A STUDY ON TONES IN JIUHE BAI PRODUCED BY NAXI SPEAKERS

Meihao Wan, Feng Wang

The Chinese University of Hong Kong, Peking University
wanmeihao@link.cuhk.edu.hk, wfwf@pku.edu.cn

ABSTRACT

Bai is a language spoken in southwestern China which uses both pitch and non-modal phonation for tone contrast. In Jiuhe, there is a group of multilingual speakers of both Bai and Naxi languages. This paper aims to investigate the tone of Bai produced by local Naxi speakers with electroglottograph. The results show that the number of tones is reduced and some tones have merged in Naxi speakers' production. The high level tone is the least affected whereas the low tense tone is prone to merge with other tones. Some Naxi speakers also use non-modal phonation for tone contrast of Bai as native Bai speakers do, even though there is no phonation contrast in the tones of Naxi. Furthermore, this case implies that tone merger is not only due to the similarity of F0 contour, phonation can also be the condition in tone merge.

Keywords: Bai, tone, Naxi, phonation

1. INTRODUCTION

Bai is a tone language of Tibeto-Burman languages and also a register language, in which both F0 and voice quality function in tone contrast. Previous studies have shown that the tense tones in Bai are associated with non-modal phonation [1], but not accompanied with specific phonation types. The non-modal phonation types in the tense tones vary across different speakers and different dialects [2], [3]. For example, in Jianchuan Bai, one speaker may use harsh voice in the tense tone T7 while another speaker may use pressed voice. However, in Meiba Bai, the phonation of the same tone can be creaky voice. Previous research focused on the production of native Bai speakers and rarely on the Bai language produced by native speakers of other languages.

The Jiuhe township is located at the border of the Bai speaking region and the Naxi speaking region, where Bai and Naxi people live together in the villages. Because of their close contact, most people in the village are multilingual and can speak at least three languages: Bai, Naxi and Mandarin Chinese. Bai and Naxi are used in most situations in daily life. Jiuhe Bai has six lexical tones (see Table 1), which can be described with Chao's five-point scale [4]. Similar to other Bai languages, there is a tense vs. lax register in Jiuhe Bai: T1, T3, and T5 are tense tones, and T2, T4, and T6 are lax tones. In Jiuhe Bai, tense tones are not associated with specific phonation as well. For example, the phonation cue for T1 can be realized as creaky, breathy or harsh when produced by different speakers.

Register	Lexical tones		
Tense	T1 low-falling [tɕi21] 'flag'	T3 mid-falling [tɕi31] 'chase'	T5 mid-level [tɕi44] 'leech'
Lax	T2 low-falling [tɕi21] 'ground'	T4 mid-level [tɕi33] 'pull'	T6 high-level [tɕi55] 'many'

Table 1. Tones of Jiuhe Bai with examples

The phenomenon of tone merger has been studied for a long while. In previous studies such as Cantonese, another language with a complex tone system, some speakers merged several tone pairs which are similar in F0 contours [5]. Whether the similarity of F0 will lead to the merging of tones in Jiuhe Bai is an important issue of concern. The investigation of Naxi Bai can shed light on the study of mechanism of tone merger.

Moreover, unlike other well-studied tone languages, Bai is one of the register languages that use phonation for tone contrast. Another question that this study hopes to answer is whether Naxi speakers can use phonation for tone contrast. Jiuhe Bai is a good case to observe phonation as a contrastive cue in tone production of non-native Bai speakers.

2. METHOD

2.1. Data sampling

The data in this paper were all collected during a field trip in Jiuhe Township in August 2021. Four native Naxi speakers participated in the production experiment, including two males (one aged 45, one aged 46) and two females (one aged 72, one aged 73), in addition to two native Bai speakers (one aged 34, one aged 51) for control involving one male and one female. All of them were farmers who have lived in the village since birth and could speak Bai proficiently.

The recording software was Adobe Audition CC 2018, and the recording was taken in dual-channel recording, with the audio signal recorded in the left channel by a condenser microphone (SONY ECM-44B), and the EGG signal recorded in the right channel by the Electroglottograph Model 7050A, and the two channels of signals were recorded through a mixer (XENYX 302 USB) and the sound card (SBX) into the computer (Huawei matebook 14 2019 model). The sampling frequency was 22050 Hz. The format of the audio file was WAV. The recording was performed in a quiet room in the villagers' house.

2.2. Bai materials

Since there is no data of Jiuhe Bai before, a corpus of Jiuhe Bai's vocabulary was created, in which eleven tonal monosyllabic minimal pairs were collected for the recording. All of the monosyllables were relatively balanced in frequency and were used frequently in daily life that Naxi speakers could respond to and pronounce instantly. For each speaker, these eleven minimal sets which contained 66 monosyllables (11 minimal sets × 6 tones) were pronounced twice. The minimal sets were identical in segmental structure but differed in tones as presented in Table 1.

2.3. Measures

In this study, three parameters were extracted from the acoustic signals and EGG signals: Fundamental Frequency (F0), Open Quotient (OQ) and Speed Quotient (SQ). F0 was extracted from the acoustic signals. OQ and SQ are extracted from EGG signals. F0 is defined as 1/period, OQ is defined as open phase/period, SQ is defined as opening phase/closing phase. These three parameters were extracted by the Matlab program 'VoiceLab', written by the Voice Laboratory of the Department of Chinese Language and Literature, Peking University.

The parameters extracted in this study were time normalized, that is, the F0, OQ, and SQ of all samples were composed of 20 points of data respectively. After the time-normalization, the first two data points and the last two points of each syllable were deleted, in that only the signals in the middle part are stable and reflect the phonetic features of the syllables. Ultimately, for each parameter, 16 stable data points were analyzed. Besides, in order to exclude the pitch range difference caused by the individual differences, we converted the F0 values into z-score for better comparison.[6]

3. RESULT

3.1. F0 contour

Figure 1 displays the F0 contours of the six lexical tones produced by native Bai speakers, which are well distinguished by pitch, in which T1 and T2 are low falling tones, T3 is a mid falling tone, T4 and T5 are mid level tones and T6 is a high level tone. The F0 contours of T1 and T2 are close but do not overlap, so are the F0 contours of T4 and T5. The F0 contour of T3 decreases more sharply with higher pitch compared to the other falling tones. The high level tone T6 is distinct from the other tones.

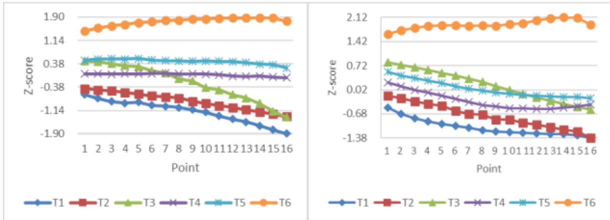

Figure 1: F0 contours of native Bai speakers (F1-B left, M1-B right)

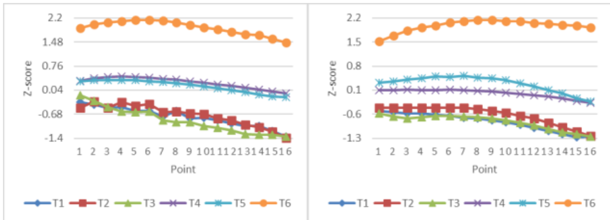

Figure 2: F0 contours of Naxi speakers (F1-N left, M1-N right)

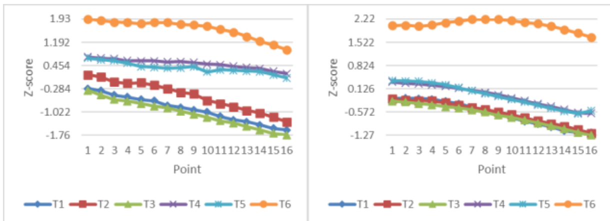

Figure 3: F0 contours of Naxi speakers (F2-N left, M2-N right)

However, some tones may merge in the production of Naxi speakers. The most common phenomena are the merging of T1 and T3 as well as the merging of T4 and T5 (see Figure 2 and Figure 3). There are also mergers of T1, T2, and T3. The affected degree varies among different tones. T6 is the least affected tone and the most affected tone is T1.

3.2. Phonation analysis

In Bai, phonation is also an important cue for tone contrast. However, there is no phonation contrast in the tones of Naxi. The phonation type can be identified by the OQ and SQ values [3], [7]. Table 2 shows the relationship between the vocalization type and the parameters based on previous studies. "+" is marked if the value of tense tone is significantly greater than that of lax tone, and "-" is marked if the value of tense tone is significantly smaller. Based on the comparison of these parameters, we can identify the phonation type of the tense tone.

	Vocal fry	Breathy	Creaky	High-pitch	Harsh
F0	-	-	-	+	-
OQ	+	+	-	-	-
SQ	+	-	+	-	-

Table 2: Phonation types and their parameters

The figures below display the phonation parameters of Bai tones produced by Naxi speakers. Since the F0 contours of T1 vs. T3, T4 vs. T5 overlap in their production, we will compare the phonation by the merged tone pair. In addition, as both T1 and T3 are tense tones with non-modal phonation in Bai, we also display the parameter of T2 which is a lax tone with modal voice for comparison. All interpretations of the OQ and SQ values are summarized in Table 3.

3.2.1 T1, T2 vs. T3

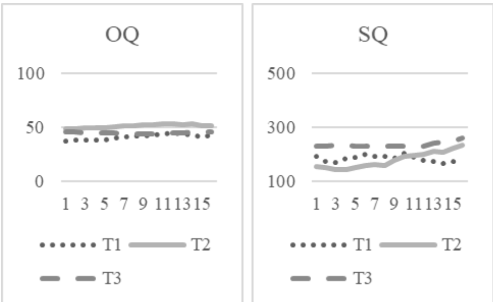

Figure 4: OQ and SQ contours of a Bai speaker (M1-B)

The OQ contours of Bai M1's T1, T2 and T3 are all relatively flat overall, with T2 slightly above T1 and T3 and the values are around 50%. The SQ contours of the three tones are also all relatively smooth, with T3 having the highest SQ at around 230%, T1 and T2 having SQ values between 150% and 230%.

Figure 5 displays the OQ and SQ contours of Naxi speaker F1. The OQ contours are flat, and the contours of the three tones are generally close to each other, with values around 50%. The SQ contours of F1 largely overlap and all show an upward trend, with values rising from about 200% to about 350%.

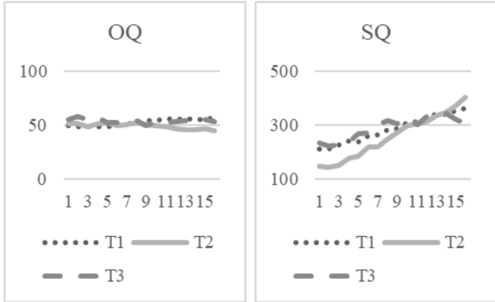

Figure 5: OQ and SQ contours of a Naxi speaker (F1-N)

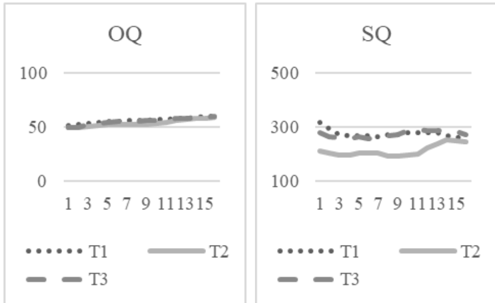

Figure 6: OQ and SQ contours of a Naxi speaker (F2-N)

The OQ and SQ contours of Naxi speaker F2 are displayed in Figure 6. The OQ contours show a slight upward trend, and the three contours are generally close, with values between 50% and 60%. In the SQ contours of F2, T1 and T3 largely overlap, with values between 250% and 300%; T2 is lower compared to T1 and T3, with lower values in the first two-thirds and slowly rising in the second third, close to the other two tones.

3.2.2 T4 vs. T5

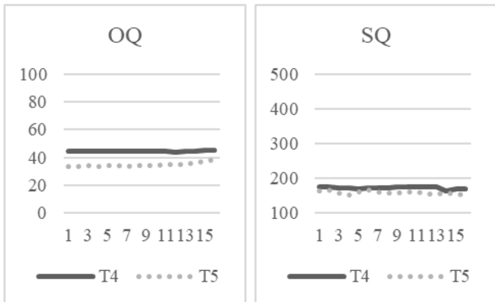

Figure 7: OQ and SQ contours of a Bai speaker (F1-B)

The OQ contour of T4 for Bai speaker F1 is relatively smooth, with values around 45%; the T5 OQ contour is also generally smooth and lower than that of T4 with values roughly around 34%. The SQ contours of the two tones are both relatively flat and largely overlap, with values around 165%.

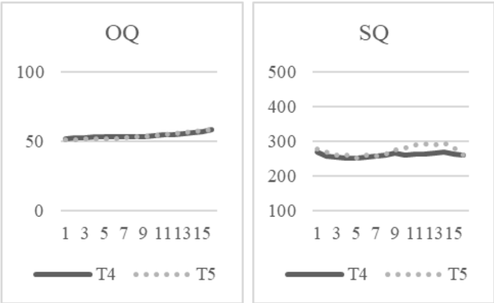

Figure 8: OQ and SQ contours of a Naxi speaker (F2-N)

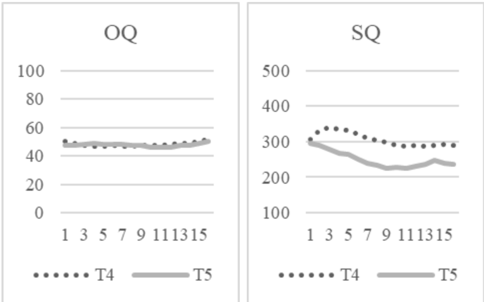

Figure 9: OQ and SQ contours of a Naxi speaker (M1-N)

The OQ contours of T4 and T5 for speaker F2 (see Figure 8) largely overlap, with an overall slight upward trend, with values ranging from about 50% to 60%. The SQ contours of T4 and T5 overlap roughly in the first half, and the curve of T5 is a little higher than that of T4 in the second half, reaching a maximum of about 290%, while T4 is around 265%.

The OQ contours of the two tones for M1 (see Figure 9) are smooth and largely overlap as well, with the value of around 50%. The SQ contours of M1's T4 is also smooth, with the SQ of T4 slightly larger than that of T5. The value of T4 is around 300%, and T5's SQ curve falls slightly, with the value dropping from 300% to around 250%.

To examine whether the differences are significant, independent sample *t*-tests are applied to the parameters point to point between the tone pairs. The results shows that there is almost no significant difference between the four Naxi speakers for each parameter of T1 vs. T3. Only the OQ values of M1's T1 and T3 have a significant difference. The results of all the speakers, including Bai and Naxi native speakers, are present in Table 3. "ns" is marked if there is no significant difference between the parameters of the two tones. The meaning of other marks is the same as in Table 2.

It can be seen from Table 3 that Bai M1 can distinguish the tone pairs by both F0 and phonation. Bai F1 can distinguish T4 and T5 with F0 and phonation but not T1 and T2, which may be due to her younger age.

Speakers	T1 vs. T2			T4 vs. T5		
	F0	OQ	SQ	F0	OQ	SQ
F1-B	-	ns	ns	+	-	ns
M1-B	-	-	ns	+	-	+
F1-N	ns	+	ns	ns	ns	ns
M1-N	-	-	-	+	-	-
F2-N	-	+	+	-	ns	+
M2-N	ns	ns	ns	ns	ns	ns

Table 3: Summary of parameter contrasts of the Bai tense vs. lax tones

For Naxi speakers, the tone merging of F1 and M2 is more radical compared to the other two speakers. They have poor performance in distinguishing the tone pairs in F0 and phonation. For the other two speakers, T2 is not only different from T1 and T3 in terms of F0, but also in terms of OQ and SQ. Based on the comparison between the parameters and Table 2, T1 of M1 is a typical harsh voice and T1 of F2 is a typical vocal fry voice. It can be seen that some Naxi speakers are also able to use phonation to distinguish different tones.

4. CONCLUSION

The results show that some Bai tone pairs have merged into one tone among the Naxi speakers. Based on the acoustic and EGG data, we can confirm that T1 vs. T3, T4 vs. T5 have merged in Naxi speakers' production, and some of the Naxi speakers have already merged T1, T2 and T3. The least affected tone is T6, which can be distinguished from other tones in all the Naxi speakers.

This case shows a different pattern of tone merging from previous studies. T4 and T5 are similar in F0 contours, which is consistent with the tone merging cases before. However, T1 and T3 are not similar in F0 contours in Jiuhe Bai. The similarity between the two tones is that they are both tense tones with non-modal phonation, which implies that tone merger is not only due to the similarity of F0 contour, phonation can also be the condition in tone merge. This also suggests that the phonation may affect the perception of tones, making them converge even if the F0 contours are not similar.

The investigation also reveals that there is a phonation contrast in the Bai tones produced by some Naxi speakers, which shows that phonation can be acquired to some extent in language contact as Jiuhe Naxi does not use phonation to distinguish different tones. Also, this may contribute to their failure in distinguishing Bai tones by phonation types.

5. ACKNOWLEDGEMENTS

This research was funded by Humanities and Social Sciences Foundation of the Ministry of Education (Grant 19JJD740001) and Major Project of Beijing Social Science Foundation (Grant 20ZDA20). We also thank Prof. Peggy Mok for helpful suggestions.

6. REFERENCES

[1] J. A. Edmondson and J. H. Esling, “The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies,” Phonology, vol. 23, no. 2, pp. 157–191, Aug. 2006, doi: 10.1017/S095267570600087X.

[2] X. Li and F. Wang, “Phonation variation and strategy of tone distinction: The case of Meiba Bai,” J. Chin. Linguist., vol. 48, no. 2, pp. 379–401, 2020, doi: 10.1353/jcl.2020.0009.

[3] 汪 鋒 (Wang, Feng), “VARIATIONS OF LARYNGEAL FEATURES IN JIANCHUAN BAI/劍川白語的嗓音變異,” J. Chin. Linguist., pp. 434–452, 2015.

[4] Y.-R. Chao, “A system of tone letters,” Maître Phon., 1930.

[5] P. P. K. Mok, D. Zuo, and P. W. Y. Wong, “Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese,” Lang. Var. Change, vol. 25, no. 3, pp. 341–370, Oct. 2013, doi: 10.1017/S0954394513000161.

[6] P. Rose, “Considerations in the normalisation of the fundamental frequency of linguistic tone,” Speech Commun., vol. 6, no. 4, pp. 343–352, Dec. 1987, doi: 10.1016/0167-6393(87)90009-4.

[7] 孔江平 (Kong, Jiangping), 论语言发声 (On Language Phonation). 中央民族大学出版社, 2001.