APICAL VOWEL IN JIXI-HUI CHINESE: AN ARTICULATORY STUDY

Bowei Shao, Rachid Ridouane

Laboratoire de Phonétique et Phonolgoie (CNRS & Sorbonne Nouvelle, Paris)
bowei.shao@sorbonne-nouvelle.fr; rachid.ridouane@sorbonne-nouvelle.fr

ABSTRACT

In this study, we analyze mid-sagittal and coronal ultrasound data from four subjects in order to investigate the articulatory properties of the apical vowel /z/ in Jixi-Hui Chinese (JHC). We seek to determine whether this unique segment, particularly well studied in Standard Chinese (SC), has articulatory characteristics of a vowel or a fricative consonant. Mid-sagittal ultrasound data show that it displays the same tongue shape as consonant [s] and may have a slightly higher tongue tip, which could explain the presence of abundant frication noise. The raising of the tongue dorsum, already observed in SC, varies in JHC across speakers. Coronal ultrasound data show that the tongue is medial-grooved and has an increased grooving corresponding to the decreased frication noise and the appearance of clear formant structure near its offset. These findings are discussed in light of the behavior of this segment within the phonological system of JHC.

Keywords: apical vowels, ultrasound, Jixi-Hui Chinese, syllabic consonant.

1. INTRODUCTION

Chinese languages are known to have a series of specific segments, often termed as "apical vowels" [12]. Some researchers showed that they could also be defined as "fricative vowels" [13], "syllabic fricatives/sibilants" [6, 19], or "syllabic approximants" [14, 15, 21].

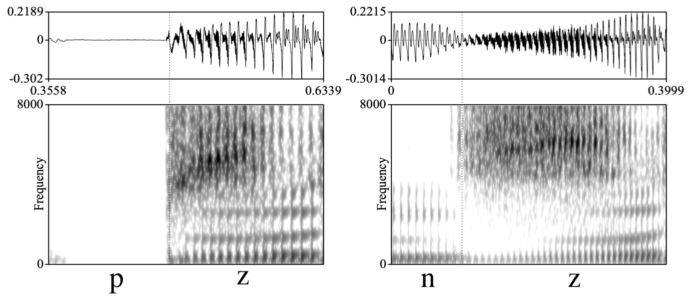
The apical vowels in Standard Chinese (SC) have been comprehensively studied. Phonologically, they are allophones to /i/, as they only occur following coronal sibilant onsets [s, ts, tsʰ, ʂ, tʂ, tʂʰ], while [i] occurs following palatal sibilant onsets [tɕ, tɕʰ, ɕ] and other consonants [3, 6]. Phonetically, the apical vowels of SC are homorganic to their sibilant onsets; the tongue tip/blade gesture remains nearly unchanged from the preceding sibilants to the voiced part of the syllable, and only a slight lowering of the tongue body is found [14]. Frication noise during apical vowels is observed [19] and considered the indicator of its sibilant property by some researches; however its presence is not systematic [8, 14].

Based on its homorganic character and its sibilant property, Dell [5] analyzed the SC apical vowel as a "voiced prolongation of the consonant", and Duanmu [6] proposed to analyze it not as an underlying phoneme, but as a segment triggered by an empty nucleus and the spreading of the feature [+fricative] from the onset. In our study, we adopt the [z] transcription following Duanmu and Dell.

While SC apical vowels are studied in much detail, less is known about these segments in other Chinese languages. This current study is based on Jixi-Hui Chinese (JHC), which is a Hui 徽 group Chinese language spoken in Jixi 绩溪 County in Anhui 安徽 province, China [10, 22, 23]. It has two major variants; our study is based on the city variant, which has 9 phonological monophthong vowels: [i, y, u, ʉ, o, ɤ, ɔ, a], in addition to the apical vowel [z].

Our previous study [17], based on acoustic data, shows that [z] in JHC has abundant frication noise on the first half of its duration, and displays a clear formant structure on the second half. This is illustrated in Fig. 1 for the items [pz^{214}] 'compare' and [nz^{214}] 'in'. In these two syllables, [p] and [n] being a labial stop and a sonorant respectively, have objectively no frication noise to propagate into the vocalic parts. The HNR pattern of [z] shows that the presence of frication noise is systematic, suggesting that this is an inherent attribute of [z] in JHC [17].

Figure 1: Examples of [z] containing frication noise on more than half of the duration, with [p] (left) and [n] onsets in syllables [pz^{214}] 'compare' and [nz^{214}] 'in' [17].



In this study, we follow up on the study of properties of [z] in JHC, through the analysis of its articulatory properties based on ultrasound data. The aim is to determine whether [z] has articulatory characteristics of a vowel or of a fricative consonant.

2. METHODS

2.1. Ultrasound data acquisition

Ultrasound data were acquired at Jixi city, in Anhui province, China. Three females and four males (age: 49±3.8) were recruited based upon the same criteria: they must be born and raised in Jixi city, by two parents born and raised in the same city, and they must speak JHC on a daily basis in both professional and family contexts. None of the speakers was reported to have speech or hearing diseases or difficulties. We had access to the sound-attenuated studio of the local television channel for our recording sessions. A word list was constructed with the following five segments [i, a, u, ʉ, z] occupying the rhyme of real monosyllabic words starting with /p_, p^h_, m_, n_, ts_, ts^h_, s_/ with different tonal structures. Each word was produced within the carrier phrase [$ki^{44}ɕɔ^{213}$_ $ɕɔ^{213}sɔ^{44}fa^{44}$] 'He writes _ three times'. The word list was repeated three times for mid-sagittal recordings and three times for coronal recordings, yielding 1218 total target syllables.

Both mid-sagittal and coronal ultrasound data were recorded with an *Ultrasound Stabilisation Headset* [1] and the *Articulate Assistant Advanced* software (AAA, V217.03) [2]. The ultrasound image was recorded with a field of view of 92°; due to speakers' morphological specificities, the depth was adjusted to have a maximum view of the tongue, causing different frame rates. The synchronization of ultrasound recordings and corresponding audio was done automatically by AAA software.

2.2. Ultrasound data analysis

Data from two female (FS1, FS2) and two male (MS1, MS2) speakers are presented in this study. Ultrasound data were segmented in AAA manually using the corresponding audio and traced manually with the help of the built-in tracing algorithm. The tongue contour of the nearest-central image of every [i, a, u, ʉ, z] token was extracted in XY coordinates; we also extracted the tongue contour of fricative [s]. The data obtained were then analyzed using smoothing-spline analysis of variance (SS ANOVA) in R environment with Package gss [4, 9]. The palatal traces were generated by averaging the palate contours extracted from six separate water-swallow tasks.

While studies based on mid-sagittal ultrasound data are rather frequent, coronal ultrasound data are rarely studied. One difficulty in coronal ultrasound data is that, in a field study, the coronal plane of the tongue can almost never be optimally positioned on the tongue [18]. Nevertheless, a comparison between coronal tongue shapes can still give a straightforward view of the possible differences between segments, namely between [s] and [z] which are the focus of this study.

3. RESULTS

3.1. Mid-sagittal ultrasound

We can see that the four speakers have two different mid-sagittal tongue contours. FS1 and MS2 have a raised tongue tip, but the tongue dorsum is also raised and higher than the tongue tip; FS2 and MS1 have the highest point of the tongue in the front, the tongue dorsum is flat, thus the tongue apex is higher than the tongue dorsum.

Mid-sagittal ultrasound images are generalized into SS ANOVA curves, presented in Fig. 2. The general observation is that [z] and [s] have the same tongue shape, but the tongue dorsum in [z] could be lower than in [s].

Based on SC apical vowel data reported by Zhou and Wu [24], Cheng [3] noted that in the production of the apical vowels, the highest point of the tongue is slightly more front, and the back of the tongue is slightly raised. He then suggested that "the SC apical vowel has two simultaneous points of articulation, one at the tongue tip and the other at the body of the tongue." (p.13). This observation on SC apical vowel is confirmed in pattern 1, but not observed in pattern 2, in which the tongue tip is indeed the highest point, but there is no raised tongue body. Interestingly, when [s] has a raised tongue body, [z] has the same gesture, as in FS1 and MS2. Similarly, if [s] does not have a raised tongue body, [z] does not either, as in FS2 and MS1. That is to say, apical vowel [z] and fricative consonant [s] display the same tongue shape in the sagittal plane for each speaker.

FS2 and MS1 seem to produce apical vowel [z] with a lifted tongue tip compared to [s]. This specific pattern could be related to one case reported by Faytak [8], where he observed a raised tongue tip in a coronal plane for one speaker; while all other speakers' tongue tips do not differ from [s] to [z]. Since the apical vowel [z] in JHC is reported to have more frication, we may consider this tongue tip constriction to be one of the factors related to the observed turbulent noise on acoustic signals.

3.2. Mid-sagittal tongue shape of [z] in different consonantal contexts

It is shown that the SC apical vowels are homorganic to their sibilant onsets [14]. Considering that in our word list of JHC, [z] is not only preceded by coronal sibilants, but also by [p, p^h, m, n], we compared the tongue contours of [z] in these different contexts to show that the resemblance between the tongue contours of [z] and [s] was not due to the influence of the preceding coronal consonants.

In Fig. 4, we can see a comparison between [z] segments preceded by labial consonants and by coronal consonants for the two speakers. Clear enough, there is no observable difference between [z] in the two contexts. This provides additional evidence that the resemblance of tongue contours between [z] and [s] is not due to the context where [z] occurs. Rather, the specific tongue shapes displayed by [z] seem to be an inherent property of this segment.

Figure 2: SS ANOVA splines of mid-sagittal ultrasound tongue contours, extracted in XY coordinates (mm) at the nearest center image of each segment, the tongue tip is on the right and tongue root on the left. Dotted lines represent [i, u, a, ʉ], dashed red lines represent [s] and solid green lines [z]; the thick grey lines represent the palatal traces; the grey areas represent 95% Bayesian confidence intervals.

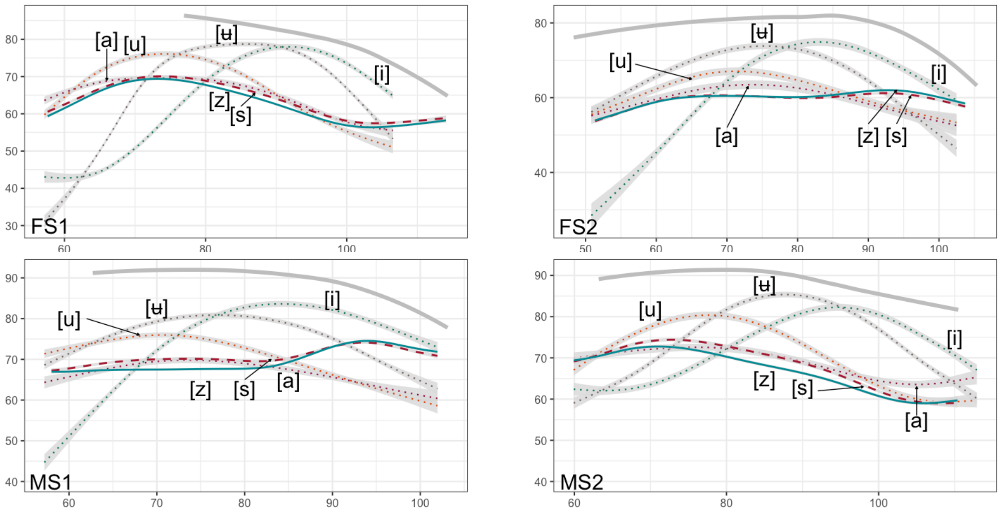
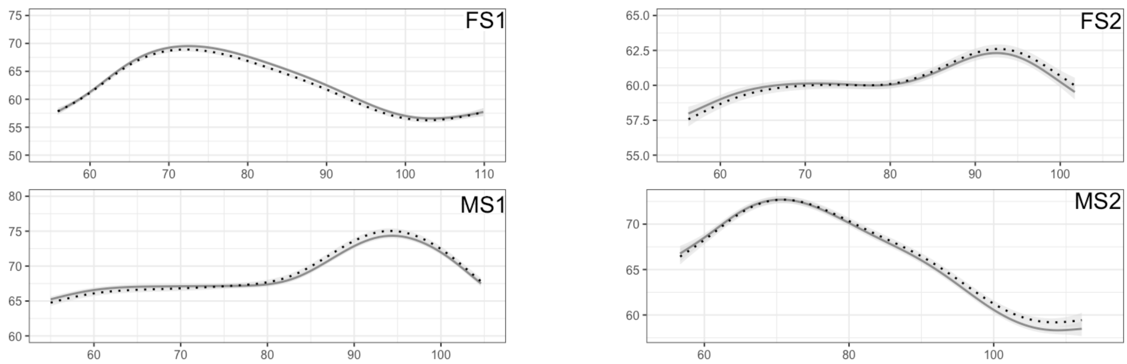


Figure 4: SS ANOVA splines (mm) of all [z] segments. Solid lines represent [z] preceded by labial consonants and dotted lines represent [z] preceded by coronal consonants. Tongue tip on the right and tongue root on the left.



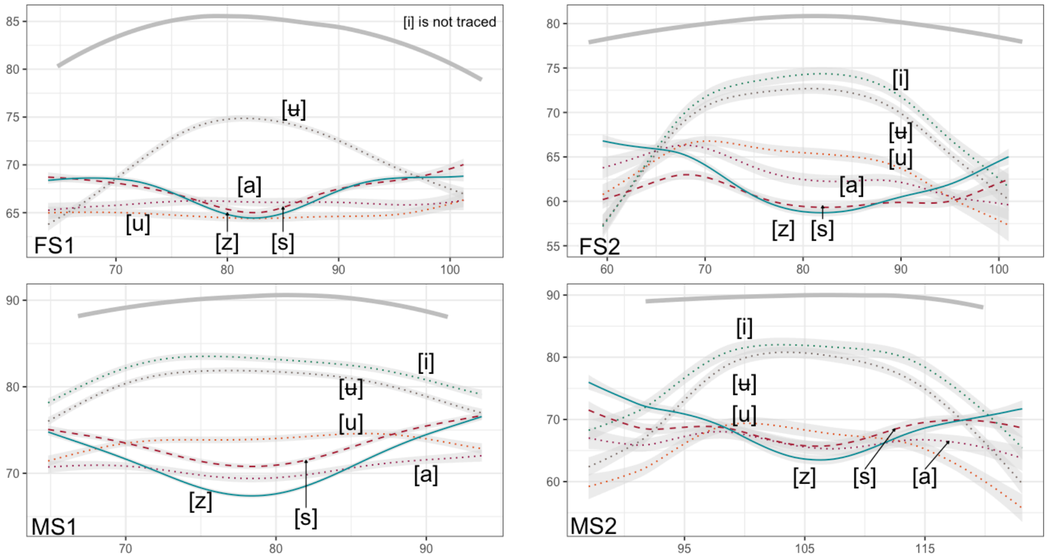
3.3. Coronal ultrasound

Coronal ultrasound tongue contours are generalized into SS ANOVA curves and shown in Fig. 3. The general pattern is also that [z] has the same tongue shape as [s]. Although we cannot determine where exactly the coronal contours of the tongue were obtained, by comparing the relationship between tongue contour lines in mid-sagittal view and coronal view, we can still see that our coronal view is not captured at the tongue tip, but rather at the frontal part of the tongue body.

The four speakers all have a medial-grooved tongue shape. For MS1 and MS2 [z] is more grooved than [s], which should correspond to the lower tongue body in our mid-sagittal contours. For FS1 and FS2, though the tongue contours are medial-grooved in both [s] and [z], we do not observe a more grooved [z] compared to [s]. This is also consistent with their mid-sagittal contours where the difference between [s] and [z] is rather small.

Compared to other vowels, this medial-grooved shape of the tongue in [z] segments can be considered as the indicator of its sibilant nature, as no vowel in JHC presented in Fig. 3 has this narrowed air passage. It is worth noting that this medial-grooved tongue shape for [z] is observed in all speakers, who have different articulatory strategies for [z] reported in mid-sagittal view.

Figure 3: SS ANOVA splines of coronal ultrasound tongue contours, extracted at the nearest center image of each segment. Line type as in Figure 2.



4. DISCUSSION AND CONCLUSION

The objective of this study was to determine whether [z] in JHC has articulatory characteristics of a vowel or of a fricative consonant. The resemblance of tongue contours between [s] and [z], together with previous results showing the presence of frication noise during [z] [17], provide evidence that this segment is best analyzed as a fricative consonant.

There are important articulatory differences between JHC [z] compared to SC [z] [8, 14]. The JHC [z] could be produced with or without a raised tongue body, and the tongue tip could be more raised as for [s]. There are also similarities between SC and JHC [z]: on the acoustic level, they display the same formant structure [17, 20]; on the articulatory level, the tongue body could be raised and medial-grooved while systematically resembling the articulatory configuration observed for [s]. The presence of frication noise in JHC [z] has also been observed in other Chinese languages [11, 16]. The absence of frication noise in SC [z] is likely an exception rather than a norm [7].

The fundamental difference between JHC [z] and SC [z] concerns their phonological status and phonotactic distribution. As already stated, SC [z] is an allophonic variant of the phoneme /i/ and is always homorganic to its sibilant onsets. In JHC it is a distinct phoneme from /i/, as both /z/ and /i/ occur after [ts, ts^{h}, s] onsets (*e.g.* [tsz^{31}] 'chicken' *vs.* [tsi^{31}] 'a family name'; [ts^{h}z^{31}] 'deception' *vs.* [ts^{h}i^{31}] 'autumn'; [sz^{31}] 'silk' *vs.* [si^{31}] 'repair'). Moreover, this apical vowel has a larger distribution than in SC, as it occurs not only after coronal sibilant onsets, but also after [p, p^{h}, m, n] onsets that have no sibilant properties to propagate. [z] in JHC is also a tone-bearing unit (TBU) and can undergo tonal sandhi processes (*e.g.* /sz^{31}/ 'west' + /ko^{31}/ 'melon' → [sz^{33}ko^{31}] 'watermelon'). Syllables containing apical vowel [z] are lexically developed, counting for 7,2% of the monosyllabic entries of JHC dictionary [23].

Studies on SC [z] tend to treat the homorganic property of [z] as being triggered by the sibilant onsets. We have shown that this analysis cannot apply to JHC [z]. Indeed, the fact that it can occur following both coronals and labials rules out the possibility of considering it as a mere prolongation of the onset consonant.

To sum up, both the articulatory and acoustic characteristics of /z/ support a consonant analysis. The fact that it can be a TBU nucleus does not rule out this analysis, since other consonants in JHC can also function as TBUs (*e.g.,* /m, n, v/). Its limited distribution – appearing only at syllable nucleus position – is most probably a consequence of its diachronic evolution from /i/ [25].

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

[1] Articulate Instruments Ltd. 2008. *Ultrasound Stabilisation Headset User's Manual: Revision 1.5*. Edinburgh, UK: Articulate Instruments Ltd.

[2] Articulate Instruments Ltd. 2012. *Articulate Assistant Advanced User Guide: Version 2.14*. Edinburgh, UK: Articulate Instruments Ltd.

[3] Cheng, C. C. 1973. *A synchronic phonology of Mandarin Chinese (Monographs on Linguistic Analysis 4)*. The Hague: Mouton.

[4] Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *J. Acoust. Soc. Am.* 120(1), 407–415.

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

[6] Duanmu, S. 2007. *The phonology of standard Chinese*. New York: Oxford University Press.

[7] Faytak, M. 2018. *Articulatory uniformity through articulatory reuse: insights from an ultrasound study of Sūzhōu Chinese*, PhD diss., UC Berkeley.

[8] Faytak, M., Lin, S. 2015. Articulatory variability and fricative noise in apical vowels. *Proc. 18th ICPhS*, Glasgow.

[9] Gu, C. 2014. Smoothing Spline ANOVA Models: R Package gss. *Journal of Statistical Software*, 58(5), 1–25. URL http://www.jstatsoft.org/v58/i05/.

[10] Hirata, S. 1998. *Huīzhōu Fāngyán Yánjiū* [Study on Huizhou dialects]. Tokyo: Kohbun Press.

[11] Hou, C. 2009. Héféi fāngyán dàicāyuányīn de shíyàn yánjiū [An experimental study on [ɿ] in Hefei Dialect]. *Journal of School of Chinese Language and Culture (Nanjing Normal University)* 1, 172–177.

[12] Karlgren, B. 1915. *Etudes sur la phonologie chinoise*. Uppsala: KW Appelberg.

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

[14] Lee-Kim, S. -I. 2014. Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *J. Int. Phon. Assoc.* 44(3), 261–282.

[15] Lee, W. S., Zee, E. 2003. Standard Chinese (Beijing). *J. Int. Phon. Assoc.* 33(1), 109–112.

[16] Ling, F. 2009. *A phonetic study of the vowel system in Suzhou Chinese.* PhD dissertation. City University of Hong Kong, Hongkong.

[17] 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.

[18] Stone, M. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical linguistics & phonetics*, 19(6/7), 455–501.

[19] Yu, A. 1999. Aerodynamic constraints on sound change: The case of syllabic sibilants. *J. Acoust. Soc. Am.* 105(2), 1096–1097.

[20] Zee, E., Lee, W. S. 2001. An acoustical analysis of the vowels in Beijing Mandarin. *Proc. EUROSPEECH-2001*, Aalborg, 643–646.

[21] Zee, E., Lee, W. S. 2007. Vowel typology in Chinese. *Proc. 16th ICPhS*, Saarbrücken, 1429–1432.

[22] Zhao, R. 1989. Ānhuī Jìxī fāngyán yīnxì tèdiǎn [Phonological characters of Jixi dialect from Anhui Province]. *Fangyan* 2, 125–130.

[23] Zhao, R. 2003. *Jìxī Fāngyán Cídiǎn [Dictionary of Jixi Dialect]*. Nanjing: Jiangsu jiaoyu chubanshe.

[24] Zhou, D., Wu, Z. 1963. *Pǔtōnghuà Fāyīn Túpǔ [Articulatory diagrams of Standard Chinese]*. Beijing: Shangwu yinshuguan.

[25] Zhu, X. 2004. Hànyǔ yuányīn de gāodǐngchūwèi [Sound changes of high vowels in Chinese dialects]. *Zhongguo Yuwen*, 302(5), 440–451.