Tonal Aspects of Languages 2016
24-27 May 2016, Buffalo, New York

Revisiting the Register Contrast in Shanghai Chinese

Jia Tian, Jianjing Kuang

Department of Linguistics, University of Pennsylvania
jiatian@sas.upenn.edu, kuangj@sas.upenn.edu

Abstract

Earlier studies claimed that the contrastive tonal registers in the Shanghai dialect of Chinese were distinguished by contrastive phonation types, with the lower register showing breathier phonation. However, more recent studies found that younger speakers of Shanghai Chinese did not use phonation cues to contrast the tonal registers. The different results between the studies may indicate an ongoing sound change: phonation contrasts are merging among younger speakers. To better understand the phonetic realization of the tonal registers in Shanghai Chinese and to investigate whether a sound change is underway, speakers from both the older generation and the younger generation were recruited to produce pairs of words in different registers. Simultaneous audio and EGG signals were collected, and extensive acoustic and articulatory measurements were extracted. Significant age difference was found. Older speakers reliably produce contrastive phonation types for different registers, as demonstrated by significant differences in the overall spectral slope and the periodicity in acoustic signals, and the contact quotient in EGG signals. The lower register is breathier. In contrast, younger speakers do not produce breathier phonation for the lower register. This study provides evidence for the suggestion that Shanghai Chinese is gradually losing the breathy voice associated with the lower register.

Index Terms: Shanghai Chinese, tonal register, phonation, breathy voice

1. Introduction

Chinese Wu dialects possess a large number of tones. Most Wu dialects have seven or eight tones. Some varieties even have nine or ten tones [1, 2]. Since Yip proposed the register feature [3], the term register has been widely used to analyze Wu dialects. Previous studies show that Wu dialects have two contrastive tonal registers, upper and lower, characterized by contrastive pitch values and phonation types [3, 4, 5]. The upper register is higher in pitch, and is produced with modal voice, while the lower register is lower in pitch, and is produced with non-modal phonation. An example of the systematic register contrasts in the Songjiang dialect (a variant of Wu spoken in the suburban area of Shanghai) is given in Table 2 [6]:

Table 1: Songjiang tones.

Register	Tones			
Upper	53	44	35	5
Lower	31	22	13	3

Since the phonation contrast in Wu was acknowledged, much discussion has been devoted to the phonetic correlates of this contrast. Among the Wu dialects, Shanghai Chinese has received the most instrumental studies. The consensus of the studies in 1980s and 1990s is that the non-modal phonation associated with the lower register is generally breathier (c.f. Ladefoged 1971's scale [7]). Acoustically, spectral measures (H1-H2 and H1-A1) are significantly different between the two registers. Articulatorily, airflow/pressure ratio is higher in the low register. Photoglottaraphy data show that maximal glottal opening is greater for the non-modal phonation [8, 9]. However, due to the limitations of technology at that time, these studies only involved very few speakers.

Recently, advanced tools have been developed for voice studies, offering us the opportunity to study the phonation contrast in Shanghai Chinese more extensively [10, 11]. However, a recent study with speakers born in the 1990s by Zhang and Yan measured H1*-H2*, H1*-A1*, H1*-A2*, H1*-A3* and CPP, but found no consistent phonation difference between registers [12]. They found that the major difference between the two registers was F0. Gao and Hallé measured OQ and H1-H2, and also found no phonation contrast for young speakers [13]. Noting that only older male speakers still produced phonation contrasts, they suggested that Shanghai Chinese is developing a new sound change in which the phonation contrast is gradually lost. However, few older speakers were investigated, making the conclusion less convincing.

To resolve the discrepancies between earlier work and the more recent work, and to better understand the acoustic and physiological properties of the register contrast in Shanghai Chinese, the current study investigates the production of speakers from different age and gender groups. We ask whether speakers from different age (older vs. younger) and gender (male vs. female) groups differ in making the register contrast, and if they do, then how different speakers make the contrast differently.

2. Methods

2.1. Language

Several dialects are spoken in the city of Shanghai. The language of our focus is the Shanghai Chinese spoken in the central area of Shanghai, as opposed to the Shanghai Chinese spoken in the suburban area of Shanghai. It is usually taken as the typical example of the Wu dialects of Chinese. It has also been variously referred to as Shanghai , Shanghai dialect and Shanghainese [8, 12, 14, 15, 16].

Shanghai Chinese has five lexical tones [17].Three of them are upper register tones, and the other two of them are lower register tones. Of the three upper register tones, two are long tones that end in nasals or vowels, while the other is a short tone that only co-occurs with closed syllables with glottal codas. One of the lower register tones is a long tone, and the other one short. Adopting the five-scale pitch system proposed by Chao [18], the five lexical tones of Shanghai Chinese are summarized in Table 2 [17]. In the system adopted here, the five tones are transcribed as 53 (high-falling), 34 (high-rising), <u>55</u> (short high-level), 23 (low-rising) and <u>12</u> (short low-rising).

 http://dx.doi.org/10.21437/TAL.2016-32

Table 2: Five lexical tones in Shanghainese. Underlines denote short tones.

Register	Tones		
Upper	53	34	55
Lower		23	12

2.2. Speech materials

The wordlist used in this study included 3 pairs of monosyllabic words for each of the 5 manners of articulation being examined. The 5 manners are stop, affricate, fricative, nasal and zero. The monosyllabic pairs contain a contrast between two rising tones that belong to two registers (the high-rising tone and the low-rising tone). Laterals were not included because they do not co-occur with the high-rising tone. In total, the dataset included about 1200 tokens (5 manners * 3 pairs * 2 words per pair * 2 repetitions * 20 speakers). None of the target syllables begins with aspirated consonant. This was done to reduce the effect of aspiration. Most of the target syllables contain mid and low vowels /a/ and /ɛ/ to eliminate the chance that a high F1 influences H2. Table 3 gives a sample of the words chosen.

Table 3: Examples of words used in this study.

Register	Monosyllabic word
Upper	/pa34/ 'visit'
Upper	/ta34/ 'put on'
Upper	/ka34/ 'false'

2.3. Speakers

Twenty native speakers of Shanghai Chinese were recorded. Speakers ranged 17-83 years of age in the year of 2015. All speakers were born in Shanghai. All speakers reported that they used Shanghai Chinese daily. Taking 45 years old in 2015 as the dividing line, speakers are evenly distributed in age and gender. Table 4 summarizes the background information of the speakers.

Table 4: Background information of the speakers.

Group	Number	Average age	Age range
Older male	5	56	45-83
Older female	5	51	45-55
Younger male	5	22	17-32
Younger female	5	25	22-28

2.4. Procedure

Simultaneous audio and electroglottographic (EGG) recordings were made for all 20 speakers. Acoustic and EGG measures were extracted automatically using VoiceSauce and EggWorks, respectively [10, 11]. Eleven spectral measures and one EGG measure were extracted for each vowel. The acoustic measures included F0, the difference between the first and second harmonics (H1*-H2*), the difference between the first harmonic and the first, second and third formants (H1*-A1*, H1*-A2*, H1*-A3*), Ceptral Peak Prominence (CPP), Subharmonic-to-Harmonic Ratio (SHR), and Harmonic-to-Noise Ratios (4 frequency bands, 0-500Hz (HNR05), 0-1500Hz (HNR15), 0-2500Hz (HNR25) and 0-3500Hz (HNR35)). Corrections were done automatically in VoiceSauce. The EGG measure was Contact Quotient (CQ). The phonation measures mentioned above have been found to be reliable indicators of phonation contrast across languages (e.g., Mazatec: [19]; Takhian Thong Chong: [20]; Southern Yi: [21]; White Hmong: [22]). VoiceSauce combined the measurements and reported the mean value of the whole vowel. Within-speaker normalization was done to minimize differences across speakers and recording conditions.

3. Results

A series of mixed-effect models (one per measure for each manner in each group) were used to evaluate main effects of register on acoustic and EGG measures, using the lmerTest package in R, with register as the main effect, and speaker and item as the random effects. The measurements for each vowel were used as independent variables. lmerTest reported the t-values and p-values for the fixed effects. P-values smaller than or equal to 0.05 were considered to be statistically significant, while p-values greater than 0.05 were considered nonsiginificant. A full report of model parameters for fixed-effect terms is provided in Table 5. The critical results are summarized in Table 6. In Table 6, a check is assigned if significant differences are found in all onset types. Fig. 1 to Fig. 4 show the mean value and standard error of F0, H1*-A1*, CPP and CQ for all four age groups investigated. These figures visualize how the register contrast is maintained or lost across speaker age and gender.

Table 5: Register contrast of monosyllabic words. Values are t-values for the fixed effect. Values greater than zero indicate higher value for the lower register. Asterisks indicate statistical signifance at $p < .05$ level.

Older male					
Measure	Stop	Affricate	Fricative	Nasal	Zero
F0	-8.402*	-8.916*	-7.334*	-10.695*	-7.213*
H1*-H2*	6.457*	-1.125	5.891*	-1.688	4.513*
H1*-A1*	8.396*	6.136*	4.679*	6.387*	5.803*
H1*-A2*	5.663*	5.263*	4.170*	7.309*	4.749*
H1*-A3*	2.945*	1.542	4.322*	-0.686	4.340*
CPP	-8.105*	-4.394*	-8.790*	-8.442*	-8.348*
CQ	-6.290*	-6.123*	-6.691*	-6.225*	-5.860*
HNR05	-13.214*	-12.010*	-7.713*	-8.224*	-9.711*
HNR15	-14.161*	-10.930*	-4.729*	-6.849*	-3.485*
HNR25	-13.082*	-10.061*	-6.020*	-4.932*	-3.341*
HNR35	-12.163*	-7.519*	-5.825*	-5.056*	-3.889*
SHR	-2.258*	-3.512*	-3.044*	-2.192*	-3.945*
Older female					
Measure	Stop	Affricate	Fricative	Nasal	Zero
F0	-8.849*	-6.781*	-5.924*	-6.195*	-4.932*
H1*-H2*	-3.643*	6.269*	-4.780*	-1.300	0.420
H1*-A1*	3.468*	0.826	2.417*	0.599	0.821
H1*-A2*	4.792*	3.757*	2.395*	3.965*	1.479
H1*-A3*	4.124*	2.073*	5.343*	0.289	2.115*
CPP	-4.836*	-5.922*	-8.795*	-4.253*	-4.598*
CQ	-5.169*	-2.686*	-2.732*	-2.686*	-4.417*
HNR05	-8.806*	-9.943*	-10.602*	-7.741*	-7.328*
HNR15	-7.338*	-7.853*	-5.961*	-7.197*	-3.106*
HNR25	-7.090*	-9.293*	-6.244*	-5.362*	-2.607*
HNR35	-7.178*	-9.195*	-7.815*	-4.562*	-3.315*
SHR	5.858*	-2.753*	-1.547	1.581	3.368*
Younger male					
Measure	Stop	Affricate	Fricative	Nasal	Zero
F0	-2.375*	-5.323*	-7.606*	-5.689*	-7.473*
H1*-H2*	0.723	-2.258*	-0.496	-2.811*	-4.153*
H1*-A1*	0.303	1.898	-0.269	1.536	2.567*
H1*-A2*	3.416*	-0.051	2.497*	1.426	2.854*
H1*-A3*	-0.649	1.118	2.007*	-0.211	0.971
CPP	-6.715*	-4.552*	-5.927*	-5.751*	-3.484*
CQ	-0.789	-1.762	-1.343	-1.277	-1.940
HNR05	-7.487*	-9.143*	-6.912*	-8.855*	-5.664*
HNR15	-1.753	-7.058*	-1.254	-8.488*	-3.455*
HNR25	-1.712	-8.509*	-1.206	-6.573*	-3.657*
HNR35	-2.076*	-8.254*	-1.022	-6.273*	-3.933*
SHR	-3.486*	-2.447*	-1.907	-3.012*	-1.916
Younger female					
Measure	Stop	Affricate	Fricative	Nasal	Zero
F0	-9.967*	-5.780*	-6.196*	-9.347*	-7.570*
H1*-H2*	-5.298*	0.656	-4.096*	1.381	-0.329
H1*-A1*	-2.107*	-0.550	-2.316*	1.988*	-0.232
H1*-A2*	0.486	1.644	-0.115	4.425*	-0.328
H1*-A3*	1.434	-0.041	-0.408	1.721	0.903
CPP	-5.015*	-4.156*	-3.086*	-2.059*	-3.150*
CQ	4.105*	-0.973	1.319	0.661	0.979
HNR05	-9.284*	-9.235*	-4.881*	-5.687*	-4.104*
HNR15	-9.904*	-6.003*	-2.414*	-5.120*	-2.362*
HNR25	-10.262*	-5.668*	-1.794	-3.032*	-1.895*
HNR35	-9.540*	-5.160*	-1.400	-2.283*	-2.359*
SHR	5.560*	1.626	0.106	2.758	3.193*

Table 6: Summary of critical results.

	Older male	Older female	Younger male	Younger female
F0	✓	✓	✓	✓
H1*-H2*				
H1*-A1*	✓			
H1*-A2*	✓			
H1*-A3*				
CPP	✓	✓	✓	✓
CQ	✓	✓		
HNR05	✓	✓		✓
HNR15	✓	✓		✓
HNR25	✓	✓		
HNR35	✓	✓		
SHR	✓			

Comparing speakers in different age and gender groups in Table 6, it was found that no two groups of speakers produced the register contrast identically the same. The general trend is that the older speakers used more cues than the younger speakers to make the contrast.

The older male speakers used ten cues in the register contrast. This is the most in all groups. The ten cues used are F0, H1*-A1*, H1*-A2*, CPP, CQ, four HNR measures, and SNR. The lower register had lower F0, higher values of spectral slopes (H1*-A1* and H1*-A2*), and lower CQ, CPP, HNR, and SHR. Of the twelve measures examined, only H1*-H2* and H1*-A3* failed to show significant main effects of register in all five manners. For both H1*-H2* and H1*-A3*, significant effects of register were found in stops, fricatives and zeroes, but not in affricates and nasals. The older female speakers used seven cues in the register contrast. The seven measures are F0, CPP, CQ, and four HNR measures. In addition to H1*-H2* and H1*-A3*, H1*-A1*, H1*-A2* and SHR also failed to show significant main effects of register in all five manners. To sum up, F0, CPP, CQ and four HNR measures were used by all the older speakers to contrast different registers.

The younger generation used much less cues than the older speakers. The younger male speakers used three cues, namely F0, CPP and HNR05 in all five manners to contrast registers. For the younger female speakers, in addition to these three cues, significant main effects of register in all five manners were also found in HNR 15. The younger female speakers thus used four cues in total. In summary, F0, CPP and HNR05 were used by all the younger speakers to make the contrast.

One phenomenon worth noticing is that H1*-H2*, a measure which has been found to successfully distinguish contrastive phonations in many of the world's languages [23, 24], is not a reliable measure for the phonation contrast here. Despite partial significant differences in several manners for some speaker groups, H1*-H2* did not show significant differences between registers in all five manners for any group of speakers. In some cases, expected direction was observed, but p-values did not reach the significant level. In other cases, reversed trends (higher H1*-H2* values in the upper register) were found.

To sum up, the two registers were realized differently by all speakers. The older speakers used more cues to make the contrast. F0 and periodicity (CPP and HNR05) show highly significant effects of register for all manners of articulation for all speakers. The lower register has lower pitch and less periodicity than the upper register. CQ showed significant effects of register for older speakers only. Spectral slopes (H1*-A1*, H1*-A2) and SHR show significant effects of register in older male speakers' production. H1*-H2* is generally less reliable.

4. Discussion and conclusion

This study is aimed to investigate the phonetic realizations of the register contrast in Shanghai Chinese. We especially focus on whether phonation contrasts are involved in the production of the tonal registers, at least for some speakers. A number of pitch and phonation parameters were measured. Significant age differences were found.

4.1. Age differences

In this set of data, older speakers used pitch, overall spectral slopes, periodicity, noise ratios and contact quotient in the register contrast, while younger speakers used pitch, periodicity and noise ratios only. The younger speakers no longer used overall spectral slopes and contact quotient. The difference between different age groups provides evidence for an ongoing sound change: phonation contrasts are merging among younger speakers.

4.2. The acoustic and physiological property of the register contrast in older speakers' production

In the older speakers' production, the lower register has lower pitch (F0) and breathier phonation.

Acoustically, the lower register has steeper spectral slopes, demonstrated by higher H1*-An* values. It was suggested that these measures correlate with the abruptness of vocal fold closure. When producing breathy phonation, vocal folds vibrate more slowly, and therefore has less abrupt vocal fold closure. As a result, the higher frequencies of the spectrum are less excited. Higher values in these measures indicate a breathy phonation type. The lower register also has higher noise ratios and less periodicity, demonstrated by smaller CPP and HNR values. Smaller values of CPP and HNR indicate less periodicity, a property of breathy phonation.

Physiologically, the lower register has smaller contact quotient (CQ). CQ is defined as the portion of time the vocal folds are closed during each vibratory cycle. Lower CQ values indicate greater opening or less constricted glottis, properties of breathy phonation.

Surprisingly, H1*-H2* is not as reliable as other measures in the register contrast in this set of data. H1*-H2* was found to be successful in distinguishing contrastive phonations in many of the world's languages. Cao and Maddieson and Ren's work in the late 1980s and early 1990s also reported that H1-H2 successfully distinguished the two registers for Shanghai Chinese speakers, at least in stops. We speculate that the non-modal phonation in Shanghai Chinese is probably a special variant of breathy voice. The reasons are as follows. First, the phonation measurements are very sensitive to the algorithms used in calculating the values, so that values given by different sources are less comparable. In this sense, our result is less comparable with Cao and Maddieson and Ren's work, and more comparable with studies that also use VoiceSauce. Second, it was reported in Jiashan Wu, another Wu dialect which is mutually intelligible with Shanghai Chinese, that H1*-H2* was not reliable in distinguishing different phonations between registers [25]. Shanghai Chinese may show some similarities with Jiashan Wu. Third, though relatively rare, H1*-H2* was also found to be unreliable in several languages. Esposito's cross-linguistic investigation of breathy vs. modal phonation contrast showed that H1*-H2* fails to distinguish breathy phonation from modal phonation in Mon and Tamang, two out of the ten languages or dialects she investigated [26]. Fourth, Cao and Maddieson and Ren mea-

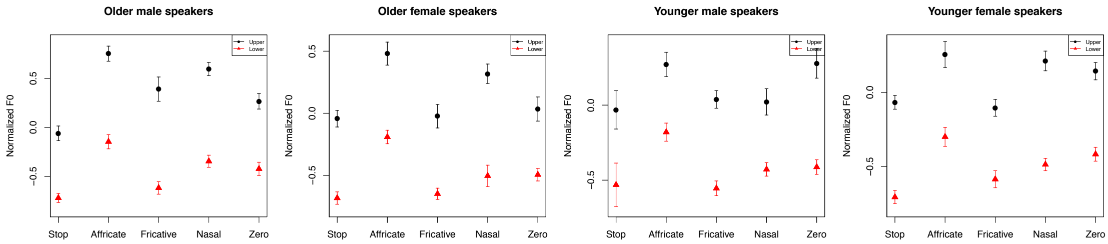

Figure 1: Mean F0 in the register contrast. Error bars indicate standard error.

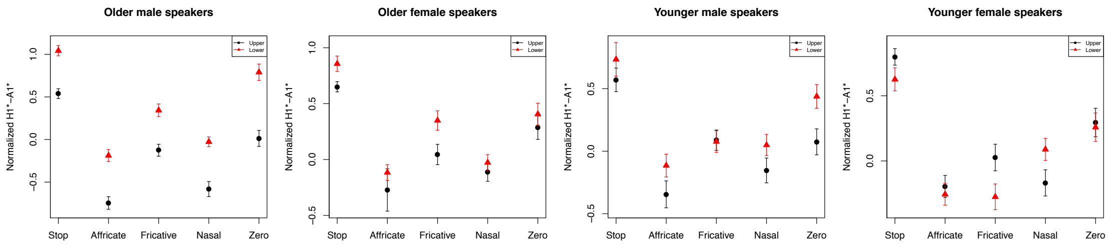

Figure 2: Mean H1*-A1* in the register contrast. Error bars indicate standard error.

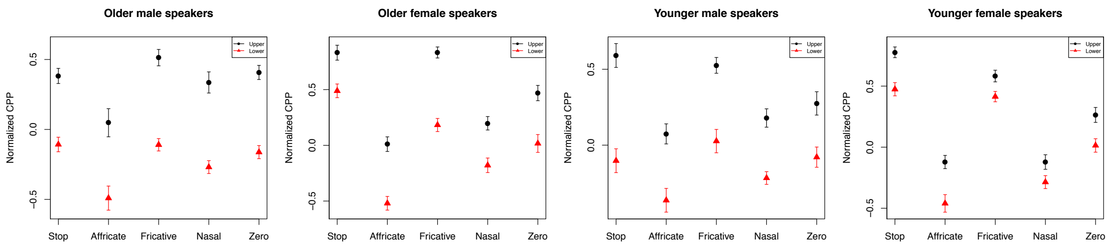

Figure 3: Mean CPP in the register contrast. Error bars indicate standard error.

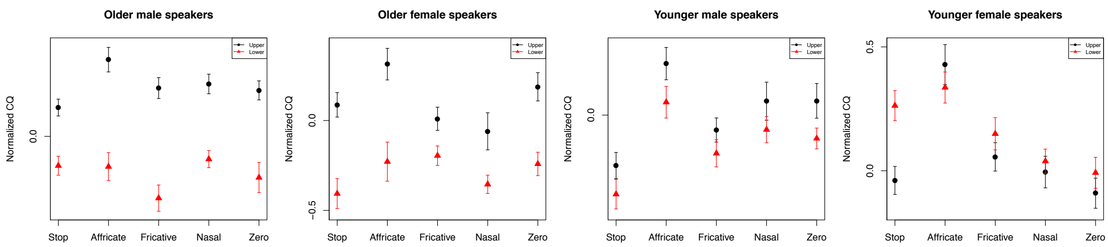

Figure 4: Mean CQ in the register contrast. Error bars indicate standard error.

sured stops only. In the current study, five manners were investigated. It was also found in this set of data that H1*-H2* is useful in the contrast for stops in the older male speakers' production. To summarize, H1*-H2* is generally less reliable, compared with other measures, in this set of data. The non-modal phonation in Shanghai Chinese is probably a special variant of breathy voice.

In summary, significant age differences were found. For the older speakers, the non-modal phonation associated with the lower register is a special variant of breathy voice. The lower register has steeper spectral slopes, less periodicity and higher noise ratios, and smaller contact quotient. H1*-H2* is not significant. For the younger generation, little evidence of breathiness was found. The age difference provides evidence for the suggestion that the Shanghai dialect of Chinese is gradually losing its lower register breathiness. This is largely consistent with Gao and Hallé and Zhang and Yan's studies [12, 13].

5. Acknowledgements

Reviews from three anonymous reviewers at the 5th International Symposium in Buffalo are gratefully acknowledged. We would like to thank Prof. Jiangping Kong and Prof. You-Jing Lin in Peking University for their assistance with equipment. We thank all the speakers for making the recordings possible. This study was supported by an Upenn faculty research fund to Jianjing Kuang.

6. References

[1] Y. R. Chao, *Studies in the Modern Wu-Dialects*. Peking: Tsing Hua College Research Institute, 1928.

[2] N. Qian, *Dangdai Wuyu Yanjiu (Studies in the Contemporary Wu dialects)*. Shanghai: Shanghai Educational Publishing House, 1992.

[3] M. Yip, "The Tonal Phonology of Chinese," Ph.D. dissertation, Massachusetts Institute of Technology, 1980.

[4] P. Rose, "Phonetics and phonology of yang tone phonation types in Zhenhai," *Cahiers de Linguistique Asie Orientale*, vol. 18, no. 2, pp. 229–245, 1989.

[5] ——, "Tonation in Three Chinese Wu Dialects," in *Proceedings of the 18th International Congress of Phonetic Science*, 2015.

[6] Z. Bao, *The Structure of Tone*. New York: Oxford University Press, 1999.

[7] P. Ladefoged, *Preliminaries to linguistic phonetics*. University of Chicago Press, 1971.

[8] J. Cao and I. Maddieson, "An Exploration of Phonation Types in Wu Dialects of Chinese," *Journal of Phonetics*, vol. 20, no. 1, pp. 77–92, 1992.

[9] N. Ren, "Phonation types and stop consonant distinctions: Shanghai Chinese," Ph.D. dissertation, The University of Connecticut, 1992.

[10] 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*, 2011, pp. 1846–1849.

[11] T. Henry. (2009) EGGWorks: a program for automated analysis of EGG signals.

[12] J. Zhang and H. Yan, "Contextually dependent cue weighting for a laryngeal contrast in Shanghai Wu," in *Proceedings of the 18th International Congress of Phonetic Science*, 2015.

[13] J. Gao and P. Hallé, "Are young male speakers losing Tone 3 breathiness in Shanghai Chinese? An acoustic and electroglottographic study," in *Proceedings of the 2nd International Conference on Phonetics of the Languages in China*, 2013, pp. 163–166.

[14] M. Sherard, "Shanghai Phonology," Ph.D. dissertation, Cornell University, 1972.

[15] E. Zee and I. Maddieson, "Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis," *UCLA Working Papers in Phonetics*, vol. 45, pp. 93–129, 1979.

[16] Y. Yao and C. B. Chang, "On the cognitive basis of contact-induced sound change: Vowel merger reversal in Shanghainese," *Language*, 2015.

[17] B. Xu and Z. Tang, *Shanghai shiqu fangyanzhi (A Description of the Urban Shanghai Dialect)*. Shanghai: Shanghai Educational Publishing House, 1988.

[18] Y. R. Chao, "A system of tone letters," *Le Maitre Phonetique*, pp. 24–27, 1930.

[19] M. Garellek and P. Keating, "The acoustic consequences of phonation and tone interactions in Jalapa Mazatec," *Journal of the International Phonetic Association*, vol. 41, no. 2, pp. 185–205, 2011.

[20] C. T. DiCanio, "The phonetics of register in Takhian Thong Chong," *Journal of the International Phonetic Association*, vol. 39, no. 2, pp. 162–188, 2009.

[21] J. Kuang, "Production and perception of the phonation contrast in Yi," 2011.

[22] 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.

[23] M. Gordon and P. Ladefoged, "Phonation types: a cross-linguistic overview," *Journal of Phonetics*, vol. 29, no. 4, pp. 383–406, 2001.

[24] P. Keating, J. Kuang, C. Esposito, M. Garellek, and S. Khan, "Multi-dimensional phonetic space for phonation contrasts," *Poster presented at LabPhon13 in Stuttart, Germany, July 2012*, 2012.

[25] J. Kuang and B. Jiang, "Consonant Effects on Tonal Registers in Jiashan Wu," in *The 90th Annual Meeting of the Linguistic Society of America*, 2016.

[26] C. M. Esposito, "The effects of linguistic experience on the perception of phonation," Ph.D. dissertation, UCLA, 2006.