Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

A Preliminary Analysis on Children's Phonation Contrast in Kunshan Wu Chinese Tones

Wenwei Xu[1], *Chunyu Ge*[1], *Wentao Gu*[2], *Peggy Mok*[1]

[1]The Chinese University of Hong Kong
[2]Nanjing Normal University
charliexu@link.cuhk.edu.hk

Abstract

Previous studies have established that phonation contrasts can be, apart from pitch, an important dimension of tonal contrasts in some languages, and modern Wu Chinese is a good example in which the lower register tones are produced with breathier phonation than the upper register tones. Nevertheless, researchers have shown that such phonation contrast is declining among young speakers in Shanghai and Suzhou Wu. This pilot study is thus motivated to investigate children's production in Kunshan Wu, a neighboring yet rather understudied dialect with more tones, in order to see if a similar trend is ongoing. Two male and two female school-age children (8;4 to 10;4) were recorded reading isolated monosyllabic words with different lexical tones, and simultaneous acoustic and electroglottographic (EGG) data were collected. Results of EGG and acoustic parameters demonstrate that at least near the onset of the vowel, glottal constriction is smaller and glottal closure is less abrupt in the lower register tones than in the upper register tones, suggesting that the lower register tones are generally produced with breathier phonation. Therefore, school-age child speakers of Kunshan Wu are still able to produce the phonation contrast between the tone registers.

Index Terms: tone registers, phonation contrasts, voice quality, breathy voice, Wu Chinese

1. Introduction

Previous studies have established that tonal contrasts can be multidimensional in some languages [1, 2], which means tones can be distinguished by not only pitch differences but also contrastive phonation types. Although there is more phonetic research on either production or perception of tone in these languages in recent years, the acquisition process remains rather understudied. In several latest studies, the development of lexical tones involving creaky voice in Cantonese [3] and Mandarin [4] was assessed, but creakiness in Cantonese T4 and Mandarin T3 is suggested to be mainly motivated by the low pitch targets [5], whereas non-modal phonation can be phonemic (e.g., Southern Yi [6]) or phonologically motivated as redundant cues (e.g., White Hmong [7], Northern Vietnamese [8]) in other languages.

Since redundant cues are more prone to sound change, phonation contrasts tend to be lost if they are not phonemic or the primary cues in perception, for which Wu Chinese is a good example. In modern Wu dialects, tone register is correlated with pitch height, onset voicing and voice quality [9, 10], and the lower register tones are usually produced with breathier phonation than the upper register tones in the word-initial position [9, 11, 12]. However, as the number of fluent Wu speakers is drastically decreasing in younger generations, the phonation contrasts have been found to be declining among younger speakers, for example, in Shanghai [13, 14, 15] and Suzhou [16]. Furthermore, many children in the Wu-speaking region are virtually like heritage speakers in that they acquire Wu almost exclusively from their primary caretakers at home while they mainly speak Mandarin elsewhere. Under such circumstances, research on children's production and perception of the phonation contrasts in Wu Chinese is urgently needed.

In addition, the acquisition of Wu Chinese tones is of great research interest to children's development of tone and voice. First, researchers have shown that the acquisition of lexical tones starts at the earliest stage [17], and children can be considered to have generally acquired the tone systems by the age of 2 or 3 years in languages such as Cantonese [18], Mandarin [17] and Thai [19] using transcription data. However, the acquisition is believed to be complete much later if adultlike production accuracy and acoustic patterns are required, for example, no earlier than age of 6 years in Cantonese [3, 20]. In fact, scholars believe that children's speech production development follows a protracted time course, which involves not only production sufficiently adultlike to be perceived and transcribed as accurate, but also adultlike speech motor control evaluated in acoustic and kinematic analysis [21]. This is important because for tone systems with phonation contrasts, the acquisition should additionally concern the development of children's voices, and indeed the anatomy of the vocal folds change gradually and is not fully developed until after puberty, which affects their vocal range as well as voice quality [22]. Another interesting point from the perspective of perception is that while creakiness has been shown to enhance tone identification [8, 23, 24], breathiness is expected to enhance register identification [15, 16, 25] but spoil periodicity and thus perception of pitch contrasts [26, 27], which is likely to pose difficulties for children to acquire the tone systems.

To sum up, the superimposition of phonation contrasts on pitch contrasts in Wu Chinese may prolong the full acquisition of the tone systems, while redundant phonation contrasts tend to be lost and a decline has already taken place in some Wu dialects. Therefore, in this pilot study, school-age children's production of Kunshan Wu, a neighboring yet rather understudied dialect of Suzhou and Shanghai Wu, is investigated to see if a similar trend is ongoing before further looking into the development of lexical tones in younger children and comparing them with the adults. Our recent acoustic study on the adult production of Kunshan Wu [28] measured various spectral tilts and noise measures, and the results confirmed that the phonation contrast between tone registers were still robust among the elderly (above 70 years

10.21437/SpeechProsody.2022-85

old) and middle-aged (35–50 years old) speakers, and the lower register tones were produced with breathier phonation than the upper register tones were.

On top of the methods previously used, analysis on electroglottography (EGG) is also incorporated in the current study, given that EGG waveforms are useful for reflecting glottal activities and have been used to study voice quality in adult speech of several languages, including Shanghai Wu [29]. Although for children EGG has been only used for research on pediatric voice disorders [30, 31], analyzing EGG data can help us constitute a more comprehensive understanding of the children's phonation contrast in Kunshan Wu tones.

2. Method

2.1. Speakers

Two female and two male Mandarin-Wu bilingual children at school-age (8;4 to 10;4, see Table 1) were recorded in this study. All the speakers were born and raised in Kunshan. They learnt Mandarin and Wu since birth and used both of them in daily life, while they spoke Kunshan Wu mostly to their primary caretakers who were all native speakers of the downtown dialect. All the child speakers were recognized by the first author who is a native Kunshan speaker to speak Kunshan Wu with native proficiency during an initial 15-minute conversation about their home and school life.

Table 1: *Age of the speakers and auditory judgment results (percentages of breathy tokens).*

Speaker	Age	L1 Upper	L1 Lower	L2 Upper	L2 Lower
F1	8;4	16.2%	87.9%	10.8%	63.6%
F2	9;10	5.3%	80.0%	2.7%	86.0%
M1	8;6	9.7%	60.3%	6.9%	52.9%
M2	10;4	4.5%	76.5%	0.9%	73.5%
(Pooled)	–	**8.1%**	**76.5%**	**4.6%**	**71.1%**

2.2. Materials

The wordlist contained 73 isolated monosyllabic words (see examples in Table 2). Each word contains an obstruent onset (/p b t d k g f v s z tɕ dʑ/), a monophthong (/a, ɛ, ɪ, ɔ/), and carries one of the five unchecked tones (T1a mid-level, T1b rising, T2a high-falling, T3a high-dipping, T3b low-dipping).

Table 2: *Examples from the wordlist*

Tonal category	Upper (a) modal	Lower (b) breathy
	/ka/	/ga/
Ping **(1)**	加 'plus'	茄 'eggplant'
Shang **(2)**	假 'fake'	解 'loosen'
Qu **(3)**	嫁 'marry'	

2.3. Procedure

Simultaneous acoustic and EGG recordings were collected for two to three repetitions of the wordlist from each speaker at their home, using the Behringer ECM 8000 Ultra-Linear Measurement Condenser Microphone and the Glottal Enterprises EG2-PCX2 EGG unit. A total of 706 tokens were analyzed. To obtain an overall picture of the children's accuracy of producing the phonation contrast, each token was auditorily judged by two listeners: a native speaker of Kunshan Wu (L1 = the first author) and a native speaker of another Wu dialect (L2 = the second author), who decided whether each token sounded breathy (see Table 1).

F0 and EGG parameters were calculated from the EGG waveforms that were band-pass filtered following [32]. The EGG parameters are the Closed Quotient (CQ, the ratio of the duration of the closed phase to the whole period in a glottal cycle) [33], the Speed Quotient (SQ, the ratio of the duration of the closing phase to the opening phase in a glottal cycle) [34], and the Peak Increase in Contact (PIC, also called Derivative-EGG Closure Peak Amplitude, the positive peak amplitude in the first derivative of EGG) [35]. The results were sampled at nine equidistant points within the vowel (excluding onset and offset).

F0, spectral tilts and noise measures were measured on the acoustic waveforms using VoiceSauce [36]. Spectral tilts are the differences between the amplitude of various harmonics (H1-H2, H1-A1, H1-A2, H1-A3, H2-H4, H4-H2K, H2K-H5K, H1-H4, H1-H2K, H1-H5K), which were corrected for formant influence (correction marked by * in Figure 2) [37]. Noise measures are the harmonics-to-noise ratios (HNR) and Cepstral Peak Prominence (CPP). The results were averaged on nine normalized-time intervals, i.e., ninths of the vowel.

For each parameter, within-speaker normalization was done, with outliers discarded, and Soothing-Spline ANOVA (SSANOVA) was performed to analyze differences between the tone registers.

3. Results

3.1. F0

The F0 calculated from the periods of the glottal cycles is normalized and the contours are plotted in Figure 1 (left), while the F0 measured on acoustic data adopting the STRAIGHT algorithm [38] is shown in Figure 1 (right). The results were highly consistent. It is obvious that the upper register tones had a higher pitch onset, and the pitch in either register did not cross the overall mean until the midpoint of the vowel, showing that the tone register contrast was well maintained during at least the first half of the vowel. Mean duration of the lexical tones across the speakers is summarized in Table 3.

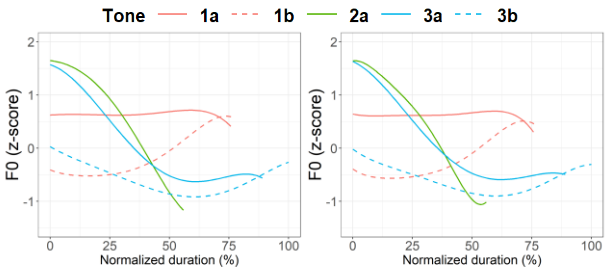

Figure 1: *Normalized F0 contours of the lexical tones (left: EGG; right: acoustic).*

Table 3: *Mean duration (ms) of the lexical tones.*

Tone	1a	1b	2a	3a	3b	Pooled
Duration	313	314	231	368	413	**322**

3.2. Auditory Judgment

The percentages of tokens that sounded breathy to each listener in each register are summarized in Table 1. The overall inter-rater reliability was moderate (Cohen's Kappa = 0.59), so two listeners' judgment results were further analyzed separately. Either listener's judgement indicates that the children were generally more likely to produce breathy tokens for the lower register tones than for the upper register tones (Boschloo's test: $p < 0.001$), which is also true for them individually (Boschloo's tests: $p < 0.001$).

In pairwise comparison, the child M1 was found to be less likely to produce breathiness for the lower register tones than F1 and F2 based on the first author's judgment, or than M2 and F2 based on the second author's judgment (Boschloo's tests: $p < 0.005$). Both listeners agreed that M1 produced some lower register tokens with remarkably harsh voice quality, which was not found in the other children's production.

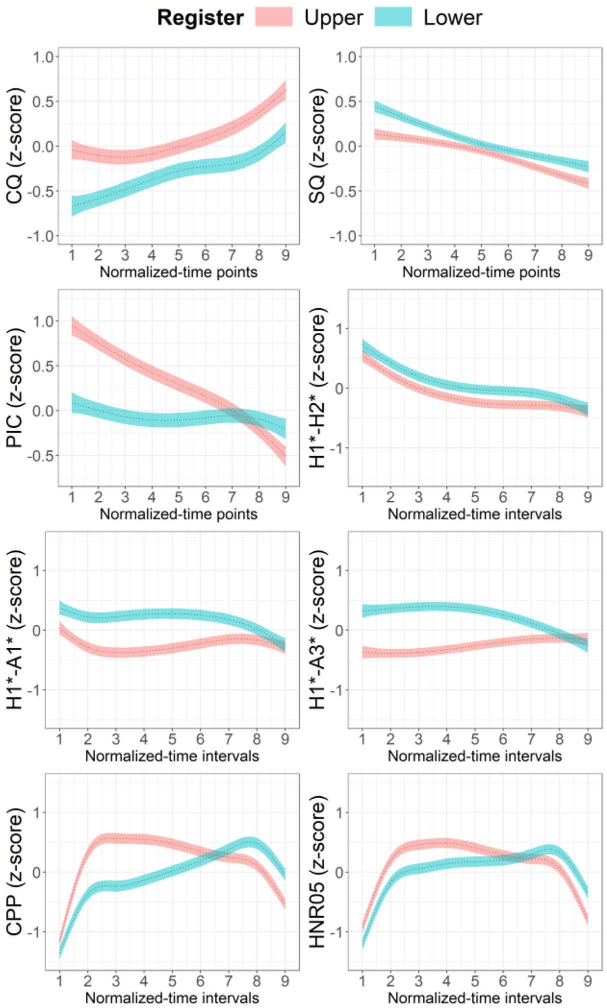

Figure 2: *SSANOVA plots of selected EGG and acoustic parameters (α = 0.01).*

3.3. EGG parameters

The results of SSANOVA on the EGG parameters are plotted in Figure 2. On the one hand, CQ was generally smaller in the lower register than in the upper register throughout the vowel, suggesting a shorter closed phase within a glottal cycle, thus smaller glottal constriction in the lower register. On the other hand, SQ was generally larger in the lower register than in the upper register throughout the vowel, suggesting a more gradual increase in vocal fold contact, thus less abrupt glottal closure in the lower register. PIC was smaller in the lower register at the first six time points, which means the peak velocity was lower at the instant of glottal closure. Although this coincides with the expectation that the glottal closure is less abrupt, the interpretation of this parameter requires attention and is discussed in Section 4.2.

3.4. Acoustic parameters

Significant differences were found in most of the spectral tilts and noise measures (i.e., H1*-H2*, H1*-A1*, H1*-A2*, H1*-A3*, H2*-H4*, H2K*-H5K, H1*-H4*, H1*-H2K*, H1*-H5K, HNR05, CPP). The spectral tilts were generally larger in the lower register than in the upper register over at least the first half the vowel. Among these parameters, a larger H1*-H2* is believed to indicate smaller glottal constriction [39], and a larger H1*-An* is believed to reflect less abrupt glottal closure, with H1*-A1* also related to a posterior glottal opening [40]. This is consistent with the findings based on the EGG parameters. Moreover, the noise measures were generally smaller in the lower register than in the upper register over at least the first half the vowel, meaning that there was comparatively more noise or higher aperiodicity in the lower register. The results of SSANOVA on these representative acoustic parameters (H1*-H2*, H1*-A1*, H1*-A3*, HNR05, CPP) are plotted in Figure 2.

4. Discussion

4.1. Pitch

Since F0 is usually measured from the acoustic data in recent phonetic studies on tone, the close resemblance between the F0 contours calculated from the EGG waveforms and those measured from the acoustic waveforms first confirms that EGG can reliably reflect the glottal activities of children. Furthermore, as the pitch contours produced by the children highly resemble those by the adults in our earlier study [28], children at school age can already produce fairly adult-like lexical tones in terms of pitch patterns.

However, there are still a few differences between the children and the adults that are even perceptible to the first author as a phonetically trained native speaker. For example, the children's T1a sounded "protracted". This is consistent with the data where the mean duration of the children's T1a was apparently longer than the adults', for whom the lower register tones were found to be generally longer than their upper register counterparts (e.g., T1a < T1b) [28]. Besides, T1a also sounded "flatter", which again concurs with the level contour of the children's T1a, in contrast with a visible rise in F0 of the adults' T1a [28]. It is very likely that the children produce T1a in a way that resembles how they produce Mandarin T1 due to the strong influence from Mandarin, but it is not impossible that these are covert differences [41] that simply reflect incomplete development and will be fine-tuned later. Data from young speakers (e.g., in 20s) who are also early Mandarin-Wu bilinguals is needed to confirm if there is a similar pattern that could have persisted into their adulthood.

4.2. Phonation

The auditory judgement results show that breathiness was highly associated with the lower tone register, indicating that children at school age have acquired the phonation contrast, while further comparison with the adults' performance is needed to determine whether they can achieve adult-like accuracy. Based on such observation, both EGG and acoustic parameters relevant to voice quality were calculated, and the results show that the glottal constriction was smaller (smaller CQ and larger H1*-H2*), and the glottal closure was less abrupt (larger SQ and larger H1*-An*) in the lower register, which altogether implies that the phonation type of the lower register tones is breathier than that of the upper register tones. Therefore, children at school age can still produce the phonation contrast between the tone registers.

Compared with adult production [28], larger spectral tilts and smaller noise measures were found in the lower register for both groups of speakers. Nevertheless, H1*-H2* and H1*-H2K* exhibited consistent differences between tone registers among the children, but not among the adults. Based on the assumption that, for example, H1*-H2* is correlated with glottal open quotient [39], speakers may have varied the magnitude of their phonation contrast and the specific breathy phonation sub-type produced (e.g., breathy voice versus whispery voice). While such differences seem to be merely covert inter-speaker variation that may not concern the native speakers as long as they all sound breathy, it is also possible that the differences constitute sociophonetic contrasts that are related to their identity, including age, gender, social class and stereotype [42, 43].

For example, in Kunshan Wu, H1*-H2* was smaller in the lower register in middle-aged female speech, as opposed to a larger H1*-H2* in the lower register in male speech [28]. This implies that the middle-aged female speakers could have produced a somewhat different phonation type, although in general it still sounded breathy to the first author. Meanwhile, although the data of young speakers were not collected, impressionistically many young female speakers sound less or even not breathy to the first author. Based on the first author's social experience, it is then hypothesized that young female speakers in Kunshan tend to produce weaker breathiness to sound less harsh and more elegant, probably because an authentic and canonical accent can be associated with a stereotypical impression of being low-ranked, provincial and poorly educated, which is not uncommon concerning the use of dialects other than the lingua franca [42, 44]. More research, including analysis on the EGG data of adult production and perception, need be done to provide more evidence. Regardless of inter-speaker variation among adults, it is still an interesting question whether and, if any, how the children will alter their phonation behavior when growing up, especially during and after puberty when there are rapid anatomical changes in the larynx.

Finally, it is worth noting that the pattern of PIC found in the current study is unusual. In Kunshan Wu, PIC was smaller while SQ was larger in the lower register, both of which suggest less abrupt glottal closure. This is consistent with our expectation that the phonation type of the lower register is breathier. In previous studies where researchers started to extensively use PIC to distinguish different phonation types, PIC was, contrary to the researchers' initial expectation, larger for breathier phonation types (e.g., in Southern Yi and White Hmong [6, 7, 45]). While a larger PIC should indicate faster glottal closure, the researchers proposed some hypotheses and accepted the usefulness of PIC [45]. However, PIC also showed no significant difference in Gujarati [46] and Shanghai Wu [29]. Therefore, at the current stage, it is still not clear whether such discrepancy is a result of cross-linguistic differences, or just because PIC is actually not a reliable parameter due to measurement or definitional issues.

5. Conclusion

In summary, the results of EGG and acoustic parameters show that at least near the onset of the vowel, glottal constriction was smaller and glottal closure was less abrupt in the lower register tones than in the upper register tones, which indicates that the lower register tones were generally produced with breathier phonation than the upper register tones were. Therefore, the school-age children in Kunshan are still able to produce the phonation contrast between tone registers in Wu Chinese, as opposed to a declining trend found among young speakers in Shanghai and Suzhou. Besides, although the children could produce fairly adultlike pitch contours and the phonation contrast, some covert or perceptible differences from adult production were also found, suggesting that the acquisition of the tone system might not be complete, or there can be age differences partly due to the strong influence from Mandarin. Future research on the production of younger children and young adults (in their 20s) and the perception of all age groups is thus called for to reveal the development of the tone system.

6. Acknowledgements

This study was supported by the RGC General Research Fund 2019/20 Project Reference 14607619 awarded to the last author, and also partially supported by the Major Program of the National Social Science Fund of China (13&ZD189) and a project from Jiangsu Higher Institutions' Excellent Innovative Team for Philosophy and Social Sciences (2017STD006) awarded to the third author.

7. References

[1] P. Rose, "Acoustic characteristics of the Shanghai-Zhenhai syllable-types," in *Papers in Southeast Asian Linguistic, No. 8: Tonation*, D. Bradley, Ed, Canberra: Australian National University, 1982, pp. 1–53.

[2] J. Gao and P. Hallé, "Phonetic and phonological properties of tones in Shanghai Chinese," *Cahiers De Linguistique-Asie Orientale,* vol. 46, no. 1, pp. 1–31, 2017.

[3] P. Mok, G. Li and & H. Fung, "Development of phonetic contrasts in Cantonese tone acquisition," *Journal of Speech, Language, and Hearing Research,* vol. 63, no. 1, pp. 95–108, 2020.

[4] N. Rhee, A. Chen and J. Kuang, "Going beyond F0: The acquisition of Mandarin tones," *Journal of Child Language,* vol. 48, no. 2, pp. 389–398, 2021.

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

[6] K., Jianjing and P. Keating, "Vocal fold vibratory patterns in tense versus lax phonation contrasts," *The Journal of the Acoustical Society of America,* vol. 136, no. 5, pp. 2784–2797, 2014.

[7] C. M. Esposito, "An acoustic and electroglottographic study of White Hmong tone and phonation," *Journal of Phonetics,* vol. 40, no. 3, pp. 466–476, 2012.

[8] M. Brunelle, “Tone perception in Northern and Southern Vietnamese,” *Journal of Phonetics*, vol. 37, pp. 79–96, 2009.
[9] M. Yip, “The tonal phonology of Chinese,” Ph.D. dissertation, MIT, Cambridge, MA, 1980.
[10] S. Duanmu, “A formal study of syllable, tone, stress and domain in Chinese languages,” Ph.D. dissertation, MIT, Cambridge, MA, 1990.
[11] Y. R. Chao, *Studies in the modern Wu-dialects*. Peking: Tsinghua College Research Institute, 1928.
[12] J. Cao and I. Maddieson, “An exploration of phonation types in Wu dialects of Chinese,” *Journal of Phonetics,* vol. 20, pp. 77–92, 1992.
[13] J. Tian and J. Kuang, “Revisiting the register contrast in Shanghainese,” in *Proceedings of the 5th Symposium on the Tonal Aspects of Languages*, Buffalo, NY, 2016, pp. 147–151.
[14] J. Gao, “Sociolinguistic motivations in sound change: On-going loss of low tone breathy voice in Shanghai Chinese,” *Papers in Historical Phonology,* vol. 1, pp. 166–186, 2016.
[15] J. Gao, P. Hallé and C. Draxler, “Breathy voice and low-register: A case of trading relation in Shanghai Chinese tone perception?” *Language and Speech,* vol. 63, no. 3, pp. 582–607, 2020.
[16] C. Ge, W. Xu, W. Gu and P. Mok, “Age differences of the non-modal phonation in Suzhou Wu Chinese,” presented at *1st International Conference on Tone and Intonation (TAI),* Sonderborg, Denmark, 2021.
[17] C. N. Li and S. A. Thompson, “The acquisition of tone in Mandarin-speaking children,” *Journal of Child Language,* vol. 4, no. 2, pp. 185–199, 1977.
[18] L. K. H. So and B. Dodd, “The acquisition of phonology by Cantonese-speaking children,” *Journal of Child Language,* vol. 22, no. 3, 473–495, 1995.
[19] H. Clumeck, “The acquisition of tone,” In *Child Phonology: Production,* G. H. Yeni-Komshian, J. F. Kavanagh and C. A. Ferguson, Eds. Cambridge, MA: Academic Press, 1980, pp. 257–275.
[20] P. S. Wong, and C. T.-T. Leung, “Suprasegmental features are not acquired early: Perception and production of monosyllabic Cantonese lexical tones in 4- to 6-year-old preschool children,” *Journal of Speech, Language, and Hearing Research,* vol. 61, no. 5, pp. 1070–1085, 2018.
[21] B. Munson, J. Edwards and M. E. Beckman, “Phonological representations in language acquisition: Climbing the ladder of abstraction,” In *The Oxford Handbook of Laboratory Phonology*, A. C. Cohn, C. Fougeron and M. K. Huffman, Eds. Oxford, UK: Oxford University Press, 2011, pp. 288–209.
[22] A. McAllister and P. Sjölander, “Children's voice and voice disorders,” *Seminars in speech and language,* vol. 34, no. 2, pp. 71–79. 2013
[23] R. X. Yang, “The phonation factor in the categorical perception of Mandarin tones,” in *Proceedings of ICPhS XVII,* Hong Kong, 2011, pp. 2204–2207.
[24] K. Yu, and H. W. Lam, “The role of creaky voice in Cantonese tone perception,” *The Journal of the Acoustical Society of America,* vol. 136, no. 3, pp. 1320–1333, 2014.
[25] M. Garellek, P. Keating, C. M. Esposito and J. Kreiman, “Voice quality and tone identification in White Hmong,” *The Journal of the Acoustical Society of America,* vol. 133, no. 2, pp. 1078–1089, 2013.
[26] D. Silverman, “Phasing and recoverability,” Ph.D. dissertation, UCLA, Los Angeles, CA, 1995.
[27] B. Blankenship, “The timing of nonmodal phonation in vowels,” *Journal of Phonetics*, vol. 30, pp. 163–191, 2002.
[28] W. Xu, and P. Mok, “The acoustic correlates and time span of the non-modal phonation in Kunshan Wu Chinese,” In *Proceedings of the 12th International Symposium on Chinese Spoken Language Processing (ISCSLP 2021)*, Hong Kong, 2021.
[29] J. Tian and J. Kuang, “The phonetic properties of the non-modal phonation in Shanghainese,” *Journal of the International Phonetic Association,* pp. 1–27, 2019.
[30] H. A. Cheyne, R. C. Nuss and R. E. Hillman, “Electroglottography in the pediatric population,” *Archives of Otolaryngology-Head & Neck Surgery,* vol. 125, no. 10, pp. 1105–1108, 1999.
[31] K. Szklanny, R. Gubrynowicz, J. Ratyńska and D. Chojnacka-Wądołowska, “Electroglottographic and acoustic analysis of voice in children with vocal nodules,” *International Journal of Pediatric Otorhinolaryngology,* vol. 122, pp. 82–88, 2019.
[32] C. T. Herbst, W. T. Fitch and J. G. Svec, “Electroglottographic wavegrams: a technique for visualizing vocal fold dynamics noninvasively,” *The Journal of the Acoustical Society of America,* vol. 128, no. 5, pp. 3070–3078, 2010.
[33] D. M. Howard, “Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers,” *Journal of Voice,* vol. 9, no. 2, pp. 163–172, 1995.
[34] K. Marasek, “Electroglottographic description of voice quality,” Ph.D. dissertation, Institute fur Maschinelle Sprachverarbeitung, Stuttgart, 1997.
[35] A. Michaud, “Final consonants and glottalization: New perspectives from Hanoi Vietnamese,” *Phonetica,* vol. 61, no. 2–3, pp. 119–146, 2004.
[36] Y. Shue, P. Keating, C. Vicenik and K. Yu, “VoiceSauce: a program for voice analysis,” in *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong*, 2011, pp. 1846–1849.
[37] M. Iseli and A. Alwan, “An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation,” in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Que.*, 2004, pp. 669–672.
[38] H. Kawahara, I. Masuda-Katsuse and A. de Cheveigne, “Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency based F0 extraction,” *Speech Communication,* vol. 27, pp. 187–207, 1999.
[39] J. Kreiman, M. Iseli, J. Neubauer, Y.-L. Shuea, B. R. Gerratt and A. Alwan, “The relationship between open quotient and H1*–H2*,” *The Journal of the Acoustical Society of America,* vol. 124, pp. 2495, 2008.
[40] H. M. Hanson, K. N. Stevens, H.-K. J. Kuo, M. Y. Chen and J. Slifka, “Towards models of phonation,” *Journal of Phonetics,* vol. 29, pp. 451–480, 2001.
[41] J. Edwards and M. E. Beckman, “Methodological questions in studying consonant acquisition,” *Clinical Linguistics & Phonetics*, vol. 22, no. 12, pp. 937–956, 2008.
[42] S. Wilhelm, “Voice quality in British English. Its nature, functions and applications,”, *Anglophonica,* vol. 27, 2019.
[43] R. J. Podesva and P. Callier, “Voice Quality and Identity,” *Annual Review of Applied Linguistics,* vol. 35, pp. 173–194, 2015.
[44] B. Liu, “Social class, language attitudes, and language use: A case study of Chinese university students,” *Chinese Language and Discourse,* vol. 11, no. 1, pp. 5–30, 2020.
[45] P. Keating, C. Esposito, M. Garellek, S. Khan and J. Kuang, “Phonation contrasts across languages,” *UCLA Working Papers in Phonetics,* vol. 108, pp. 188–202, 2010.
[46] S. ud D. Khan, “The phonetics of contrastive phonation in Gujarati,” *Journal of Phonetics,* vol. 40, no. 6, pp. 780–795, 2012.