'Apical vowels' are not vowels: acoustic and ultrasound evidence from Jixi-Hui Chinese

Bowei Shao, Rachid Ridouane

▶ **To cite this version:**

Bowei Shao, Rachid Ridouane. 'Apical vowels' are not vowels: acoustic and ultrasound evidence from Jixi-Hui Chinese. 12th International Seminar on Speech Production, Dec 2021, On line, United States. hal-03598741

HAL Id: hal-03598741

https://hal.science/hal-03598741

Submitted on 5 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

‘Apical vowels’ are not vowels: acoustic and ultrasound evidence from Jixi-Hui Chinese

Bowei Shao, Rachid Ridouane

Laboratoire de Phonétique et Phonologie
Université Sorbonne Nouvelle - CNRS (UMR7018), France

bowei.shao@sorbonne-nouvelle.fr, rachid.ridouane@sorbonne-nouvelle.fr

Abstract

This study focuses on the ‘apical vowel’ as attested in Jixi-Hui Chinese. The objective is to determine its acoustic and articulatory characteristics when it occurs in the context of labial onsets. Phonologically, this segment is a distinct phoneme that opposes to /i/ from which it is diachronically derived. It is exclusively attested in the syllable nucleus position where it constitutes a tone-bearing unit. On the articulatory level, the ultrasound data show that when preceded by the bilabial consonants [p p^h m], the ‘apical vowel’ displays an articulatory configuration which is virtually identical to that of the alveolar fricative [s]. The tongue configuration of [z] is realized in an anticipatory way during the bilabials. One consequence of this overlap is that the release phase of [p^h] exhibits similar acoustic characteristics as the friction noise of the fricative [s], as evidenced by the resemblance of their centre of gravity. These results provide additional evidence that the apical vowel in Jixi-Hui Chinese is better defined, at least phonetically, as a voiced alveolar fricative.

Keywords: apical vowel, Jixi-Hui Chinese, ultrasound, centre of gravity

1. Introduction

Chinese languages are known to have a specific set of segments, termed ‘apical vowels’ (Lee & Zee 2003, Duanmu 2007, Lee-Kim 2014). These segments, according to Ladefoged and Maddieson (1996: 314), are ‘*made with the tongue in essentially the same position as in the corresponding fricatives*’ and refer to these segments as ‘fricative vowels’. Trubetzkoy (1969: 171) describes them as ‘*a type of vowel with a much lesser degree of aperture and with a much more fronted position of articulation than, for example, i, so that a frictionlike noise resembling a humming is audible in its production*’.

‘Apical vowels’ do not display the same phonetic characteristics nor the same phonological behaviour across Chinese languages (Shao 2020). Standard Chinese (SC), where these segments have been mostly studied, has two apical vowels, both analyzed as allophonic variants of /i/: [ɿ] occurs after dental sibilants, and [ʅ] occurs after retroflex sibilants (R. Cheng, 1966; C. Cheng, 1973). The two variants are always homorganic to the preceding coronal fricatives or affricates [s ts ts^h ʂ tʂ $tʂ^h$], while [i] occurs following palatal sibilant onsets [tɕ $tɕ^h$ ɕ] and other consonants. Based on the homorganicity between the ‘apical vowels’ and the preceding sibilant consonants, these segments, occupying the nucleus position of a syllable, are interpreted as the voiced prolongation of the sibilant onsets (Dell 1994), derived through a feature spreading process (Duanmu 2007). Following this analysis, the ‘apical vowels’ are not the manifestations of an underlying phoneme but are triggered by an empty nucleus. The feature [+fricative] spreads from the sibilant onsets into the nucleus, resulting in a fricative nucleus homorganic to the onsets.

In this study, we are concerned with the ‘apical vowel’ in Jixi-Hui Chinese (JHC) 绩溪话, a Chinese language of Hui 徽 group spoken in southern Anhui 安徽 province. Its ‘apical vowel’, noted /z/ or [z] here (traditionally as [ɿ]), has two structural properties that make it different from the SC variant: (i) Unlike in SC, the JHC apical vowel is a distinct phoneme which contrasts with /i/ and the other vowels (e.g., [si^{31}] 修 ‘repair’ vs. [sz^{31}] 丝 ‘silk’, [tsi^{31}] 周 ‘a family name’ vs. [tsz^{31}] 鸡 ‘chicken’), (ii) It appears not only following the coronal sibilants [s ts ts^h] but also following labial stops and nasals (e.g., [pz^{214}] 比 ‘compare’, [p^hz^{214}] 被 ‘quilt’, [mz^{214}] 米 ‘rice’, [nz^{214}] 里 ‘in’).

Our previous acoustic study (Shao & Ridouane 2018) showed that [z] in JHC is produced predominantly with frication noise superposed on voicing. The frication noise never continues throughout the entire duration of [z], and stops earlier or later in the second half of the segment. When frication noise diminishes, the formant structure becomes clearer, resulting in an approximant-like configuration. This dynamic evolution of [z] during its time course makes it often realized as a hybrid segment with the first part fricative-like and the second part more approximant-like.

In Shao and Ridouane (2019), we presented a study on the articulatory configuration of [z] based on ultrasound data. Results showed that the tongue configuration of [z] resembles that of an alveolar sibilant [s] on both mid-sagittal and coronal planes. On the mid-sagittal plane, when speakers adopt different strategies for [s], they also adopt the same strategies for [z]; so that the two segments always display similar tongue configurations. Similarly, on the coronal plane, [z] displays a medial-grooved tongue shape, similar to the fricative [s]. This medal-grooving indicates a narrowed air passage which is typical of an alveolar fricative consonant.

In the present study, we provide further phonetic analyses of [z] in JHC, and focus on its acoustic and articulatory characteristics when it occurs in the context of labial onsets. The idea is that the characteristics of [z] will be negligibly influenced by the onset consonants in the context of labials, assumed to be minimally coarticulatory. We specifically examine the nature of the gestural overlap between [z] and the labial onsets [p p^h m] and evaluate its acoustic consequences. The aim is to show that [z] in JHC is fully specified for its phonetic features and can’t be analyzed as a continuation of the preceding consonant.

2. Methods

We collected acoustic data from 10 JHC speakers (5 females and 5 males) and ultrasound tongue imaging data from 7 speakers (3 females and 4 males). A word list was constructed with the following 5 segments /i a u ʉ z/ occupying the nucleus of monosyllabic real words starting with /p_ pʰ_ m_ n_ ts_ tsʰ_ s_/ with different tones. Each word was pronounced within a carrier phrase ([ki˦ ɕɔ˧˩˦ _ ɕɔ˧˩˦ sɔ˧ fa˧] 'He writes _ three times'), with 5 repetitions for acoustic recording and 6 repetitions for ultrasound recording (3 repetitions for the mid-sagittal plane and 3 for the coronal plane). Mid-sagittal and coronal data were all obtained using an Ultrasound Stabilisation Headset (Articulate Instruments Ltd. 2008) and the Articulate Assistant Advanced software (AAA, V217.03) (Articulate Instruments Ltd. 2012). The coronal data were obtained by turning the probe by 90º when the mid-sagittal recording is completed. The ultrasound recording of one male speaker is excluded due to technical reasons.

The ultrasound study focuses on the gestural coordination between the labial onsets and [z], on both mid-sagittal and coronal planes. The tongues shapes of different time-points were extracted in x/y coordinates and generalized into SS ANOVA splines (Davidson 2006). By doing so, a dynamic tongue shape evolution can be observed during the [pz pʰz mz] syllables. The advantage of SS ANOVA is that it can show if two splines are significantly different and where the significant differences lie. If the interaction term of the SS ANOVA model is statistically significant, then the groups have different shapes. Since the interaction may be significant even if only a small section of the curves is different (e.g., the tongue root is the same, but the tip of one group is raised), Bayesian confidence intervals are used to determine which sections of the curves are statistically different (Davidson, 2006).

The acoustic experiment provides a comparison of the spectral centre of gravity (COG) of the aspiration phase of [pʰ] and that of the frication noise of [s] in the context of [z] (i.e., [pʰz] vs. [sz]). The COG, which reflects on average how high the frequencies are in a spectrum (Stevens 1998), was obtained with Praat by averaging the COG of the entire duration and over the entire frequency domain of the targets (Boersma & Weenink 2018).

3. Results

3.1. Tongue configuration of [z] in the context of labials

Figure 1 presents the mid-sagittal tongue configuration of [z] when preceded by the aspirated labial stop [pʰ] for the six speakers (note that the first four speakers and the last two speakers have different tongue configurations; see Shao & Ridouane (2019) for more details). The tongue shapes were taken from three different points: the mid-point image of the closure phase of [pʰ], the mid-point image of the aspiration phase of [pʰ] and the mid-point image of nucleus [z]. As the figure clearly shows, the fricative alveolar gesture for [z] is achieved as early as during the closure phase of [pʰ]. The tongue keeps this gestural configuration during the entire syllable [pʰz]. This is shown by the fact that confidence intervals overlap for all speakers (though in a lesser degree for MS3 where the tongue dorsum seems to be higher during the closure of [pʰ]). The same dynamic tongue shape evolution is also observed for the forms [pz] and [mz], suggesting that when the apical vowel [z] is preceded by a labial consonant, the fricative alveolar gesture largely overlaps with the labial gesture.

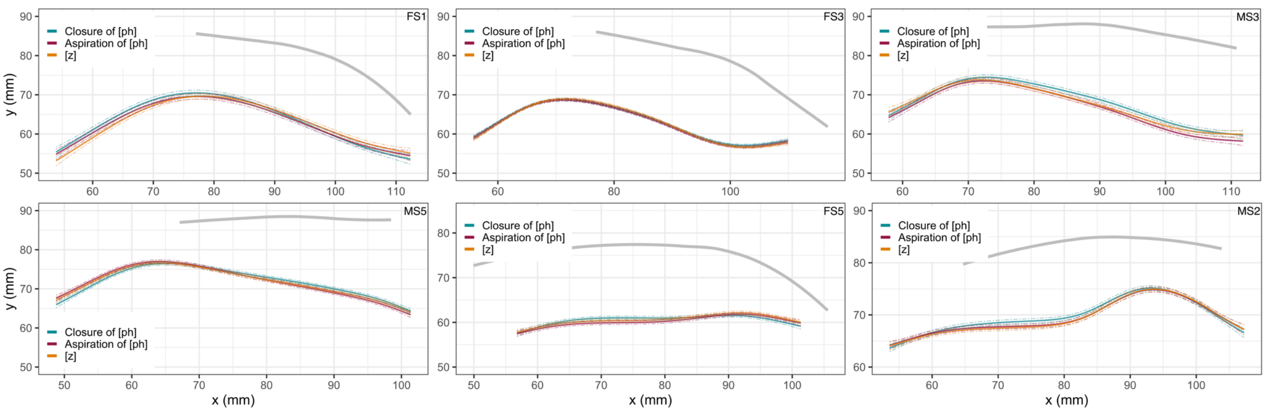

Figure 1: *SS ANOVA splines (mm) of mid-sagittal ultrasound tongue contours, extracted in x/y values (mm) (tongue front on the right and tongue root on the left). Blue lines represent the nearest mid-point images of the closure of [pʰ], red lines represent the nearest mid-point images of the release phase of [pʰ], and yellow lines represent the nearest mid-point images of [z]. The splines are presented with 95% Bayesian confidence intervals. The thick grey lines represent the palatal traces.*

The coronal tongue shape evolution for the syllable [pʰz] is shown in **Figure 2**. The tongue shapes were also taken from three different points: nearest mid-point image of the closure phase of [pʰ], the aspiration phase of [pʰ] and the nucleus [z], generated in smoothing-splines. The medial grooving of [z] is achieved during the closure phase of the aspirated labial plosive (note that FS1 does not show medial grooving neither for [z] nor for [s]). This medial-grooved tongue shape is maintained during the release phase of [pʰ] and continues into the nucleus [z]. The same pattern is also observed for the forms [pz] and [mz].

The tongue configuration for [z] indicates that this segment has articulatory characteristics of a fricative consonant (i.e., the [s]-like tongue configuration and the presence of medial grooving). The labial consonants, with no lingual gesture, have minimal influence on this articulatory configuration, suggesting that this alveolar gesture is an inherent property of [z], and can't be considered a continuation of the preceding consonant.

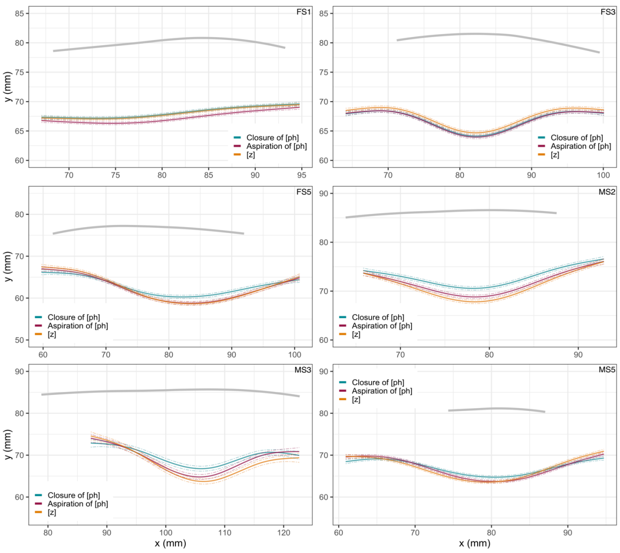

Figure 2: *SS ANOVA splines of coronal ultrasound tongue contours extracted in x/y values (mm). Same legend as Figure 1.*

3.2. Centre of gravity: comparing the aspiration in [p^h] to frication in [s]

The articulatory study shows that the alveolar fricative gesture for [z] is achieved as early as during the closure phase of the preceding labial consonants. One consequence of this gestural overlap is a modification of the acoustic characteristics of the aspiration phase of the labial stop [p^h]. The release phase of [p^h] in [p^{h}z] displays different acoustic characteristics compared to the aspiration phase of the same segment when preceding genuine vowels. This is illustrated in **Figure 3**, where the release phases of [p^h] are compared in [p^{h}u] and [p^{h}z]. Unlike in [p^{h}u], the release of [p^h] in [p^{h}z] presents energy concentration at higher frequencies, much similar to a fricative [s], suggesting that the frication noise generated by the alveolar constriction for [z] dominates the glottal sourced aspiration noise of [p^h].

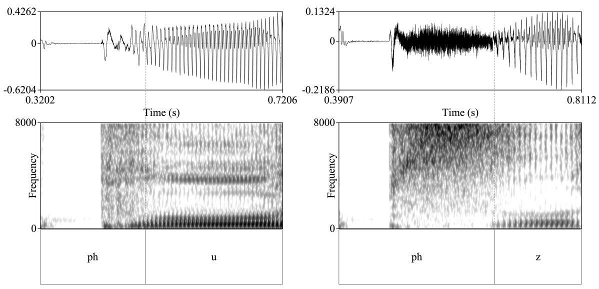

Figure 3: *Acoustic waveforms and spectrograms of one realization of [p^{h}u] (left) and one realization of [p^{h}z] by MS2.*

COG measurement was used to quantify the acoustic similarity between the aspiration phase of [p^h] and the frication noise of [s] in the same nucleic contexts. The results are presented in **Figure 4**. They show that when followed by [ʉ u a], the COG of [s] and the COG of the release phase of [p^h] are different. However, when followed by [z], the release of [p^h] displays the same COG as the frication noise of [s]. Welch's *t*-tests conducted on each vocalic context confirm this similarity from a statistical point of view (see **Table 1**).

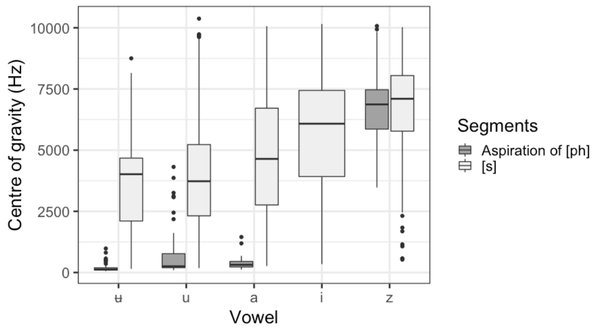

Figure 4: *COG of the release phase of [p^h] and the frication noise of [s] in the contexts of [i ʉ u a z]. Data obtained from all speakers ([i] vowel cannot be preceded by [p^h]).*

Table 1: *Mean COG values (Hz) of [s] and the release phase of [p^h]. Welch's* t*-test results conducted on the two groups of COG values.*

Nuclei	ʉ	u	a	z
Mean COG of the release phase of onset [p^h]	176	711	384	6735
Mean COG of onset [s]	3670	4003	4728	6682
Welch's *t*-test	$t(50.26)=-12.12$, $p<.0001$	$t(123.09)=-11.84$, $p<.0001$	$t(50.33)=-11.75$, $p<.0001$	$t(164.54)=0.19$, $p=.85$

4. Discussion and conclusion

This study presented an acoustic and articulatory investigation of the JHC apical [z], focusing on the labial context. The [z] has phonetic characteristics of an alveolar fricative. The articulatory gesture of this segment is virtually identical to that of an alveolar [s]. This similarity was observed on both mid-sagittal and coronal planes. The medial-grooved tongue shape of [z] is particularly important, as it is a fundamental indicator of a narrowed air channel typical of a fricative gesture. The fact that this fricative-like tongue shape was observed in the context of labial onsets is important since it shows that this tongue configuration is not a mere consequence of gestural overlap between homorganic sibilants and apical vowels (as argued for in other Chinese languages such as SC).

When preceded by labial onsets, the gesture for [z] is achieved anticipatorily during the labial closure and is maintained during the nucleus [z]. A consequence of this anticipatory gesture is that the release phase of [p^h] is realized with similar acoustic characteristics as the frication noise of [s], as shown by the COG measurement.

While the phonetic characteristics of [z] provide evidence that this segment is a fricative, the way it phonologically patterns may suggest that it is best defined as a vowel. Indeed, similar to vowels, [z] functions as a carrier of prosodic information (i.e., it occupies the nucleus position of a syllable and can be a tone-bearing unit (TBU)). This point of view is phonologically convenient since it complies with the habitual syllable structure of Chinese languages, in which the nucleus of a syllable should almost always contain a vowel.

4.1. Apical vowel: finding its place in the vowel-consonant continuum

There exists no clear-cut boundary between the vowel category and the consonant category, similar to any categorization based on complex physical phenomena. Between the 'most vocalic sounds', such as [a e], to the 'most consonantal sounds', such as [p k], there is a 'grey area' where sonorants [l r], approximants [j w] and even voiced fricatives [z ʐ] can be found. Thus, the speech sounds, when considered as physical entities, form a continuum on both acoustic and articulatory plane, with the 'most vocalic sounds' on one side and the 'most consonantal sounds' on the other side (see Saussure 1916; Pike 1947; Greenberg 1962; among others). A phonetic vowel may have the phonological function of a consonant and a phonetic consonant may have the phonological function of a vowel. The semi-vowel/semi-consonant/glide is in the former case, and the apical vowel in JHC is in the latter case. The phonetic categories and the phonological categories thus have different specificities: the vowel-consonant ensemble in phonetics may form a continuum, while in phonology, the functions of speech sounds are categorical.

The presence of apical vowels in Chinese languages has always been related to a historical vowel /i/ in some stage of the evolution (Zhu 2004; Zhao 2007). The voiced fricative realization in JHC could be one step within this evolution, while in other Chinese languages an approximant-like sound may be observed. To account for the apparent mismatch between the phonetic manifestation of the apical vowel and its phonological function, we argue that /z/ should be analyzed as a syllabic fricative, and thus call for an extension of the nucleic inventory of JHC.

This analysis is certainly unusual and calls into question the nature of syllable nuclei in Chinese languages (in SC for example only vowels can arguably be syllable nuclei). The situation in JHC is different, however. In addition to /z/ three other consonants can be syllable nuclei and thus tone-bearing units: /m n v/. Although these consonantal TBUs are not equally productive in the lexis, it is still a fact that they function as syllable nuclei (e.g., [m^{214}] 母 'female animal'; [n^{214}] 你 'you'; [fv^{214}] 虎 'tiger'). This analysis thus implies that the segments in JHC that can serve as TBUs are not only phonetic vowels, but also nasal consonants and voiced fricatives. All these segments pattern as a natural class which can be defined using the features [+ continuant, + voiced].

The nucleic inventory proposed here of JHC is not typologically uncommon. Various other languages in the world allow segments other than vowels to be syllable nuclei (e.g., /l/ in English, [bɒ.tl] 'bottle'; /r/ in Czech, [br.no] 'Brno'; /n/ in German, [ha.bn] 'to have'). The JHC nucleic inventory is only one step further (than English, for example) into more consonantal sounds. The case of Tashlhiyt (Dell & Elmedlaoui 1988; Ridouane 2008) is yet another example where a nucleic inventory even extends into all the consonantal sounds. As shown by Greenberg (1962), the existence of friction continuant nuclei (i.e., /z v/) implies the existence of frictionless continuant nuclei (i.e., /m n/), and the existence of frictionless continuant nuclei implies the existence of vowel nuclei. This is the case for, a priori, all languages. JHC, with its syllabic consonants /m n v z/, makes no exception to this typological generalization.

5. Acknowledgements

We thank all speakers who have kindly participated in all recording sessions. This study is part of the French program 'Investissements d'Avenir' run by the Agence Nationale de la Recherche ANR-10-LABX-0083 (Labex EFL).

6. References

Articulate Instruments Ltd. (2008). Ultrasound Stabilisation Headset User's Manual: Revision 1.5. Edinburgh, UK: Articulate Instruments Ltd.

Articulate Instruments Ltd. (2012). Articulate Assistant Advanced User Guide: Version 2.14. Edinburgh, UK: Articulate Instruments Ltd.

Boersma, P., & Weenink D. (2018). PRAAT: Doing phonetics by computer. Version 6.0.43, http://www.praat.org (accessed November 2018).

Cheng, C. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.

Cheng, R. L. (1966). Mandarin phonological structure. *Journal of Linguistics*, 2(2), 135–262.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America*, 120(1), 407–415.

Dell, F. (1994). Consonnes à prolongement syllabique en Chine. *Cahiers de linguistique–Asie orientale*, 23(1), 87–94.

Dell, F., & Elmedlaoui, M. (1988). Syllabic consonants in Berber: Some new evidence. *Journal of African Languages and Linguistics*, 10(1), 1–17.

Duanmu, S. (2007). *The phonology of Standard Chinese*. New York: Oxford University Press.

Greenberg, J. H. (1962). Is the Vowel–Consonant Dichotomy Universal? *Word*, 18(1-3), 73–81.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford & Malden, MA: Blackwell.

Lee-Kim, S.-I. (2014). Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association*, 44(3), 261–282.

Lee, W.S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, 33(1), 109–112.

Pike, K.L. (1947). *Phonemics. A Technique for Reducing Languages to Writing.* Ann Arbor, MI: The University of Michigan Press.

Ridouane, R. (2008). Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology,* 25(2), 321–359.

Saussure, F. de. (1916). *Cours de linguistique générale*. Lausanne & Paris: Payot.

Shao, B. (2020). *The apical vowel in Jixi-Hui Chinese: phonology and phonetics,* PhD dissertation. Université Sorbonne Nouvelle, Paris, France.

Shao, B., & Ridouane, R. (2018). La « voyelle apicale » en chinois de Jixi : caractéristiques acoustiques et comportement phonologique. *Proc. XXXIIe Journées d'Études sur la Parole*, 685–693.

Shao, B., & Ridouane, R. (2019). Apical vowel in Jixi-Hui Chinese: an articulatory study. *Proc. ICPhS2019*, Melbourne, Australia.

Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Trubetzkoy, N. S. (1969). *Principles of phonology*. Berkley and Los Angeles: University of California Press.

Zhao, R. (2007). Hànyǔ fāngyán zhōng de [i] > [ɿ] [Sound change of [i] > [ɿ] in Chinese dialects]. *Zhongguo yuwen*, 1, 46–54.

Zhu, X. (2004). Hànyǔ yuányīn de gāodǐngchūwèi [Sound changes of high vowels in Chinese dialects]. *Zhongguo yuwen*, 5, 440–451.