US English and Mandarin Chinese spoken in Mainland China. Both languages have appeared in previous studies of creaky voice, with mixed results. In this study, we created two parallel speech corpora with sentence recordings from native speakers of English and Mandarin, each reading in their respective native language.

10.21437/Interspeech.2026-3235

Our central research question concerned how talker gender and language would affect creak production in relatively careful speech. To address these questions, we analyzed four acoustic correlates of creaky phonation (F0, H1–H2, HNR, CPP) in the recordings and assessed the effect of gender on these correlates across languages.

2. Methods

2.1. Corpus construction

The corpora consist of read-speech recordings produced by 66 native Mandarin Chinese speakers (34F, 32M; age 28–40) and 61 native US English speakers (28F, 33M; age 20–40). Three other English speakers (1F, 2M) finished the reading task but were not included in the corpus: one male due to a non-US accent, and one male and one female due to the absence of identifiable speech in their recordings. The Mandarin speakers were born and raised in Mainland China and completed at least primary and secondary education in Mandarin. Consistent with the general linguistic profile of this demographic group, all Mandarin participants reported speaking English as a second language (L2), and many reported knowledge of additional non-Mandarin varieties of Chinese as home dialects or heritage languages. All the US English participants, on the other hand, self-reported as monolinguals with no additional languages. Both groups were relatively young and roughly balanced in gender.

All the recordings were made on the Gorilla platform at a sampling frequency of 48 kHz. Mandarin participants attended the experiment in person in a sound-treated room in Hong Kong, while the English participants participated remotely from various US locations due to practical constraints. To ensure the recording quality of remote participants, we implemented pre-study device checks, which blocked participants without required devices, and post-study quality screening, which filtered out recordings with sub-par audio quality and/or audible background noise.

Table 1. *Sentences for recording.*

Language	Sentence
Mandarin sentences (with English glosses)	小李去商店买了很多东西 (“Xiaoli bought many things from the store.”) 老师让我们明天交论文草稿 (“The teacher asked us to submit the paper draft tomorrow.”) 丽丽经常去咖啡店写作业 (“Lili often goes to a café to do her homework.”) 小明特别喜欢在阳台前读书 (“Xiaoming especially likes reading on the balcony.”) 珍珍可没说她一定会来 (“Zhenzhen didn’t say she would come.”)
English sentences	The man left his keys on the kitchen table. She likes to read before going to bed. My parents moved into a new apartment last week. I plan to study biology at the university. The weather is getting warmer every day.

Each participant read five sentences in their native language conveying emotion-neutral content (see Table 1 for the full sentence list). Across the two languages, the sentence sets were carefully controlled so that the two groups produced materials with comparable length, similar syntactic structures, and a balanced distribution of consonants and vowels (and tones, for Mandarin). Thus, for both languages, the sentences were 10–15 syllables in length (M = 11.7), limited to one matrix clause with a nominal subject, and contained a variety of phones/tones. All participants were instructed to read the sentences as naturally and clearly as possible in a self-paced experiment with no interlocutor present (either in person or on screen). Sentence recordings with serious speech errors or audio issues impacting intelligibility were excluded from analysis; this resulted in the exclusion of 0.3% of the Mandarin sentence tokens and 2.6% of the English sentence tokens.

2.2. Data preprocessing and acoustic measurement

Raw recordings were first converted to 16-bit mono-channel WAV files at a sampling rate of 16 kHz and then scaled to 70 dB in average intensity. Forced alignment with the Montreal Forced Aligner [20] was then applied, using language-specific acoustic models and pronunciation dictionaries (Mandarin: Mandarin (China) MFA dictionary v3.0.0; English: English (US) MFA dictionary v3.1.0). The resulting annotations at word and phoneme levels were manually inspected and adjusted by the authors as needed. With phoneme-level annotations, we identified all intervals of vowels in each language comprising low, mid, and high vowels (Mandarin: [a], [i], [y], [u], [ej], [ə], [ɛ], [o]; English: [æ], [a], [ɑ], [i], [ɪ], [u], [ej], [ə], [ɛ]; N > 3,600 in each language).

Given the relationship of creak with F0, F0 was central to our acoustic analyses and was the primary basis for excluding a subset of vowel tokens. F0 was automatically tracked in Praat [21] in 1-ms steps, with the possible F0 range set to 70–450 Hz for female talkers and 50–300 Hz for male talkers. We then excluded tokens with (1) extremely short (< 50 ms) or long (> 500 ms) durations (Mandarin: N = 258; English: N = 315), which possibly reflected disfluencies or speech errors, or (2) untrackable F0 in the 70% mid-portion (Mandarin: N = 430; English: N = 615). The final dataset contained 3,002 vowel tokens in Mandarin and 2,872 in English. The proportion of tokens with untrackable F0 was similar across languages (English: 13%; Mandarin: 17%) and also genders. In English, the male-female breakdown of excluded tokens was 52% vs. 48%; in Mandarin, it was 50.5% vs. 49.5%.

Apart from F0, three additional acoustic measures were obtained for the 70% mid-portion of each vowel token with VoiceSauce [22]: the corrected amplitude difference between the first and second harmonics of the voice spectrum (H1–H2c), the harmonics-to-noise ratio in the 0–3500 Hz range (HNR35), and the cepstral peak prominence (CPP). Because potentially unreliable measurements arising from severe F0 tracking failures were already excluded through the F0-based screening procedure described above, no additional acoustically-based outlier removal was performed beyond the above token-level exclusions.

2.3. Data analysis

To investigate the effects of language and gender on creaky phonation, we fitted separate linear mixed-effects models on the four acoustic measures (F0, H1–H2c, HNR35, CPP) using the *lme4* package [23] in R [24]. For each model, the dependent variable comprised the by-token mean values of the acoustic measure over the central 70% of each vowel token. In all

models, the critical fixed effects were Language, Gender, and their interaction. Following [15], we also included VowelHeight, a categorical variable that distinguishes low vowels (Mandarin: [a]; English: [a], [ɑ], [æ]) and high/mid (i.e., non-low) vowels (Mandarin: [i], [y], [u], [ej], [ə], [e], [o]; English: [i], [ɪ], [u], [ej], [ə], [ɛ]), as a fixed effect to control for potential acoustic differences between low vowels and high/mid vowels. Random intercepts for Talker and Word were included in all the models, with additional random slopes (VowelHeight by Talker, Gender by Word) where applicable [25-26]. All the categorical predictors were coded with sum contrasts, with the levels ordered as follows: Language: English, Mandarin; Gender: Female, Male; Vowel Height: Nonlow, Low. Significant Language × Gender interactions were followed up with post-hoc pairwise comparisons using estimated marginal means (implemented by the *emmeans* package [27]) with the Bonferroni correction.

3. Results

3.1. Effects of language and talker gender on F0

The results for F0 were consistent with the predicted gender difference in F0, but no significant cross-language variations in F0 were found (see Figure 1). The mixed-effects model of F0 revealed only a significant effect of Gender (β = 38.65, SE = 2.13, p < .001), with female talkers producing higher F0 than male talkers. There was no effect of Language or VowelHeight, nor was there a Language × Gender interaction (all ps > .1).

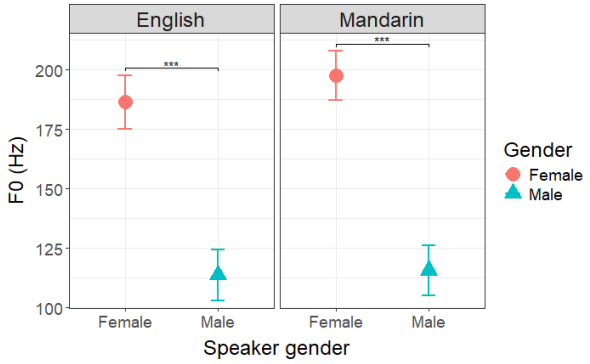

Figure 1: *Estimated F0 values by language group and talker gender. Error bars indicate 95% confidence intervals. Significance code:* *** p < .001.

3.2. Effects of language and talker gender on H1–H2c

Figure 2 shows gender differences in H1–H2c across English and Mandarin. The model of H1–H2c revealed a significant effect of Gender (β = 2.09, SE = 0.38, p < .001), showing that female talkers exhibited significantly **higher** H1–H2c values than male talkers. There was also a significant effect of Language (β = -1.40, SE = 0.43, p = .001), reflecting English talkers' **lower** H1–H2c values as compared to Mandarin talkers. The Language × Gender interaction was not significant (β = -0.40, SE = 0.36, p = .27), consistent with the visually similar gender difference across languages. The effect of VowelHeight was not significant, either (β = 0.09, SE = 0.16, p = .58). Overall, the H1–H2c model suggested that, by virtue of their lower H1–H2c values, male talkers tended to be creakier than female talkers, in both Mandarin and English.

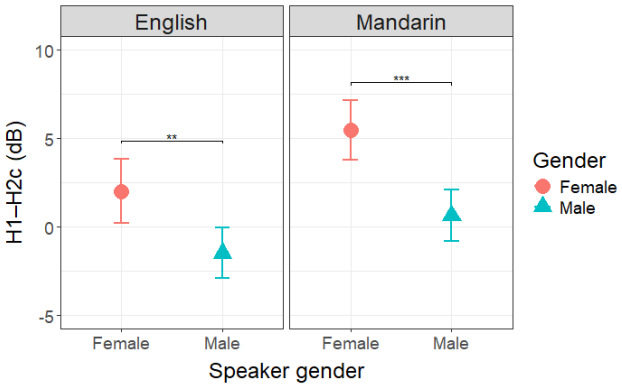

Figure 2: *Estimated H1-H2c values by language group and talker gender. Error bars indicate 95% confidence intervals. Significance codes:* ** p < .01, *** p < .001.

3.3. Effects of language and talker gender on HNR35

Figure 3 shows estimated HNR35 values by language and gender. While the HNR35 model showed no significant effect of Language (β = 0.97, SE = 0.76, p = .21), it revealed a significant effect of Gender (β = 2.77, SE = 0.37, p < .001), indicating higher values for female talkers than male talkers. Furthermore, the Language × Gender interaction was significant (β = -0.95, SE = 0.37, p = .012). Post-hoc comparisons showed that the effect of Gender was significant for both languages but larger in Mandarin (*est.* [female – male] = 7.26, p < .001) than in English (*est.* = 3.61, p < .0001). While female talkers were not significantly different across languages (p = .49), Mandarin male talkers had significantly lower HNR35 than English male talkers (p = .004). Additionally, there was a significant effect of VowelHeight (β = 0.91, SE = 0.22, p < .001), with high/mid vowels showing higher HNR35 values than low vowels.

In summary, the HNR35 results showed a pattern of male talkers producing creakier speech (as suggested by lower HNR35 values) than female talkers in both languages. The gender gap was larger for Mandarin than English, mainly due to the especially low values of Mandarin male talkers.

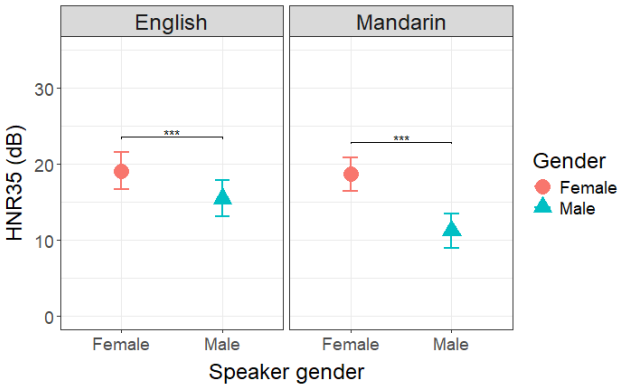

Figure 3: *Estimated HNR35 values by language group and talker gender. Error bars indicate 95% confidence intervals. Significance code:* *** p < .001.

3.4. Effects of language and talker gender on CPP

Figure 4 shows estimated CPP values for female and male talkers in English and Mandarin. The model of CPP did not reveal any significant effects of Gender or Language, nor a

Language × Gender interaction (all *p*s > .1). Only the effect of VowelHeight was significant (β = -0.37, SE = 0.14, p = .008): high/mid vowels showed lower CPP values than low vowels.

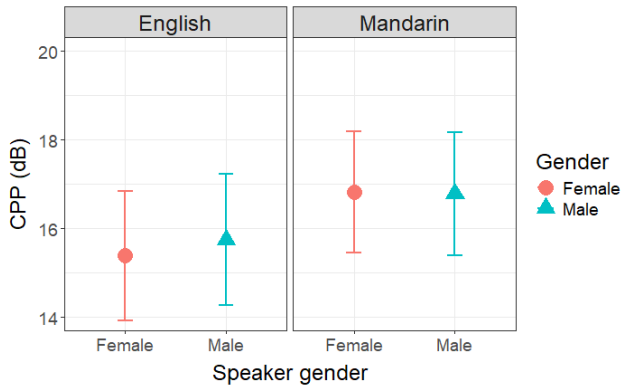

Figure 4: *Estimated CPP values by language group and talker gender. Error bars indicate 95% confidence intervals.*

4. Discussion

Our results demonstrate a consistent gendered pattern for Mandarin and English: **men are creakier than women**. The most important acoustic evidence for this pattern came from the results for H1–H2c and HNR35. Results for F0 were also in line with this pattern; however, since gender differences in F0 are more majorly influenced by physiology, we do not consider the F0 results as the main evidence for the gender effect on creakiness. Only the CPP results did not show any significant effect of gender; this may be because CPP primarily reflects the overall periodic organization of speech in the cepstral domain and is therefore less sensitive to glottal irregularities that are characteristic of creaky voice.

The observation of men being creakier than women in US English is consistent with recent corpus findings on Canadian English [15], but it contrasts with the preexisting notion—largely derived from social perception studies on US English—that creakiness is predominantly a feature of young women's speech. Our results provide further evidence of a dissociation between the perceived gendered use of creaky voice and its actual production patterns. The dissociation may reflect the complex nature of social perception, in which social meanings may be constructed at least partly independently of production patterns. Alternatively, it may signal an ongoing change in speech production, with younger men catching up with—or even exceeding—younger women in their adoption of creaky voice. In fact, these potential explanations are not mutually exclusive.

Crucially, the pattern of men being creakier than women exists for Mandarin as well, and to an even larger extent. A significant Language × Gender interaction was observed for HNR35, indicating larger gender differentiation in Mandarin than in English. These findings contrast with previous reports of no gender difference in creak use by Mandarin speakers [18]. The discrepancy between our findings and previous findings may be due to differences in speech materials, acoustic measures, analytic procedures, and/or population focus. Notably, in [18], creak was quantified in terms of the overall creak *rate* (calculated over tokens), whereas this study examined four individual *acoustic correlates* of phonation at the token level. Moreover, while the analysis in [18] was based on talker-level summaries with the Wilcoxon rank-sum test, the current study employed mixed-effects modeling to account for talker- and word-level variability. In addition, the Mandarin speakers in [18] were from Taiwan, whereas those in the current study were recruited from Mainland China. It is thus possible that the discrepancy between our results and those in [18] reflect methodological differences and/or variation in gendered patterns of creak production across different varieties of Mandarin.

That said, our results on Mandarin creak are compatible with previous perceptual work showing negative perception of creaky male voices by Mandarin listeners, particularly male listeners [19]. The coexistence of robust acoustic correlates of creakiness in male talkers and negative evaluations of such voices suggests an emerging creaky speech style among Mandarin men that is, unfortunately, subject to unfavorable social evaluations. Creak in Mandarin may help construct an affected, mature-sounding, masculine persona, which is particularly unwelcome by heterosexual male listeners.

Taken together, the current crosslinguistic corpus findings suggest that there are indeed gendered production disparities in creakiness across languages. These gender effects on creak production may encourage creaky voice to carry gendered socio-indexical meanings in a variety of languages; however, as we observed here in the difference in effect size between Mandarin and English, we expect that future research will reveal substantial crosslinguistic variation in the manifestation of gender effects, as well as social perception patterns that do not necessarily align with gendered production disparities (cf. [10] vs. [15]). By contributing suggestive evidence of variability in gender effects across languages, these results attest to the importance of linguistic and cultural diversity in sociophonetic research and can serve as a basis for future crosslinguistic work exploring variation in creak production.

In closing, we would like to point out some limitations of the current findings, which highlight several avenues for further research. First, there was variation in recording devices in the English group only. Because recording equipment is known to influence acoustic data [28-29], we are cautious about overinterpreting within-gender language differences, and have kept our focus on gender effects within each language. Second, our analysis excluded tokens with untrackable F0, which were disproportionately creaky. Our findings may therefore underestimate the overall extent of creak production in the current speaker sample. Finally, our results are based on read speech, which may not reflect patterns of creak production in more spontaneous speech styles where social meaning is more directly negotiated. Future research could address these limitations by more strictly controlling recording devices across languages, adopting an analytic approach that accounts for untrackable F0 proportions (allowing otherwise excluded tokens to be retained and analyzed), and using naturalistic speech production tasks.

Another direction for future work on cross-language differences is to focus on bilingual speakers, such as Mandarin-English bilinguals. A bilingual speech corpus can provide a complementary perspective on whether gendered patterns of creaky phonation in a given language are truly specific to that language. If, for example, a bilingual's creak patterns were to shift systematically across their languages, this would provide further evidence that voice quality is shaped by language-specific phonological structures and sociophonetic norms. Bilingual speech will thus shed unique light on the extent to which cross-language variation in creaky voice reflects linguistic conditioning versus broader speaker-level tendencies.

5. Generative AI Use Disclosure

The authors used ChatGPT (OpenAI) only for language editing and proofreading. All scientific content, analyses, interpretations, and conclusions were developed and verified by the authors, who take full responsibility for the manuscript.

6. References

[1] P. Keating, J. Kuang, M. Garellek, C. M. Esposito, and D. Khan, "A cross-language acoustic space for vocalic phonation distinctions," *Language*, vol. 99, no. 2, pp. 351–389, 2023.

[2] P. Keating, M. Garellek, and J. Kreiman, "Acoustic properties of different kinds of creaky voice," *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, 2015, paper 0821.

[3] M. Garellek, "The phonetics of voice," in *The Routledge Handbook of Phonetics*, W. F. Katz and P. F. Assmann, Eds. London, U.K.: Routledge, 2019, pp. 75–106.

[4] M. Garellek, "Theoretical achievements of phonetics in the 21st century: Phonetics of voice quality," *Journal of Phonetics*, vol. 94, article 101155, 2022.

[5] L. Davidson, "The effects of pitch, gender, and prosodic context on the identification of creaky voice," *Phonetica*, vol. 76, no. 4, pp. 235–262, 2019.

[6] M. Garellek, "Perception of glottalization and phrase-final creak," *Journal of the Acoustical Society of America*, vol. 137, no. 2, pp. 822–831, 2015.

[7] M. Garellek and S. Seyfarth, "Acoustic differences between English /t/ glottalization and phrasal creak," in *Proceedings of INTERSPEECH 2016*, 2016, pp. 1054–1058.

[8] K. Becker, S. Khan, and L. Zimman, "Beyond binary gender: creaky voice, gender, and the variationist enterprise," *Language Variation and Change*, vol. 34, no. 2, pp. 215–238, 2022.

[9] K. Dallaston and G. Docherty, "The quantitative prevalence of creaky voice (vocal fry) in varieties of English: A systematic review of the literature," *PLoS ONE*, vol. 15, no. 3, article e0229960, 2020.

[10] R. C. Anderson, C. A. Klofstad, W. J. Mayew, and M. Venkatachalam, "Vocal fry may undermine the success of young women in the labor market," *PLoS ONE*, vol. 9, no. 5, article e97506, 2014.

[11] S. D. F. Greer and S. J. Winters, "The perception of coolness: Differences in evaluating voice quality in male and female speakers," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, 2015, paper 0833.

[12] C. Ligon, C. Rountrey, N. V. Rank, M. Hull, and A. Khidr, "Perceived desirability of vocal fry among female speech communication disorders graduate students," *Journal of Voice*, vol. 33, no. 5, pp. 805.e21–805.e35, 2019.

[13] I. P. Yuasa, "Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women?," *American Speech*, vol. 85, no. 3, pp. 315–337, 2019.

[14] R. J. Podesva, "Gender and the social meaning of non-modal phonation types," *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, vol. 37, no. 1, 2011, pp. 427–448.

[15] J. Brown and M. Sonderegger, "A sociophonetic study of creaky voice across language, gender and age in Canadian English-French bilinguals," *Journal of Phonetics*, vol. 112, article 101431, 2025.

[16] A. Li, W. Lai, and J. Kuang, "Creaky voice identification in Mandarin: The effects of prosodic position, tone, pitch range and creak locality," *Journal of the Acoustical Society of America*, vol. 154, no. 1, pp. 126–140, 2023.

[17] J. Kuang, "Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice," *Journal of the Acoustical Society of America*, vol. 142, no. 3, pp. 1693–1706, 2017.

[18] J. Kuang, "The influence of tonal categories and prosodic boundaries on the creakiness in Mandarin," *Journal of the Acoustical Society of America*, vol. 143, no. 6, pp. EL509–EL515, 2018.

[19] Y. Yao, M. Li, S. Li, and C. B. Chang, "Perceiving the social meanings of creaky voice in Mandarin Chinese," in *Proceedings of Speech Prosody 2024*, 2024, pp. 652–656.

[20] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable text-speech alignment using Kaldi," in *Proceedings of INTERSPEECH 2017*, 2017, pp. 498–502.

[21] P. Boersma and D. Weenink, *Praat: Doing phonetics by computer*, Version 6.3, 2023.

[22] Y.-L. Shue, P. Keating, C. Vicenik, and K. Yu, "VoiceSauce: A program for voice analysis," in *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 2011)*, 2011, pp. 1846–1849.

[23] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.

[24] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2023.

[25] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[26] D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily, "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *Journal of Memory and Language*, vol. 68, no. 3, pp. 255–278, 2013.

[27] R. V. Lenth, "emmeans: Estimated marginal means, aka least-squares means," *R package version 1.8.8*, 2023.

[28] J. Penney, A. Gibson, F. Cox, M. Proctor, and A. Szakay, "A comparison of acoustic correlates of voice quality across different recording devices: A cautionary tale," in *Proceedings of INTERSPEECH 2021*, 2021, pp. 1389–1393.

[29] C. Sanker, S. Babinski, R. Burns, M. Evans, J. Kim, S. Smith, N. Weber, and C. Bowern, "(Don't) try this at home! The effects of recording devices and software on phonetic analysis," *Language*, vol. 97, no. 4, pp. e360–e382, 2021.