A preliminary ultrasonic investigation of tenseness in Northern Yi

Shuwen Chen

Institute of Linguistics, Chinese Academy of Social Sciences
chensw@cass.org.cn

Abstract

This study investigates the articulatory and acoustic features of tense-lax vowel contrast in Northern Yi. Northern Yi has five pairs of phonologically-defined tense and lax vowels. Previous studies have shown that the difference between tense and lax vowels lies mainly in phonation type, while the vowel quality difference between the two types of vowels remains controversial. The current study provided tongue shape data and acoustic data of all tense and lax vowels in a Northern Yi dialect, Suondi Yi, using ultrasound imaging and acoustic analysis. The ultrasound data showed that tense vowels were consistently articulated with a more retracted tongue root and a lower tongue front compared to lax vowels. Acoustically, tense vowels have a higher F1 and a lower F2 compared to lax vowels. The data suggests that there are consistent vowel quality differences between all tense-lax vowel pairs in Northern Yi, and the vowel quality difference involves advanced tongue root/retracted tongue root ([ATR]/[RTR]) contrast.

Index Terms: Northern Yi, tenseness, ultrasound imaging

1. Introduction

The contrast between tense and lax vowels is a feature of many Tibeto-Burman languages spoken in Southwest China. One of the languages is Northern Yi (also called Nuosu language), a member of the Lolo-Burmese subgroup of the Tibero-Burman family [1]. It is spoken in southern Sichuan Province and northern Yunnan Province in China [2]. Previous studies have shown that tense and lax vowels in Northern Yi language differ in a number of phonetic parameters: phonation type, vowel quality, duration, intensity, f0, and the aspiration and voicing of the initial consonants [3, 4, 5, 6, 7]. Different from the tense-lax distinction in Germanic languages, the differences between tense and lax vowels in Northern Yi lie mainly in phonation type and laryngeal constriction [2, 3, 4, 5].

The most salient and well-investigated phonetic difference between tense and lax vowels is the phonation type. Tense vowels are produced with a creaky voice, and lax vowels are produced with a modal or breathy voice [3, 5]. Acoustically, the amplitude difference between F0 and H2 (F0-H2) is larger for tense vowels [7]. F0-H2 is a measure that differentiates various phonation types [8, 9, 10, 11]. A larger F0-H2 value in tense vowels indicates that the larynx is more restricted when producing those vowels. By measuring the airflow out of the mouth, Maddision & Ladegoged [3] suggested that tense vowels were produced with a greater glottal constriction because there was less airflow in lax vowels. Using electroglottograph (EGG), Kuang [4] showed that the main difference between tense and lax vowels in Yi is the contact quotient in the EGG signal, a measure that is defined as the duration of the vocal fold contact during each single vibratory cycle. Edmondson et al. [13] argued that the production of tense vowels involves laryngeal sphincter, with the larynx raising and epiglottic root backing [13]. Another phonetic cue that is related to phonation type is the f0. Some studies found higher f0 in tense vowels [3,14], while other studies found no differences in f0 [4, 15].

Tenseness in Yi is also realized with distinct patterns of duration and intensity. Tense vowels have shown to be shorter than lax vowels in falling tone in Liangshan Yi and low falling tone in Xinping Yi (Southern Yi), but no difference has been found in high and mid tones [3]. A similar effect was reported in intensity [7]. Tense vowels in Xinping Yi have lower intensity compared to lax vowels in the low tone but not in other tones [7]. The pattern in Wuding Yi is consistent with Liangshan and Xinping Yi [6]. Wang and Kong [6] showed that tenseness in Wuding Yi was realized with different phonetic correlates for high and low tones. For syllables with a high tone, the distinction mainly lies in phonation type, while for syllables with a low tone, the major difference is in duration. In addition, aspirated or voiced consonants before tensed vowels have prolonged aspiration and voicing [3, 14].

Another important difference between tense and lax vowels is vowel quality. Shi and Zhou [7] reported that tense vowels had a consistently larger F1 compared to their lax counterparts, suggesting that tense vowels were lower in tongue height. Similar phenomenon has been reported in Xinping and Jiangcheng Yi [4]. Kuang [4] reported that tense vowels had a higher F1 than the corresponding lax vowels, but no significant differences were found in the F2. Edmondson et al. [13] showed that tense vowels has a larger F1 and a smaller F2, suggesting the tense vowels were not only lower, but also backer. Besides the acoustic evidence, Lan and Wu [16] provided the first articulatory data using ultrasound imaging. They compared two pairs of tense and lax vowels [ɿ ɿ̠] and [ʅ ʅ̠] spoken in Xide County. Based on their visual inspection of the raw ultrasound images, they suggested that lax vowels [ɿ ʅ] are higher and more fronted than the tense vowels [ɿ̠ ʅ̠]. The limitation of their study is that their ultrasound data is based on only two vowel pairs, and the ultrasound data are not subject to any statistical analysis. Some other studies, however, did not find vowel quality differences in Yi. For example, Maddision and Ladegoged [3] reported that tense and lax vowels did not differ in vowel quality in Liangshan Yi. Wang and Kong [6] demonstrated that in Wuding Yi, only /ɒ/ and /ɑ/ demonstrated some quality difference, and no difference was observed in other vowel pairs.

While the tense-lax vowel differences in phonation type have been well investigated, the vowel quality differences remain controversial. To further clarify the vowel quality differences between tense and lax vowels in Northern Yi, the current study aims to examine the lingual gesture and acoustic characteristics of all tense and lax vowel pairs in a Northern Yi dialect, Suondi Yi, using ultrasound imaging. Compared to acoustic data, ultrasound images can provide more direct evidence of the vowel quality difference between tense and lax vowels.

2. Method

One male speaker of Suondi Yi was recorded with ultrasound imaging while reading Suondi Yi sentences. The speaker was a native speaker of Suondi Yi, and learnt Mandarin and English at school. Suondi Yi and Mandarin were his dominant languages, and he used Suondi Yi to communicate with his families. The speaker was in his forties at the time of recording and did not report any history of speech and hearing disorders.

2.1. Stimuli

The stimuli included ten vowel phonemes which grouped as five tense-lax pairs: /i/ (lax) and /ɛ/ (tense); /ɯ/ (lax) and /a/ (tense); /o/ (lax) and /ɔ/ (tense); /u/ (lax) and /u̱/ (tense), and two allophones of the apical vowels – [ɿ] (lax) and [ɿ̱] (tense) and [ʅ] (lax) and [ʅ̱] (tense) (Table 1). The 12 vowels were embedded in the /pV#/ and /p^{h}V#/ contexts. The bilabial stop /p/ and /p^{h}/ were chosen to be the initial consonants because bilabial sounds did not have any coarticulatory influences on the tongue. The lingual difference between tense and lax vowels can be better revealed without any influences from flanking consonants. The tense and lax apical vowels in /tsV#/, /ts^{h}V#/, /tʂV#/ and /tʂ^{h}V#/ contexts were also included to compare the difference between the apical vowel allophones (lax apical vowel [ɿ] and tense apical vowel [ɿ̱]) and the retroflex vowel allophones (lax retroflex vowel [ʅ], and tense retroflex vowel [ʅ̱]). The 28 words were read in the carrier phrase “ŋa33__ pi33” (“I read __”).

2.2. Procedure

The ultrasound video and the audio signal were recorded with Telemed Echo B system at the Chinese University of Hong Kong. The ultrasonic transducer was placed under the chin of the speaker and the mid-sagittal view of the tongue was obtained. The participant wore a light-weight headset made by Articulate Instruments Ltd. which can stabilize the ultrasonic transducer under the speaker’s chin. The ultrasound video was recorded at the frame rate of 60 fps with Articulate Assistant Advanced (AAA) software [17]. The participant was asked to read through the stimuli list before the experiment to make sure that he can read the words correctly and fluently. In the recording session, the stimuli were randomized and the participant was asked to read the prompt list for five times.

2.3. Data analysis

The ultrasound video and the audio signal were automatically synchronized by AAA software. The target vowels were first labeled according to the acoustic information in AAA. In the current study, each vowel contained 9-16 frames of images. Among the frames corresponding to the vowel, the middle frame was chosen for further analysis, assuming that the middle frame represented the target tongue shape of a monophthong.

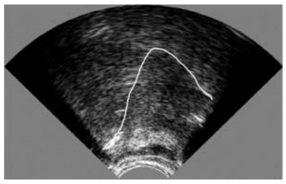
Table 1: *Stimuli used in the experiment.*

Ten vowels (5 tense and 5 lax) in /pV#/ and /p^{h}V#/ contexts

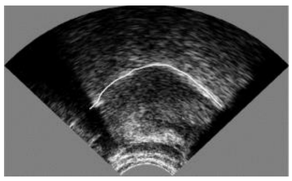
Vowels	Word	Gloss	Word	Gloss
i	pi^{33}	to read	p^{h}i^{33}	push aside; open
ɛ(i̱)	pɛ^{33}	to kick	p^{h}ɛ^{33}	carve
ɯ	pɯ^{33}	to exclaim	p^{h}ɯ^{33}	(no meaning)
a(ɯ̱)	pa^{33}	to exchange	p^{h}a^{33}	bladder
o	po^{33}	to rummage	p^{h}o^{33}	run away
ɔ(o̱)	pɔ^{33}	to split	p^{h}ɔ^{33}	servant
ɿ	pɿ^{33}	eagle call	p^{h}ɿ^{33}	painful
ɿ̱	pɿ̱^{33}	to poop	p^{h}ɿ̱^{33}	swing
u	pu^{33}	river deer	p^{h}u^{33}	price
u̱	pu̱^{33}	to go back	p^{h}u̱^{33}	translate; counterattack

Four apical vowels (2 tense and 2 lax) in /tsV#/, /tʂV#/, /ts^{h}V#/ and /tʂ^{h}V#/ contexts

Vowels	Word	Gloss	Word	Gloss
ɿ	tsɿ	hatch	ts^{h}ɿ	save (someone from the water)
ɿ̱	tsɿ̱	crack	ts^{h}ɿ̱	clench
ʅ	tʂʅ	urge (someone to go faster)	tʂ^{h}ʅ	becoming an orphan
ʅ̱	tʂʅ̱	pull out	tʂ^{h}ʅ̱	sprout, coming out

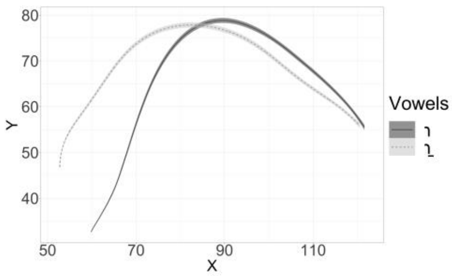


(1) A raw ultrasound image of the lax vowel /u/

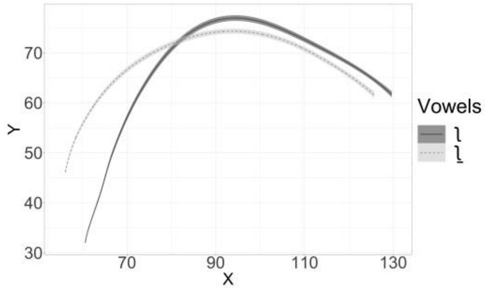


(2) A raw ultrasound image of the tense vowel /u̱/

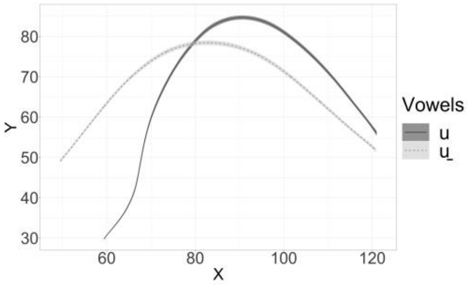
Figure 1: *Raw ultrasound images of lax vowel /u/ and tense vowel /u̱/ with the tracked tongue curves. The white lines in the ultrasound images represent the tongue-air interface and the superimposed tongue curves. The tongue tip is on the right side of the image, and the tongue root is on the left.*



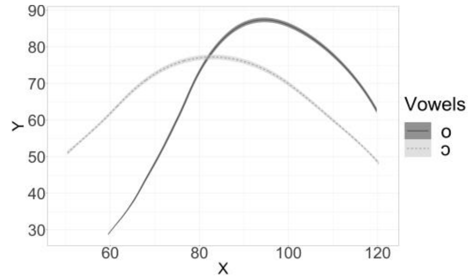
(1) Tongue shape differences between [ɿ] and [ɿ̱]



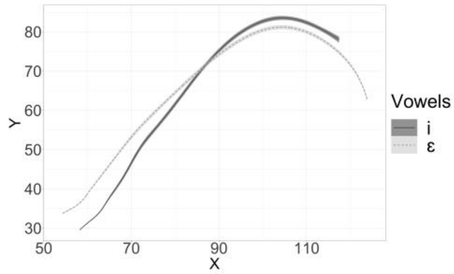
(2) Tongue shape differences between [ʅ] and [ʅ̱]



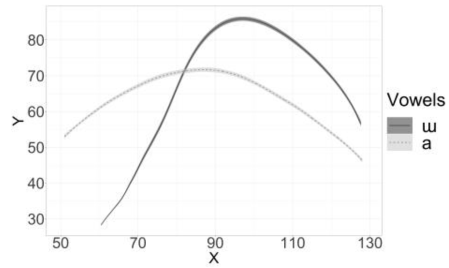
(3) Tongue shape differences between /u/ and /u̱/



(4) Tongue shape differences between /o/ and /ɔ/.



(5) Tongue shape differences between /i/ and /ɛ/.



(6) Tongue shape differences between /a/ and /ɯ/.

Figure 2: *The smoothing spline estimate and 95% Batesian confidence interval for each pair of tense and lax vowels. The darker lines represent the tongue curves of lax vowels, and the lighter lines represent the tongue curves of tense vowels. The x axis represents the tongue backness and the y axis represents the tongue height. The tongue tip is on the right side of the image.*

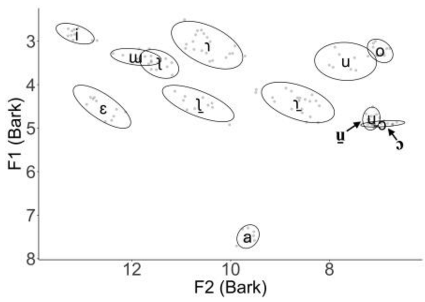
The tongue curves in the selected frames were manually tracked using AAA. The lighter line in the ultrasound images represents the tongue-air interface, so the tongue splines were drawn on the lower boundary of the light line (Figure 1). The tongue splines were exported from AAA as 124 data points for further data analysis. Smoothing spline ANOVA (SS-ANOVA) [17], contained in the R “gss” package, was used to compare tongue curves between tense ad lax vowels. SS-ANOVA is a statistical model that has been widely adopted in ultrasound studies to compare the differences between two or more curves. The formants were tracked using linear predictive coding (LPC) in Praat [19]. The F1 and F2 of each tense-lax vowel pair were compared using pair-wise t-tests. In total, there are 10 tokens for each vowel condition except for the vowel [ɿ] and [ɿ̱] which have 20 tokens in each condition.

3. Results

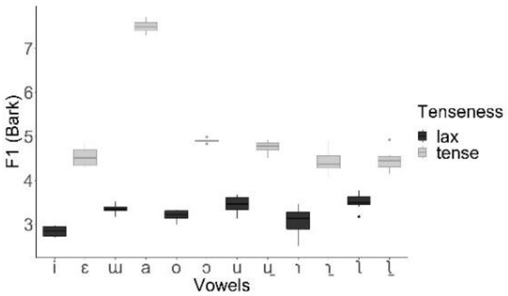
The tongue shape differences between each pair of tense and lax vowels modeled by SSANOVA were shown in Figure 2. In a SSANOVA figure, if the confidence intervals of two curves do not overlap, it means that the two curves are significantly different. As shown in Figure 2, for each pair of tense and lax vowels, the tongue back was significantly higher and the tongue root was more retracted in tense vowels. The anterior part of the tongue in tense vowels was lower. For high front vowels ([ɿ] and [ɿ̱], [ʅ] and [ʅ̱], /i/ and /ɛ/), the difference in the anterior part of the tongue was small, but still statistically significant.

In general, the position of the tongue root is the cause of the major differences between the tense and lax vowels, with the tongue root in tense vowels being substantially more retracted than that in lax vowels, resulting in a lower tongue dorsum. This pattern is consistent for all vowel pairs.

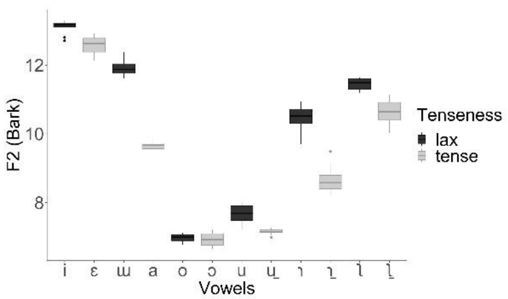
Figure 3 shows the acoustic vowel space of Suondi vowels, and the F1 and F2 of tense-lax vowel pairs. In the vowel chart, the five tense vowels were lower and to the right of their corresponding lax vowels, showing that tense vowels have a higher F1 and a lower F2 compared to lax vowels. This pattern was consistent for all vowel pairs except /o/ and /ɔ/. For the vowel pair /o ɔ/, the lax vowel /o/ has a significantly higher F1 than the tense vowel /ɔ/ (t (11.63) = -44.384, p < 0.001), but no significant difference was found in F2 (t = 0.64401, df = 14.739, p-value = 0.5295). For other pairs of tense and lax vowels, the F1 of tense vowels was significantly higher than that of lax vowels ([ɿ]-[ɿ̱]: t (37.993) = -16.919, p < 0.001; [ʅ]-[ʅ̱]: t(16.239) = -10.829, p<0.001; /u/-/u/: t (12.604) = -17.226, p < 0.001; /i/-/e/: t (12.47) = 20.882, p < 0.001; /ɯ/-/a/:t (13.145) = 72.736, p < 0.001). The F2 of tense vowels was significantly lower than lax vowels except the /o/-/ɔ/ pair ([ɿ]-[ɿ̱]: t (37.542) = 15.28, p < 0.001; [ʅ]-[ʅ̱]: t(14.633) =6.7217 , p< 0.001; /u/-/u̱/: t (7.9103) = 5.1434, p = 0.001; /i/-/e/: t (17.261) = -5.203, p<0.001; /ɯ/-/a/:t (11.141) = -24.229, p<0.001).



(1) The vowels of Suondi Yi.



(2) The F1 of Suondi Yi vowels.



(3) The F2 of Suondi Yi vowels.

Figure 3: *Vowels in Suondi Yi. The upper panel demonstrates the vowel space of Suondi Yi. The middle and lower panel show the F1 and F2 of all tense and lax vowels in Suondi Yi.*

4. Discussion

The ultrasound data in the current study demonstrated that tense vowels are articulated with a more retracted tongue root and lower tongue front compared to the corresponding lax vowels. This is consistent with previous acoustic studies which reported a higher F1 in tense vowels [4, 7, 13, 16]. The ultrasound data complement previous studies by providing direct evidence of the tongue shape differences between tense and lax vowels.

Previous studies using laryngoscopy have shown that Yi tense vowels were articulated with a raised larynx and constricted epiglottic root [13]. Our ultrasound images further showed that the articulation of tense vowels also involves a retracted tongue root. Taken together, to articulate tense vowels, the epilaryngeal tube is constricted and the larynx is raised, while the tongue is retracted, with the posterior part of the tongue raised and the anterior part lowered. Therefore, when articulating the tense vowels, there is a great constriction and tension in the tongue root and the throat.

Our articulatory data also showed that tenseness distinction in Northern Yi is related to advanced tongue root [ATR] and retracted tongue root [RTR] contrast. The ART and RTR contrast is often seen in many West African languages, such as Igbo and Akan [20]. It has been shown that the articulation of ART usually involves the widening of pharynx walls and the lowering of the larynx, whereas the production of RTR vowels involves pharyngeal constrictions and larynx raising [21, 22]. This pattern is similar to the articulation of tense and lax vowels in Yi. In Northern Yi, the tense vowels have a more retracted tongue root [+RTR] and a widened pharynx, while the lax vowel is characterized by the advanced tongue root [+ATR] and a constricted pharynx. According to Ladefoged and Maddieson [19], the major difference between ATR/RTR contrast in African languages and tense-lax distinction in Germanic languages is that tongue root position is correlated with tongue height in the tense-lax contrast, while the tongue root position in ATR/RTR contrast is independent of tongue height. For example, the difference between Akan ATR/RTR vowels in tongue height is rather small compared to the difference in tongue root and pharyngeal region. In Northern Yi, the pattern is mixed. For vowel pairs /u u̱/, /ɯ a/, and /o ɔ/, the tongue height difference between tense and lax vowels is correlated with tongue root position, which is similar to tense-lax contrast in Germanic languages. For high front vowels /i e/ and apical vowels [ɿ ɿ̱ ʅ ʅ̱], the tongue height difference seems to be less related to tongue root position. The tongue height difference between tense and lax vowels is much smaller than the difference in tongue root, which resembles the pattern of ATR/RTR differences in African languages.

Our acoustic data also confirms that tense and lax vowels have distinct vowel qualities. Previous studies had controversial results regarding the acoustic correlates of tenseness. Edmondson et al. [13] and Shi and Zhou [7] suggested that there were differences in F1 and F2 between tense and lax vowels. On the contrary, Maddison and Ladefoged [3] found no difference in vowel quality, while Wang and Kong [6] found formant differences only for /a/ and /ɒ/ vowel pairs. Our acoustic result is consistent with Edmondson et al. [13] and Shi and Zhou [7]. We found consistent vowel quality differences between tense and lax vowels in both F1 and F2. According to the Perturbation Theory [23], higher F1 represents lower tongue position, and lower F2 indicates that the tongue is backer. This is in accordance with our articulatory data where a lower and more retracted tongue position is found for tense vowels.

5. Conclusion

In summary, this study provides preliminary ultrasonic and acoustic data on tense and lax vowels in Northern Yi. The articulatory data showed that tense vowels are articulated with retracted tongue root and lower tongue front, while lax vowels are articulated with advanced tongue root and higher tongue front. Acoustically, tense vowels have higher F1 and lower F2 than lax vowels. The data in the current study suggests that tenseness in Yi is not only related to phonation types or laryngeal constriction, but also involves vowel quality differences that are similar to retracted or advanced tongue root ([ATR]/[RTR]) contrast. As a preliminary study, the data in the current study is from only one speaker, but it provides direct data revealing consistent vowel quality differences in all tense-lax vowel pairs. More speakers will be needed to confirm the findings from the current study.

6. Acknowledgement

Thank Prof. Lama Ziwo from Southwest Minzu University for his help in data collection and his insightful comments on this paper.

5. References

[1] B. David, *Proto-Loloish (Scandinavian Institute of Asian Studies, Monograph Series No. 39)*. Curzon Press, London, 1979.

[2] J. Edmondson, J. Esling, and L. Ziwo, "Nuosu Yi," *Journal of the International Phonetic Association*, vol. 47, no.1, pp. 87-97, 2017.

[3] I. Maddieson and P. Ladefoged, " 'Tens' and 'lax' in four minority languages of China," *Journal of Phoneitcs*, no. 13, pp. 433-454, 1985.

[4] J. Kuang, "Phonation contrast in two register contrast languages and its influence on vowel and tone," *Proceedings of the 17th International Congress of Phonetics Sciences*, pp. 1146-1149, 2011.

[5] J. Kong. "An acoustic study of tense and lax vowels in Liangshan Yi," Studies of Yi-Burman Languages, Sichuan Renmin Press, 1997.

[6] F. Wang and J. Kong, "A study on tenseness in Wuding Yi," *China Lingistics*, pp. 98-118, 2009.

[7] F. Shi and D. Zhou, "The acoustic realization of tense and lax vowels in Southern Yi," , *Studies of Language and Linguistics*, 1, 2005.

[8] C. Bickley, "Acoustic analysis and perception of breathy vowels," *Speech Communication Group Working Papers*, no.1, pp.71-81. 1982.

[9] P.L. Kirk, P. Ladefoged, and J. Ladefoged, "Using a spectrograph for measures of phonation types in natural language," *UCLA Working Papers in Phonetics*, no. 59, pp. 102-113, 1984.

[10] P. Ladefoged, "The relative nature of voice quality," *The Journal of the Acoustical Society of America*, 69(S1), S67–S67, 1981.

[11] P. Ladefoged, "The linguistic use of different phonation types," In D. Bless & J. Abbs (Eds.), *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, pp. 351–360. San Diego: College Hill Press, 1983.

[12] P. Ladefoged and N. Antoñanzas-Barroso, "Computer measures of breathy voice quality," *UCLA Working Papers in Phonetics*, 61(S1), pp. 79–86, 1985.

[13] J.A. Edmondson, Z. Lama, J. Esling, J.G. Harris, and S. Li, "The aryepiglottic folds and voice quality in the Yi and Bai languages: Laryngoscopic case studies," *Monkhmer Studies*, pp. 83-100, 2001.

[14] I. Maddieson and S. Hess. " 'Tense' and 'lax' revisited: more on phonation type and pitch in minority languages in China," *UCLA Working Papers in Phonetics*, no. 63, pp. 103-109, 1986.

[15] Z. Ye, "A study on the tenseness in Xinpin Yi," *Peking University Speech Prosody Research Report*, 2006.

[16] Z. Lan and X, Wu, "Physiological mechanisms of tenseness in Yi vowels, " *Minzu Yuwen*, no. 4, pp. 24-34, 2017.

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

[18] C. Gu, "Smoothing spline ANOVA models: R package gss," *Journal of Statistical Software*, vol. 58, no. 5, pp. 1-25, 2014.

[19] P. Boerma, Praat, a system for doing phonetics by computer. *Glot International*, no. 5, 2002.

[20] P. Ladefoged and I. Maddieson, *The sounds of the world's languages.* Oxford : Blackwell, 1995

[21] M. Lindau. The feature expanded. *Journal of Phonetics*, vol. 7, no. 2, pp. 163-176, 1979.

[22] M. Tiede, An MRI-based study of pharyngeal volume contrasts in Akan and English. *Journal of Phonetics*, vol. 24, no. 4, pp. 399 -421, 1996.

[23] K.N. Stevens, *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1998.

[This paper was published at ISCSLP 2022]