Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

Investigating Glottal Stop Coda Loss During Sound Change of Checked Syllables Based on Speech-EGG Voice Offset Alignment

Bingliang Zhao[1,2], *Xiyu Wu*[1,2]

[1]Laboratory of Language Sciences, Peking University, China
[2]Department of Chinese Language and Literature, Peking University, China
zhaobingliang@pku.edu.cn, xiyuwu@pku.edu.cn

Abstract

The glottal stop coda in Wu Chinese checked syllables diachronically weakens into non-modal phonation on the preceding vowel and finally disappears. It's crucial to capture the initial stage of this sound change when the glottal stop coda is realized as an abrupt glottal closure. Aiming to clarify the age-related differences in the phonetic realizations of the glottal stop coda, we detected the presence of the abrupt glottal closure at voice offset using speech and EGG signals from Shengzhou Wu speakers. Results show that the voice offsets in checked syllables produced by most old speakers remain aligned, indicating the sound change has not been initiated. However, the glottal stop coda is no longer realized as abrupt glottal closure in the younger generation. This study directly measures the glottal stop coda loss through the physiological basis of articulation, potentially providing insights applicable to abrupt glottal closure detection for other linguistic purposes.

Index Terms: voice offset, checked syllables, glottal closure instant, Shengzhou Wu

1. Introduction

The cross-linguistic variability in the phonological status and phonetic realization of glottal stops has long intrigued phonetic researchers. In Wu Chinese, checked syllables exhibit glottal stop coda and short vowel duration, maintaining a phonemic contrast with their unchecked counterparts. Previous studies have reported that checked syllables in Wu Chinese are in the process of an ongoing sound change, which is marked by glottal stop coda loss and vowel duration lengthening [1, 2]. This could result in the merger of checked and unchecked syllables. During this sound change, the glottal stop coda may weaken to a creaky voice on preceding vowels or disappear entirely. Different speakers, grouped by factors such as age and sex, may experience different stages of glottal stop coda weakening. Developing robust parameters to detect the presence of glottal stop coda in checked syllables is essential to reconstruct the micro-stages of this sound change and enhance the understanding of the interplay between the phonological structure and sociophonetic variation.

The phonetic analysis of glottal stops is inherently complex due to the mismatch between their phonological representations and phonetic realizations. The IPA symbol ʔ in the existing literature is frequently applied not only to voiceless glottal stop with abrupt glottal closure but also to non-modal phonation (e.g., creaky voice which occurs partway through vowels) [3]. Current detection methods for glottal stops in speech signals mainly rely on measuring irregular vocal fold vibrations and longer closed phases in voiced segments. The periodicity is quantified via subharmonic-to-harmonic ratio (SHR), harmonic-to-noise ratio (HNR) [4], and short-time peak autocorrelation [5] while longer closed phase indicating more glottal constriction can be measured through parameters like dominant resonance frequency (DRF) and strength of excitation (SoE) [6]. When applied to sound change research of the checked syllables, these approaches primarily identify non-modal phonation derived from *weakened* glottal stop but fail to capture the abrupt glottal closure that characterizes the *initial* stage of the glottal stop loss. Notably, a previous study reveals that neighboring voiced segments of fully realized glottal stops with abrupt closure exhibit minimal glottalization [7]. Therefore, new methods that directly capture the articulatory characteristics of glottal closure are critically required to precisely determine whether the glottal stop loss in checked syllables remains at its initial stage or has progressed in different populations.

We propose analyzing the abrupt glottal closure of the glottal stop coda in checked syllables from the perspective of voice offset. The voice offsets of vowels preceding voiceless stops (including true glottal stops) are linguistically constrained, which is marked by abrupt cessation of phonation. On the other hand, linguistically unconstrained voice offsets (e.g., vowels in open syllables like unchecked syllables in Wu Chinese) allow gradual decay of vocal fold vibration [8]. The different physiological events in the laryngeal system at voice offsets may serve as a reliable marker of the presence of true glottal stops in Wu Chinese checked syllables.

Compared to speech signals, glottographic technologies perform better and more precisely when observing the oscillatory movements of vocal folds at voice offsets. Some studies using costly methods like high-speed videoendoscopy have revealed that linguistically unconstrained voice offsets are characterized by progressive vocal fold contact area reduction [9, 10]. Moreover, electroglottography (EGG) has also been used to explore voice offsets with simultaneously recorded speech signals. Vocal Release Time (VRT), measuring the time lag between the falling of the speech and EGG signals at the voice offset, effectively quantifies linguistically unconstrained voice offsets of sustained vowels [8]. Although this approach is non-invasive and more accessible, its application to tonal languages like Wu Chinese is limited. As described by [11], VRT extraction depends on bandpass filtering centered on F0, which may be unreliable under dynamic F0 contours, such as the rising or falling tones in Wu Chinese.

To overcome this limitation, we employed glottal closure instant (GCI) detection algorithms, which measure the vocal fold movements. GCIs mark the moments of maximum excitation when the vocal folds contact initially in each glottal cycle [12]. In speech signals, GCIs align with peaks in the Linear Prediction (LP) residual, reflecting abrupt vocal tract excitation during glottal closure [13]. GCIs also correspond to the positive

 10.21437/Interspeech.2025-38

peaks in the EGG derivative [14]. By comparing the locations of the last GCIs in speech and EGG signals, we can directly detect the abrupt glottal closure or the gradually reduced vocal fold vibration in checked and unchecked syllables, respectively.

This study aims to investigate the articulatory properties of the glottal stop coda in checked syllables, taking Shengzhou Wu as an example. We recorded speech and EGG signals from native Shengzhou Wu speakers. The alignment of the last GCIs in speech and EGG signals was used to measure the abrupt glottal closure that occurs at the voice offset of checked syllables. Since the checked syllable in Shengzhou Wu is undergoing glottal stop coda loss, age-related differences in the realizations of checked syllables were also discovered. This paper not only puts forward a method for directly detecting glottal stop codas in tonal languages but also provides a framework for tracking ongoing sound change and sociophonetic variation in speech production data in endangered languages or dialects.

2. Method

We collected a comprehensive dataset of checked and unchecked monosyllabic tokens in Shengzhou Wu with simultaneously recorded speech and EGG signals. Based on the GCI detection algorithms described below, we estimated the voice offset time difference between the speech and EGG signals. The originally developed MATLAB code mentioned in Section 2.2 and the detailed results of the statistical analyses in R are publicly accessible at `https://github.com/Dzaau/Interspeech2025`.

2.1. Data collection

Thirty-nine native speakers from Chongren Town (first documented in [15]) participated in the production experiment. All participants were physically healthy with normal hearing and vision. Most of them were bilingual in Shengzhou Wu and Mandarin. Younger speakers exhibited greater fluency in Mandarin due to better educational opportunities. Informed consent forms were obtained, and participants received monetary compensation after the experiment.

Table 1: *Age groups and sex composition of participants.*

Group	N	Mean (S.D.) of age
Old	10 Female; 11 Male	57.09 (5.02)
Young	9 Female; 9 Male	26.83 (3.78)

The recording material included monosyllabic words in CV(ʔ) structure, selected based on phonotactic constraints. Target syllables were designed to cover full factorial combinations of three vowel types (/a, e, o/), four onset consonants, and eight lexical tones. Notably, these tones are divided into two registers: four tones of lower register co-occurring with voiced onsets (voiced plosives and nasals) and four tones of higher register co-occurring with voiceless onsets (voiceless unaspirated/aspirated plosives). This yielded a 3 (vowels) $\times$ 4 (onsets) $\times$ 4 (tones per register) design with 48 stimuli in total. Checked syllables (bearing T7 and T8 tones, historically derived from Middle Chinese *Rusheng* category) were explicitly marked by their glottal stop codas, making up 25% of the stimuli.

Recordings were conducted in a sound-attenuated room using Adobe Audition 2025 on a laptop computer with a sampling rate of 44,100 Hz and 16-bit quantization. The SONY ECM-77B microphone and the EGG D200 model manufactured by Laryngograph were utilized for recording signals with an external mixer XENYX 302USB and a sound card XONAR U5. Material presentation and response timing were controlled via a PsychoPy program [16]. Each trial displayed a target syllable in a single Chinese character on the screen, and participants were required to produce it twice with enough pauses to prevent tonal sandhi. The production experiment included 60 randomized trials, consisting of 48 targets and 12 fillers.

We used MATLAB R2023a and PRAAT [17] to inspect, annotate, and segment the collected signals. Data from four female participants and one male participant were excluded from the analysis due to the poor quality of their EGG signals. We employed an automated PRAAT script to detect syllable boundaries [18], followed by manual adjustments to accurately annotate the vowel segments for each syllable. Each annotated vowel segment, along with an additional 200 ms before and after, was saved as a single file. We ensured that each file contained only one target segment, and 176 files were discarded due to incidental noise.

2.2. Parameter estimation

We quantified voice offset lag (VOL) for each syllable as the temporal difference between the last detected GCIs in speech and EGG signals. The positive value of VOL implies that the last GCI in the EGG signal occurs later than the last GCI in the speech signal, and vice versa. Therefore, VOL values are expected to be close to zero with a small variance if glottal stop codas exist, but show larger absolute values with significant variance in unchecked syllables. GCI detection was implemented via customized MATLAB routines combining filtering and peak detection. Figure 1 illustrates two examples of GCI detection and VOL calculation.

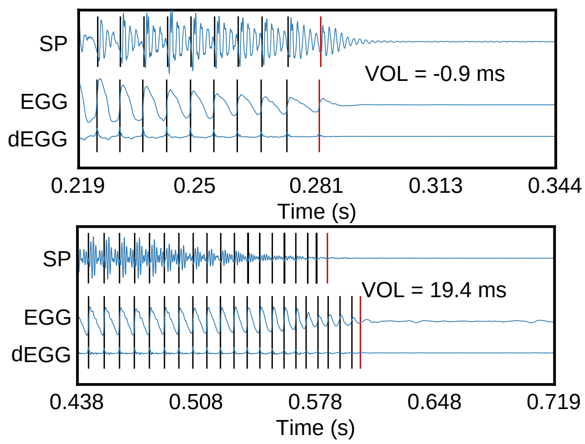

Figure 1: *Examples of VOL calculation of one checked token (top) and one unchecked token (bottom) produced by an old male speaker. The detected GCIs are marked in black lines, and the last GCIs are highlighted in red.*

A publicly available algorithm was employed for GCI detection in the speech signals [19]. We chose this algorithm because of its robustness when applied to our data. It applies a raised cosine filter (RCF) to suppress vocal tract resonances in the speech signals. Voiced segments were identified, followed by adaptive peak detection using dynamic thresholding to extract GCIs.

A custom MATLAB code for EGG signal processing was developed by the first author (accessible on GitHub). First, a 10 Hz high-pass filter was employed to eliminate baseline drift. Then, wavelet-based decomposition removed high-frequency noise. The positive peaks in the derivative of EGG (dEGG) were used to identify candidate GCIs. Finally, the Hilbert transform-derived amplitude envelope was calculated to detect syllable boundaries. The nearest GCI before the syllable endpoint marks the last GCI in the EGG signals.

We used Contact Quotient (CQ) to validate the performance of VOL. CQ, defined as the ratio of the closed phase to the glottal cycle in the EGG signals, has proven effective when measuring glottal constriction in a Wu dialect [20]. Higher CQ indicates greater glottal constriction. CQ was computed via the 30% amplitude threshold method in Kay3700. For each token, the mean CQ was extracted from the final third of the annotated vowel segment, a temporal window sensitive to coarticulation. Participants were considered to retain glottal stop codas in checked syllables if they exhibited significantly higher CQ compared to their unchecked counterparts. It's anticipated that VOL will correlate with CQ in these participants.

2.3. Statistical analysis

All statistical analyses were conducted in R using the `lme4` package [21]. Outliers in VOL and CQ measurements were removed using the 3-standard deviation criterion, resulting in 3,027 valid tokens for subsequent analysis.

To validate the reliability of VOL to detect glottal stop coda in checked syllables, general linear models (GLMs) were fitted separately for each participant. Pearson correlation analyses then evaluated the relationship between model goodness-of-fit (adjusted R^2) for VOL and CQ across speakers.

Linear mixed-effects models were employed to examine VOL differences between checked and unchecked syllables. The independent variable `syllable type` was coded as a binary variable with `unchecked` as the reference level. `Age` and `sex` and their interactions with `syllable type` were also included. Grouping speakers by age is a common practice for analyzing an ongoing sound change, so `age` was binary coded to highlight group-level differences, with `old` as the reference level.

3. Results

3.1. The correlation of VOL with CQ

To assess the efficacy of VOL in detecting glottal stop codas in checked syllables, we analyzed the correlation between VOL and CQ for each participant. As described in Section 2.2, checked syllables with glottal stop coda tend to have higher CQ than unchecked syllables. Therefore, GLMs were applied to all participants, with `syllable type` as the independent variable and `VOL` or `CQ` as the dependent variable.

Eighteen participants showed higher CQ in checked versus unchecked syllables (βs > 0, *p*s < 0.05), indicating retained glottal stop coda in their productions. Figure 2 (left) displays the scatterplot of R^2 values of CQ and VOL for these 18 participants. In contrast, the CQ values of the remaining participants didn't reveal any significant differences between the checked and unchecked syllables ($R^2 \approx 0$, *p*s > 0.05).

A strong positive correlation was discovered between R^2 values of VOL and CQ ($r = 0.74$, $p < 0.001$), indicating that participants with greater CQ differentiation between syllable types also exhibited stronger VOL contrasts. For example, Figure 2

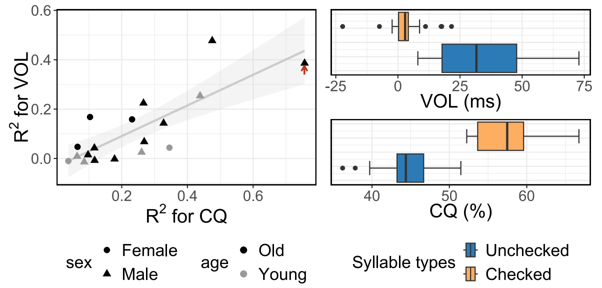

Figure 2: *Left: scatterplot showing the correlation of VOL with CQ. Right: the boxplots of VOL and CQ from one participant marked with a red arrow in the left panel.*

(right) shows that there is a significant difference in the distribution of CQ and VOL between the two syllable types for one old male participant (marked with a red arrow in the scatterplot), which is reflected by high R^2 values for both VOL and CQ.

Notably, 12 of the 18 participants in Figure 2 (left) were from the old group. This indicates age-related differences in the preservation of glottal stops in checked syllables, which were further investigated through the analyses of VOL in Section 3.2.

3.2. Age-related differences revealed by VOL

The distributions of VOL of checked and unchecked syllables are different across age groups (Figure 3). To quantify potential age and sex effects, a linear mixed-effects model was fitted with `VOL` as the dependent variable. The fixed factors included `syllable type`, `age`, `sex`, and their interactions. Based on model comparisons, the random effects structure was `(syllable type|participant) + (age|item)`. The VOL of checked syllables is significantly lower than unchecked syllables, as indicated by the fixed effect of `syllable type` (β = -13.48, $p < 0.001$). `sex` showed no fixed effect or significant interaction with other factors (*p*s > 0.05). Although `age` itself lacked a fixed effect, its interaction with `syllable type` was significant (β = 12.28, $p < 0.05$). With `unchecked` and `old` as the reference levels of the fixed factors, the coefficient implies: (1) Younger speakers produced checked syllables with significantly higher VOL than older speakers; (2) No age-based VOL difference existed for unchecked syllables.

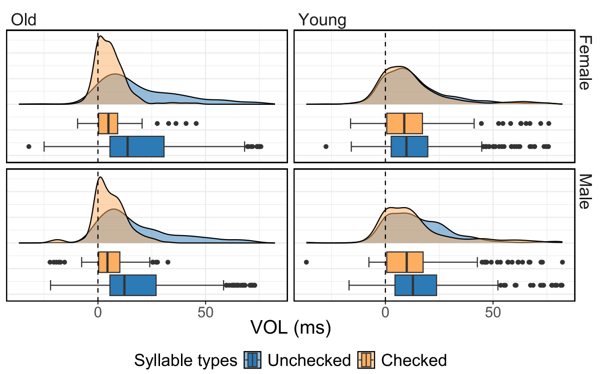

Figure 3: *The distribution of VOL in different ages and sexes. The dashed line marks the zero value of VOL.*

To further compare the differences in VOL among different age groups, the mean and standard deviation of VOL for checked and unchecked syllables were separately calculated for each participant (Figure 4). Notably, in the older group, both the mean and standard deviation of VOL for checked syllables were significantly smaller than those for unchecked syllables (Mean: β = -13.16, $p < 0.001$; S.D.: β = -8.71, $p < 0.001$). However, in the younger group, these measures showed no significant difference between checked and unchecked syllables (ps>0.05).

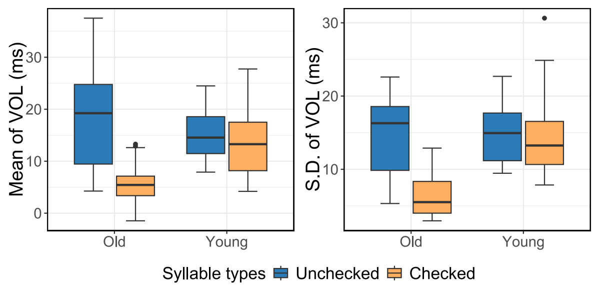

Figure 4: *The mean and S.D. of VOL for participants in different age groups.*

4. Discussion

We propose VOL to quantify the temporal difference between the last GCIs in speech and EGG signals. The results show that VOL is a robust parameter to track the glottal stop coda loss in checked syllables across generations.

The near-zero mean and reduced variance of VOL in typical checked syllables reflect abrupt phonation cessation enforced by glottal stop codas. When a full glottal closure occurs, vocal fold vibration terminates instantaneously, synchronizing the final GCIs in speech and EGG signals. Notably, unchecked syllables mainly exhibit large positive VOL values, indicating a later offset in EGG than in speech signals. This finding is consistent with [22] but contrasts with previous reports of negative VRT for linguistically unconstrained voice offsets, where the offset time in EGG signals precedes that in speech signals [8]. While advanced EGG offsets may arise from small-amplitude oscillations of the vocal fold, the delayed EGG offsets remain physiologically unexplained [23]. We hypothesize that in unchecked syllables, reduced vocal tract resonance and airflow attenuation at syllable endings result in undetectable GCIs in speech signals due to small amplitudes. In contrast, abrupt glottal closure in checked syllables prevents such syllable-end adjustments, preserving synchrony between speech and EGG signals. Our GCI-based method prevents limitations of VRT in tonal languages, where dynamic pitch contours invalidate F0-based bandpass filtering. Besides, VOL directly measures the abrupt glottal closure via speech-EGG alignment, identifying the earliest stage of checked syllable sound change. The physiological precision of VOL makes it superior to established measures like the CQ, which effectively quantifies glottal constriction during phonation but lacks sensitivity to abrupt glottal closure at voice offsets.

The age-related differences in VOL patterns mirror the loss of glottal stop coda during the ongoing sound change of Shengzhou Wu checked syllables. Compared to the older group, the distribution range of VOL for checked syllables in the younger group significantly shifts closer to that of unchecked syllables, suggesting that these two groups may be at different stages of sound change. The VOL for checked syllables in the older group is relatively concentrated near the value of zero, reflecting the synchronized voice offset in speech and EGG signals for checked syllables with a glottal stop coda. In contrast, the VOL distribution for checked syllables in the younger group is more dispersed, largely overlapping with the VOL of unchecked syllables, which indicates that young speakers tend to lose the glottal stop coda when producing checked syllables. This aligns with our perceptual study, where younger listeners rely more on vowel duration rather than the glottal stop coda to distinguish between checked and unchecked syllables [24].

The alignment of voice offsets in speech and EGG signals reflects the abrupt glottal closure inherent to the glottal stop coda in checked syllables. However, the relative importance of this articulatory property compared to other cues in maintaining phonemic contrasts between the checked and unchecked syllables is still unclear. Many languages contrast syllables with and without glottal codas, often using additional acoustic cues like short vowel duration [4, 25, 26] or different vowel quality [27]. This may account for the relatively low values of R^2 shown in Figure 2. Further systematic studies on the relationship between production and perception will not only help identify the primary contrastive cues but also shed light on the micro-mechanisms of sound change [28], especially as checked syllables evolve into unchecked ones. This sound change is particularly significant in the evolution of many southern Chinese dialects [26].

5. Conclusion

This study introduces a new method to directly detect the abrupt glottal closure of the glottal stop coda in checked syllables. The voice offset lag (VOL) was established as the time difference between the last detected GCIs in speech and EGG signals to identify if abrupt glottal closure existed. Checked syllables with glottal stop coda exhibit near-aligned GCIs in speech and EGG signals, while the speech-EGG voice offsets of unchecked syllables are misaligned. Although the VOL of unchecked syllables remains stably large across age groups, most younger participants have lost the abrupt glottal closure at voice offsets of checked syllables compared to older speakers. This indicates an extremely rapid sound change where the glottal stop coda loss has not been initiated in many old speakers, while the glottal stop coda has been almost lost in the younger generation. Therefore, the checked and unchecked syllables become less distinguishable and tend to result in an entire merger if other cues also disappear. These findings deepen our understanding of the special articulatory and acoustic characteristics of the glottal stop coda during the ongoing sound change of checked syllables. This study also provides an effective framework to identify the presence of abrupt glottal closure in various studies of sound change, prosody, and emotional speech.

Although VOL effectively detects the abrupt glottal closure when glottal stop codas are fully realized, its reliance on high-quality speech/EGG signals may limit its applications. Future work could integrate air pressure measurements or high-speed imaging to validate the glottal closure events of checked syllables. Additionally, comparative studies of dialects like Cantonese, where checked syllables retain unreleased oral stop codas (-p, -t, -k), could help to explore whether similar voice offset synchrony patterns exist. Such efforts will advance cross-dialectal models of the sound change of checked syllables and enrich typological perspectives on laryngeal contrasts.

6. Acknowledgements

This study was supported by the National Social Science Foundation of China (No.22&ZD213). We thank Xuejun Cai and Yiwen Lou for their assistance in recruiting participants. We thank all the participants for participating in this study.

7. References

[1] X. Zhu, L. Jiao, Z. Yan, and Y. Hong, "入声演化三途 (Three ways of Rusheng sound change)," *Studies of the Chinese Language*, no. 4, pp. 324–338+383–384, 2008.

[2] Z. Cao, "吴徽语入声演变的方式 (On the change of entering tones in Wu and Hui dialects)," *Studies of the Chinese Language*, no. 5, pp. 441–446+479, 2002.

[3] M. Garellek, Y. Chai, Y. Huang, and M. V. Doren, "Voicing of glottal consonants and non-modal vowels," *Journal of the International Phonetic Association*, vol. 53, no. 2, pp. 305–332, 2023.

[4] X. Gao and J. Kuang, "Phonation variation as a function of checked syllables and prosodic boundaries," *Languages*, vol. 7, no. 3, p. 171, 2022.

[5] M. Ashby and J. Przedlacka, "Measuring incompleteness: Acoustic correlates of glottal articulations," *Journal of the International Phonetic Association*, vol. 44, no. 3, pp. 283–296, 2014.

[6] S. Kalita, W. Lalhminghlui, L. Horo, P. Sarmah, S. M. Prasanna, and S. Dandapat, "Acoustic Characterization of Word-Final Glottal Stops in Mizo and Assam Sora," in *Interspeech 2017*. ISCA, 2017, pp. 1039–1043.

[7] C. T. DiCanio, "Coarticulation between tone and glottal consonants in Itunyoso Trique," *Journal of Phonetics*, vol. 40, no. 1, pp. 162–176, 2012.

[8] B. C. Watson, R. M. Roark, and R. Baken, "Vocal release time: A quantification of vocal offset," *Journal of Voice*, vol. 26, no. 6, pp. 682–687, 2012.

[9] R. R. Patel, R. Walker, and M. Döllinger, "Oscillatory onset and offset in young vocally healthy adults across various measurement methods," *Journal of Voice*, vol. 31, no. 4, pp. 512–e17, 2017.

[10] P. H. Dejonckere and J. Lebacq, "Damping of vocal fold oscillation at voice offset." *Biomed. Signal Process. Control.*, vol. 37, pp. 92–99, 2017.

[11] R. M. Roark, B. C. Watson, R. Baken, D. J. Brown, and J. M. Thomas, "Measures of vocal attack time for healthy young adults," *Journal of Voice*, vol. 26, no. 1, pp. 12–17, 2012.

[12] R. Smits and B. Yegnanarayana, "Determination of instants of significant excitation in speech using group delay function," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 5, pp. 325–333, 1995.

[13] T. Drugman and T. Dutoit, "Glottal closure and opening instant detection from speech signals," in *Proc. Interspeech 2009*, 2009, pp. 2891–2894.

[14] J. Z. Tomaszewska and A. Georgakis, "Electroglottography in medical diagnostics of vocal tract pathologies: A systematic review," *Journal of Voice*, 2023.

[15] Y.-R. Chao, 现代吴语的研究 *(Studies of the Modern Wu Dialects)*. Beijing: Tsinghua Xuexiao Yanjiuyuan, 1928.

[16] J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger, H. Sogo, E. Kastman, and J. K. Lindeløv, "Psychopy2: Experiments in behavior made easy," *Behavior research methods*, vol. 51, no. 1, pp. 195–203, 2019.

[17] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]," https://www.praat.org, 1992–2022.

[18] N. H. de Jong, J. Pacilly, and W. Heeren, "Praat scripts to measure speed fluency and breakdown fluency in speech automatically," *Assessment in education: Principles, policy & practice*, vol. 28, no. 4, pp. 456–476, 2021.

[19] P. Rengaswamy, G. Reddy, K. S. Rao, and P. Dasgupta, "A robust non-parametric and filtering based approach for glottal closure instant detection." in *INTERSPEECH*, 2016, pp. 1795–1799.

[20] J. Tian and J. Kuang, "The phonetic properties of the non-modal phonation in shanghainese," *Journal of the International Phonetic Association*, vol. 51, no. 2, pp. 202–228, 2021.

[21] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[22] R. R. Patel, K. Forrest, and D. Hedges, "Relationship between acoustic voice onset and offset and selected instances of oscillatory onset and offset in young healthy men and women," *Journal of Voice*, vol. 31, no. 3, pp. 389–e9, 2017.

[23] R. Zhang, "Vocal release time of mandarin vowels at different pitch levels," *Journal of Voice*, vol. 35, no. 5, pp. 804–e1, 2021.

[24] B. Zhao, J. Kong, and X. Wu, "Age-related differences in acoustic cues for the perception of checked syllables in Shengzhou Wu," in *Proc. Interspeech 2024*, 2024, pp. 4233–4237.

[25] C. DiCanio, "Cue weight in the perception of Trique glottal consonants," *The Journal of the Acoustical Society of America*, vol. 135, no. 2, pp. 884–895, 2014.

[26] Y. Chai and S. Ye, "Checked syllables, checked tones, and tone sandhi in Xiapu Min," *Languages*, vol. 7, no. 1, 2022.

[27] Y. Yang and Y. Chen, "Effects of entering tone on vowel duration and formants in Nanjing dialect," in *Studies on Speech Production: 11th International Seminar, ISSP 2017, Tianjin, China, October 16-19, 2017, Revised Selected Papers 11*. Springer, 2018, pp. 146–157.

[28] J. Schertz and E. J. Clare, "Phonetic cue weighting in perception and production," *WIREs Cognitive Science*, vol. 11, no. 2, p. e1521, 2020.