THE ROLE OF CREAKY VOICE IN CANTONESE TONAL PRODUCTION

Chu-Yan Ho

City University of Hong Kong; The University of Hong Kong
chuyanho2-c@my.cityu.edu.hk

ABSTRACT

Cantonese is a tonal language in which its low tones are reported to be produced with creaky voice inconsistently. The presence of the allophonic creaky voice in Cantonese has driven two hypotheses regarding its source: it is (1) a specific cue for Tone 4 identification or (2) a covariate when producing a low F0 [1]. To test whether creaky voice is tone-specific, native speakers were asked to produce Cantonese lexical tones under normal condition and low-pitch condition. Results revealed a correlation between creaky voice and low F0, and between creaky voice and prosodic positions. When participants were asked to speak in a low-pitch condition, creak occurrence increased significantly for tones other than Tone 4. The results shows that the source of Cantonese creaky voice is likely to be triggered by low F0, and is less likely to be phonemically tied to Tone 4.

Keywords: creaky voice, lexical tones, F0, Cantonese, tone production.

1. INTRODUCTION

Tonal languages use pitch contour and height as the primary acoustic cues for lexical tone identification, while register language refers to languages where phonation is lexically contrastive. However, the distinction has become less clear with the evidence that some tonal languages utilise both pitch and phonation for tone identification, such as Mandarin [2,3], Cantonese [1,4] and Hmong [5]. Notwithstanding the co-occurrence of phonation and pitch for lexical contrasts, tonal languages differ from register languages in that the former mainly relies on pitch for perceptual differences in lexical contrasts and phonation is only adopted as a secondary cue.

Traditionally, Hong Kong Cantonese (HKC) is studied as a tonal language where pitch has been assumed to be the main, if not the only, cue for tonal contrast. It is generally recognized that there are six contrastive tons with three entering tones, which are allotones of the level tones in HKC (See Table 1 for tone descriptions and Figure 2 for tone trajectories). Yu and Lam were the first to identify the relatively high occurrence of creaky voice in the low-falling Tone 4 (24.6%) in HKC and investigate the use of creaky correlates for the identification of T4-T6 pair [1]. They found an enhanced Tone 4 identification when Tone 4 was realized with creaky voice and that creaky voice biased participants toward Tone 4 even with contextual pitch cues. Zhang and Kirby disagreed on the extent of the effect of creaky voice in Tone 4 identification and questioned Yu and Lam's methodology as their manipulated stimuli may belong to double pulsing creak with extra low F0, thus directing the participants to identify Tone 4 [4]. Nonetheless, the two studies acknowledged that creaky voice enhances perception in low pitch range.

Tone no.	Tone category	Tone letter
T1	High-level	55
T2	High-rising	25
T3	Mid-level	33
T4	Low-falling	21
T5	Low-rising	23
T6	Low-level	22
T7 (T1)	High-stopped	5
T8 (T2)	Mid-stopped	3
T9 (T3)	Low-stopped	2

Table 1: The Cantonese tonal system in [6]

Other sociolinguistic studies documented regional, gender and educational differences in voice quality in HKC production in general, but none considered the role of tonal categories [7,8]. Fung verified that HKC is creakier than Guangzhou Cantonese when measuring the H1*-H2*, H2*-H4*, H1*-A1*, H1*-A2*, H1*-A3*, and CPP [7] when calculating mean F0 across tonal categories. Notably, her research documented a gender difference between the mean values of the acoustic parameters of phonation – nearly all HKC female participants showed creakier production in all five mean values of the acoustic parameters than male, regardless of age (except for CPP in middle-aged participants). Fung and Lee found a similar gender asymmetry in HKC voice quality [8]. They measured the pitch dynamics, spectral tilt, and periodicity of 60 HKC speakers varying in age, gender, and education level in two passage-reading tasks; their results revealed that female speakers adopted an extremely low pitch floor in general and higher-educated females exhibited creakier voice quality. These sociolinguistic findings are in accordance with creaky voice research in other languages as seen in English [9].

What remains unresolved is the source of creaky voice in Cantonese tonal production. Yu and Lam proposed two hypotheses in [1]; because of the relatively high occurrence of creaky Tone 4 in their study, they hypothesize that creaky voice (1) is a phonemic cue for Tone 4, or (2) a by-product of low F0. The current research intends to address this question by examining the occurrence of creaky voice among Cantonese tonal categories and gauge the effect of low F0 on non-modal phonation of Cantonese tones, especially when the pitch range is manipulated. As limited research has considered the role of tonal categories in Cantonese creaky voice production, this study is the second to document the prevalence of creaky phonation in Cantonese tonal production.

2. EXPERIMENT I: DESIGN

2.1. Participants and stimuli

12 Cantonese native speakers (6M 6F, age: 18-24) with no history of speech or listening disorders were recruited for the current study. As education level and language background have been found to influence Cantonese speakers' voice quality [8], the participants were selected to be local Hong Kong undergraduate students with less than one year of living experience in English speaking countries.

Participants were asked to read a list of Cantonese monosyllables /fu/, /sɛ/, /si/, / soeng/, /fan/, / jau/ and /ma/ in all six Cantonese tones using a SONY A10 Linear PCM Recorder in a sound-attenuated booth. As this experiment intends to answer whether creaky voice is a T4 specific cue, it is crucial that participants accurately produce all six citation forms of Cantonese tones. To avoid contaminating of recorded data by HKC tone mergers [10, 11], the utterances which deviated from the citation form of HKC were discarded. A total of 2016 tokens were collected in this experiment (2 syllables * 7 vowels * 6 tones * 2 repetitions * 12 participants), of which 11 tokens were discarded.

2.2. Segmentation criteria

All speech outputs were segmented and annotated manually using Praat v.6.0.37 [12] to retrieve F0 values. The onset of the vowels was chosen to be the onset point of the tokens, meaning that all initial consonants, regardless of whether it is voiced or voiceless, were discarded in the recordings. All offsets, such as /n/ in [fɐn] and /ng/ in [sœːŋ], were not discarded because the lowering of F0 value that usually accompanies phrase-final creak is of interest to the study.

2.3. Measurements

The presence of creak was defined through both visual and audio inspection by the author, a native speaker of Cantonese. Tokens were coded as creaky if they had the auditory percept of creaky voice, as determined by the authors and if: (1) there were alternating cycles of amplitude and/or frequency or irregular glottal pulses in the wide-band spectrogram and waveform or, (2) there were missing values or discontinuities in the F0 track determined by Praat's autocorrelation algorithm with default settings, or (3) there was the appearance of strong subharmonics or a lack of harmonic structure in the narrow-band spectrogram.

3.EXPERIMENT I: RESULTS AND DISSCUSSION

3.1. Creaky voice occurrence

The results of experiment 1 are shown in Table 2, alongside with the findings given by Yu and Lam's study [1]. It is found that, in addition to 227 Tone 4 tokens (61.0%), 75 Tone 5 tokens (21.7%), 55 Tone 2 tokens (16.9%), and 57 Tone 6 tokens (16.7%) showed creaky phonation. Tone 3 had relatively fewer creaky tokens – 26 (7.5%) while Tone 1 had none. The results suggest that it is unlikely that creaky voice is tied to Tone 4 production to ease perception, but rather, creaky phonation is sensitive to all low pitch tones. If creakiness in Tone 4 is specifically produced to enhance identification as proposed by Yu and Lam [1], then it would be unlikely for Tone 5 and Tone 2 to show 21.7% and 16.9% of creak occurrence in the present study. Consequently, it is unlikely that creakiness attribute in Cantonese tone is specific to any tone for perceptual ease, as Tone 2 and Tone 5 exhibit noticeably similar creak occurrence in this study, and they are argued to be most susceptible to mixed perception as they are both rising tones in Cantonese [10, 11].

Tone	Current Study	Yu and Lam [1]
1	0%	NG
2	16.9%	2.6%
3	7.5%	NG
4	61.0%	24.2%
5	21.7%	NG
6	16.7%	NG

Table 2: Percentage of creak in Cantonese tonal production found in current study and [1]

It should be noted that the distribution of creak occurrence generally matches the pitch heights of the six tones. For example, Tone 4, the lowest tone in

Cantonese tonal categories, generally has the highest creak occurrence. Tone 5 and Tone 2, bearing the tonal letter of 23 and 25 respectively, have the second and third highest creaky tones, which suggest that F0 is the major trigger of creaky voice production.

In addition, creak appears at similar F0 values for the recorded creak occurrences, around 90 Hz for male speakers and 170 Hz for female speakers in Figure 1. The value of pitch threshold is recorded by measuring the lowest F0 value before and/or after the occurrence of creak calculated by Praat. The unity in terms of creak threshold across tonal categories suggests that when a speaker's F0 values fall below a certain threshold, creaky phonation is triggered. This finding is similar to the result presented by Kuang [2], in which she defines 95Hz and 180Hz in her study as the threshold for creak to naturally take place when the pitch values hit a bottom line in Mandarin.

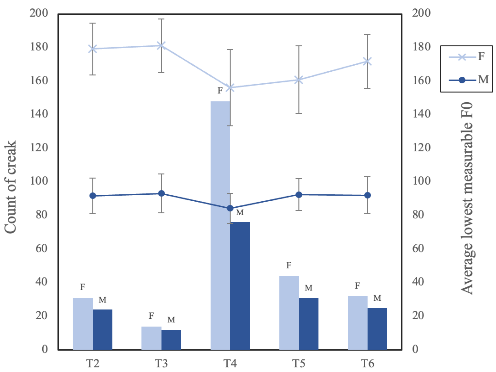

Figure 1: Pitch threshold values – measured by lowest measurable F0 and count of creaks for female and male participants in experiment 1.

Two binomial logistic regression were performed separately for male and female participants to analyse the relationship between creak occurrence and F0 values. Results showed that creak occurrence is correlated with low F0 for both male (β=-0.0852, $p<0.001$) and female (β=0.0311, $p<0.001$) and showed a negative relationship between F0 and creak occurrence for both male and female participants. The higher the F0, the less likelihood for a token to be produced with creaky voice.

3.2 Prosodic positions

The results of experiment 1 also show a significantly higher frequency of creak occurrence on the second syllable ($\chi2$=9.68, p=0.002). The second syllable recorded 246 creaky tokens, compared to 191 in the first syllable. A paired t-test was also performed to show the significant difference between the mean F0 values of the first syllable and the second (t=3.07, $p<0.001$). These results matched the previous study of Cantonese creaky voice production [1]. Although no previous research has documented the use of creaky voice as a sentence/syllable ending prosodic cue for Cantonese, some studies confirmed the use of creaky to indicate prosodic boundary in Mandarin [13,14]. Future studies should explore the effect of creak on eliciting prosodic cues in Cantonese on different prosodic boundaries.

It should be made clear that because all test words in Experiment 1 were spoken in isolation in a laboratory setting, the reason for the occurrence of creak, especially for the second syllable, could be caused by sentence-final creak accompanied by the lowering of F0. Such limitation could be easily erased if one could replicate the analysis of this experiment with a bigger corpus of daily continuous discourse, especially when tones that are less likely to creak are put at the end of a sentence to gauge the effect of prosodic position on triggering creak.

The difference of creak occurrences based on prosodic positions combined with the results of the logistic regression support that creaky voice is likely to be triggered by producing a low F0 and the percentages of creaky Tone 2 and Tone 5 hint that creaky voice is less likely to be tied to Tone 4. In experiment 2, the claim will be further supported by a manipulated production task asking experiment participants to read the same materials in two conditions: in normal speech and a lowered pitch speech with a lowered pitch range. It is hypothesised that if creak is triggered by low F0 values, when speakers are asked to speak in a low pitch range, more creaky tokens will be observed for tones that are produced in lower pitch ranges other than Tone 4.

4. EXPERIMENT II: EXPERIEMENT DESIGN

The same group of participants was recruited in experiment 2. Following the experiment set up by Kuang [2], Subjects were asked to produce the same set of materials in experiment 1 in two conditions. (1) normal pitch (words are produced in speakers' most comfortable pitch ranges); (2) low pitch (participants were instructed to speak in their lower-than-normal pitch range and were reminded not to whisper). Each token was repeated two times in two conditions monosyllabically.

The segmentation and creak coding were done the same as experiment 1 in which 12-time intervals of each token were extracted for their F0 values. Six seperate paired-sample t-tests were performed to test whether the participants were able to produce a pitch contrast under the two conditions.

5. EXPERIMENTII: RESULTS AND DISCUSSION

The results of experiment 2 are shown in Table 3. Overall, a significantly higher number of creaks is recorded under the low-pitch condition (χ2=74.7, p<0.001). Female participants also produced a significantly higher number of creaks (χ2=33.3, p<0.001). While other tones have recorded a significant increase in creak occurrence, the number of creaky T4 only differs slightly under the low-pitch condition (-4% for female participants and 2% for males. Results of the t-tests confirmed that for Tone 1 (W=10910, p<0.001), Tone 3 (W= 8594, p=0.005), and Tone 4 (W=5464, p=0.049), the average F0 values of the two conditions are significantly different. As for Tone 2 (W=7366, p=0.399), Tone 5 (W=8353, p=0.091), and Tone 6 (W=6993, p=0.426), there is no significant difference in mean F0 values between the two conditions. Figure 2 shows that while the F0 trajectories of female participants are compressed and lowered under the low-pitch condition, male participants' F0 trajectories are only compressed but not lowered.

Tone	Normal		Low pitch		Δ%	
	F	M	F	M	F	M
1	0	0	6	3	-	-
2	30	20	45	30	50%	50%
3	8	9	27	23	238%	156%
4	77	42	74	43	- 4%	2%
5	19	16	42	38	121%	138%
6	21	8	28	32	33%	300%
Total	155	95	222	169	43%	78%

Table 3: Count of creaky phonation in all 6 tones under normal and low-pitch condition by gender

The conflicting results on the mean F0 under two conditions could be caused by two reasons: (1) participants did not fully comprehend the requirements of experiment 2, or (2) some participants were not able to produce an even lower pitch range due to physiological constraints. This could explain why Tone 4 has the lowest percentage change in terms of creak occurrence in the low-pitch condition in Table 3. As Tone 4 generally hits the bottom of the pitch range of the participants, their occurrences in the low-pitch condition are similar to that of the normal condition. For future studies, it is recommended to give clearer guidance to participants in eliciting creaky production. For example, instead of asking the participants to speak in a lower-than-normal pitch range, researchers could ask the participants to speak in their lowest possible pitch range until their voices creak.

The average lowest measurable F0 for creaky tokens was recorded at 165 Hz for female participants and 90 Hz for male participants for experiment 2 for both conditions, as shown in Figure 3. These values are similar to that of experiment 1. As we manipulated speakers' pitch range in experiment 2, having similar F0 values to that of experiment 1 and higher creak occurrences hint that the pitch threshold values we obtained are possible F0 triggers for creaky phonation, regardless of tonal categories.

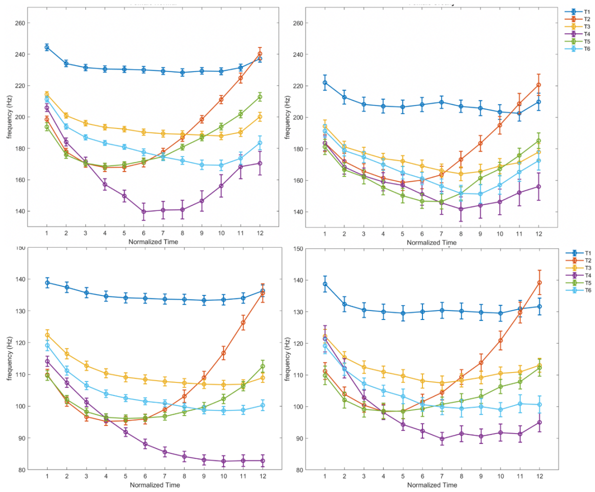

Figure 2: Mean F0 values in two production conditions (left: normal; right: low pitch) for female (upper) and male participants (lower) across time.

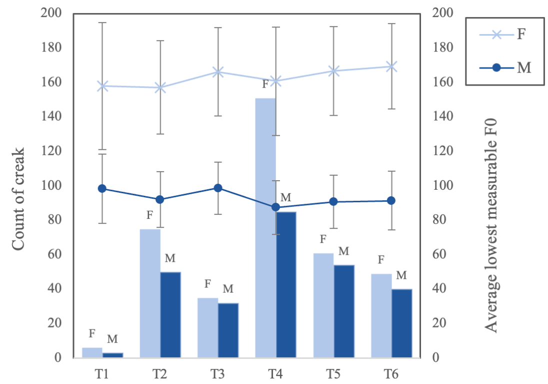

Figure 3: Pitch threshold values – measured by lowest measurable F0 and count of creaks for female and male participants in experiment 2.

6. CONCLUSION

In conclusion, this study demonstrates that creaky voice in HKC is under the influence of covariation of phonation and pitch range, and it is triggered when the speaker's pitch reaches an extremely low pitch, regardless of tonal categories. Some findings could be made from the results of the current study: (1) Tone 4 is not the only tone that creak in tonal production; Tone 2 and Tone 5 also shows considerable a number of creak occurrence and therefore Tone 4 is unlikely a language-specific cue for tonal identification. (2) Creaky voice is more likely to be produced at the second syllable, probably due to F0 declination. (3) A speaker's pitch range manipulation can trigger his/her creaky voice production.

7. ACKNOWLEDGEMENT

The author would like to thank Dr. Vanti Lee for providing access to the sound booth and equipment at CityUHK, as well as her guidance during the revision process. The author also thanks Dr. Jonathan Havenhil's continuous support and encouragement throughout the research process.

8. REFERENCES

[1] Yu, K. M., Lam, H. W. 2014. The role of creaky voice in Cantonese tonal perception. *J. Acoust. Soc. Am*. 136, 1320-1333.
[2] Kuang, J. 2017. Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *J. Acoust. Soc. Am*. 142, 1693-1706.
[3] Huang, Y. 2019. The role of creaky voice attributes in Mandarin tonal perception. *Proc. 19th ICPhS.* Melbourne.
[4] Zhang, Y., Kirby, J. 2020. The role of F0 and phonation cues in Cantonese low ton perception. *J. Acoust. Soc. Am*. 148, EL40-EL45
[5] Garellek, M. 2012. The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics*, 40, 152-161
[6] Bauer, R. S. 1998. Hong Kong Cantonese tone contours, in S. Matthews (ed.) *Studies in Cantonese Linguistics.* Hong Kong: Linguistic Society of Hong Kong.
[7]Fung, S. Y. R. 2015. Voice Quality: A Preliminary Study on the Phonetic Distinctions of Two Cantonese Accents. *Proc. 18th ICPhS*, Glasgow.
[8] Fung, S.Y.R., Lee, C.K.C. 2019. Voice quality and identity: The case of Hong Kong Cantonese. *The 30th North American Conference on Chinese Linguistics.* The Ohio State University, 16-33.
[9] Yuasa, I. P. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85(3), 315–337.
[10] Fung, R. S. Y., Lee, C. K. C. 2019. Tone mergers in Hong Kong Cantonese: An asymmetry of production and perception. *J. Acoust. Soc. Am*. 146, EL424-EL430.
[11] Mok, P., Zuo, D., Wong, P. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change*, 25(3), 341-370.
[12] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.
[13] Kuang, J. 2018. The influence of tonal categories and prosodic boundaries on the creakiness in Mandarin. *J. Acoust. Soc. Am* 143, EL509-EL515.
[14] Huang, Y. 2019. The source of creak in Mandarin utterance. *Proc. 19th ICPhS.* Melbourne.