Tonal split and laryngeal contrast of onset consonant in Lili Wu Chinese

Menghui Shi, Yiya Chen, and Maarten Mous

Citation: The Journal of the Acoustical Society of America **147**, 2901 (2020); doi: 10.1121/10.0001000

View online: https://doi.org/10.1121/10.0001000

View Table of Contents: https://asa.scitation.org/toc/jas/147/4

Published by the Acoustical Society of America

JASA ARTICLE

Tonal split and laryngeal contrast of onset consonant in Lili Wu Chinese

Menghui Shi,[1,a)] Yiya Chen,[2,b)] and Maarten Mous[1]
[1]Leiden University Centre for Linguistics, Leiden University, Van Wijkplaats 4, Witte Singel-complex, Leiden, 2300 RA, The Netherlands
[2]Leiden Institute for Brain and Cognition, Leiden University, Wassenaarseweg 52, Leiden, 2300 RC, The Netherlands

ABSTRACT:
This study examines the acoustic properties concerning tonal split and stop onsets in an under-documented Wu Chinese variety, Lili Wu, using speech production data collected from field research. Lili Wu Chinese has been reported to demonstrate an unusual tonal split phenomenon known as "aspiration-induced tonal split" (ATS). ATS refers to the distinct lowering of *f0* of a lexical tone over syllables beginning with a voiceless aspirated obstruent, compared to that of syllables beginning with an unaspirated obstruent. Two debates lingering in the existing literature are discussed: (i) is ATS an on-going change or a completed change? and (ii) is it onset aspiration or vowel breathiness that directly triggers ATS? Results suggest that ATS is a completed change, which, however, is conditioned by tonal contexts. Regarding the second debate, results suggest that neither aspiration nor breathiness serves as the direct trigger for tonal split. Moreover, one unexpected on-going sound change was observed: The breathiness of vowels after voiced onsets seems to be disappearing among the younger generation. These findings extend the understanding of the acoustic properties of tonal development in a complex system and highlight the importance of experimental methods in understanding the sound structure and changes of under-documented languages.

https://doi.org/10.1121/10.0001000

(Received 26 April 2019; revised 4 September 2019; accepted 26 November 2019; published online 30 April 2020)

[Editor: Richard A. Wright] Pages: 2901–2916

I. INTRODUCTION

Mergers and splits are generally believed to be the two main processes of phonemic change in the development of language (Labov, 1994; Campbell, 2013). In previous studies, mergers have drawn a great deal of attention to a variety of linguistic properties [e.g., Labov *et al*. (2006) for vowels in North American English; Harris (1969) for consonants in Latin American Spanish; and Yu (2007) for tones in Cantonese]. The converse process, namely, *splits*, defined as "the division of a preexisting phoneme to create a new phonemic distinction" (Labov, 1994, p. 331), is commonly argued as a conditioned, complicated, and unusual event. Nevertheless, tonal split is considered to be a highly remarkable change observed in many Asian languages with an already established tonal system (Haudricourt, 1972; Brown, 1975; Brunelle and Kirby, 2016). A great deal of work has been done in two aspects: One is of diachronic reconstructions on how the tonal inventories of proto-languages evolved into those of modern languages (e.g., Haudricourt, 1972; Li, 1977); the other is of phonetic explanations for mechanisms of tonal split using acoustic or perceptual experiments (e.g., Abramson and Erickson, 1978; Rischel, 1986; House and Svantesson, 1996; Thurgood, 2007). However, very few studies have access to a living language where the implementation of tonal split can be systematically and synchronically observed without indirect assumptions or inferences.

Fortunately, the language of focus in this paper, Lili Wu Chinese (Sino-Tibetan, Sinitic branch, Wu), provides us with just such an opportunity to fill in the neglected hiatus for our understanding of tonal split. The language has been reported to show fundamental frequency (*f0*) lowering after voiceless aspirated onsets, and this lowering effect has been argued to be phonologized and to have resulted in the splitting of an existing lexical tone and the forming of new tonal categories (Chao, 1928).

A. Wu Chinese and tonal split with aspiration onsets in Lili Wu

Wu Chinese[1] is commonly classified as one of the ten major dialect groups within the Sinitic branch of the Sino-Tibetan language family (Wurm *et al.*, 1987). The most prominent feature of Wu Chinese is the existence of a three-way laryngeal contrast in obstruents, known as voiceless unaspirated, voiceless aspirated, and voiced, respectively (Chao, 1967). The three-way laryngeal contrast has different manifestations in initial as opposed to medial position (see Chen, 2011 and references therein). In the initial position, these obstruents vary in their phonation from clearly modal (voiceless unaspirated), to aspirated with breathiness

---

[a)] Electronic mail: m.shi@hum.leidenuniv.nl, ORCID: 0000-0003-1663-8079.
[b)] Also at: Leiden University Centre for Linguistics, Leiden University, Van Wijkplaats 4, Witte Singel-complex, Leiden, 2300 RA, The Netherlands.

(voiceless aspirated), to breathy (voiced). In the medial position, however, voiced obstruents are fully voiced, leading to a three-way laryngeal distinction in voice onset time (VOT).

Another distinct feature of modern Wu dialects is that the majority of them are known to still preserve, to a large extent, an eight-way tonal system that developed from Middle Chinese (MC). MC is a sound system reconstructed based mainly on written records such as rhyme dictionaries (see a comprehensive introduction in Norman, 1988, Chapter 2). In the development of the tonal system from MC to the modern Wu dialects, the four-way tonal contour contrast of MC, traditionally labeled as *Ping* (level), *Shang* (rising), *Qu* (departing), and *Ru* (entering), split into a new dual-register, eight-tone system, conditioned by onset consonants, which is evident in modern Wu varieties[2] (Pulleyblank, 1978; Ting, 1984; Norman, 1988). Generally speaking, syllables with voiceless initials are argued to be produced within a higher *f0* range, referred to as the high register, while voiced initials condition a lower *f0* range, known as the low register.

Let us take Lili Wu Chinese as an example. Lili Wu Chinese is a Northern Wu dialect spoken in the town of Lili, one of the ten major towns in the Wujiang district, which belongs to the prefectural-level municipality Suzhou City in Jiangsu Province. There are eight lexical tones in Lili Wu Chinese. Figure 1 illustrates the *f0* contours of the eight tones (T1–T8) uttered in isolation by a male speaker who was born in 1947. As illustrated in Fig. 1, lexical tones marked as odd numbers start within the high register (above 160 Hz), while those marked as even numbers start within the low register (under 160 Hz). T1 is a high-register level tone (h-level) while T2 is a low-register rising tone (l-rising). T3 starts within the high register and falls (h-falling). T4 is a low-register level tone (l-level). T5 has a convex contour which starts within the high register, falls and ends with a slight rise (h-dipping). T6 is realized with a similar *f0* contour to that of T5 but starts within the low register (l-dipping). Both T7 and T8 are associated with syllables that have a much shorter duration than the other tone-bearing syllables. T7 starts within the high register and despite the slight falling contour, sounds like a high-register level tone (short-h-level). T8 is a low-register level tone (short-l-level).

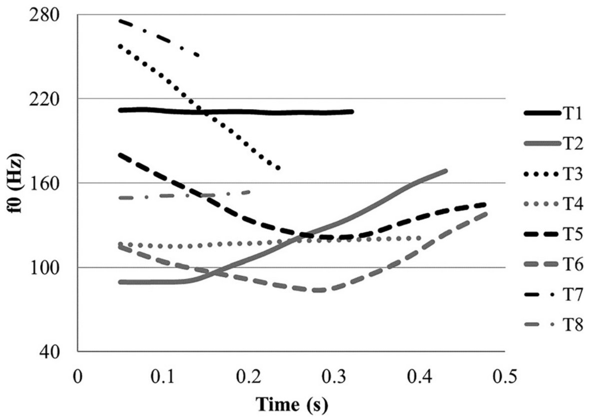

FIG. 1. *f0* contours (in Hz) of the eight lexical tones.

TABLE I. Co-occurrence constraints on consonantal onset and lexical tone in Lili Wu Chinese. Transcription of lexical tonal system based on the tonal transcription system developed by Chao (1930). This system divides a speaker's pitch range into 5 levels, with 5 indicating the highest end and 1 the lowest. A single number refers to cases where the tone-carrying syllables have short duration and only co-occur with the coda /ʔ/.

MC	*Ping*	*Shang*	*Qu*	*Ru*
Onset				
Voiceless unaspirated	h-level	h-falling	h-dipping	Short h-level
(High register)	44 (T1)	53 (T3)	423 (T5)	5 (T7)
Voiced	l-rising	l-level	l-dipping	Short l-level
(Low register)	13 (T2)	22 (T4)	213 (T6)	3 (T8)

As shown in Table I, the eight lexical tones of Lili Wu Chinese exhibit a remarkable co-occurrence pattern between consonantal onset and lexical tone. Syllables with voiceless onsets only license tones that start in the high register while those with voiced onsets allow tones that are in the low register. In Lili Wu Chinese, as shown in Table II, syllables beginning with voiceless aspirated obstruents in three of the MC tonal categories (i.e., *Shang*, *Qu*, and *Ru*), are reported to introduce distinctively lower tones than syllables beginning with unaspirated obstruents (see a comprehensive review in Wang, 2008). This unusual tonal phenomenon is termed "aspiration-induced tonal split" by Sinologists[3] (ATS hereafter). ATS is not considered to be a widespread phenomenon which occurred across Chinese languages/dialects. Existing literature rather indicates that ATS has been reported for varieties spoken in only 39 cities/counties of China (Xu, 2013), mainly including dialects of the Wu, Gan, and Xiang groups as well as some languages belonging to the Tai-Kadai and Hmong-Mien language families (Ho, 1989; Shi, 1998; Chen, 2005). Note that in Lili Wu Chinese, this phenomenon is absent in words within the MC tonal category *Ping*, where syllables with voiceless aspirated obstruents still bear a high tone.

B. Two debates on ATS

Most studies on ATS in Chinese languages/dialects are impressionistic descriptions. To our knowledge, Lili is thus far the only dialect which has been investigated in a number of studies. According to Wang (2008), the general consensus on the condition of the occurrence of the ATS phenomenon is that ATS is only present in words within non-*Ping* tonal categories (i.e., *Shang*, *Qu*, and *Ru*). However, researchers differ greatly in their analyses of the ATS phenomenon, which have resulted in various debates. Among them, two debates have long been a focus. The first debate regards the

TABLE II. ATS in Lili Wu Chinese.

MC	*Ping*	*Shang*	*Qu*	*Ru*
Onset				
Voiceless unaspirated	h-level	h-falling	h-dipping	Short h-level
Voiceless aspirated	Still high	Lower	Lower	Lower

progress of lexical tones beginning with aspirated onsets—Is tonal split an on-going change or a completed one? The second debate concerns the trigger of ATS—Is onset aspiration or vowel breathiness the direct trigger for ATS? The two debates focus on different aspects of ATS in Lili Wu Chinese. From a broader perspective, both, however, can be regarded as a facet of what Weinreich *et al.* (1968) call the "constraints" in their foundational work on language change. A possible approach for investigating sound-change constraints is to study asymmetries between sound change and phonetic patterns. As argued in Garrett and Johnson (2013), asymmetries in sound change usually reflect asymmetries in sound patterns. In Lili Wu Chinese, given the asymmetry of ATS conditioned by different MC tonal categories, it motivated us to pay more attention to the incongruent patterns of lexical tones between tonal categories (debate one) and its possible phonetic biases (debate two). Exploring both debates can not only further sharpen our understanding of the lexical tonal system of Lili Wu Chinese but also is pivotal for answering general issues on constraints of sound change, especially of the tonal-split phenomenon.

1. Debate one

Sound change in progress can be synchronically observed (Labov, 1994). However, the estimate of the stage of completion at which one particular linguistic change finds itself at any given time is always debatable. The first point of contention concerns the independence of lexical tones beginning with voiceless aspirated obstruents, namely, whether tonal split is an on-going change or a completed one. Two opposing views have been proposed.

On the basis of two speakers (one male and one female without exact ages), Shi (1992) argues that tonal split of syllables beginning with aspirated onsets is an on-going change (hereafter referred to as the "on-going change view"). Specifically, Shi (1992) claims that lexical tones beginning with aspirated onsets are independent of lexical tones beginning with unaspirated onsets and are merging toward lexical tones beginning with voiced onsets. Shi (2008, p. 227) further posits a stepwise lowering of lexical tones beginning with aspirated onsets across three MC tonal categories (i.e., *Shang*, *Qu*, and *Ru*) which Shi assumes should be observed from speakers of different generations within the speech community. This assumption is then confirmed by Zhu and Xu (2009), based on acoustic data collected from speakers of two generations (three old speakers whose ages are 60, 66, and 61 yr; and three young speakers whose ages are 35, around 30 and 28 yr) in Songling Wu, a variety spoken in the administrative-level town of Wujiang area with a similar phenomenon of tonal split to Lili Wu. They claim that two old speakers show a three-way contrast of lexical tones conditioned by initial onsets in the MC *Ru* tonal category, while the merging of lexical tones beginning with aspirated and voiced onsets is happening in the speech of one young speaker.

In contrast to the on-going change view, Shen (1994) maintains that there is no so-called on-going change, but rather a completed merger between lexical tones beginning with voiced and aspirated onsets in Lili Wu Chinese (hereafter referred to as the "completed change view"). Based on acoustic data obtained from two young speakers (high school students without exact ages) of Lili Wu Chinese collected in 1985, Shen (1994) claims that a completed merger is observable in three MC tonal categories, namely, *Shang*, *Qu*, and *Ru*.

2. Debate two

The phonetic trigger in the production of sound changes is also a widely discussed issue. As most phonetic studies of tonal systems show, the development of tones may result from different articulatory reinterpretations of segmentally induced perturbations in intrinsic *f0* (Hombert *et al.*, 1979). The second debate concerns various proposals regarding the trigger of ATS in Lili Wu Chinese form the second debate, which fall into two general views.

One view is that tonal split in Lili Wu Chinese is directly due to onset aspiration (Ye, 1983; Wang, 2008) (hereafter the "aspiration view"), since synchronically speaking, onset aspiration seems to be the most prominent feature to have actuated the change. This view has been widely adopted by Sinologists after Chao (1928), who was the first to report this phenomenon of Lili Wu Chinese but without mentioning its phonetic substances and mechanisms. It is also in line with the view that voice quality of initial consonants plays an important role in the process of tonal split via the phonologization of phonetic perturbation effects caused by initial onsets (see Thurgood, 2007; Chen *et al.*, 2017 for reviews of such work).

The alternative view argues for a phonation-based account, which emphasizes an important role of breathiness during the process of tonal split. (Sagart, 1981; Ho, 1989; Zhu and Xu, 2009; Hirayama, 2010; Chen, 2014) (hereafter the "breathiness view"). This view seems to have been initiated by Sagart (1981), who assumes a correlation between breathy phonation and low tonal onset. Subsequent studies attempt to provide more elaborate interpretations. For example, Zhu and Xu (2009) report that the magnitude of breathiness at the 30–40 ms interval of vowels after aspirated onsets is higher than that after unaspirated onsets. Hirayama (2010) further argues that a higher magnitude of breathiness can be observed throughout the entire vowel (or part of the vowel). Recently, Chen (2014) attempts to explain the correlation between breathier phonation and lower tone as being due to the intrinsic aerodynamic property of an aspirated stop release suggested by Ohala (1978). This view is in line with the observation that non-modal phonation types and *f0* contours correlate with one another (e.g., Laver, 1994; Gordon and Ladefoged, 2001; see a comprehensive review in Esling and Harris, 2005). Increasing evidence also shows that breathier phonations[4] are commonly associated with *f0* lowering cross-linguistically (e.g., Gordon and Ladefoged, 2001; see also a detailed review in Kuang, 2013b).

C. The current study

The overarching goal for this current study is to shed light on the two aforementioned debates concerning tonal split in Lili Wu Chinese. Generally speaking, previous studies on both debates suffer from manifold inadequacies. For the first debate, the on-going change view fails to rule out the speaker-specific possibility of a three-way contrast of lexical tones as a function of the tone-bearing syllable onset, due to the small sample size recruited in each generation (e.g., Zhu and Xu, 2009). The completed change view however draws its conclusion on data from speakers with a limited age range (Shen, 1994).

With respect to the second debate, adherents to the aspiration view base their analyses of tonal categories on impressionistic observations only (after Chao, 1928). Furthermore, we know that the *f0* lowering effect of aspirated onsets varies across languages and speakers of the same language (Thavisak, 2004; Chen, 2011). The breathiness view also lacks empirical evidence. The study by Zhu and Xu (2009) includes the results of breathiness, but lacks control of the lexical properties of the stimuli as well as details on how the measurements were taken. Moreover, although breathier voices are commonly associated with lower *f0*, such an interaction is not inevitable. For example, Hirano *et al.* (1970) demonstrate that vowels within very high pitches tend to be breathy due to the relaxation of muscles. Kuang (2013a) reports that vowels with the tone /33/ are significantly breathier than vowels with two lower tones /11/ and /22/ in Black Miao, a Hmong-Mien language.

In brief, none of the existing studies provide comprehensive and empirically sound data for assessing the two debates. Consequently, the tonal system of Lili Wu Chinese as well as the acoustic properties of lexical tones and tone-bearing syllables require further investigation. To this end, multiple acoustic measures are needed to gain insights into the above issues. The current study was therefore designed to elicit data from a large sample of speakers of different generations. Studying linguistic variables across age groups in apparent time is commonly considered to be the first and most straightforward approach to studying a linguistic change across decades in real time (Labov, 1994, pp. 45–46; also see Labov *et al.*, 2013 for a comprehensive review). Based on the literature, the following evidence is expected to be observed to support the competing views of each debate.

As for the first debate, namely, the on-going or completed change of tonal split with aspirated onsets, we predict that if ATS has indeed been an on-going change within the speech community, different stages of this change should be reflected by generational data. Figure 2 shows the scenario assumed by studies holding the on-going change view. In Stage I, identical to most Wu varieties, lexical tones beginning with unaspirated and aspirated onsets in Lili Wu Chinese have the same *f0* contours. In Stage II, lexical tones beginning with aspirated onsets bifurcate from those beginning with unaspirated counterparts and become independent as new tonal categories, distinct from both *f0* contours of the other two types. In Stage III, the *f0* lowering trend continues and finally leads to the merging of contours beginning with aspirated onsets with those beginning with voiced onsets. If the on-going change view is true, the three stages of tonal categorization and the stepwise lowering of *f0* contours beginning with aspirated onsets are expected across different age groups (from old to middle-aged to young). On the contrary, according to the completed change view, a merger of lexical tones beginning with aspirated and voiced onsets is expected for all three generations.

For the second debate, the trigger of tonal split, the aspiration view predicts that similar patterns of onset aspiration will lead to similar patterns of tonal split. That is to say, if ATS does not occur in the MC *Ping* tonal category, onset aspiration in the *Ping* category is expected to show a significant difference from that in the other three MC tonal categories (i.e., *Shang*, *Qu*, and *Ru*) where ATS is observed. As for the breathiness view, studies have shown that vocalic breathiness tends to be sustained for longer and is not localized at the onset or part of the adjacent vowel (Gordon and Ladefoged, 2001; Blankenship, 2002; Esposito and Khan, 2012). Moreover, Hirayama (2010) argues a higher magnitude of breathiness throughout the entire vowel (or most part of the vowel). Therefore, as long as ATS happens, a

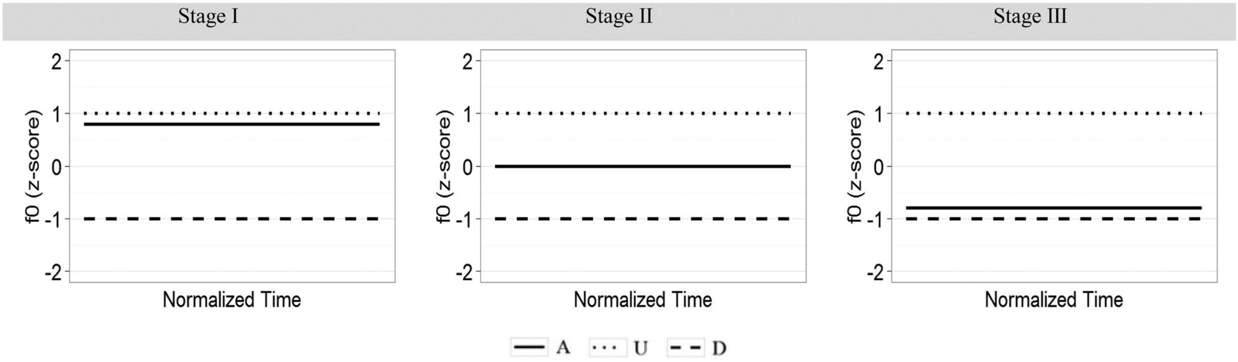

FIG. 2. Three expected stages of ATS based on the “on-going change view.” “A,” “U,” “D” represent *f0* contours with voiceless aspirated onsets, voiceless unaspirated onsets, and voiced onsets, respectively.

consistent higher magnitude of breathiness is expected over the whole vowel or at least the majority of the vowel interval.

II. METHOD

A. Stimuli

The stimulus list consists of a minimal set of 36 real monosyllabic words (3 consonant onsets × 3 vowels × 4 MC tonal categories) with three laryngeal-alveolar contrasts /t t^h d/ (voiceless unaspirated, voiceless aspirated, and voiced) combined with three vowels /ɑ ɛ i̟/ (low, middle, and high) (but /a ʌ ɪ/ for the *Ru*-category words). Each syllable is further associated with four tonal categories, covering all potential CV (CVʔ in *Ru*) combinations between consonantal onsets and lexical tones. All words were selected from the Questionnaire of Character for Dialect Surveys (*Fangyan Diaocha Zibiao*, 方言调查字表) compiled by the Institute of Linguistics of the Chinese Academy of Social Sciences (1988). The characters in the questionnaire are listed according to their MC pronunciations, which guaranteed the reflex of the tonal system from modern Lili Wu Chinese to MC (see a detailed introduction in Kurpaska, 2010, Chap. 7). The full stimulus list is provided in the Appendix (Table VI). The place of articulation of the plosive onset is restricted to alveolar so as to minimize variation due to the intrinsic effect of different places of articulation on VOT (Cho and Ladefoged, 1999; Lai *et al.*, 2009). All stimuli were confirmed to be frequent and familiar words in Lili Wu Chinese by an educated native speaker who has spent most of his life living in the Lili town and also took part in the experiment.

B. Participants

A total of 68 native speakers were recruited in the experiment. However, only 60 participants were selected as qualified participants as the speech production of eight participants turned out to be problematic either due to too much disfluency or equipment failure. So, we only included 60 participants' data for further acoustic analysis. Among the selected participants (Mean: 47 yr; SD: 17 yr), there were 20 participants for each age group (old: 12 males and 8 females born between 1933 and 1956, Mean: 67 yr, SD: 6 yr; middle-aged: 11 males and 9 females born between 1961 and 1976, Mean: 48 yr, SD: 6 yr; young: 8 males and 12 females born between 1978 and 1994, Mean: 27 yr; SD: 6 yr). In addition to Lili Wu Chinese, all participants were able to speak Standard Chinese but with different levels of proficiency. Younger participants generally achieved higher level of proficiency than older participants. However, according to their self-reports, all thought Lili Wu Chinese as the first and dominant language.

C. Procedure

The recordings were conducted for all participants in a quiet room in Lili town. The utterances were recorded on an external sound card (Cakewalk UA-1G) with a Sennheiser PC 151 Headset condenser microphone. The signal was digitized at a 22 050 Hz sampling rate. Stimuli were presented twice as differently randomized lists via the Field Phon program (Pan *et al.*, 2015). The participants first heard a pre-recorded question in Lili dialect which was read by a native male speaker and then answered the question verbally with the target words on the screen. The pre-recorded question was "What is it called in Lili Wu dialect?"[5] In this way, we controlled the discourse context (Lea, 1973) and made sure that all target words were uttered as information elicited by a *wh*-question in the same controlled discourse context (Chen, 2011). In total, 4320 tokens were collected (36 target words × 2 repetitions × 60 speakers). The participants were asked to pronounce each word at their normal speaking rate. To make sure that the task was correctly understood, all participants undertook five practice trials (with no target words included) to become familiar with the procedure before the real recording, but none knew the purpose of the experiment. All were paid the equivalent of 10 euros in local currency for their participation.

D. Acoustic measurements

Segments were identified manually with Praat (Boersma and Weenink, 2016) based on the periodicity in the acoustic waveform, supplemented by spectrographic analyses (Lehiste and Peterson, 1961; Turk *et al.*, 2006). To explore the two debates on tonal split, three sets of acoustic measurements were extracted.

For the exploration of whether ATS is an on-going change, *f0* in Hz was measured at 20 equidistant points over syllables with a long duration (i.e., *Ping*, *Shang*, and *Qu*), but ten points over syllables with a short duration (i.e., *Ru*) starting from the first regular vocal pulsing to the end of the syllable using a custom-written script (Chen, 2011). Furthermore, in order to eliminate the pitch range difference due to individual variation and to plot *f0* contours for visual inspection, the raw *f0* values at all points were normalized using the within-speaker z-score (Rose, 1987). The plotted tonal contours were then averaged across speakers in each group on the basis of the mean z-score.

With respect to the debate on the trigger of tonal split, two points are worth noting. First, since the aspiration view argues that onset aspiration is the obvious trigger of tonal split, we measured the absolute duration of onset release (DOR). DOR was defined as the interval between the beginning of the release burst of the aspirated onset and the onset of voicing for the following vowel (identified as the onset of the low-frequency voicing energy on the spectrogram) (Francis *et al.*, 2003). It has been known that speaking rate can affect the production of VOT of stop consonant, especially for long lag VOT (Kessinger and Blumstein, 1997). For better control of the speaking rate, we further divided the raw DOR values by the duration of the tone-carrying syllable (DOS). Both the raw DOR as well as the DOR/DOS ratio served as the dependent variable for statistical analysis, respectively. Second, since the breathiness view maintains that the breathiness of vowels is the direct trigger for tonal split, we took the corrected H1*-H2* values (see Hanson, 1997; Iseli *et al.*, 2007 for formant corrections denoted with an asterisk). While for the lexical tonal realization, the exact trajectory of the *f0* contours can be crucial, for the breathy contrast here, our main interest was whether the difference was maintained throughout the whole vowel. To this end, three points were taken (i.e., the one-third, middle, and two-thirds of the vowel). This was automatically obtained by using VoiceSauce (Shue *et al.*, 2011) with a 25 ms window size. The acoustic cue H1*-H2*, which is the amplitude difference between the first and second harmonics, has been widely adopted as an indication of the phonatory state across languages with higher H1*-H2* values signaling breathier phonation (e.g., Gordon and Ladefoged, 2001; Blankenship, 2002; Keating *et al.*, 2010; Kuang, 2013a).

E. Statistical analysis

In order to obtain a better understanding of the time course data of normalized *f0* contours, growth curve analysis (GCA) (Mirman, 2014) was employed with the *lme4* package (Bates *et al.*, 2016) in *R* (*R* Core Team, 2016). GCA is a multilevel regression technique which has been argued to be appropriate and powerful for analyzing non-linear time-varying data such as *f0*. The advantage of this technique is that overall successive data are taken into consideration. (See the usage of this method for tonal contours in other Chinese dialects such as Li and Chen, 2016 for tonal realization in Tianjin Mandarin; and Zhang and Meng, 2016 for Shanghainese tones; see also the comparison between GCA and other statistical methods on tone analysis in Chen *et al.*, 2017.) The basic idea of GCA is to build higher-order polynomials including multiple polynomial terms for capturing the time-varying changes in real data. For example, a second-order polynomial model ($y = \alpha + \beta \times \text{Time} + \gamma \times \text{Time}^2 + \Delta$)[6] would have three time terms: intercept (α), linear (β), and quadratic (γ). These terms index the overall mean of the curve, the direction of curve change, such as rising vs falling, and the steepness of curve rising or falling, respectively (Li and Chen, 2016). If two curves differ from each other, we should expect a statistical difference in at least one of the three terms.

In order to choose the best polynomial order and avoid overfitting the data, we first determined the best shape for capturing the changes of overall *f0* contours. Both practical and statistical reasons were taken into consideration (Mirman, 2014, pp. 46–47). According to the tonal system of Lili Wu Chinese, the most complex *f0* contour has only a convex contour shape. Therefore, we then compared the model having a simple linear shape with the one having a curved shape. Following the method suggested in Winter and Wieling (2016), two base models with different time terms (i.e., linear shape: ot1 vs curved shape: ot1 + ot2) and individual participants (i.e., Speaker) varying in the random intercept were built for model comparisons within each MC tonal category.

After choosing the polynomial order for analyzing *f0* contours, separate mixed-effects models were used to investigate the effects of Consonant (aspirated vs unaspirated vs voiced), Generation (old vs middle-aged vs young), and their interaction within each MC tonal category. The base model included the time terms in fixed factor structure and the Speaker random effect on all time terms. If a significant effect of the interaction was observed, separate models were built in order to further explore the simple effect of Consonant within each generation. In such cases, the dataset was one of the 12 subsets (4 categories × 3 generations) according to each of the MC tonal categories and three generations. The base model of each dataset was first established containing only the time terms in fixed structure and the random structure of Speaker on all time terms. The fixed predictor Consonant was then added. Vowel (high vs middle vs low), Repetition (first vs second) and Gender (male vs female) as control variables were further entered in a stepwise fashion, since all of them were known to have an effect on *f0* realization (e.g., Jacewicz and Fox, 2015 for vowel intrinsic *f0*; Lam and Watson, 2010 for repetition effect, and Simpson, 2012 for gender effect). In addition, Speaker by Consonant and Item were also tested as random effects on all time terms via model comparisons.

With regard to the analysis of DOR-related (i.e., raw DOR and ROR/DOS ratio) and H1*-H2* data, we built linear mixed-effects models with the *lme4* package in *R*. For the analysis of the DOR-related data of aspirated onsets, as fixed effects, we entered Category, Generation, and their interaction into the models in a stepwise fashion. We also took the intercept for Speaker as a random structure. If the interaction between Category and Generation showed a significant effect, separate linear mixed-effects models were further built based on each of the generations. Each base model was first built only with the random intercept of Speaker. Category and additional control factors (i.e., Vowel, Repetition, and Gender) were then introduced stepwise for model comparisons. The intercept for Item as well as the slope for Speaker by Consonant were also tested as random effects via model comparisons.

The analysis of H1*-H2* was similar to the DOR-related data. We first entered Consonant, Generation, Position (one-third vs middle vs two-thirds), Category, and their interactions as fixed effects into the models in a stepwise fashion. All models kept the random intercept of Speaker consistent. If there was interaction of the four factors, data were further divided to explore the H1*-H2* of vowels as a function of different onsets. To this end, we carried out separate linear mixed-effects models according to each of MC tonal categories, generations, and time positions. Each base model contained the random intercept of Speaker only. Consonant and additional control factors were then introduced stepwise via model comparisons. The operation of other random effects was developed in the same manner applied to the analysis of the DOR-related data.

In comparisons of all model data, an improvement of model fit was obtained through likelihood ratio tests indicted

TABLE III. Results (χ^2) of model comparisons for the effect of Consonant, Generation, and Consonant * Generation on *f0* contours. Parameter-specific *p*-values (superscript) are indicated by Pr(>Chisq). n.s.: not significant.

Predictor	Time term	I	II	III	IV
Consonant	Intercept	$6578.4^{<0.001}$	$14814.28^{<0.001}$	$5017.56^{<0.001}$	$11663.88^{<0.001}$
	Linear	$2163.19^{<0.001}$	$2252.9^{<0.001}$	$2680.76^{<0.001}$	$161.99^{<0.001}$
	Quadratic	$190.63^{<0.001}$	$50.44^{<0.001}$	$11.86^{<0.01}$	$2.37^{\text{n.s.}}$
Generation	Intercept	$3.46^{\text{n.s.}}$	$1.01^{\text{n.s.}}$	$16.03^{<0.001}$	$1.14^{\text{n.s.}}$
	Linear	$0.55^{\text{n.s.}}$	$3.81^{\text{n.s.}}$	$5.81^{\text{n.s.}}$	$7.61^{<0.05}$
	Quadratic	$17.81^{<0.001}$	$1.45^{\text{n.s.}}$	$5.34^{\text{n.s.}}$	$1.57^{\text{n.s.}}$
Consonant * Generation	Intercept	$6871.14^{<0.001}$	$15115.7^{<0.001}$	$5428.58^{<0.001}$	$11911.77^{<0.001}$
	Linear	$2200.92^{<0.001}$	$2374.6^{<0.001}$	$2761.41^{<0.001}$	$187.16^{<0.001}$
	Quadratic	$241.64^{<0.001}$	$63.2^{<0.001}$	$26.02^{<0.001}$	$10.97^{\text{n.s.}}$

by Pr(>Chisq) (χ^2) for each model with the effect in question against the model without the effect in question. Under any circumstances where the model failed to converge, the newly added structure was then dropped. All data were plotted in *R* using the *ggplot2* package (Wickham, 2009).

III. RESULTS

In the following, the four MC tonal categories (i.e., *Ping*, *Shang*, *Qu*, and *Ru*) were labeled as I to IV, respectively. "A," "U," and "D" represented measured values with voiceless aspirated, voiceless unaspirated, and voiced onsets, respectively.

A. F0 contour

Results showed that the second-order polynomial model significantly improved the model fit for all four MC tonal categories [I: $\chi^2 = 187.35$, $p < 0.001$; II: $\chi^2 = 75.68$, $p < 0.001$; III: $\chi^2 = 1638$, $p < 0.001$; IV: $\chi^2 = 20.9$, $p < 0.001$]. We therefore applied the second-order polynomial to all data for further analyses.

As shown in Table III, except for Consonant in quadratic in IV, both Consonant and its interaction with Generation had significant effects on all three time terms across all MC tonal categories. However, the Generation factor failed to show a significant main effect except for quadratic in I, intercept in III, and linear in IV. Given the across-the-board significance of the interaction between Consonant and Generation across categories, we further decomposed the data according to each of the three generations.

Figure 3 displays the *f0* contours of each MC tonal category (I–IV) within the factor Generation. Detailed results of *f0* contours tested by GCA and final models for calculating

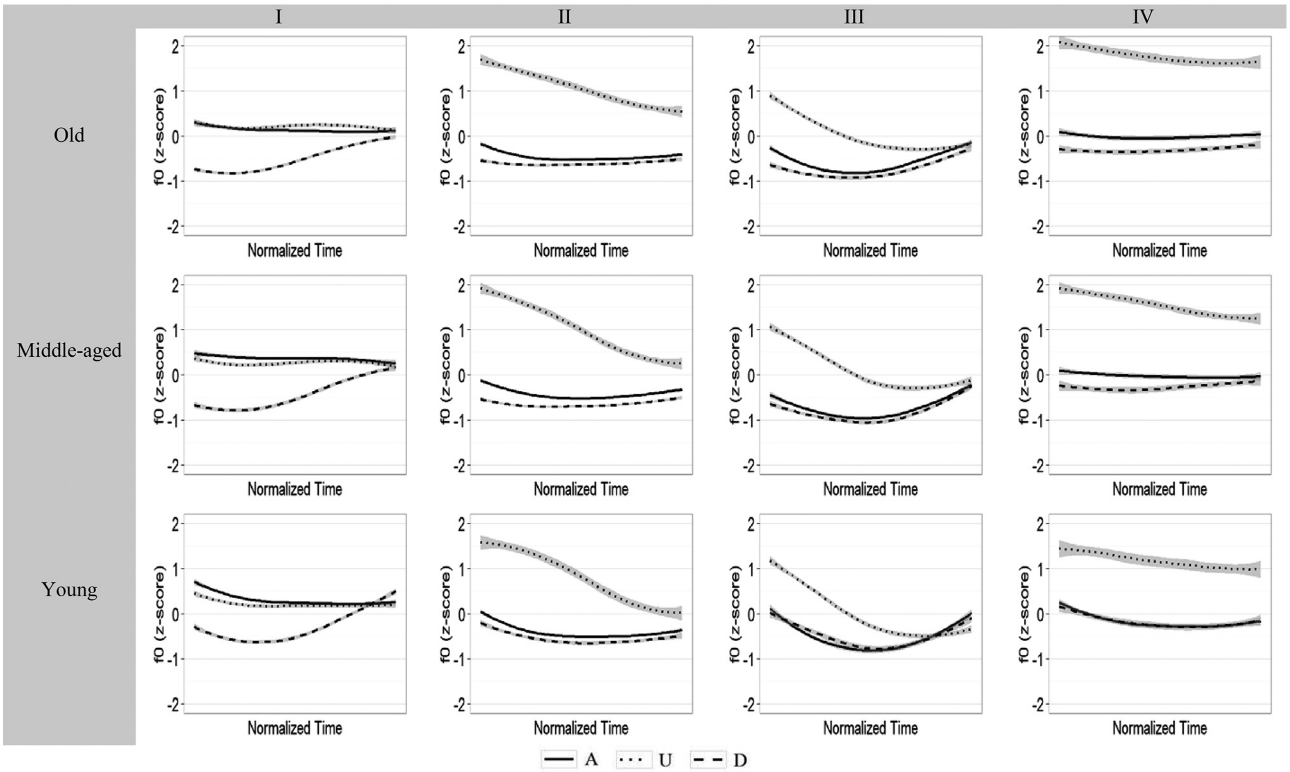

FIG. 3. The *f0* realization within Generation. Mean normalized *f0* (Gray areas indicate ± SE) averaged across speakers and repetitions. X-axis is normalized time and Y-axis is normalized *f0*.

them are attached in the Appendix (Tables VII and VIII) for interested readers.

As shown in Fig. 3-I, across generations, both aspirated (I-A) and unaspirated (I-U) onsets introduce comparable *f0* contours within the high register, but voiced onsets (I-D) introduce low-rising contours. Within the factor Generation, compared to I-A, no time terms of I-U showed any significant effect. Irrespective of generation, both intercept and linear terms of I-D consistently showed significant differences from those of I-A, as supported by the significant results of I-D. Moreover, compared to I-A, the *f0* contour with I-D presents a slightly concave trajectory produced by the middle-aged and young speakers, as indicated by the significant quadratic term of I-D in both groups.

As shown in Fig. 3-II, aspirated (II-A) and voiced (II-D) onsets lead to more comparable *f0* contours, which are realized within the low register, while unaspirated onsets (II-U) introduce falling contours within the high register. Results indicated that there were significant differences between *f0* contours with II-A and II-U in all three time terms. However, no significant difference was observed in any time term between II-A and II-D. Such a pattern held across generations.

Figure 3-III suggests that contours beginning with aspirated (III-A), voiced (III-D), and unaspirated (III-U) onsets are consistently realized as concave trajectories. We did not observe any significant difference between *f0* contours beginning with III-A and III-D. However, both are produced with a lower *f0* than III-U is. This difference was reflected by the significant results in both intercept and linear terms. Again, this pattern held across generations.

Finally, as shown in Fig. 3-IV, similar low-level contours beginning with aspirated (IV-A) and voiced (IV-D) onsets are again observed within the factor Generation. This was reflected by the lack of significant results in all time terms. Both, however, are significantly different from the *f0* contours beginning with unaspirated onsets (IV-U), which basically show high-level trajectories. It is worth noticing that the contour for IV-U produced by young speakers, as compared to that for I-A, shows a less concave trajectory as indicated by the significant quadratic.

These findings confirmed descriptions in the existing literature that tonal split did not happen in the MC *Ping* (I) tonal category.[7] This implies that ATS was not an across-the-board phenomenon in Lili Wu Chinese, but rather that its appearance was conditioned by certain tonal contexts (i.e., MC tonal categories). More importantly, in those tonal contexts where ATS occurred (i.e., MC *Shang*, *Qu*, and *Ru*), the *f0* contours beginning with voiceless aspirated and voiced onsets completely merged. Both were significantly lower than the *f0* contours beginning with unaspirated onsets. Such a pattern of ATS was stable across all three generations.

B. Raw DOR and DOR/DOS ratio

For the raw DOR, there was a significant main effect of Category ($\chi^2 = 145.98$, $p < 0.001$). However, both the main effect of Generation ($\chi^2 = 6.19$, $p > 0.05$) and its interaction with Category ($\chi^2 = 2.74$, $p > 0.05$) failed to show a significant effect. The insignificant interaction impeded us from dividing the data. For the DOR/DOS ratio, results showed both a significant main effect of Category ($\chi^2 = 477.76$, $p < 0.001$) and significant interaction of Category and Generation ($\chi^2 = 50.14$, $p < 0.001$), but there was no significant main effect of Generation ($\chi^2 = 3.88$, $p > 0.05$). A subset of data was then generated for each generation. Separate models were run for each subset in order to examine the difference in the DOR/DOS ratio between MC *Ping* and the other three MC tonal categories. Figure 4 depicts the DOR/DOS ratio of MC tonal categories (I–IV) of each generation. Although there was no statistical significance for the factor

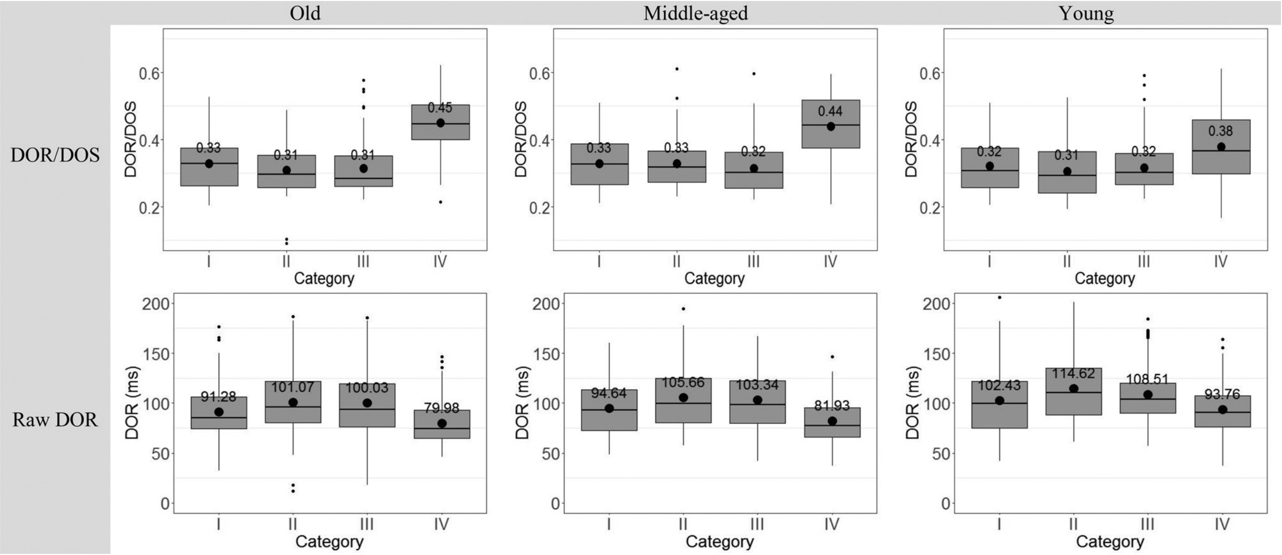

FIG. 4. Boxplots of the DOR/DOS ratio and the raw DOR in target syllables of each MC tonal category for three generations. The solid point in the box represents the mean and the line within the box the median.

Generation, we did observe a trend of difference as plotted in Fig. 4. Detailed results of the DOR/DOS ratio tested by linear mixed-effects models and final models for calculating them are attached in the Appendix (Tables IX and X) for interested readers.

As shown in Fig. 4, the mean value of the DOR/DOS ratio of IV remains highest across all generations (old: 0.45; middle-aged: 0.44; young: 0.38). Correspondingly, the mean value of the raw DOR of IV is the shortest (old: 79.98 ms; middle-aged: 81.93 ms; young: 93.76 ms). This visual inspection was further supported by a consistently significant difference between I and IV across generations (old: $\beta = 0.12$, $p < 0.001$; middle-aged: $\beta = 0.11$, $p < 0.001$; young: $\beta = 0.06$, $p < 0.001$). This result was thought to be attributed to the short tone-carrying syllable of IV, which reduced the duration of vowels and increased the ratio accordingly. The expected difference between I and the other three categories, however, was not observed. Within each generation, I failed to show any significant difference from the other two counterparts, namely, II and III. This pattern held across generations.

C. H1*-H2*

As summarized in Table IV, our results indicated that, except for Generation, all other factors (i.e., Consonant, Category, and Position) showed a significant main effect on H1*-H2*. Moreover, four factors significantly interacted in all orders (i.e., two-way, three-way, and four-way). There were multiple scenarios to further quantify the interactions. Given the purpose of comparing H1*-H2* values of vowels beginning with different onsets, we divided the dataset into 12 subsets according to each of the generations, MC tonal categories and time positions where H1*-H2* values were measured. To help visualize the interactions, Fig. 5 plots the mean H1*-H2* measured over the three positions for all 12 subsets. A series of linear mixed-effects models were run over each subset. Detailed results and final models for calculating them are attached in the Appendix (Tables XI and XII) for interested readers.

Figure 5 shows that in I, II, and III of both old and middle-aged groups, there is little H1*-H2* difference between vowels after A and D. Both however, shows higher H1*-H2* values than vowels after U. This observation was supported by the results of Table XI in the Appendix. On the one hand, H1*-H2* of vowels after A (old and middle-aged: I-A, II-A, and III-A) was consistently different from that after U (old and middle-aged: I-U, II-U, and III-U). On the other hand, it did not differ significantly from that after D for the old and middle-aged speakers (I-D, II-D, and III-D). However, the young-generation speakers showed a very different pattern. As shown in Fig. 5-I/II/III, it is quite clear that the H1*-H2* of vowels after A is much higher than that after U and D. Very different from the pattern of old and middle-aged speakers, H1*-H2* of vowels after D of young speakers tends to be lower. It leads to an approximation of H1*-H2* of vowels after U and D. Significant differences existed between A (young: I-A, II-A, and III-A) and its two counterparts (young: I-U, II-U, III-U, I-D, II-D, and III-D).

The situation of IV was different from all other tonal categories (i.e., I, II, and III). As shown in Fig. 5-IV, across generations, H1*-H2* of vowels after A is always higher than that for its two counterparts (i.e., U and D). A significant effect was found between IV-A and IV-U as well as between IV-A and IV-D across all generations. Moreover, as observed from Fig. 5, the difference of H1*-H2* between vowels after U and D is also obvious within each generation. H1*-H2* of vowels after U is lower than that after D in speakers of old and middle-aged generations, but higher than that after D in the young-generation speakers.

When we focus on the middle (P2) and two-thirds points (P3), as shown in Fig. 5, all differences presented at P1 tend to be diminished across generations. This pattern was also confirmed by the results of Table XI in the Appendix. In the majority of cases, Consonant did not significantly improve the model fit (indicated by "—"), which suggested that there was no significant difference of H1*-H2* of vowels after the three onsets. In six cases of P2 (old: I, III, and IV; middle-aged: I, II, and III), Consonant did help to improve the model fit. However, five of them did not

TABLE IV. Results (χ^2) of model comparisons for the effect of Consonant, Generation, Category, Position, and their interactions on H1*-H2*. Parameter-specific p-values (superscript) are indicated by Pr(>Chisq). n.s.: not significant.

One-way	χ^2	Two-way	χ^2	Three-way and four-way	χ^2
Consonant	$250.88^{<0.001}$	Consonant *Generation	$56.65^{<0.001}$	Consonant *Generation*Category	$88.23^{<0.001}$
Generation	$0.05^{\text{n.s.}}$	Generation *Category	$17.4^{<0.01}$	Generation *Category*Position	$184.16^{<0.001}$
Category	$47.43^{<0.001}$	Category *Position	$153.91^{<0.001}$	Category *Position* Consonant	$255.04^{<0.001}$
Position	$191.95^{<0.001}$	Position *Consonant	$104.72^{<0.001}$	Position *Consonant*Generation	$177.1^{<0.001}$
		Consonant *Category	$59.01^{<0.001}$	Consonant *Generation*Category*Position	$314.73^{<0.001}$
		Generation *Position	$24.45^{<0.001}$		

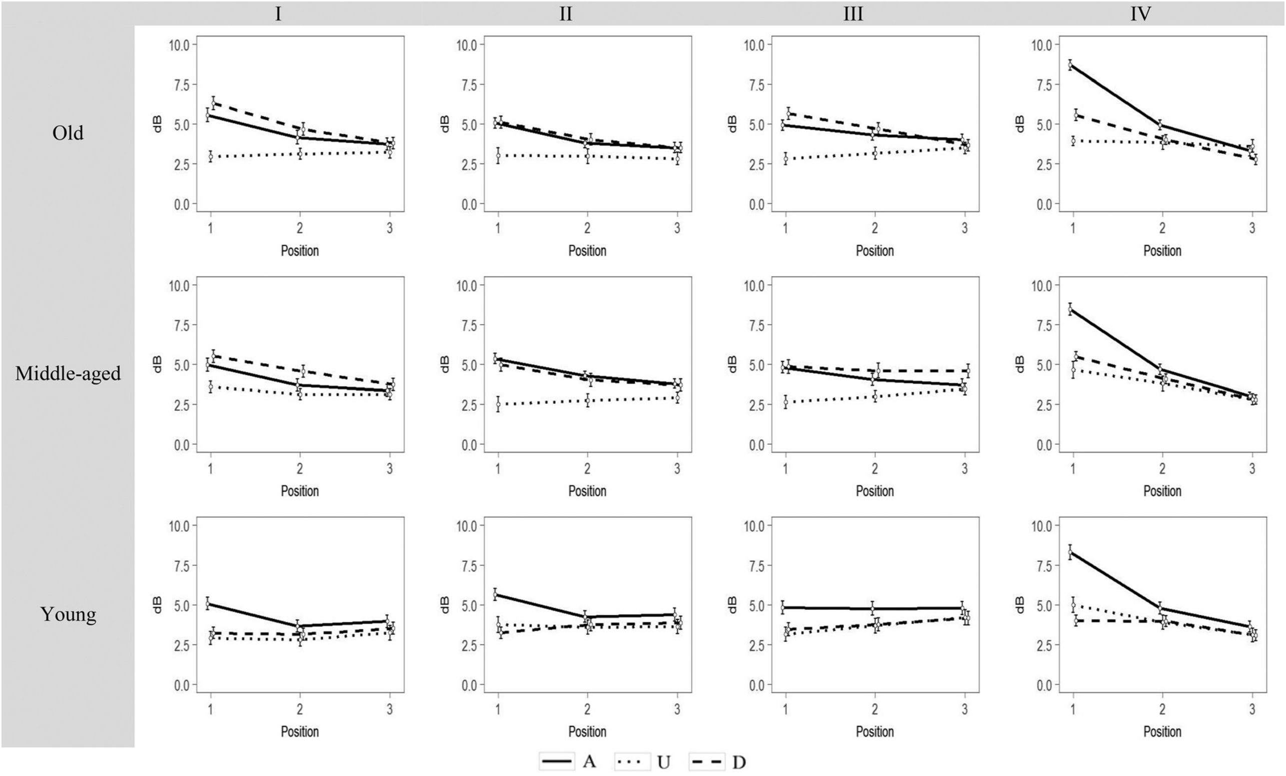

FIG. 5. Mean H1*-H2* of three positions (1: one-third; 2: middle; 3: two-thirds) from the onset of the first regular vocal pulse till the end of the rhyme in target syllables of each MC tonal category. Error bars represent the standard error of the mean.

show significant results between A and its two counterparts. The only significant result was found in I, where H1*-H2* after A was higher than that after U.

The consistently higher H1*-H2* of vowels after aspirated onsets was not observed. It was only present at the beginning of adjacent vowels (i.e., one-third position), but vanished after the midpoint regardless of whether ATS happened (i.e., *Shang*, *Qu*, and *Ru*) or not (i.e., *Ping*). This pattern held across generations. An interesting finding is that across all MC tonal categories, the two older groups showed more comparable patterns of H1*-H2* of vowels after voiced and aspirated onsets, whereas the young group showed a minimized H1*-H2* difference. This suggests that the phonatory state of vowels after voiced stops, is experiencing an on-going change across generations.

IV. DISCUSSION

A. New light on the two existing debates

The primary goal of this study is to examine the two long-standing debates, namely (i) is "aspiration-induced tonal split" (ATS) an on-going change or a completed change? and (ii) is onset aspiration or vowel breathiness the direct trigger for ATS?

With respect to the first debate, a stepwise lowering of lexical tones beginning with aspirated onsets as described in Fig. 2 was not observed across generations. The on-going change view therefore, is challenged. The results of GCA instead tend to favor the completed change view. A two-way categorization of *f0* contours conditioned by MC tonal categories was consistently observed across generations. For the MC *Ping* tonal category, both voiceless onset types (i.e., aspirated and unaspirated) introduced similar high-level *f0* contours, while voiced onsets introduced low-rising contours. For the remaining three MC tonal categories (i.e, *Shang*, *Qu*, and *Ru*), *f0* contours of aspirated onsets exactly patterned with contours of voiced onsets. Both, however, differed from *f0* contours of unaspirated onsets. This pattern held across generations.

As regards the second debate, our results lend no support to either claim. Specifically, results of the DOR/DOS ratio for aspirated onsets were similar in MC *Ping*, *Shang*, and *Qu*. All, however, were significantly lower than that of the *Ru* category across generations. This raises doubts about the "aspiration view," which argues that the aspirated onset is the obvious trigger of tonal split. If it was true, the DOR/DOS ratio for *Ping* should be significantly different from that for the other three MC tonal categories (i.e., *Shang*, *Qu*, and *Ru*). On the other hand, results of H1*-H2* showed that a higher magnitude of breathiness was only observed at the beginning of adjacent vowels (i.e., one-third position), but vanished after the midpoint across all MC tonal categories regardless of whether there was tonal split (i.e., *Shang*, *Qu*, and *Ru*) or not (i.e., *Ping*). Again, this pattern remained constant across generations. This finding is inconsistent with the expected results of the breathiness view, which predicts higher breathiness throughout the vowel for the MC tonal categories showing ATS. Taken together, we may safely draw the conclusion that synchronically speaking, there is no direct and consistent link between either aspiration or breathiness and tonal split in Lili Wu Chinese. The relationship between them is less transparent and more complex.

B. Tonal split and onset aspiration

Although aspiration is not likely to be the direct trigger of tonal split given the similar DOR/DOS ratio for the three MC tonal categories (i.e., *Ping*, *Shang*, and *Qu*), we believe that it played a crucial role in the process of tonal split. On the one hand, our results showed that irrespective of whether tonal split occurred or not, aspiration had a consistent perturbation effect on its following vowel, as evident in the strong degree of breathiness (indexed via the H1*-H2* differences). Such a relationship between aspiration and breathiness has been reported in some languages, such as Swedish (Gobl and Chasaide, 1988), English (Löfqvist and McGowan, 1992), and German (Chasaide and Gobl, 1993). The perceived greater breathiness has been attributed to a delayed laryngeal adjustment after the release of an aspirated onset. At the release of an aspirated onset, the glottis may have a more abducted posture, which consequently causes the vocalis muscle's effect either to be weak or not as effective as after an unaspirated stop, as Chen (2011) argued for the mechanism of *f0* lowering in Shanghainese. In addition, aspiration-induced greater aperiodic noise may also be at play. It has been argued that aspiration is typically followed by a considerably greater airflow (Stevens, 1971), which can result in a breathier transition between aspiration and vowel voicing (Sagart, 1981; Ren, 1992; Zhu and Xu, 2009). Given the widely observed compatibility between breathier phonation and lower *f0* (see Sec. I B 2), phonatory breathiness, together with aperiodic noise provides a possible pivot for linking onset aspiration to *f0* lowering. Aspirated onsets hence have the potential to behave like voiced onsets in introducing *f0* lowering. For example, similar lowering effect after aspirated and voiced onsets have been reported for Standard Thai (Gandour and Maddieson, 1976), Ikalanga (a Bantu language, Mathangwane,1996), and Tsua (a Khoisan language, Mathes and Chebanne, 2018). In Shanghai Wu Chinese, Chen (2011) found that the perturbation effect of the aspirated category was more similar to that of the voiced one in non-initial position. Moreover, lower *f0* and breathier phonation are found to correlate with both voiceless and voiced aspirated stops in Nepali (Clements and Khatiwada, 2007; Khatiwada, 2008; Mazaudon, 2012) and Bengali (Mikuteit and Reetz, 2007). However, it is also worth noting that in Lili Wu Chinese, such a breathier phonation can be generally observed at the onset of the vowel after all aspirated onsets regardless of whether ATS happened or not. This suggests that there should be no inevitable correlation between breathier voice and lower *f0*. As *f0* range is related to the degree of stiffness of the vocal cords while breathier voice is related to the glottal constriction and noise component, it is not difficult to imagine that the same glottal constriction and noise component can vary via different rates of vibration of the vocal folds (Ladefoged, 1973; Kuang, 2013b).

To summarize, in Lili Wu Chinese, not only voiced onsets, but also aspirated onsets can introduce a breathier phonation to the beginning of the following vowels, leading to a lowered *f0* contour. This produces a stable tonemic pattern across all generations of Lili speakers. However, it is also clear that this pattern is conditioned by certain tonal contexts (i.e., MC tonal categories).[8] This fact is reminiscent of the statement in Chen (2011, p. 622): "(…) [S]peakers may use different strategies to produce aspirated stops in different languages which lead to different perturbation effects." Based on the findings from Lili Wu Chinese, we may add that different strategies could also be adopted by speakers even within the same language.

C. An on-going change: The phonatory state of voiced onsets

One serendipitous finding is the reduced breathiness of vowels following voiced onsets as an on-going sound change. Consequently, we are interested in two questions: (i) why did this change happen? and (ii) what is the effect of this change on the phonological system? We will approach both questions from the perspectives of cue redundancy and robustness for signaling phonological contrasts.

In the Northern Wu dialects, there are a variety of acoustic cues for signaling the three-way laryngeal contrast, demonstrating a robustly encoded phonological contrast. For example, in Shanghai Wu Chinese, breathiness has been argued to act as a secondary cue for enhancement on vowels after voiced onsets, while the *f0* contour of a lexical tone is taken to be the primary cue for the contrast between syllables beginning with voiced and voiceless unaspirated onsets (Gao, 2015; Chen and Gussenhoven, 2015). A similar pattern can also be found in Lili Wu Chinese. As demonstrated in Table V, breathiness on vowels after voiced onsets has a superfluous role in cueing the three-way laryngeal contrast. First, for the contrast between voiced and voiceless aspirated onsets, VOT combined with *f0* suffices as a robust cue in the MC *Ping* category and VOT suffices as a robust cue in the MC *Shang*, *Qu*, and *Ru* categories. Second, with regard to the contrast between voiced and voiceless unaspirated onsets, lexical tonal contours serve as a robust cue. Finally, for the contrast between voiceless aspirated and unaspirated onsets, VOT serves as the prominent cue in the MC *Ping* category and both VOT and lexical tonal contour serve as primary cues for the MC *Shang*, *Qu*, and *Ru* categories. Given the superfluous role of breathiness in signaling any laryngeal contrasts of Lili Wu Chinese, it is not difficult to

TABLE V. Acoustic cues used for signaling the three-way laryngeal contrast in Lili Wu Chinese.

	Voiceless aspirated		Voiceless unaspirated	
Versus	*Ping*	*Shang*/*Qu*/*Ru*	*Ping*	*Shang*/*Qu*/*Ru*
Voiced	VOT & *f0*	VOT	*f0*	
Voiceless aspirated	—		VOT	VOT & *f0*

understand the reduced degree of breathiness in vowels following voiced onsets produced by young speakers.

With respect to the second question, generally speaking, the decrease of breathiness of vowels with voiced onsets can potentially threaten the three-way laryngeal consonant contrast. The contrast between voiced and voiceless unaspirated onsets is in jeopardy of losing its cue robustness. The weakening of cue robustness, to a large extent, introduces more bias, which then can reduce the precision of the contrast (Kirby, 2013). On the other hand, a contrast is more likely to survive when more cues signal it (Steven *et al.*, 1986; de Jong, 1995; Wright, 2004). Predictably, if no strategy of enhancement is taken by younger speakers, the loss of the redundant cue can eventually lead to the three-way laryngeal contrast becoming less robust. Very likely, with the weakening of cues for distinguishing the voiced vs voiceless unaspirated contrast, a neutralization of the three-way contrast may be triggered. Such a tendency has already been observed in some of the younger female speakers of Shanghainese (Gao, 2016) and in some speakers of Tamang dialects (Mazaudon, 2012).

V. CONCLUSION

This study provides a substantial amount of experimental data collected from Lili Wu Chinese to examine two debates of ATS in previous literature. In conclusion, our results suggest that ATS in Lili Wu Chinese is a completed sound change but conditioned by certain tonal contexts, namely, MC tonal categories. A direct link between either aspiration or breathiness and tonal split is not tenable in Lili Wu Chinese. One on-going sound change we observed serendipitously is that the breathiness of vowels after voiced onsets seems to be disappearing among the younger generation of Lili speakers. This is probably due to its superfluous role in cueing the three-way laryngeal contrast which makes it a less robust cue for the laryngeal contrast in Lili Wu Chinese. We may expect that this on-going language change can lead to the loss of this cue for the three-way laryngeal contrast in the future.

Tonal development is a complex sound change. More research on the issue of the integrality and the hierarchy of the phonetic properties associated with tonal contrasts, is urgently needed. In the future, we may consider comparing the acoustic properties (i.e., aspiration and breathiness) of aspirated onsets in Lili Wu Chinese to those of neighboring dialects such as Shanghainese where there is no ATS. It is still unclear how aspiration in other Wu dialects such as Shanghainese may differ from that in Lili Wu. If they show similarity, the question that follows is why ATS takes place in Lili Wu but not in Shanghai Wu. Answers to this question would be related to the "actuation" problem, which has been argued as one of the three long-standing questions in the study of sound change (Garrett and Johnson, 2013). Another unsolved issue is why ATS is conditioned by MC tonal categories and absent in the MC *Ping* tonal category only. Sagart (1981) assumes the tone-categories where ATS took place were "glottalized tones" (i.e., tones in which the vowel had creaky phonation) historically. More evidence from the reconstruction of proto-Wu Chinese is hence needed in order to support or overrule this view, which however, is beyond the scope of this paper. Another worthy topic for further exploration, pointed out by one of the reviewers, is the effect of speaking style on the realization of phonological contrasts. Kang and Guion (2008) show that phonological contrasts between Korean stops are enhanced in clear speech production compared with conversational or citation-form speech. In clear speech production, there are further generational differences in the use of VOT and *f0* cues: Older speakers tend to enhance VOT differences whereas younger speakers tend to enhance *f0* differences. What remains for further research is then to test whether similar contextual/situational variations in the phonetic realization of phonological contrasts can also be observed in Lili Wu Chinese and what their implications are on on-going sound changes.

ACKNOWLEDGMENTS

We would like to thank all our participants for making this study possible. We would also like to thank Yifei Bi, Feng Ling, and Ruiqing Shen for valuable comments on earlier versions of our paper, and thanks to Voeten Cesko, Zhongmin Chen, He Huang, James Kirby, Qian Li, Min Liu, Zhongwei Shen, Yiming Sheng, Rujie Shi, Huan Tao, Ping Wang, Qing Yang, and Dan Yuan for sharing their thoughts with us on various aspects. In addition, we are grateful to the Guest Associate Editor Richard Wright and the two anonymous reviewers for their helpful comments. Portions of this work were presented on different occasions including the 9th International Wu Dialect Conference (2016, Suzhou, China) and the 4th Workshop on Sound Change (2017, Edinburg, UK). We thank, as well, the audience for their feedback. The proofreading assistance from Kate Bellamy and Andrew Wigman is gratefully appreciated. This work is supported by China Scholarship Council (CSC) and Leiden University Centre for Linguistics (LUCL) to M.S. Neither the individuals and institutions cited herein nor the funding agency, however, should be held responsible for the views expressed in this paper.

APPENDIX

TABLE VI. Stimulus list.

	Unaspirated (U)	Aspirated (A)	Voiced (D)
Ping (I)	低ti̟ "low"	梯tʰi̟ "ladder"	提di̟ "to mention"
	单 tɛ "sole"	瘫tʰɛ "paralysis"	台dɛ "platform"
	刀 tɑ "knife"	涛tʰɑ "billow"	桃dɑ "peach"

TABLE VI. (*Continued*)

	Unaspirated (U)	Aspirated (A)	Voiced (D)
Shang (II)	底ti̤ “bottom”	体tʰi̤ “body”	弟di̤ “younger brother”
	胆tɛ “gallbladder”	毯tʰɛ “mat”	淡dɛ “light”
	岛tɑ “island”	讨tʰɑ “to ask for”	稻dɑ “rice”
Qu (III)	渧ti̤ “to drop”	替tʰi̤ “to replace”	地di̤ “ground”
	对tɛ “right”	退tʰɛ “to retreat”	代dɛ “dynasty”
	到tɑ “to arrive”	套tʰɑ “case”	盗dɑ “robber”
Ru (IV)	滴tɪʔ “drop”	贴tʰɪʔ “to paste”	敌dɪʔ “enemy”
	得tʌʔ “to get”	脱tʰʌʔ “to take off”	特dʌʔ “specially”
	搭taʔ “and”	塔tʰaʔ “pagoda”	达daʔ “to extend”

TABLE VII. Results (β) of model testing for the effect of Consonant on *f0* contours. Parameter-specific *p*-values (superscript) were estimated using the normal approximation. Baseline = voiceless aspirated onset (A). n.s.: not significant.

Generation	I-A	Intercept	Linear	Quadratic	II-A	Intercept	Linear	Quadratic
Old	I-U	$0.07^{\mathrm{n.s.}}$	$0.14^{\mathrm{n.s.}}$	$-0.14^{\mathrm{n.s.}}$	II-U	$1.51^{<0.001}$	$-1.45^{<0.001}$	$-0.25^{<0.05}$
	I-D	$-0.64^{<0.001}$	$1.45^{<0.001}$	$0.16^{\mathrm{n.s.}}$	II-D	$-0.15^{\mathrm{n.s.}}$	$0.28^{\mathrm{n.s.}}$	$-0.19^{\mathrm{n.s.}}$
Middle-aged	I-U	$-0.1^{\mathrm{n.s.}}$	$0.2^{\mathrm{n.s.}}$	$0.01^{\mathrm{n.s.}}$	II-U	$1.43^{<0.001}$	$-2.38^{<0.001}$	$-0.3^{<0.01}$
	I-D	$-0.78^{<0.001}$	$1.64^{<0.001}$	$0.42^{<0.001}$	II-D	$-0.23^{\mathrm{n.s.}}$	$0.28^{\mathrm{n.s.}}$	$-0.17^{\mathrm{n.s.}}$
Young	I-U	$-0.09^{\mathrm{n.s.}}$	$0.24^{\mathrm{n.s.}}$	$-0.13^{\mathrm{n.s.}}$	II-U	$1.15^{<0.001}$	$-1.98^{<0.001}$	$-0.48^{<0.05}$
	I-D	$-0.54^{<0.001}$	$1.59^{<0.001}$	$0.54^{<0.001}$	II-D	$-0.12^{\mathrm{n.s.}}$	$0.1^{\mathrm{n.s.}}$	$-0.07^{\mathrm{n.s.}}$
Generation	III-A	Intercept	Linear	Quadratic	IV-A	Intercept	Linear	Quadratic
Old	III-U	$0.6^{<0.001}$	$-1.84^{<0.001}$	$-0.09^{\mathrm{n.s.}}$	IV-U	$1.78^{<0.001}$	$-0.61^{<0.05}$	$0.03^{\mathrm{n.s.}}$
	III-D	$-0.16^{\mathrm{n.s.}}$	$0.19^{\mathrm{n.s.}}$	$-0.22^{\mathrm{n.s.}}$	IV-D	$-0.29^{\mathrm{n.s.}}$	$0.18^{\mathrm{n.s.}}$	$-0.02^{\mathrm{n.s.}}$
Middle-aged	III-U	$0.84^{<0.001}$	$-2.12^{<0.001}$	$-0.01^{\mathrm{n.s.}}$	IV-U	$1.56^{<0.001}$	$-0.84^{<0.01}$	$-0.01^{\mathrm{n.s.}}$
	III-D	$-0.1^{\mathrm{n.s.}}$	$0.15^{\mathrm{n.s.}}$	$-0.03^{\mathrm{n.s.}}$	IV-D	$-0.24^{\mathrm{n.s.}}$	$0.39^{\mathrm{n.s.}}$	$0.06^{\mathrm{n.s.}}$
Young	III-U	$0.48^{<0.001}$	$-2.08^{<0.001}$	$-0.26^{\mathrm{n.s.}}$	IV-U	$1.47^{<0.001}$	$-0.41^{\mathrm{n.s.}}$	$-0.32^{<0.01}$
	III-D	$0.01^{\mathrm{n.s.}}$	$-0.17^{\mathrm{n.s.}}$	$-0.18^{\mathrm{n.s.}}$	IV-D	$-0.002^{\mathrm{n.s.}}$	$0.16^{\mathrm{n.s.}}$	$-0.09^{\mathrm{n.s.}}$

TABLE VIII. Final models for calculating the results presented in Table VII.

Generation-Category	Fixed structure	Random structure
Old-I&II; Middle-I;	(ot1 + ot2) * Consonant + Vowel + Repetition + Gender	(ot1 + ot2 | Speaker) + (ot1 + ot2 | Speaker: Consonant) + (ot1+ ot2 | Item)
Young-I	(ot1 + ot2) * Consonant + Vowel + Repetition + Gender	(ot1 + ot2 | Speaker) + (ot1 + ot2 | Speaker: Consonant)
Middle-II; Young-III	(ot1 + ot2) * Consonant + Vowel + Repetition	(ot1 + ot2 | Speaker) + (ot1 + ot2 | Speaker: Consonant) + (ot1 | Item)
Young-II; Old-III; Middle-IV; Young-IV	(ot1 + ot2) * Consonant + Vowel + Repetition	(ot1 + ot2 | Speaker) + (ot1 + ot2 | Speaker: Consonant) + (ot1+ ot2 | Item)
Middle-III	(ot1 + ot2) * Consonant + Vowel	(ot1 + ot2 | Speaker) + (ot1 | Speaker: Consonant) + (1 | Item)
Old-IV	(ot1 + ot2) * Consonant + Vowel	(ot1+ ot2 | Speaker) + (ot1 + ot2 | Speaker: Consonant) + (ot1 + ot2 | Item)

TABLE IX. Results of linear-mixed effects model fit to the DOR/DOS ratio of each generation. Baseline = I. n.s.: not significant.

	Old			Middle-aged			Young		
	estimate (β)	*t*	*p*	estimate (β)	*t*	*p*	estimate (β)	*t*	*p*
Intercept	0.32	23.55	<0.001	0.33	23.8	<0.001	0.31	23.06	<0.001
II	−0.02	−1.44	n.s.	−0.01	−0.07	n.s.	−0.02	−1.38	n.s.
III	−0.02	−1.21	n.s.	−0.01	−1.08	n.s.	−0.01	−0.53	n.s.
VI	0.12	8.13	<0.001	0.11	6.42	<0.001	0.06	3.3	<0.001

TABLE X. Final models for calculating the results presented in Table IX.

Generation	Fixed structure	Random structure
Old and Middle-aged	Category + Vowel + Repetition	(1 + Category | Speaker) + (1 | Item)
Young	Category + Vowel	(1 + Category | Speaker) + (1 | Item)

TABLE XI. Results (β) of model testing for the effect of Consonant on H1*-H2*. "P1" to "P3" represent the three time positions where the H1*-H2* measurements were made. Parameter-specific p-values are superscripted. Baseline = voiceless aspirated onset (A). —: Consonant factor did not improve the model fit. n.s.: not significant.

Generation	I-A	P1	P2	P3	II-A	P1	P2	P3
Old	I-U	$-2.63^{<0.001}$	$-1.05^{<0.05}$	—	II-U	$-2.04^{<0.05}$	—	—
	I-D	$0.74^{\mathrm{n.s.}}$	$0.51^{\mathrm{n.s.}}$	—	II-D	$0.04^{\mathrm{n.s.}}$	—	—
Middle-aged	I-U	$-1.37^{<0.01}$	$-0.59^{\mathrm{n.s.}}$	—	II-U	$-2.85^{<0.05}$	$-1.52^{\mathrm{n.s}}$	—
	I-D	$0.55^{\mathrm{n.s.}}$	$0.87^{\mathrm{n.s.}}$	—	II-D	$-0.39^{\mathrm{n.s.}}$	$-0.25^{\mathrm{n.s.}}$	—
Young	I-U	$-2.17^{<0.001}$	—	—	II-U	$-1.89^{<0.05}$	—	—
	I-D	$-1.85^{<0.001}$	—	—	II-D	$-2.41^{<0.01}$	—	—

Generation	III-A	P1	P2	P3	IV-A	P1	P2	P3
Old	III-U	$-2.12^{<0.001}$	$-1.16^{\mathrm{n.s.}}$	—	IV-U	$-4.75^{<0.001}$	$-1.12^{\mathrm{n.s.}}$	—
	III-D	$0.72^{\mathrm{n.s.}}$	$0.38^{\mathrm{n.s.}}$	—	IV-D	$-3.12^{<0.001}$	$-0.94^{\mathrm{n.s.}}$	—
Middle-aged	III-U	$-2.16^{<0.05}$	$-1.06^{\mathrm{n.s.}}$	—	IV-U	$-3.81^{<0.001}$	—	—
	III-D	$0.05^{\mathrm{n.s.}}$	$0.55^{\mathrm{n.s.}}$	—	IV-D	$-2.97^{<0.001}$	—	—
Young	III-U	$-1.67^{<0.01}$	—	—	IV-U	$-3.3^{<0.001}$	—	—
	III-D	$-1.36^{<0.01}$	—	—	IV-D	$-4.3^{<0.001}$	—	—

TABLE XII. Final models for calculating the results presented in Table XI.

Generation-Category-Position	Fixed structure	Random structure
Old-I-P1	Consonant + Vowel	(1 | Speaker) + (1 | Item)
Old-I-P2; Middle-I-P1; Middle-IV-P1	Consonant + Vowel	(1 | Speaker)
Middle-I-P2; Young-I-P1; Young-III-P1	Consonant + Gender	(1 | Speaker)
Old-II-P1; Middle-II-P1	Consonant	(1 | Speaker) + (1 | Item)
Young-II-P1	Consonant + Gender	(1 | Speaker) + (1 | Item)
Middle-II-P2	Consonant + Vowel+ Gender	(1 | Speaker) + (1 | Item)
Old-III-P1&P2; Old-IV-P1	Consonant + Vowel	(1 + Consonant | Speaker)
Middle-III-P1	Consonant + Vowel	(1 + Consonant | Speaker) + (1 | Item)
Middle-III-P2	Consonant + Vowel + Gender	(1 + Consonant | Speaker)
Old-IV-P2	Consonant + Vowel + Repetition	(1 + Consonant | Speaker)
Young-IV-P1	Consonant + Vowel + Gender	(1 | Speaker)
Others	Consonant did not significantly improve the model fit.	

[1] Wu Chinese is spoken by approximately 70×10^6 people who reside in the area of Southern Jiangsu, Shanghai and Zhejiang provinces, as well as some part of Anhui, Jiangxi, and Fujian provinces in the People's Republic of China (Zhengzhang and Zheng, 2015).

[2] For the vast majority of the Wu dialects including Lili Wu Chinese, tone-carrying syllables in the MC *Ru* category are relatively short compared to those in the other three MC tonal categories and can co-occur with the glottal coda /ʔ/ only.

[3] In the existing literature, the same phenomenon has been referred to as "送气分调" [lit. aspiration divides tones] (e.g., Ho, 1989), "气流分调" [lit. airflow divides tones] (Xu, 2006), or "次清分调" [lit. secondary voiceless divides tones] (Zhu and Xu, 2009). The English appellations include "aspiration-conditioned tone-lowering" (Sagart, 1981), "tone-split by aspiration" (Ho, 1989), "aspirated tones" (Shen, 1994), "tonal split based on the aspiration" (Shi, 1998), and "tonal split following voiceless aspirated stop onsets" (Chen, 2011).

[4] In the existing literature, three non-modal phonation types, namely, slack/lax voice, breathy voice and whispery voice have been argued to be produced with a larger glottal aperture and less glottal constriction. All are regarded to be relatively "breathier," hence are further classified into the so-called "breathier voice" (see a comprehensive review in Tian and Kuang, 2019). However, these different types of "breathier voice" differ in the size of glottal aperture and the rate of flow of air. For example, breathy voice is argued to have a greater glottal aperture and a higher rate of flow of air than slack voice has (Ladefoged and Maddieson, 1996, pp. 57–66). Whispery voice, however, is produced with a substantial amount of aperiodic noise (Catford, 1977). Tian and Kuang (2019) argue that the non-modal phonation in Shanghai Wu Chinese would be better categorized as "whispery voice." In this study, we have no intention of exploring the differences among them, hence do not distinguish them strictly due to their similar distribution of energy in the fundamental and higher frequencies (Ladefoged and Maddieson, 1996, p. 317).

[5] IPA transcription: /kɛ^{44} kʌʔ joŋ^{213} li̟^{13} li̟ tʰu^{22} u naʔ^{3} haŋ u^{213}/. Tones are marked for each prosodic unit on the basis of the tone of the initial syllable due to tone sandhi.

[6] Δ stands for any random factor.

[7] Except for Chao (1928), which reports the absence of ATS in the MC *Shang* tonal category. A further question to be discussed is whether this is due to change over the MC *Shang* category not having started yet at that time.

[8] In Lili Wu Chinese, synchronically speaking, it seems that ATS is conditioned by the shape of tonal contours, namely, it cannot co-occur with the high-level contour. However, when we focus on other Wu varieties bearing ATS, the shape of tonal contours does not help to predict ATS. For example, in Jiaxing Wu (Yu, 1988), a Northern Wu dialect spoken in the city of Jiaxing, ATS is also observed in syllables with the high-level tone /44/, which developed from the MC *Shang* tonal category. The high-falling tone /51/ developed from the MC *Ping* tonal category, however, is not reported to show ATS. This pattern is consistent with what we have found in Lili Wu Chinese.

Abramson, A. S., and Erickson, D. M. (**1978**). "Diachronic tone splits and voicing shifts in Thai: Some perceptual data," in *Haskins Laboratories: Status Report on Speech Research SR-53 Vol. 2*, edited by Haskins Laboratories (Haskins Laboratories, New Heaven), pp. 85–96.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (**2016**). "Lme4: Linear mined-effects model using Eigen and S4 (package version 1.1-6)," http://CRAN.R-project.org/package=lme=4 (Last viewed August 15, 2019).

Blankenship, B. (**2002**). "The timing of nonmodal phonation in vowels," J. Phon. **30**, 163–191.

Boersma, P., and Weenink, D. (**2016**). "Praat: Doing phonetics by computer (version 6.0.18) [computer program]," http://www.praat.org/ (Last viewed August 15, 2019).

Brown, J. M. (**1975**). "The great tone split: Did it work in two opposite ways," in *Studies in Tai Linguistics in Honor of William J. Gedney*, edited by J. G. Harris and J. R. Chamberlain (Central Institute of English Language, Bangkok), pp. 33–48.

Brunelle, M., and Kirby, J. (**2016**). "Tone and phonation in Southeast Asian languages," Lang. Linguist. Compass **10**, 191–207.

Campbell, L. (**2013**). *Historical Linguistics: An Introduction*, 3rd ed. (Edinburgh University Press, Edinburgh).

Catford, J. C. (**1977**). *Fundamental Problems in Phonetics* (Indiana University Press, Bloomington, IN).

Chao, Y.-R. (**1928**). *Xiandai Wuyu de Yanjiu* (*Studies of the Modern Wu Dialects*) (Tsing Hua College Research Institute, Beijing).

Chao, Y.-R. (**1930**). "ə sistim əv 'toun-letəz' (A system of tone letters)," Le Maitre Phonetique **8**, 24–27; reprinted in (1980), Fangyan (Dialects) **2**, 81–83.

Chao, Y.-R. (**1967**). "Contrastive aspects of the Wu dialects," Language **43**, 92–101.

Chasaide, A. N., and Gobl, C. (**1993**). "Contextual variation of the vowel voice source as a function of adjacent consonants," Lang. Speech **36**, 303–330.

Chen, L.-Z. (**2005**). "Hanyu fangyan shengdiao Songqifenhua xianxiang chutan (An investigation on the tone-aspiration-division in Chinese dialects)," Hanyu Xuebao (Chinese Linguistics) **4**, 31–39.

Chen, S., Zhang, C.-C., McCollum, A. G., and Wayland, R. (**2017**). "Statistical modelling of phonetic and phonologised perturbation effects in tonal and non-tonal languages," Speech Commun. **88**, 17–38.

Chen, Y.-Y. (**2011**). "How does phonology guide phonetics in segment-*f0* interaction?," J. Phon. **39**, 612–625.

Chen, Y.-Y, and Gussenhoven, C. (**2015**). "Shanghai Chinese," J. Int. Phonetic Assoc. **45**, 321–337.

Chen, Z.-M. (**2014**). "Jiangzhehu hiaojiechu fangyan Songqifendiao jiqi xiangguan wenti (Aspiration-induced tonal split in dialects of the juncture area of Jiang-Zhe-Hu and its related questions)," in *Chengzetang Fangyan Luncong: Wang Futang Jiaoshou Bazhi Shouqi Lunwenji* (*Essays on Dialects: A Festschrift for Prof. Futang Wang*), edited by X.-F. Li, and M.-B. Xiang (Language and Literature Press, Beijing), pp. 123–138.

Cho, T., and Ladefoged, P. (**1999**). "Variation and universals in VOT: Evidence from 18 languages," J. Phon. **27**, 207–229.

Clements, G. N., and Khatiwada, R. (**2007**). "Phonetic realization of contrastively aspirated affricates in Nepali," in Proc. of ICPhS *16*, Saarbrücken, Germany, pp. 629–632.

De Jong, K. (**1995**). "On the status of redundant features: The case of backing and rounding in American English," in *Phonology and phonetic evidence: Papers in laboratory phonology IV*, edited by B. Connell, and A. Arvaniti (Cambridge University Press, Cambridge), pp. 68–86.

Esling, J. H., and Harris, J. G. (**2005**). "States of the glottis: An articulatory phonetic model based on laryngoscopic observations," in *A Figure of Speech: A Festschrift for John Laver*, edited by W. J. Hardcastle, and B. J. Mackenzie (Routledge, New York), pp. 347–383.

Esposito, C. M., and Khan, S. D. (**2012**). "Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong," J. Int. Phon. Assoc. **42**, 123–143.

Francis, A. L., Ciocca, V., and Yu, J. M. C. (**2003**). "Accuracy and variability of acoustic measures of voicing onset," J. Acoust. Soc. Am. **113**, 1025–1032.

Gandour, J., and Maddieson, I. (**1976**). "Measuring larynx movement in standard Thai using the cricothyrometer," Phonetica **33**, 241–267.

Gao, J.-Y. (**2015**). "Interdependence between tones, segments and phonation types in Shanghai Chinese: Acoustics, articulation, perception and evolution," Ph.D. dissertation, Université Sorbonne Nouvelle–Paris III, Paris, France.

Gao, J.-Y. (**2016**). "Sociolinguistic motivations in sound change: On-going loss of low tone breathy voice in Shanghai Chinese," Papers Hist. Phonology **1**, 166–186.

Garrett, A., and Johnson, K. (**2013**). "Phonetic bias in sound change," in *Origins of Sound Change: Approaches to Phonologization*, edited by A. C. L. Yu (Oxford University Press, Oxford), pp. 51–97.

Gobl, C., and Chasaide, A. N. (**1988**). "The effects of adjacent voiced voiceless consonants on the vowel voice source: A cross language study," STL-QPSR **2-3**, 23–59.

Gordon, M., and Ladefoged, P. (**2001**). "Phonation types: A cross-linguistic overview," J. Phon. **29**, 383–406.

Hanson, H. M. (**1997**). "Glottal characteristics of female speakers: Acoustic correlates," J. Acoust. Soc. Am. **101**, 466–481.

Harris, J. (**1969**). *Spanish Phonology* (MIT Press, Cambridge, MA).

Haudricourt, A. G. (**1972**). "Two-way and three-way splitting of tonal systems in some Far Eastern languages," translated by C. Court, in J. G. Harris, and R. B. Noss, *Tai Phonetics and Phonology* (Central Institute of English Language, Bangkok), pp. 58–86.

Hirano, M., Vennard, W., and Ohala, J. J. (**1970**). "Regulation of register, pitch and intensity of voice," Folia Phoniatr. Logop. **22**, 1–20.

Hirayama, H. (**2010**). "Ciyindiao xingcheng de shengxue yuanyin—Yi Wujiang fangyan weili (On the phonetic cause of the evolution of the ci-yin tones—The case of the Wujiang dialects)," Zhongguo Fangyan Xuebao (Bull. Chin. Dialects) **2**, 17–23.

Ho, D.-A. (**1989**). "Songqifendiao ji xiangguan wenti (The Tone-split by Aspiration and other related problems)," BIHP **60**, 765–778.

Hombert, J-M., Ohala, J. J., and Ewan, W. G. (**1979**). "Phonetic explanations for the development of tones," Language **55**, 37–58.

House, D., and Svantesson, J. O. (**1996**). "Tonal timing and vowel onset characteristics in Thai," in *Proc. of Pan-Asiatic Linguistics 4*, Bangkok, Thailand, pp. 104–113.

Iseli, M., Shue, Y. L., and Alwan, A. (**2007**). "Age, sex, and vowel dependencies of acoustic measures related to the voice source," J. Acoust. Soc. Am. **121**, 2283–2295.

Jacewicz, E., and Fox, R. A. (**2015**). "Intrinsic fundamental frequency of vowels is moderated by regional dialect," J. Acoust. Soc. Am. **138**, EL405–EL410.

Kang, K. H., and Guion, S. G. (**2008**). "Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies," J. Acoust. Soc. Am. **124**, 3909–3917.

Keating, P. A., Esposito, C., Garellek, M., Khan, S., and Kuang, J. (**2010**). "Phonation contrasts across languages," UCLA Work. Pap. Phon. **108**, 188–202.

Kessinger, R. H., and Blumstein, S. E. (**1997**). "Effects of speaking rate on voice-onset time in Thai, French, and English," J. Phon. **25**, 143–168.

Khatiwada, R. (**2008**). "Some acoustic cues in the detection of the Nepalese aspiration," J. Acoust. Soc. Am. **123**, 3889–3889.

Kirby, J. (**2013**). "The role of probabilistic enhancement in phonologization," in *Origins of Sound Change: Approaches to Phonologization*, edited by A. C. L. Yu (Oxford University Press, Oxford), pp. 228–246.

Kuang, J.-J. (**2013a**). "The tonal space of contrastive five level tones," Phonetica **70**, 1–23.

Kuang, J.-J. (**2013b**). "Phonation in tonal contrasts," Ph.D. dissertation, UCLA, Los Angeles, CA.

Kurpaska, M. (**2010**). *Chinese Language(s): A Look through the Prism of the Great Dictionary of Modern Chinese Dialects* (Walter de Gruyter, Berlin).

Labov, W. (**1994**). *Principles of Language Change: Internal factors* (Blackwell, Oxford).

Labov, W., Ash, S., and Boberg, C. (**2006**). *Atlas of North American English: Phonology and Phonetics* (Mouton de Gruyter, Berlin).

Labov, W., Rosenfelder, I., and Fruehwald, J. (**2013**). "One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis," Language **89**, 30–65.

Ladefoged, P. (**1973**). "The features of the larynx," J. Phon. **1**, 73–83.

Ladefoged, P., and Maddieson, I. (**1996**). *The Sounds of the World's Languages* (Blackwell, Oxford).

Lai, Y., Huff, C., Sereno, J., and Jongman, A. (**2009**). "The raising effect of aspirated prevocalic consonants on F0 in Taiwanese," in *Proc. of East Asian Linguistics 2*, available at: https://kuppl.ku.edu/sites/kuppl.ku.edu/files/docs/Lai_EtAl.pdf (Last viewed August 15, 2019).

Lam, T. Q., and Watson, D. G. (**2010**). "Repetition is easy: Why repeated referents have reduced prominence," Mem. Cogn. **38**, 1137–1146.

Laver, J. (**1994**). *Principles of Phonetics* (Cambridge University Press, Cambridge).

Lea, W. A. (**1973**). "Segmental and suprasegmental influences on fundamental frequency contours," in *Consonant Types and Tone 1*, edited by L. H. Hyman (University of Southern California, California), pp. 15–70.

Lehiste, I., and Peterson, G. (**1961**). "Transitions, glides, and diphthongs," J. Acoust. Soc. Am. **33**, 268–277.

Li, F.-K. (**1977**). *A Handbook of Comparative Tai* (The University Press of Hawaii, Honolulu, HI).

Li, Q., and Chen, Y.-Y. (**2016**). "An acoustic study of contextual tonal variation in Tianjin Mandarin," J. Phon. **54**, 123–150.

Löfqvist, A., and McGowan, R. S. (**1992**). "Influence of consonantal environment on voice source aerodynamics," J. Phon. **20**, 93–110.

Mathangwane, J. T. M. (**1996**). "Phonetics and phonology of Ikalanga: A diachronic and synchronic study," Ph.D. dissertation, University of California, Berkeley, CA.

Mathes, T. K., and Chebanne, A. M. (**2018**). "An SS ANOVA (smoothing spline analysis of variance) study of high tone lowering in Tsua," S. Afr. J. Afr. Lang. **38**, 137–148.

Mazaudon, M. (**2012**). "Paths to tone in the Tamang branch of Tibeto-Burman (Nepal)," in *The Dialect Laboratory: Dialects as a Testing Ground for Theories of Language Change*, edited by G. Vogelaer and G. Seiler (John Benjamins Publishing Company, Amsterdam), pp. 139–178.

Mikuteit, S., and Reetz, H. (**2007**). "Caught in the ACT: The timing of aspiration and voicing in East Bengali," Lang. Speech **50**, 247–277.

Mirman, D. (**2014**). *Growth Curve Analysis and Visualization Using R* (CRC Press, Boca Raton, FL).

Norman, J. (**1988**). *Chinese* (Cambridge University Press, Cambridge).

Ohala, J. J. (**1978**). "Production of tone," in *Tone: A linguistic survey*, edited by V. A. Fromkin (Academic, New York), pp. 5–39.

Pan, W.-Y., Li, L., and Han, X. (**2015**). "Field Phon (Version F2.0.2) [computer program]," http://www.eastling.org/ (Last viewed 1 December 2016).

Pulleyblank, E. G. (**1978**). "The nature of the Middle Chinese tones and their development of early Mandarin," J. Chin. Linguist. **6**, 173–203.

R Core Team (**2016**). "R: A language and environment for statistical computing (Version 3.2.2) [computer program]," http://www.R-project.org/ (Last viewed 18 August 2019).

Ren, N.-Q. (**1992**). "Phonation types and stop consonant distinctions: Shanghai Chinese," Ph.D. dissertation, University of Connecticut, Storrs, CT.

Rischel, J. (**1986**). "Can 'the great tone split' in Thai be phonetically explained?," Annu. Rep. Inst. Phon. Univ. **20**, 79–98.

Rose, P. (**1987**). "Considerations in the normalization of the fundamental frequency of linguistic tone," Speech Commun. **6**, 343–352.

Sagart, L. (**1981**). "Aspiration-conditioned tone-lowering in Chinese dialects," in *Proc. Sino-Tibetan Languages and Linguistics 14*, Gainesville, FL, https://www.academia.edu/38233622/Aspiration-conditioned_tone_lowering_in_Chinese_dialects (Last viewed 22 August 2019).

Shen, Z.-W. (**1994**). "The tones in Wujiang dialect," J. Chin. Linguist. **22**, 279–315.

Shi, F. (**1992**). "Wujiang fangyan shengdiao geju de fenxi (The analysis of the tonal space in Wujiang dialects)," Fangyan (Dialects) **3**, 189–194.

Shi, F. (**1998**). "Songqi shengmu dui shengdiao de yingxiang (The influence of aspiration on tones)," J. Chin. Linguist. **26**, 126–145.

Shi, F. (**2008**). *Yuyin Geju: Yuyinxue he Yinxixue de Jiaohuidian* (*Phonetic Space: The Intersection between Phonetics and Phonology*) (The Commercial Press, Beijing).

Shue, Y.-L., Keating, P. A., Vicenik, C., and Yu, K. (**2011**). "VoiceSauce: A program for voice analysis [computer program]," http://www.ee.ucla.edu/~spapl/voicesauce/ (Last viewed November 1, 2017).

Simpson, A. P. (**2012**). "The first and second harmonics should not be used to measure breathiness in male and female voices," J. Phon. **40**, 477–490.

Stevens, K. N. (**1971**). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," J. Acoust. Soc. Am. **50**, 1180–1192.

Stevens, K. N., Keyser, S. J., and Kawasak, H. (**1986**). "Toward a phonetic and phonological theory of redundant features," in *Symposium on invariance and variability of speech processes*, edited by J. Perkell and D. H. Klatt (Lawrence Erlbaum Associates, Hillsdale, MI), pp. 426–449.

Thavisak, A. (**2004**). "F0 behavior of vowels influenced by aspirated and unaspirated initials in Southeast Asian languages: Implications for tonogenesis theories," in *Proc. of the Southeast Asian Linguistics Society 11*, Tempe, Arizona, pp. 721–730.

The Institute of Linguistics of the Chinese Academy of Social Sciences. (**1988**). *Fangyan Diaocha Zibiao (Questionnaire of Character for Dialect Surveys)* (Commercial Press, Beijing).

Thurgood, G. (**2007**). "Tonogenesis revisited: Revising the model and the analysis," in *Studies in Tai and Southeast Asian Linguistics*, edited by J. G. Harris, S. Burusphat, and J. E. Harris (Ek Phim Thai Co., Bangkok), pp. 263–291.

Tian, J., and Kuang, J.-J. (**2019**). "The phonetic properties of the non-modal phonation in Shanghainese," J. Int. Phonetic Assoc. (published online).

Ting, P.-H. (**1984**). "Wuyu shengdiao zhi yanjiu (Reconstruction of proto-Wu tones)," BIHP **55**, 755–788.

Turk, A., Nakai, S., and Sugahara, M. (**2006**). "Acoustic segment durations in prosodic research: A practical guide," in *Methods in Empirical Prosody Research*, edited by S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schließer (Walter de Gruyter, Berlin), pp. 1–28.

Wang, P. (**2008**). "Wujiang fangyan shengdiao zaitaolun (Tones in the Wujiang dialects: A revisit)," Zhongguo Yuwen (Studies of the Chinese Language) **5**, 455–461.

Weinreich, U., Labov, W., and Herzog, M. I. (**1968**). "Empirical foundations for a theory of language change," in *Directions for Historical Linguistics: A Symposium*, edited by W. P. Lehmann (University of Texas Press, Austin), pp. 95–195.

Wickham, H. (**2009**). *Ggplot2: Elegant Graphics for Data Analysis* (Springer, New York).

Winter, B., and Wieling, M. (**2016**). "How to analyze linguistic change using mixed models, Growth curve analysis and generalized additive modeling," J. Lang. Evol. **1**, 7–18.

Wright, R. (**2004**). "A review of perceptual cues and cue robustness," in *Phonetically Based Phonology*, edited by B. Hayes, R. Kirchner, and D. Steriade (Cambridge University Press, Cambridge), pp. 34–57.

Wurm, S. A., Li, R., Baumann, T., and Lee, M. W. (**1987**). *Language Atlas of China* (Longman, Hong Kong).

Xu, Y. (**2006**). "Jiangshan fangyan de Qiliufendiao (The Tone-Split by aspiration of Jiashan dialect)," Yuyan Yanjiu (Studies in Language and Linguistics) **3**, 77–80.

Xu, Y. (**2013**). "Hanyu fangyan zhong de Qiliufendiao xianxiang (Tone-split by aspiration in Chinese dialects)," Stud. Chin. Lang. **3**, 265–272.

Ye, X.-L. (**1983**). "Wujiang fangyan shengdiao zaidiaocha (A revisit to tones in the Wujiang dialects)," Fangyan (Dialects) **1**, 32–35.

Yu, A. C. L. (**2007**). "Understanding near mergers: The case of morphological tone in Cantonese," Phonology **24**, 187–214.

Yu, G.-Z. (**1988**). "Jiangxiang fangyan tongyin zihui (A syllable for Jiangxing dialect)," Fangyan (Dialects) **3**, 195–208.

Zhang, J., and Meng, Y.-L. (**2016**). "Structure-dependent tone sandhi in real and nonce disyllables in Shanghai Wu," J. Phon. **54**, 169–201.

Zhengzhang, S.-F., and Zheng, W. (**2015**). "Wu dialect," in *The Oxford Handbook of Chinese Linguistics*, edited by W. S.-Y. Wang, and C.-F. Sun (Oxford University Press, Oxford), pp. 189–199.

Zhu, X.-N., and Xu, Y. (**2009**). "Chihua: Tansuo Wujiang Ciqingfendiao de yuanyin (Slack voice as the cause of tone split in Wujiang: A phonetic investigation into the pitch contour of the syllables with a voiceless aspirated onset in the Wujiang (Songling) Wu dialect)," Zhongguo Yuwen (Stud. Chin. Lang.) **4**, 324–332.