Tonal and phrasal distributions of sub-phonemic creaky voice in Mandarin

Yaqian Huang

University of California Los Angeles
huangyq@g.ucla.edu

Lexical tones are primarily characterized by contrastive pitch patterns. Recent studies on the interaction between tone and phonation have enriched our understanding of non-contrastive phonetic dimensions that additionally capture tonal space [1,2]. In Mandarin Chinese, a tone language with contrastive pitch and without contrastive phonation, creaky voice is often found in tones that contain a low pitch target, most frequently with the lowest dipping Tone 3 (21[4]) and occasionally with rising Tone 2 (35) and falling Tone 4 (51) [3,4,5]. Sub-phonemic creak includes period doubling and vocal fry, which were found to have different gender distributions [6] and whose phonatory features have been detailed in [7]. It still remains unclear whether period doubling and vocal fry carry different linguistic functions in tone production and perception. This paper thus investigates how the production of period doubling and vocal fry may vary by prosodic factors such as tone and phrasing as they exist as different realizations of sub-phonemic creaky voice in Mandarin.

Methods. The materials are from a Mandarin corpus of simultaneous audio and electroglottography (EGG) recordings collected to document contextual tonal variation. The sentences consist of a fixed carrier phrase with varying stimuli of trisyllabic compounds, and the words are coded by phrasal positions:

'I	teach	you		STIMULUS		how-to	say'	(gloss)
wo3	*tɕau1*	*nʲi3*	*Syll1*	*Syll2*	*Syll3*	*tsən3-mɤ0*	*ʂʷo1*	(sentence)
UI	PS2	PS1	PI	PM	PF	AS	UF	(phrasal positions)

Trisyllabic compounds have 64 (4 tones x 4 tones x 4 tones) varying tonal combinations. Three sets of 64 sentences with two repetitions (384 sentences in total) were elicited per recording. 20 native Mandarin speakers (10 F) in college participated in the production experiment. The corpus has 7680 sentences. We labeled the lexical tones and excluded neutral tones based on their phonetic realization. Fig. 1 shows examples of creak from the corpus: period doubling, characterized by alternating pulses between amplitudes and/or periods [8], and vocal fry, characterized by low f0 and high glottal constriction [9]. Both creak subtypes were identified anywhere in the phrase – during the trisyllabic stimuli or the carrier phrase – and coded with tone and phrasal positions, using the EGG signal, to avoid possible formant-induced interferences with the voicing signal. A total of 5848 tokens of period doubling and 1574 tokens of vocal fry have been identified and used in the prosodic analysis.

Results. The distribution of period doubling and vocal fry by tone, shown in Fig. 2, was analyzed for the compound stimuli only because tones do not vary elsewhere in the corpus. Overall, Tones 2 and 3 have more tokens of creak (either period doubling or vocal fry) than Tones 1 and 4. Vocal fry is rarely observed for Tone 1. For Tones 2 and 3, period-doubled tokens gradually increase as a function of the phrasal positions from the left edge (PI) to the right edge (PF) in trisyllabic sequences. This pattern is not attested in Tones 1 and 4 with an even distribution, probably due to fewer occurrences of creak for those tones. In contrast, the distribution of vocal fry does not seem to be conditioned by sentence-medial phrasal positions.

The majority of the creaky tokens are associated with the after-stimulus (AS) and utterance-final (UF) positions, which is consistent with the findings in [3] that creaky voice in Mandarin frequently occurs in the final and penultimate positions. Interestingly, for both women and men, a linear increase is only observed in period-doubled voice starting from the second word (PS2) to UF. While period doubling is mostly found in UF, vocal fry mostly occurs immediately after the stimulus (AS), which is a post-focal position. Tonal influences are also observed in these environments; e.g. fewer occurrences of vocal fry than period doubling are found in the UF position associated with a high-pitched Tone 1, and PS1 with the low Tone 3 shows more instances of vocal fry at least for women. These are consistent with the findings of tonal distribution in Fig. 2. However, UI also with Tone 3 has substantially fewer occurrences of vocal fry than PS1, possibly due to the phrasal effect. Thus, prosodic position seems to be a stronger driving factor than tone in favoring the occurrences of the two creaky voice subtypes.

Discussion. We hypothesize that period doubling is more strongly driven by utterance edges than is vocal fry, because the former reflects unstable voicing: towards the end of the utterance, voicing becomes progressively less stable [10]. Consequently, the occurrences of period doubling in utterance-final position could be a byproduct of downdrift (declination) as f0 progressively lowers towards the utterance edge without necessary constriction (see [7] on non-constricted quality during period doubling). The increasing trend of period doubling towards the end of the utterance is comparable to a similar phonation-ending

gesture that [10] has found for both irregular and regular phonation at the end of an utterance. The utterance-final irregular phonation is usually produced with short intervals of adduction followed by longer intervals of abduction and/or with incomplete closure of the vocal folds, rather than adducted like vocal fry [10]. This hypothesis, if proved correct, has ramifications for speech production studies, such that data elicitation is recommended in non-final contexts to avoid conflation of modal and non-modal phonation in both segmental and suprasegmental units. For example, the realization of period doubling in Mandarin Tone 3 or other tones will result in similar kinds of articulation and voicing instability. In contrast, vocal fry is typically triggered by a low and compressed f0 range with more probable constriction, associated with the post-focal position [11]. The fact that vocal fry occurs more restrictively in the penultimate position suggests that it could signal a stronger linguistic role such as marking a weak prosodic element post-focally. These results help clarify subcategories within creaky voice and have implications for the taxonomy of creaky voice as a refined phonetic category with potentially different linguistic functions.

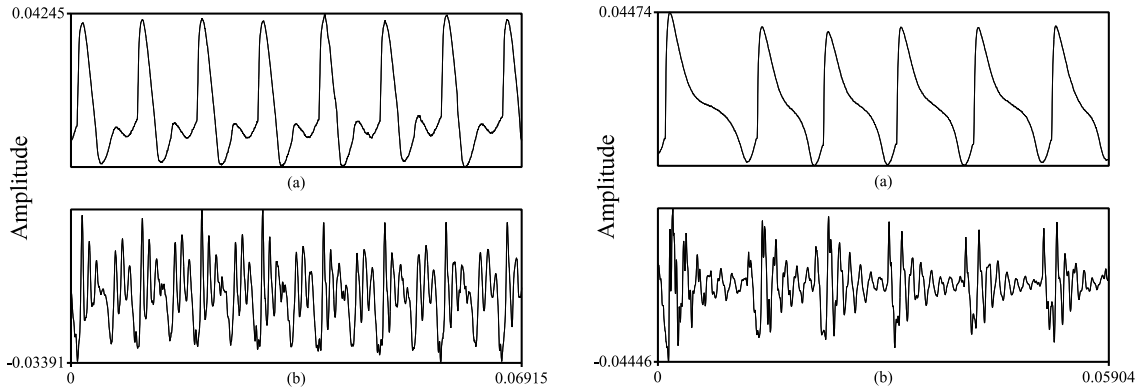

Figure 2. EGG (a) and audio (b) waveforms of period doubling (left) and vocal fry (right).

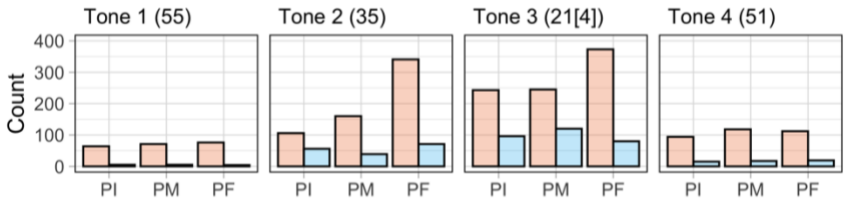

Figure 1. Raw count of period doubling and vocal fry across tones in sentence-medial positions. PI = phrase initial, PM = phrase medial, PF = phrase final.

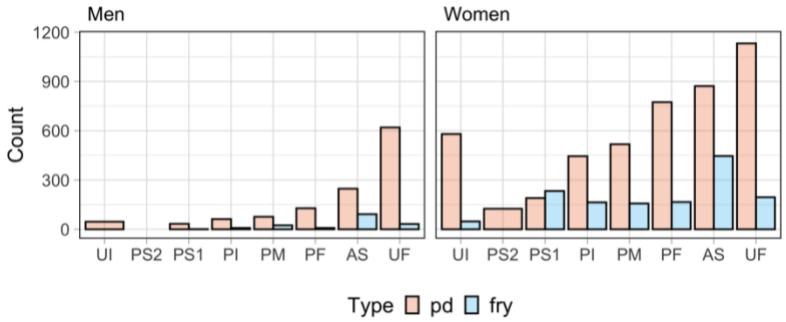

Figure 3. Raw count of period doubling and vocal fry across different phrasal positions throughout the sentence by gender.

[1] Kuang, J. (2013). *Phonation in tonal contrasts* (Doctoral dissertation, UCLA). [2] Garellek, M., Keating, P., Esposito, C. M., & Kreiman, J. (2013). Voice quality and tone identification in White Hmong. *The Journal of the Acoustical Society of America*, *133*(2), 1078-1089. [3] Belotel-Grenié, A., & Grenié, M. (2004). The creaky voice phonation and the organisation of Chinese discourse. In *International symposium on tonal aspects of languages: With emphasis on tone languages*. [4] Kuang, J. (2017). Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *The Journal of the Acoustical Society of America*, *142*(3), 1693-1706. [5] Huang, Y., Athanasopoulou, A., & Vogel, I. (2018). The effect of focus on creaky phonation in Mandarin Chinese tones. *University of Pennsylvania Working Papers in Linguistics*, *24*(1), 12. [6] Yu, K. M. (2010). Laryngealization and features for Chinese tonal recognition. In *Eleventh Annual Conference of the International Speech Communication Association*. [7] Huang, Y. (2022). Articulatory properties of period-doubled voice in Mandarin. *Proc. Speech Prosody 2022*, 545-549. [8] Kreiman, J., Gerratt, B. R., Precoda, K., & Berke, G. S. (1993). Perception of supraperiodic voices. *The Journal of the Acoustical Society of America*, *93*(4):2337–2337. [9] Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. In *ICPhS*, volume 2015, pp. 2–7. [10] Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of voice*, *20*(2), 171-186. [11] Xu, Y. (2011). Post-focus compression: Cross-linguistic distribution and historical origin. In *ICPhS*, volume 2011, pp. 152–155.