

Gender effects in the social perception of creaky voice in Mandarin Chinese

Yao Yao¹, Meixian Li¹, Shiyue Li¹, and Charles B. Chang²

¹The Hong Kong Polytechnic University (Hong Kong), ²Boston University (USA)

The social meanings of voice quality (i.e., phonation type) have garnered much interest from linguists in recent years. For example, creaky voice has been reported to be a cue to various talker properties, including age, sex, gender identity, body size, ethnicity, attractiveness, employability, and diverse personality traits [1]–[3]. However, most of the research in this area focuses on English, and uses speech stimuli that are either naturally produced, such that voice quality covaries with other variables such as pitch and speech rate, or fully synthesized, leading to small repertoires of talker voices and speech samples.

In this study, we address the above limitations by examining **how creaky voice is socially perceived in Mandarin**, using a large set of resynthesized stimuli. Crucially, the investigation of Chinese languages is situated in a different linguistic and socio-cultural context than previous research on English and other European languages. For example, in the perception of Chinese languages, creaky voice was found to interact with lexical tones [4]–[6], which do not occur in English. The few studies of voice quality perception in Mandarin reported overall lower ratings of talker attractiveness for creaky voice than modal voice [7]–[9], resembling previous findings on English [10] but with no interaction with gender.

Applying a novel method of resynthesizing modal and creaky speech stimuli from the same production of short sentences, the current study conducted a detailed social perception experiment with 60 native Mandarin listeners (33F, 27M; ages 18–35). The resynthesis was done with a Klatt synthesizer, targeting all the voiced portions of the original production; creakiness was achieved by adding double pulsing points (i.e., multiply-pulsed creak [11]; Fig. 1), ensuring that modal and creaky versions of the same base token differed only in voice quality and that creaky stimuli contained a type of creakiness found across all Mandarin tones [12]. Using this method, we generated 9600 potential stimuli (120 sentences × 40 talkers: 20F, 20M × 2 voice qualities), which were vetted for intelligibility and naturalness to yield 6–10 pairs of sentence stimuli per talker. In the experiment (which included, for each listener, stimuli from 14 talkers balanced for gender and voice quality), listeners were told that the voices were “created by speech engineers” and provided subjective evaluations of the virtual talkers. For each talker, the listener judged 4 demographic characteristics (age, gender, sexuality, education level) and rated 15 personality traits (e.g., confident, professional, charismatic) and 4 aspects of interactive potential (e.g., engagingness, potential for further interaction). These responses were then analyzed using factor analysis [13] of the personality and interactive potential ratings and mixed-effects models (with random intercepts and slopes by talker and listener) of the factor scores and of the demographic judgments.

Results showed several effects of, and interactions between, voice quality, talker gender, and listener gender. First, the factor analysis revealed three latent factors related to perceived “competence”, “likeability”, and “warmth”. While creaky voice had no effect on competence or likeability, it significantly decreased warmth for male talkers, especially for male listeners. By contrast, creaky voice significantly increased warmth for female talkers, but only for male listeners (Fig. 2). Second, creaky voice led to significantly more gender identification errors on female talkers by female listeners (but not by male listeners) and made male talkers sound older, patterns likely due to effects of creaky voice on pitch perception. These findings point toward **multifaceted social meanings of creaky voice in Mandarin, which extend beyond talker attractiveness and are closely linked to gender**, both the talker’s and the listener’s.

Crucially, the holistic evaluation of creaky voice in Mandarin differs from its evaluation in English [2], highlighting the need for further research on the social perception of non-modal voice qualities in diverse languages that make use of voice quality phonologically in different ways. Additional research is also needed to establish the degree to which the above patterns generalize to the perception of other types of creaky voice, such as vocal fry, that occur in Mandarin [12, 14]. Ongoing work is examining the relationship of listeners’ social perceptions to their own speech patterns and to their attitudes on gender-related issues.

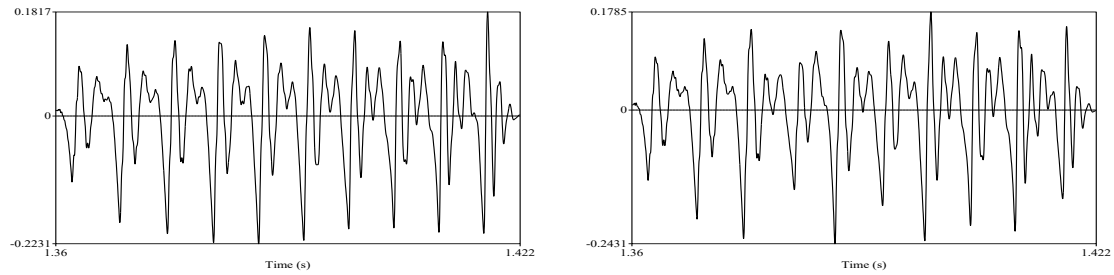


Fig. 1. Modal (left) and creaky (right) resynthesis of an [i] vowel produced by a female talker.

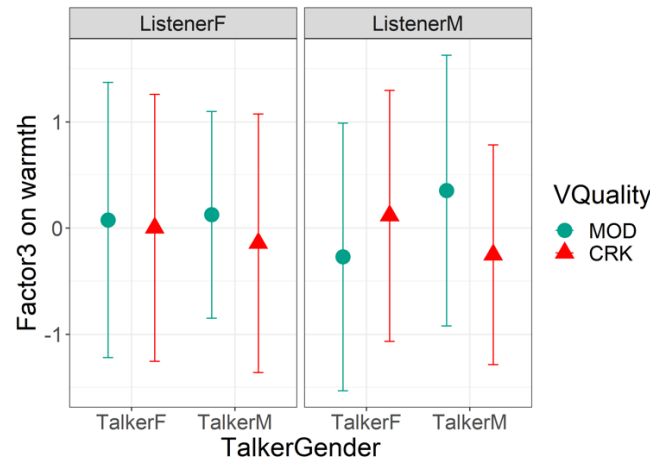


Fig. 2. Mean scores on Factor 3 (warmth) by Talker Gender, Voice Quality (modal/MOD, creaky/CRK), and Listener Gender. Error bars mark 95% CIs of the mean.

References

- [1] 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*, 15, 2020.
- [2] L. Davidson, "The versatility of creaky phonation: Segmental, prosodic, and sociolinguistic uses in the world's languages," *WIREs Cognitive Science*, 12, 2021.
- [3] R. J. Podesva and P. Callier, "Voice quality and identity," *Annual Review of Applied Linguistics*, 35, 2015.
- [4] A. Li, W. Lai, and J. Kuang, "Creaky voice identification in Mandarin: The effects of prosodic position, tone, pitch range and creak locality," *JASA*, 154, 2023.
- [5] K. M. Yu and H. W. Lam, "The role of creaky voice in Cantonese tonal perception," *JASA*, 136, 2014.
- [6] J. Kuang and M. Liberman, "Influence of spectral cues on the perception of pitch height," in *Proceedings of the 18th ICPHS*, 2015.
- [7] A. Li and W. Lai, "How do listeners evaluate creak: A matched-guise study in Mandarin Chinese," Paper presented at the 2023 LSA Annual Meeting, June 5–8, Denver, CO, 2023.
- [8] Q. Li and P. Mok, "A perception study on voice quality and stance in Mandarin Chinese," in *Proceedings of the 19th ICPHS*, 2023.
- [9] A. Xu and A. Lee, "Perception of vocal attractiveness by Mandarin native listeners," in *Proceedings of the 9th International Conference on Speech Prosody*, 2018.
- [10] R. C. Anderson, C. A. Klotz, W. J. Mayew, and M. Venkatachalam, "Vocal fry may undermine the success of young women in the labor market," *PLoS One*, 9, 2014.
- [11] P. Keating, M. Garellek, and J. Kreiman, "Acoustic properties of different kinds of creaky voice," in *Proceedings of the 18th ICPHS*, 2015.
- [12] Y. Huang, "Articulatory properties of period-doubled voice in Mandarin," in *Proceedings of Speech Prosody 2022*, 2022.
- [13] J. K. Sakaluk and S. D. Short, "A methodological review of exploratory factor analysis in sexuality research," *Journal of Sex Research*, 54, 2017.
- [14] Y. Huang, *Phonetics of period doubling*, Ph.D. dissertation, University of California, San Diego, 2023.