Consonant and lexical tone interaction: Evidence from two Chinese dialects

Shi, M.

Author: Shi, M
Title: Consonant and lexical tone interaction: Evidence from two Chinese dialects
Issue Date: 2020-06-04

Chapter 4 The sound system of Shuangfeng Xiang Chinese

4.1 Introduction

Shuangfeng Xiang Chinese (双峰方言) is a Xiang dialect (湘语, ISO 639-3; code: hsn) spoken in the town of Yongfeng (永丰镇). According to the census in 2000, there are approximately 88,000 residents in Yongfeng, which is the administrative town of Shuangfeng county (双峰县) (indicated by the solid circle in Figure 4.1). Shuangfeng county belongs to the prefectural-level municipality of Loudi city (娄底市) in Hunan province (湖南省), the People's Republic of China. It is located at the hinterland of Hunan province and is about 165 kilometers from to Changsha city (长沙市) (indicated by the hollow circle in Figure 4.1), the provincial capital of Hunan, about 1,500 kilometers to Beijing, the capital of China.

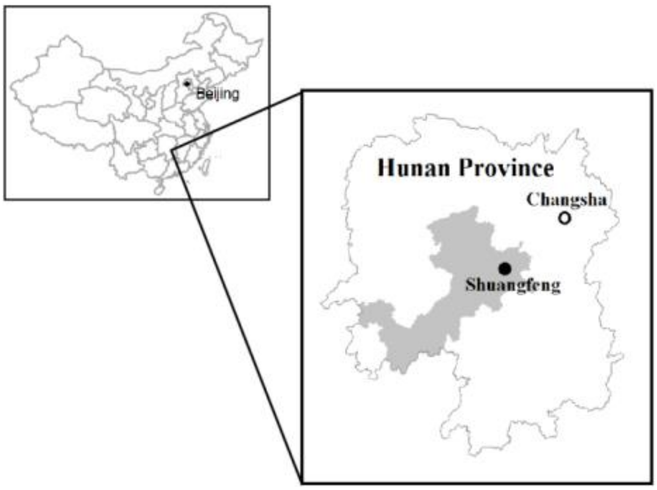

Figure 4.1 Location of Shuangfeng county in Hunan province.

Shuangfeng Xiang Chinese is commonly considered part of the Xiangshuang cluster (湘双小片), which in turn is classified as a member of the Loushao subgroup (娄邵片) of the Xiang dialect group, a Sinitic branch (汉语支) within the Sino-Tibetan family (Bao & Chen, 2005). The Xiang dialect group is conventionally believed to be one of the seven major dialect groups in China (Yuan, 1960). As shown by the dark gray area in Figure 4.1, the Loushao subgroup dialect is also known as the Old Xiang Chinese (老湘语) in the majority of sinological studies, which is further differentiated from the Changsha dialect, known as one of the New Xiang Chinese (新湘语) (after Yuan, 1960: 102).[27] What distinguishes these two varieties is attributed mainly to their obstruent voicing contrasts. The Old Xiang group shows a three-way contrast (i.e., voiceless unaspirated vs. voiceless aspirated vs. voiced) while the New Xiang group has only a two-way obstruent contrast (i.e., voiceless unaspirated vs. voiceless aspirated) (Yuan, 1960; Yang, 1974; Tsuji, 1979; Bao & Yan, 1986; Wurm et al., 1987; Bao & Chen, 2005).

Shuangfeng Xiang Chinese has attracted much attention over the last six decades, which has resulted in a handful of descriptive works on not only Shuangfeng Xiang but also closely-related dialects in the Loushao subgroup which appear to have a similar three-way voicing contrast. Representative works include Yuan (1960), which provides the first introduction of the sound system of Shuangfeng Xiang Chinese in comparison to Changsha Xiang Chinese; Xiang (1960) on the first comprehensive description of Shuangfeng Xiang Chinese. *Hanyu fangyin zihui* (汉语方音字汇 [Lexicon of Chinese dialect pronunciations], *Zihui* hereafter) compiled by the Department of Chinese Language and Literature of Peking University (1989) provides a comprehensive reference work on the pronunciations of Chinese characters in Shuangfeng Xiang Chinese, with recordings of 2,961 monosyllabic words. Zhou (2005) and

[27] In some studies, the Old Xiang Chinese is also labeled as Southern Xiang Chinese and the New Xiang Chinese as Northern Xiang Chinese (e.g., Zhou & You, 1985).

Yuan (2005) are two MA theses, of which the first illustrates the historical sound change of the dialect, and of which the latter concentrates on dialectal variations across Shuangfeng county. Chen (2006) describes the typological features of the sound system of the Shuangfeng dialect in comparison with other Xiang dialects. Worth noting is Yang (1974), which reports the sound system of dialects spoken in 75 cities/counties of Hunan province. Unfortunately, Shuangfeng is not included, but the word list, together with its auditory recordings, provides a useful resource for comparative research on Shuangfeng Xiang within the Xiang dialect family.

It is important to note that the aforementioned studies tend to analyze the sound system of the Shuangfeng dialect based on auditory impression, and are prone to explore the synchronic system or variations from a historical perspective. Consequently, the existing literature remains obscure to researchers with limited knowledge of the research conventions within sinology.

This description is accompanied by recordings of a sixty-five-year-old female native speaker, who was born in 1953 and was raised in Yongfeng town. She spent most of her life living in Yongfeng and speaking Shuangfeng Xiang Chinese, except for five years when she was engaged in farming at a nearby village, where the local dialect shared a high level of intelligibility with Shuangfeng Xiang. According to her self-report, before her retirement, she sometimes spoke accented Standard Chinese with colleagues and customers who cannot understand Shuangfeng Xiang Chinese. In addition, she is also able to speak some Xiangtan Xiang Chinese (湘潭方言) (i.e., a Xiang variety spoken in the city center of Xiangtan) when the situation requires her to do so (e.g., in conversations with relatives who speak Xiangtan Chinese only).[28] Her primary language

[28] Xiangtan Xiang Chinese is classified as a dialect of the Changyi group (长益片) of the Xiang Chinese (Bao & Chen, 2005).

used at work and home is otherwise exclusively Shuangfeng Xiang Chinese.

4.2 Lexical tones

Figure 4.2 illustrates the five *f0* contours of the lexical tones in Shuangfeng Xiang Chinese, with each contour averaged over 20 words. The choice of these words was made by taking into consideration their segmental compositions; they constitute 20 near-minimal quintuplets. Example words are listed in Table 4.1. There are two level tones (i.e., T1 and T5), two rising tones (i.e., T2 and T4), and one falling tone (i.e., T3). Generally speaking, both T1 and T4 start in a higher *f0* range (above 180 Hz, high register hereafter), while both T2 and T5 start in a lower *f0* range (under 180 Hz, low register hereafter). Both T1 (black solid line) and T5 (gray long dashed line) are level tones, contrasting mainly in pitch height (high–level vs. low–level). T2 (gray solid line) and T4 (black long dashed line) are both rising, but differ in terms of not only their pitch registers but also their rising slopes. Specifically, T2 has a low-rising or low-dipping shape, whose pitch contour falls slightly within the lower pitch range initially, and then rises from the low to the mid-level (low–rising). Compared to T2, T4 starts from the mid-pitch range and stays mid till after the midpoint of the syllable and then rises sharply from the mid to the upper end of the speaker's pitch range (high–rising). T3 (black dotted line) is a falling tone, which falls from the upper end to the lower of the pitch range. Adopting the five-scale tonal descriptive system developed by Chao (1930), the five lexical tones can be broadly transcribed as /55/ (T1), /13/ (T2), /41/ (T3), /35/ (T4), and /22/ (T5).

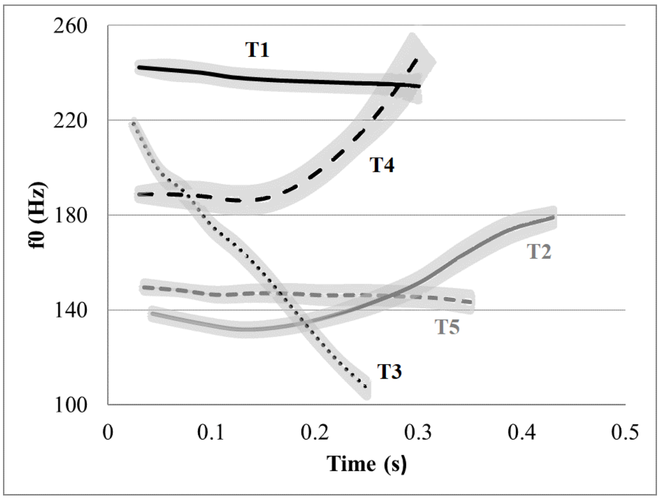

Figure 4.2 *f0* contours of the lexical tones in Shuangfeng Xiang Chinese. For each lexical tone, mean *f0* (light gray areas indicate ± *SE*) of 10 equidistance points averaged over 20 words is plotted.

Table 4.1 Examples of the lexical tones in Shuangfeng Xiang Chinese.

Lexical tone	Tonal contour	Example	Orthography	Gloss
Tone 1 (T1)	high–level	/ti¹/	低	'low'
Tone 2 (T2)	low–rising	/di²/	啼	'to crow'
Tone 3 (T3)	falling	/ti³/	底	'bottom'
Tone 4 (T4)	high–rising	/ti⁴/	帝	'emperor'
Tone 5 (T5)	low–level	/di⁵/	地	'ground'

There are some co-occurrence constraints on consonantal onsets and lexical tones, as shown in Table 4.2. In general, the five lexical tones can be divided into three subgroups, namely, i) T1 and T5, ii) T3 and T4, and iii) T2. The two level tones (i.e., T1 and T5) are in complementary distribution, where T1 can only co-occur with voiceless onsets (e.g., /pɘ¹/ 'bag'), while T5 occurs exclusively with voiced and sonorant onsets (e.g., /bɘ⁵/ 'to embrace'; /ljɘ⁵/ 'to predict'). T3 and T4 both can co-occur with voiceless and sonorant onsets, but not with voiced obstruent onsets. For example, T3 can appear in /pɘ³/ 'treasure', /pʰjɘ³/ 'to bleach', and /lɘ³/ 'old', while T4 in /pɘ⁴/ 'newspaper', /pʰɘ⁴/ 'bubble', and /mɘ⁴/ 'to emit'. T2 is a unique lexical tone in Shuangfeng Xiang Chinese, as there is no co-occurrence constraint as illustrated in /ti²/ 'target', /tʰi²/ 'to kick', /di²/ 'to mention', and /li²/ 'to leave'.

Table 4.2 Co-occurrence constraints on onset-tone combinations in Shuangfeng Xiang Chinese. +: Yes; –: No.

Onsets	T1	T2	T3	T4	T5
Voiceless unaspirated	+	+	+	+	–
Voiceless aspirated	+	+	+	+	–
Voiced	–	+	–	–	+
Sonorant	–	+	+	+	+

4.3 Consonants

Shuangfeng Xiang Chinese has 30 consonants. Corresponding key words/bound morphemes are provided below the consonant chart.

	Bilabial			Labio-palatal	Alveolar			Postalveolar			Alveolo-palatal			Velar		
Plosive	pʰ	p	b		tʰ	t	d							kʰ	k	g
Affricate					tsʰ	ts	dz	tʃʰ	tʃ	dʒ	tɕʰ	tɕ	dʑ			
Nasal			m				n									ŋ
Fricative					s			ʃ			ɕ			x		ɣ
Approximant			w	ɥ								j				
Lateral approximant							l									

pʰ	pʰə^{1}	抛	‘to throw’	tʰ	tʰə^{1}	涛	‘billow’	kʰ	kʰə^{1}	敲	‘to knock’				
p	pə^{1}	包	‘bag’	t	tə^{1}	刀	‘knife’	k	kə^{1}	高	‘high’				
b	bə^{5}	抱	‘to embrace’	d	də^{5}	道	‘path’	g	gen^{2}	乾	‘heaven in eight trigrams’				
tsʰ	tsʰə^{1}	操	‘to hold’	tʃʰ	tʃʰɻ̩^{1}	痴	‘infatuation’	tɕʰ	tɕʰə^{3}	悄	‘quiet’				
ts	tsə^{1}	糟	‘terrible’	tʃ	tʃɻ̩^{1}	知	‘to know’	tɕ	tɕə^{1}	焦	‘scorched’				
dz	dzə^{2}	曹	‘surname, Cao’	dʒ	dʒɻ̩^{2}	池	‘pond’	dʑ	dʑə^{2}	桥	‘bridge’				
m	mə^{2}	毛	‘fur’	n	njə^{5}	尿	‘urine’	ŋ	ŋə^{2}	熬	‘to endure’	l	lə^{2}	劳	‘labor’
s	sə^{1}	骚	‘coquettish’	ʃ	ʃɻ̩^{1}	诗	‘poem’	ɕ	ɕə^{1}	消	‘to vanish’	x	xə^{1}	好	‘good’
												ɣ	ɣə^{2}	毫	‘fine long hair’
w	wa^{2}	挖	‘to dig’	ɥ	ɥa^{2}	曰	‘to say, literary form’	j	ja^{2}	压	‘to press’				

Generally speaking, Shuangfeng Xiang Chinese features a three-way contrast in plosives across three places of articulation (i.e., bilabial vs. alveolar vs. velar), known as voiceless unaspirated, voiceless aspirated, and voiced, respectively. Table 4.3 shows the mean VOT values and their standard deviation of the three-way contrast of plosives in three places of articulation. The measurements were obtained from 306 monosyllabic morphemes consisting of 102 sets of triplets. As shown in Table 4.3, irrespective of the place of articulation, the voiceless aspirated plosives have longer VOT values than the other two counterparts. The voiced plosives have negative VOT values. For further details, see below in the section on the three-way laryngeal contrast of obstruents (Section 4.4.1).

Table 4.3 VOT of unaspirated vs. aspirated vs. voiced plosives in different places of articulation, based on 306 monosyllabic morphemes with plosive onsets.

	Bilabial (35 triplets)		Alveolar (64 triplets)		Velar (3 triplets)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Voiceless unaspirated	14 ms	7 ms	15 ms	5 ms	23 ms	8 ms
Voiceless aspirated	85 ms	17 ms	105 ms	21 ms	107 ms	14 ms
Voiced	−73 ms	18 ms	−68 ms	16 ms	−60 ms	2 ms

Identically to the plosives, there is also a three-way contrast in affricates across three places of articulation. /tsʰ ʦ ʣ/ are alveolars showing a more dental contact area than those in Standard Chinese. Alveolar-palatals /ʨʰ ʨ ʥ/ have a contact area of the alveolar ridge and the forward part of the palatal region; postalveolars consisting of /tʃʰ tʃ dʒ/. The postalveolar affricates in Shuangfeng Xiang Chinese are similar to those in both Beijing (Lee & Zee, 2003) and Tianjin Mandarin (Li et al., 2019). All are produced with the tongue tip raised against the postalveolar region and tend to be apical. More work on palatograms and linguograms can be informative to confirm the impressionistic observations.

Shuangfeng Xiang Chinese has five fricatives /s ɕ ʃ x ɣ/. /s ɕ ʃ/ have a slightly more anterior contact area than their corresponding affricates, respectively. Acoustically speaking, the center of gravity (COG) is argued to be an important indication of differentiating fricatives' places of articulation (Gordon et al., 2002). Therefore, COG values of ten minimal pairs of /s/ vs. /ɕ/ as well as ten additional morphemes which consist of the /ʃɻ̩/ syllable with various lexical tones were measured. As the stimuli contain different vowels, in order to exclude the coarticulation effect of the vowel, the mean of COG was measured within the first 30 ms interval from the beginning of each fricative. The results are plotted in Figure 4.3. The mean of /s/ is the highest (4999 Hz) while the mean of /ʃ/ the lowest (3198 Hz). A one-way ANOVA revealed a main effect for Place [$F(2, 27) = 56.42$, $p < .001$]. The results of *post hoc* comparisons (with Tukey HSD) suggested significant differences among the three places of articulation. It has been known that COG correlates well with the front cavity length: the lower COG, the larger front cavity (Ladefoged & Maddieson, 1996: 163). The results thus echo well with findings reported for /s ʃ/ in English (Jongman et al., 2000) and for /s ɕ ʃ/ in Standard Chinese (Svantesson, 1986; Li, 2008) and Swedish (Lindblad, 1980).[29]

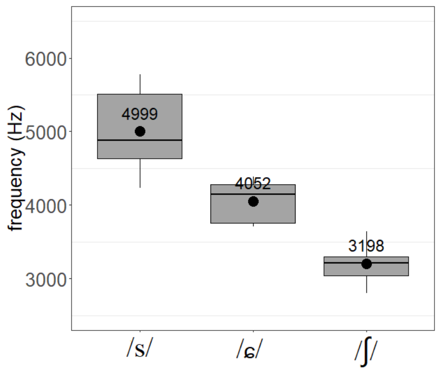

Figure 4.3 Mean COG of /s/ vs. /ɕ/ vs. /ʃ/ measured within the first 30 ms interval from the beginning of each fricative.

[29] /ʃ/ is transcribed as /ʂ/ in both studies of Standard Chinese.

It is important to note that there is a two-way voicing contrast in velar fricatives but no voicing contrast in fricatives with other places of articulation. Such asymmetry has been proposed to be triggered by a parallel merger between voiced fricatives and voiced affricates, where the earlier non-velar /*z *ʑ *ʒ/ fricative onsets had merged with their affricate counterparts /dz dʑ ʤ/ (Xiang, 1960).

In a lot of Chinese dialects, a phonemic distinction between labiodental fricatives and velar/glottal fricatives can be recognized, such as /f/ vs. /x/ in Standard Chinese (Lee & Zee, 2003) or /v/ vs. /ɦ/ in Shanghai Wu Chinese (Chen & Gussenhoven, 2015).[30] However, in Shuangfeng Xiang Chinese, such a contrast in place of articulation has been lost. Words with the labiodental fricatives /f v/ have merged with their velar counterparts (i.e., /x ɣ/). For example, both 'wrong' and 'brightness' are pronounced as /xwi^{1}/ in Shuangfeng Xiang Chinese, while in Standard Chinese, they are /fei^{1}/ and /xwei^{1}/, respectively. In addition, /x ɣ/ are pronounced as glottal fricatives followed by a non-high vowel, as illustrated in /xɔn^{1}/ [χɔ̃^{1}] 'desolate' and /ɣan^{5}/ [ʁãn^{5}] 'slit'.

The realization of the voiced fricative /ɣ/ in Shuangfeng Xiang Chinese is contingent on the contexts. When /ɣ/ is pronounced in initial position, it is seldom realized as fully voiced. In most cases, the voicing is only partly observed, as shown in /ɣan^{5}/ 'slit', /ɣan^{5}/ (one morpheme of) 'phoenix', and /ɣjan^{5}/ 'lucky' in Figure 4.4, where the proportions of the voicing over the entire fricative vary from more to less to almost none. However, when /ɣ/ appears in non-initial position, voicing is consistently observed throughout the entire fricative sound. As shown in Figure 4.5, in initial position, /ɣ/ in the syllable /ɣjɔn^{2}/ 'frequent' fails to show voicing; while in non-initial position as in /tʰan^{1} ɣjɔn^{2}/ 'usual', /ɣ/ is realized as a completely voiced sound.

[30] In some modern varieties of Wu Chinese, /ɦ/ is a phonetic segment only co-occurring with mid and open vowels and tends to disappear before high vowels (see Chen & Gussenhoven, 2015 and Section 2.8 for more information).

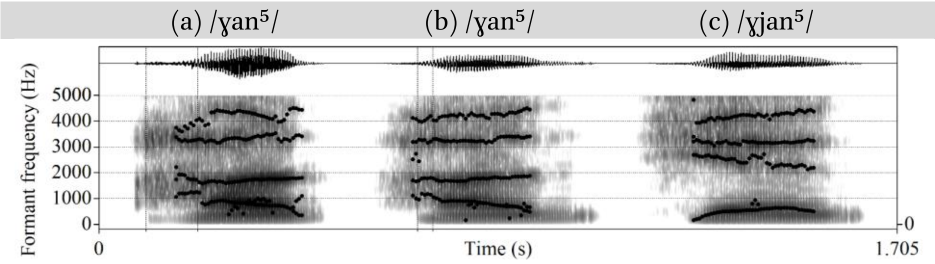

Figure 4.4 Waveforms and spectrograms of (a) /ɣan⁵/ 'slit', (b) /ɣan⁵/ (one morpheme of) 'phoenix', and (c) /ɣjan⁵/ 'lucky'. Intervals indicate the voicing over the entire fricative.

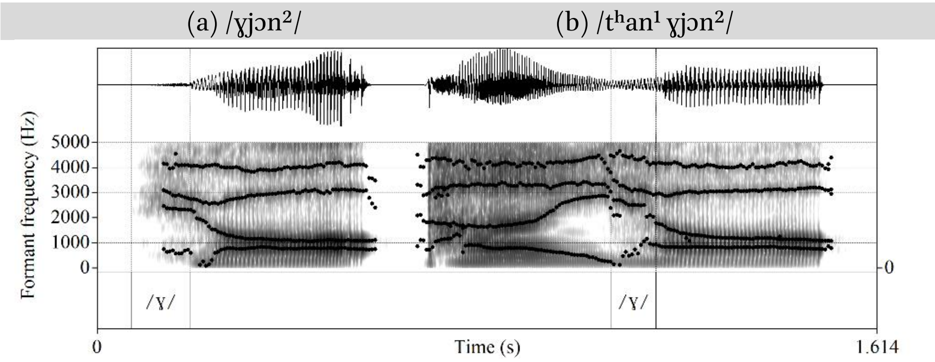

Figure 4.5 Waveforms and spectrograms of (a) /ɣjɔn²/ 'frequent' and (b) /tʰan¹ ɣjɔn²/ 'usual'.

4.4 The three-way laryngeal contrast of obstruents

4.4.1 Phonetic realization

The three-way laryngeal contrast of obstruents in varieties of the Old Xiang group is commonly regarded as a crucial feature that makes the Xiang dialect group related to the Wu Chinese dialect group (e.g., Ting, 1982), the majority of which also have a three-way laryngeal contrast (after Chao, 1928). The phonetic manifestation of the three-way contrast in Shuangfeng Xiang Chinese, however, is very different from that observed in the Wu dialects. For example, in Shanghai Wu and Lili Wu Chinese, the three-way contrast in initial position cannot be distinguished via VOT. Instead, phonation has been argued to better distinguish from clearly modal (voiceless unaspirated), aspirated (voiceless aspirated), to breathy (voiced) (see Chen, 2011; Chen & Gussenhoven, 2015; Section 2.3 and references therein for further details). In Shuangfeng Xiang Chinese, voiced obstruents tend to be fully voiced, leading to a full-fledged distinction in terms of VOT, namely short-lag VOT (voiceless unaspirated), long-lag VOT (voiceless aspirated), and lead VOT (voiced). This is exemplified by the triplet /ti^2/ 'target', /t^{h}i^2/ 'to kick', and /di^2/ 'title' in Figure 4.6, respectively.

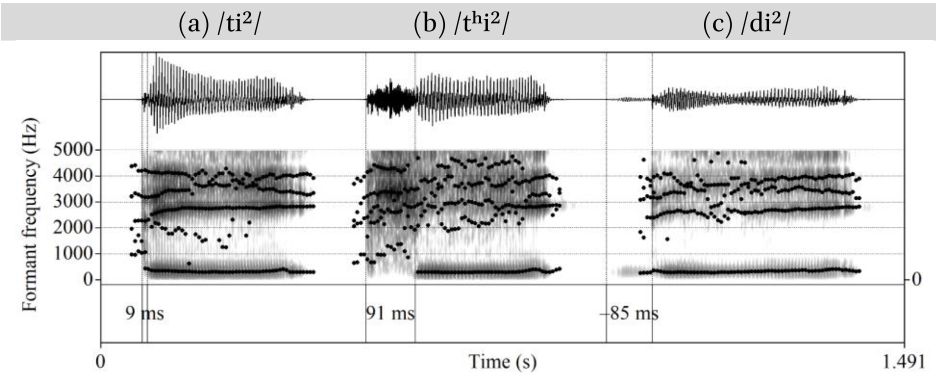

Figure 4.6 Waveforms and spectrograms of (a) /ti^2/ 'target', (b) /t^{h}i^2/ 'to kick', and (c) /di^2/ 'title'. Intervals indicate VOT values.

Worth noting is that the acoustic realization of the voiced category also teems with variation. It has been observed that a voiced consonant may be produced with a relatively short VOT lead (/dų^2/ [d̪əu^2] 'pathway', –33 ms) or even a short VOT lag (/dɘ^2/ 'to escape', 13 ms). Even for the same word, voicing can be optional. Figure 4.7 shows two tokens of the syllable /dja^2/ 'jar' produced by the consultant. There is a lead VOT of –77 ms in the first instance, but only a short-lag VOT of 9 ms in the second.

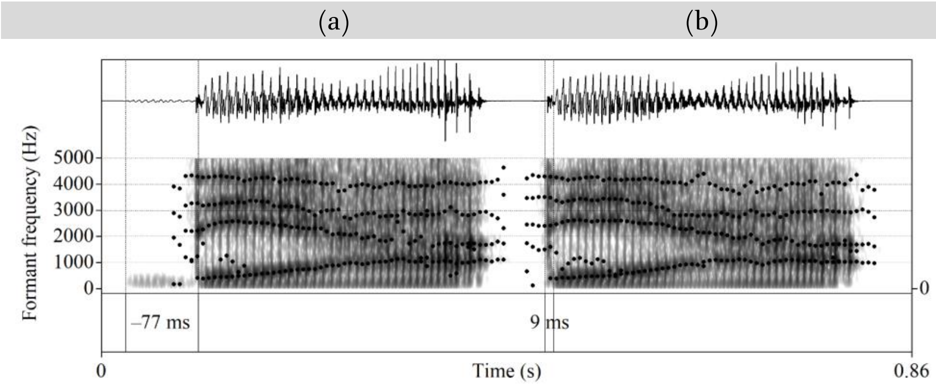

Figure 4.7 Waveforms and spectrograms of two tokens of /dja^{2}/ 'jar'. Intervals indicate VOT values. /d/ is realized with a lead VOT (–77 ms) in the first instance (a), but as a short-lag VOT (9 ms) in the second instance (b).

Voiced plosives in Shuangfeng Xiang may also be realized as implosives. As shown in Figure 4.8, /d/ is realized as a voiced plosive in /den^{2}/ [d̪ẽĩ^{2}] 'farmland', but an implosive [ɗ] in /djan^{5}/ [ɗjãn^{5}] 'starch'. In the implosive [ɗ], the amplitude of vibrations gradually increases during the oral closure, in contrast with the voiced plosive [d], where one can identify a more typical biphasic and decreased amplitude during the closure (Lindau, 1984; Ladefoged & Maddieson, 1996). Such differences, to a large extent, can be attributed to the characteristic airflow during the articulatory closure in the production of obstruents. As explained by Ladefoged and Johnson (2010: 140), '[i]n the production of implosives, the downward moving larynx is not usually completely closed. The air in the lungs is still being pushed out, and some of it passes between the vocal cords, keeping them in motion.' Consequently, this particular airflow mechanism generates increasing amplitude visible in the waveforms. With regard to voiced plosives, the pulmonic airflow cannot continue through the glottis to maintain voicing. As more air flows to the mouth through the glottis, the pressure in the lungs returns to normal, and the vibration of vocal cords attenuates, resulting in a decrease of the amplitude over the closure time. Commonly, voiced implosives are regarded as allophones of voiced plosives in many languages, such as

Vietnamese and Khmer (Kirby, 2018), though implosives can also contrast with plosives in few languages as in Sindhi (Ladefoged & Johnson, 2010). In Shuangfeng Xiang Chinese, there is no phonemic contrast between voiced plosives and voiced implosives.[31]

(a) /den^{2}/	(b) /djan^{5}/

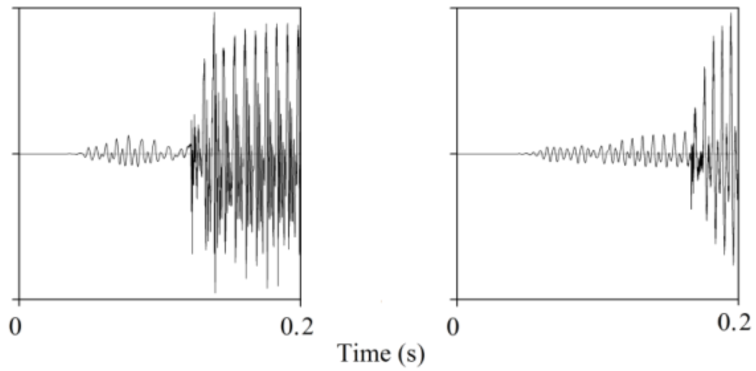

Figure 4.8 Oscillograms of /d/ in (a) /den^{2}/ 'farmland' and (b) /djan^{5}/ 'starch'.

4.4.2 Historical development

It is important to note that the three-way contrast in modern Shuangfeng Xiang seems to be conditioned by the reconstructed MC tonal categories (e.g., Xiang, 1960; Zhou, 2005). Generally speaking, the contrast can be observed in syllables on the MC *Ping*, *Shang* and *Qu* tonal categories, but only a two-way obstruent contrast (i.e., voiceless unaspirated vs. voiceless aspirated) is maintained in syllables associated with the MC *Ru* tonal category. My corpus contains 899 high-frequency syllables selected from the Questionnaire of character for dialect surveys (*Fangyan diaocha zibiao*, 方言调查字表) with voiced obstruent onsets in MC. As indicated by

[31] As argued by Ladefoged and Maddieson (1996: 82), the relationship between voiced plosives and the so-called 'true' implosives is more gradient, rather than a clearly defined dichotomous distinction. This kind of relationship, to a certain extent, is reflected in diachronic changes, where the present-day implosives likely originated from earlier-time voiced plosives (Greenberg, 1970: 135; Cun, 2009; Ladefoged & Johnson, 2010: 143).

Table 4.4, none of the 149 syllables originating from the MC *Ru* tonal category is phonetically realized with voicing. With respect to the 750 syllables associated with the other three MC tonal categories (i.e., *Ping*, *Shang*, and *Qu*), the presence of the voice bar seems to be contingent upon the manner of articulation. Both plosives (48%) and affricates (50%) show higher percentages of the tokens with the voice bar than fricatives (30%). Compared to plosives, the higher devoicing-proclivity of voiced fricatives can be explained by the problematic mechanism of its articulation (Stevens, 1971; Ohala, 1983; Smith, 1997). The production of voiced fricatives is aerodynamically challenging, due to the opposing efforts required. On the one hand, a greater subglottal pressure with a weaker airflow should be kept for the maintenance of vibration of the vocal folds; on the other hand, a sufficient airflow above the glottis is required in order to generate frication noise. It is worth noting that typically partial voicing exists only for some voiced fricatives which do show the voice bar. In this description, I assume the voicing-present case (i.e., phonetically voiced) as long as the voice bar presents in the duration of the frication.

Table 4.4 Percentage of the voiced category with the presence of the voice bar based on MC tonal categories.

Voice bar	MC *Ru* (0/149)			MC *Ping*, *Shang*, and *Qu* (315/750)		
Presence	plosive	affricate	fricative	plosive	affricate	fricative
	0	0	0	48%	50%	30%
	(0/33)	(0/440)	(0/76)	(132/275)	(97/193)	(86/282)

Interestingly, when taking a closer look at those items without a voice bar, two patterns can be observed. As shown in Table 4.5, in the MC *Ru* tonal category, 79% of the MC voiced plosives or affricates are produced as voiceless aspirated counterparts, while 21% become voiceless unaspirated counterparts, but their place of articulation remains the same. By contrast, plosives or affricates associated with the MC *Ping*, *Shang*, and *Qu* tonal categories are more likely produced as their voiceless unaspirated counterparts (unaspirated 92% vs. aspirated 21%) while maintaining their place of articulation.

Table 4.5 Percentage of the voiced category without the presence of the voice bar based on the MC tonal categories.

Voice bar	MC *Ru* (149/149)			MC *Ping*, *Shang*, and *Qu* (435/750)		
	plosive & affricate		fricative	plosive & affricate		fricative
	aspirated	unaspirated		aspirated	unaspirated	
Absence	79%	21%	100%	8%	92%	100%
	(58/73)	(15/73)	(76/76)	(18/239)	(221/239)	(196/196)

Given these observations, it may conclude that in Shuangfeng Xiang Chinese, the three-way obstruent contrast is no longer a consistent feature but shows strong instability. First, voiced obstruents are realized with multiple variants. Furthermore, the three-way obstruent contrast seems to be largely conditioned by MC tonal categories. Onsets with a lexical tone belonging to the MC *Ru* tonal category, due to the ubiquitous loss of the voicing realization, can be regarded to show a two-way contrast only (i.e., voiceless unaspirated vs. voiceless aspirated). Onsets with a lexical tone belonging to non-*Ru* tonal categories (i.e., *Ping*, *Shang*, and *Qu*), to a large extent, still maintain a full-fledged three-way contrast (i.e. voiceless unaspirated vs. voiceless aspirated vs. voiced), although various realizations of the voiced category also exist. A similar phenomenon has also been reported for neighboring dialects of Shuangfeng Xiang Chinese, such as Xiangxiang (湘乡) and Loudi (娄底) (Chen, 2006: 27). Needless to say, further research, especially with data elicited from a large sample of speakers of different generations, are needed for a more precise description of their realizations and distributions.

4.5 Sonorants

All nasals (i.e., /m n ŋ/) can occur in onset position. The alveolar nasal /n/ is palatalized before high front segments (i.e., /i j ʏ/) such as in /ni^{2}/ [ɲi^{2}]

'Buddhist nun'. Sometimes, /ŋ/ tends to give rise to a strong nasalized quality over the following vowel due to progressive nasal assimilation, as shown in /ŋo^2/ [ŋõ^2] 'tooth'. /m n/ can also form syllable nuclei (e.g., /m̩^{1} mja^N/ 'mother' and /n̩^{3}/ 'you').

/n l/ contrast only before three high segments (i.e., /i ʏ j/), as in /ni^2/ [ɲi^2] 'muddy' vs. /li^2/ 'to leave'; /nʏ^3/ [ɲʏ^3] 'female' vs. /lʏ^3/ 'to travel'; and /njɔn^2/ [ɲjɔ̃^{2}] 'a form of address for an elderly married woman' vs. /ljɔn^2/ [ljɔ̃^{2}] 'grain'. Words distinguished by /n/ (e.g., /nau^3/ 'brain') and /l/ (e.g., /lau^3/ 'old') in Standard Chinese therefore become homophones /lɘ^3/ in Shuangfeng Xiang Chinese. Due to this merger, the consultant failed to distinguish /n/ and /l/ in the non-/i j ʏ/ conditions. One observed exception is the loss of /n l/ contrast before /ja/. For example, 'difficulty' (/*n/) and 'blue' (/*l/) are homophones (i.e., /lja^2/). This is due to the fact that these syllables (with /ja/ in modern Shuangfeng Xiang) have likely developed from the earlier /ã/ (Xiang, 1960) or /æ̃/ (*Zihui*, 1989), without the glide /j/.[32] Diachronically speaking, it is likely that the glide /j/ in /ja/ is an epenthetic segment that emerged after the merger of /n/ and /l/ before non-high-frontal segments, and therefore, was exempted from the condition which contributed to the merger of /n/ and /l/. The relative chronology of the observed exceptions can be delineated as the following: /*nã/ (e.g., 'difficulty') and /*lã/ (e.g., 'blue') first merged to /*lã/, and then became /lja/.

Typologically speaking, the lack of distinction between /n/ and /l/ seems to be rarely reported in the world's languages. However, such a merger has been commonly observed in Jianghuai and Southwest Mandarin (see Chen & Guo, forthcoming and references therein) as well as the Min and Xiang varieties (Xiang, 1960) (see Huang, 2007 for further

---

[32] In both studies, researchers also point out that /ã/ (/æ̃/) and /ja/ (/jɛ/) are in free variation for some speakers. My current consultant, however, tends to consistently produce all /ã/ as /ja/.

details on the different patterns of the /n l/ merger across Chinese dialects).

/n/ can serve as a coda which varies in its phonetic realization due to the articulatory height of the vowel. After /a/ and /ɜ/, /n/ is realized as a clear coda as in /tan^1/ [t̪ãn^1] 'east' and /dwɜn^1/ [d̪wɜ̃n^2] 'suckling pig'. By contrast, after an open-mid vowel /ɔ/, /n/ is not realized as a clear coda but triggers vowel nasalization (e.g., /tɔn^1/ [t̪ɔ̃^{1}] 'when'). Perhaps most prominently, when /n/ follows the closer vowel /e/ (e.g., /ten^1/ 'jolt'), the alveolar nasal seems to introduce no complete oral closure near the alveolar ridge and is realized as [t̪ẽĩ^1] where a nasalized [ĩ] element can be heard.

4.6 Approximants

Shuangfeng Xiang Chinese has three approximants in onset position, as indicated by the triplet /ja^2/ 'to press', /wa^2/ 'frog', and /ɥa^2/ 'to say, the form in classical Chinese'. All three approximants can serve as a glide, as in /tja^1/ 'to delay', /twa^1/ 'to hold something level with both hands', and /ɕɥe^2/ 'snow'. It is important to note that, sometimes, when /j/ occurs as a glide, compared to Standard Chinese, its quality appears to be characterized by a more rapid transition to the following vowel. Xiang (1960) argues that syllables containing the glide /j/ may have a closer relationship to the onset, such that /tja^1/ 'to delay' is realized as [tʲa^1].[33] However, I hasten to note that in the consultant, a very clear /j/ in /tja^1/ 'to delay' has been observed. After the alveolo-palatals (i.e. /tɕʰ tɕ dʑ ɕ/), only a brief transition from the alveolo-palatal onset to the following vowel (/ɕɘ^1/ 'to vanish') was observed, which suggests /j/-quality. However, given three reasons, I have adopted the analysis of Chen and Gussenhoven (2015) for Shanghainese and the analysis for Lili Wu Chinese (Section 2.4) and posited no underlying /j/ after alveolo-palatal

[33] In Xiang's proposal (1960), the glide /j/ is transcribed as /i/.

onsets. The first reason is the briefness of such transitive /j/ quality, which is much less than that in /tja¹/. The second is due to the predictability of such transitive /j/ quality after alveolo-palatal onsets. The third reason is the lack of contrasts like /ɕɘ/ vs. /ɕjɘ/ in Shuangfeng Xiang Chinese.

In Shuangfeng Xiang Chinese, there exist two syllabic approximants /ɹ̩ ɻ̩/ as in /tsɹ̩³/ [t̪s̪ɹ̪̩³] 'purple' and /tʃɻ̩⁴/ 'to make'. The production of /ɹ̩ ɻ̩/ in Shuangfeng Xiang Chinese is similar to that in Standard Chinese. /ɹ̩/ only occurs after alveolar consonants /tsʰ ts dz s/ and /ɻ̩/ after postalveolars /tʃʰ tʃ dʒ ʃ/. The tongue configuration of /ɹ̩/ is almost homorganic with the preceding consonant and therefore tends to be dental. For more information on both syllabic approximants, interested readers are referred to Lee-Kim (2014).

4.7 Vowels

Vowels in open syllables are plotted in Figure 4.9(a), and those in closed syllables with a nasal coda in Figure 4.9(b). The vowel ellipses were calculated based on 50 tokens of each vowel except for /ʉ̜/ (18) and /ʏ/ (30). The measurements of formant values are identical to Section 2.5. Example keywords/bound morphemes are provided below the plots. In open syllables, there are eight monophthongs (/i ʏ u ʉ̜ e ɘ o a/) in Shuangfeng Xiang Chinese. These eight vowels constitute a four-way distinction (i.e., close, near-close, close-mid, and open) in height and a three-way distinction (i.e., front, central, and back) in backness. Compared to the vowels in open syllables, the number of vowels in closed syllables is reduced, and so is their acoustic vowel space. Phonemically, there are four vowels (/e ɜ ɔ a/) contrasting in height and backness.

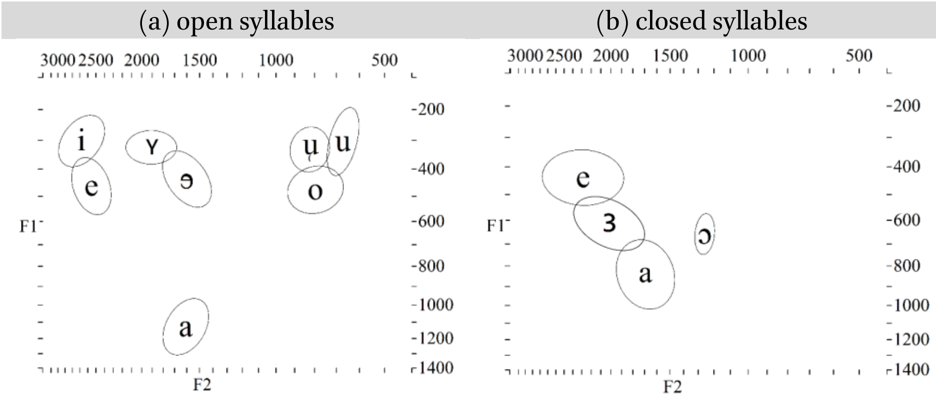

Figure 4.9 Relative F1/F2 formant values of monophthongs produced in (a) open syllables and in (b) closed syllables.

Monophthongs in open syllables

i	bi^2	皮	'skin'	u	bu^2	婆	'old woman'
ʏ	ʏ^2	鱼	'fish'	ų	bų^2	蒲	'calamus'
e	be^5	陪	'to accompany'	o	bo^2	爬	'to creep'
ɘ	bɘ^2	袍	'robe'				
a	ba^2	牌	'card'				

Monophthongs in closed syllables

e	ten^1	颠	'jolt'	ɜ	tjɜn^1	真	'real'
a	tan^1	东	'east'	ɔ	tɔn^1	当	'when'

/ɘ/ is commonly transcribed as /ɤ/ (e.g., Xiang, 1960). In the consultant's productions, however, this vowel tends to be much more central. Hence, the IPA symbol /ɘ/ has been opted for.

/ʏ/ is consistently transcribed as /y/ in all previous studies. However, it is worth noting that there are remarkable differences between /ʏ/ in Shuangfeng Xiang Chinese and /y/ in Standard Chinese, as shown in Figure 4.10, where /y^2/ 'fish' in Standard Chinese (4.10a) and /ʏ^2/ 'fish' in Shuangfeng Xiang Chinese (4.10b) are plotted. Compared to /y^2/ in Standard Chinese, /ʏ^2/ in Shuangfeng Xiang Chinese has much lower formant values (F2: 2505 Hz vs. 1995 Hz; F3: 3088 Hz vs. 2708 Hz; F4: 4350 Hz vs. 3549 Hz), indicating an articulatory retraction. In this current study,

/ʏ/ instead of the widely-used /y/ is posited accordingly. Moreover, when /ʏ/ follows /tʃʰ tʃ dʒ/, due to coarticulation, the tongue body position of /ʏ/ as in /dʒʏ^2/ [dʒʏ̄^{2}] 'kitchen' is slightly further back than that in /ʏ^2/ 'fish', indicated by a lower F2 in /dʒʏ^2/ (1808 Hz, compared to 1995 Hz in /ʏ^2/).

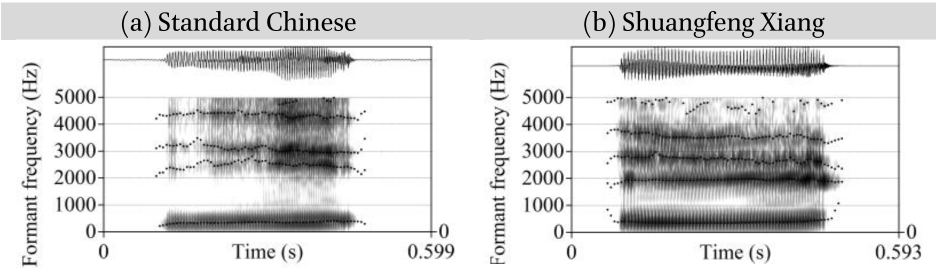

Figure 4.10 Waveforms and spectrograms of 'fish' in (a) Standard Chinese and (b) Shuangfeng Xiang Chinese. The sound file of Standard Chinese was adopted from Lee and Zee (2003).

Shuangfeng Xiang Chinese has an interesting and rarely observed three-way contrast in high back vowels, as illustrated by the triplet of /bo^2/ 'to climb' vs. /bu^2/ 'old woman' vs. /bu̜^{2}/ 'calamus'. As shown in Figure 4.11, /o/ is produced with higher F1 and lower F2 values. The difference between /u/ and /u̜/ is mainly in F2 with /u/ having lower values than /u̜/. What is more salient is the lip gesture differences among these three back vowels as shown in Figure 4.12. The lip aperture is markedly smaller for the higher vowel /u/ than for the lower vowel /o/. However, both /o/ and /u/ differ from /u̜/ in having a more rounded and protruding lip constriction. /u̜/, in contrast, is produced with greater lip compression and less protuberance. The slightly higher F2 of /u̜/ probably results from the compressed lip gesture which shortens the length of the vocal tract accordingly. In some languages, vertical compression and horizontal protrusion are two lip-position parameters for vowel contrasts (Ladefoged & Maddieson, 1996: 295). For example, /y:/ is argued to differ from /ʉ:/ mainly with a more protruded lip position in Swedish (See Fant, 1983 and references therein). In Japanese, /u/ is usually phonetically transcribed as [ɯ] due to its unrounded and compressed lips (Okada, 1991). Compared to [ɯ] in Japanese, /u̜/ in Shuangfeng Xiang Chinese tends to be produced with similarly compressed lips. Given such a similarity, Zeng (2019) hence adopts /ɯ/ to transcribe this phoneme in Xiangxiang Xiang, an Old Xiang variety closely related to Shuangfeng Xiang. According to the results in Keating and Huffman (1984), the ratio of F1 and F2 for [ɯ] pronounced in word lists is around 0.33 (400 Hz: 1200 Hz, based on recordings of 7 young male speakers). However, the ratio of /u̜/ produced by the consultant is around 0.41 (350 Hz: 850 Hz), which suggests a different vowel quality to [ɯ] in Japanese. It is quite close to the ratio of /u/ in American English (0.42, based on average formant frequencies of 48 female speakers presented in Hillenbrand et al., 1995). Given these reasons, /u̜/ instead of /ɯ/ with the symbol ̜ has been adopted to highlight the less rounded and compressed lips of /u̜/.

In addition, /u̜/ is realized as a monophthong after bilabial onsets only, but with diphthong quality [əu] after other onsets or appearing without an onset, e.g., /du̜^{2}/ [dəu^{2}] 'pathway', /ku̜^{1}/ [kəu^{1}] 'alone', and /u̜^{1}/ [əu^{1}] 'to pollute'. Such a complementary realization of /u̜/ has also been found in some Northern Wu Chinese, such as Shanghainese (Chen & Gussenhoven, 2015) and Lili Wu Chinese (Section 2.5). In contrast, /u/ can appear as a monophthong not only after bilabial onsets, but also after other places of articulation, such as /tu^{1}/ 'many', /tsu^{3}/ 'left', /ku^{1}/ 'song', and /ɣu^{2}/ 'river'.

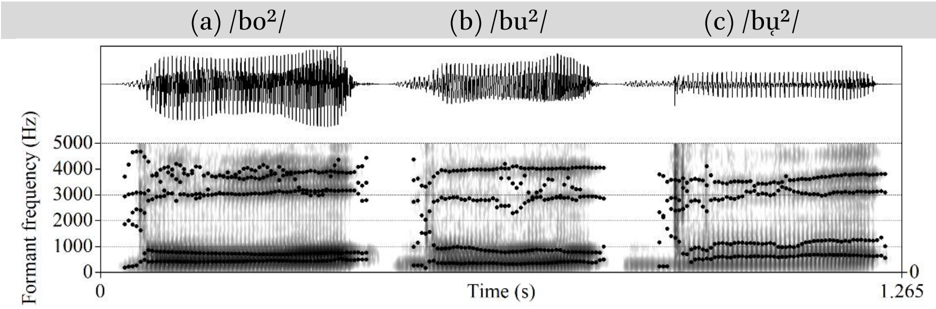

Figure 4.11 Waveforms and spectrograms of (a) /bo^2/ 'to climb', (b) /bu^2/ 'old woman', and (c) /bu̜^{2}/ 'calamus'.

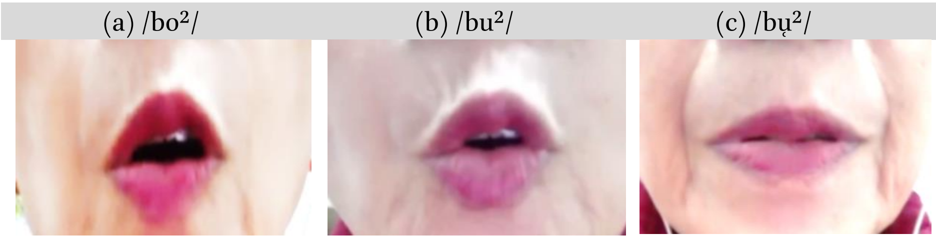

Figure 4.12 Pictures of the maximal gesture of the lips in (a) /bo^2/ 'to climb', (b) /bu^2/ 'old woman', and (c) /bu̜^{2}/ 'calamus'.

There are four vowels that occur in closed syllables. Generally speaking, vowels in closed syllables are more central than those in open syllables. /ɔ/ is a back rounded open-mid vowel as in /tɔn^1/ [t̪ɔ̃^{1}] 'when'. /e/ is a close-mid unrounded front vowel as in /ten^1/ 'jolt' with a more salient diphthong quality, as discussed in Section 4.5 on sonorants. /a/ is a front low vowel as in /tan^1/ 'east'. /ɜ/ is a central open-mid vowel and is licensed after glides only. However, its actual realization is highly contingent upon the preceding glide due to coarticulation. As shown in Figure 4.13, it is pronounced as [ɛ] after the glide /j/, such as in /tjɜn^1/ [tjɛ̃n^1] 'real'; but [ɜ] after the more rounded glides (i.e., /w ɥ/), such as in /twɜn^1/ [twɜ̃n^1] 'army' and /ɕɥɜn^1/ [ɕɥɜ̃n^1] 'elder brother'. Each realization of /ɜ/ is plotted based on the measurements of ten real words produced by the consultant.

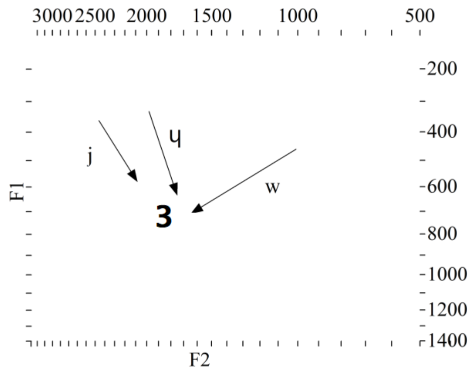

Figure 4.13 Gliding trajectories of the phoneme /ɜ/ produced after glides /j/, /w/, and /ɥ/.

At last, /ja/ and /e/ occurring in open syllables are sometimes in free variation. For example, the morpheme 'north' can be pronounced either /pja²/ or /pe²/, and the morpheme 'to separate' can be either /kja²/ or /ke²/. As told by the consultant, the two pronunciations do not yield different lexical meanings.[34]

4.8 Syllable structure

Identical to a lot of Sinitic varieties, the canonical Shuangfeng syllable minimally consists of an obligatory nucleus (V) and a lexical tone as in /i¹/ 'clothes' and /u¹/ 'an old name of Japan'. V stands for vowel or syllabic consonant (i.e., /ɹ̩ ɻ̩ m̩ n̩/). It may contain up to three optional elements in

[34] It is worth pointing out that the exchangeable pronunciations of /ja/ and /e/ do not universally appear in Shuangfeng Xiang Chinese. /ja/ can be alternatively pronounced as /e/ only in syllables originating from the MC *Ru* tonal category diachronically, the majority of which became either /ei/ (e.g., /pei³/ 'north') or /ɤ/ (e.g., /kɤ²/ 'to separate') in Standard Chinese. The /e/ pronunciation of Shuangfeng Xiang probably emerged as an adaptation under the influence of Standard Chinese. Syllables with /ja/ but not developing from the MC *Ru* tonal category, which did not become /ei/ or /ɤ/ in Standard Chinese, do not bear the /e/ alternation. For example, /pja¹/ 'class' developed from the MC *Ping* tonal category. It is pronounced /pan¹/ in Standard Chinese and does not have the alternative pronunciation */pe¹/.

the following linear sequence: (C1)(G)V(C2), where C1 can be any consonant in the consonant inventory; G is either /j/, as in /ja^{2}/ 'to press', /w/, as in /wi^{1}/ 'mighty', or /ɥ/, as in /ɥɐn^{2}/ 'cloud'; C2 contains one element only, namely /n/, as in /an^{1}/ 'surname, Weng'. All combinations are demonstrated in Table 4.6.

Table 4.6 Syllabic combinations in Shuangfeng Xiang Chinese.

Combination	Example					
V	/i^{1}/	衣	'clothes'	/u^{1}/	倭	'an old name of Japan'
GV	/wi^{1}/	威	'mighty'	/ja^{2}/	压	'to press'
C1V	/ti^{1}/	低	'low'	/ta^{2}/	搭	'to put over'
VC2	/en^{1}/	烟	'smoke'	/an^{1}/	翁	'surname, Weng'
C1VC2	/ten^{1}/	颠	'jolt'	/tan^{1}/	东	'east'
GVC2	/wen^{1}/	冤	'wronged'	/ɥɜn^{2}/	云	'cloud'
C1GVC2	/twen^{1}/	专	'specialized'	/tjɜn^{1}/	真	'real'

There are some constraints of onset-rhyme combinations in Shuangfeng Xiang Chinese, as illustrated in Table 4.7.

For open syllables, bilabials /pʰ p b m/ are not licensed before /ʏ/. Alveolars /tʰ t d n l/ have the widest distribution, which can co-occur with the majority of vowels (except for /ɹ̩ ɻ̩/). The syllabic approximants /ɹ̩ ɻ̩/ only follow after /ʦʰ ʦ ʣ s/ and /tʃʰ tʃ dʒ ʃ/, respectively. /ʦʰ ʦ ʣ s/ and /ʨʰ ʨ ʥ ɕ/ do not occur in the environments where postalveolars /tʃʰ tʃ dʒ ʃ/ occur. The distribution of /kʰ k g ŋ x ɣ/ is quite similar to that of /ʦʰ ʦ ʣ s/, except that /kʰ k g ŋ x ɣ/ are not allowed before /ɹ̩/.

Both /i ʏ/ are high front vowels. The distribution of /ʏ/, however, is more restricted. It can form an onsetless syllable, as in /ʏ^{2}/ 'fish' or co-occur with alveolar sonorants /n l/ (e.g., /nʏ^{3}/ 'female' and /lʏ^{3}/ 'to travel'). Interestingly, /i/ is forbidden after /tʃʰ tʃ dʒ ʃ/, while /ʏ/ is allowed (/dʒʏ^{2}/ 'kitchen'). The three back vowels /u u̜ o/ behave similarly. However, compared to /u/, /o/ cannot combine with /ɣ/ or form an onsetless syllable, while /u̜/ is impossible to appear after alveolo-palatals or the glide /j/. Both glides /w ɥ/ cannot be followed by central and back vowels (i.e., /u ʉ̜ o ɘ/) in open syllables. /j/, however, can precede the majority of them (except /ʉ̜/), as in /ju^{2}/ 'to swim', /jo^{3}/ 'also', and /jɘ^{2}/ 'to shake'.

For closed syllables, /en an ɔn/ can combine with alveolars, velars, and form onsetless syllables as in /en^{1}/ 'smoke', /an^{1}/ 'I', and /ɔn^{1}/ 'surname, Weng'. /en/ and /ɔn/ can follow after /w/ and /j/ as in /wen^{2}/ 'round' and /jɔn^{2}/ 'central'. /ɔn/ after bilabials (i.e., /pʰ p b m/), while /en/ after alveolo-palatals (i.e., /ʨʰ ʨ ʥ ɕ/) is forbidden. Worth noting is that the distribution of /ɜn/ is highly restricted. As discussed in Section 4.7, it can only co-occur with glides.

Table 4.7 Observed onset-rhyme combinations in Shuangfeng Xiang Chinese. +: Yes; –: No.

Onset	e a	i	ɤ	u	ų	o	ɘ	ɹ̩	ɻ̩	en	an	ɔn	ɜn
Bilabial (pʰ p b m)	+	+	–	+	+[35]	+	+	–	–	+	+	–	–
Alveolar (tʰ t d n l)	+	+	+[36]	+	+	+	+	–	–	+	+	+	–
Alveolar sibilants (tsʰ ts dz s)	+	–	–	+	+	+	+	+	–	+	+	+	–
Postalveolar (tʃʰ tʃ dʒ ʃ)	–	–	+	–	–	–	–	–	+	–	–	–	–
Alverlo-palatal (tɕʰ tɕ dʑ ɕ)	–	+	–	+	–	+	+	–	–	–	+	+	–
Velar (kʰ k g ŋ x ɣ)	+	–	–	+	+	+[37]	+	–	–	+	+	+	–
j	+	+	–	+	–	+	+	–	–	–	–	+	+
w	+	+	–	–	–	–	–	–	–	+	–	–	+
ɥ	+	–	–	–	–	–	–	–	–	–	–	–	+
Zero onset	–	+	+	+	+	–	–	–	–	+	+	+	–

[35] /m/ are not included.

[36] Alveolar plosives /tʰ t d/ are excluded. In some proposals (e.g., Xiang, 1960), /ɤ/ (/y/ in their transcriptions) can also co-occur with alveolar plosives, as in /dɤ²/ 'kitchen'. However, my consultant tends to produce these plosives as affricates.

[37] /ɣo/ cannot be observed.

4.9 Onset pitch perturbations

It is worth emphasizing the fact that the actual pitch contours of lexical tones in Shuangfeng Xiang Chinese can vary significantly due to the onset perturbation effects (i.e., C*F*0).

Take T2 as an example. It has already been noted in several impressionistic studies that voiceless unaspirated onsets introduce a higher *f*0 contour compared to voiced onsets (Xiang, 1960; Yuan, 1960; *Zihui*, 1989). Data in this study also suggest perturbation effects after voiceless aspirated and nasal onsets, as shown in Figure 4.14. Here, the four pitch contours of T2 are introduced by four types of onsets, namely voiceless unaspirated obstruents (solid line, 1-VU), voiceless aspirated obstruents (dotted line, 2-VA), voiced obstruents (dashed line, 3-D), and sonorants (dash-dotted line, 4-SN). Each contour was averaged over 50 samples, 10 of which form minimal quartets of onsets. Given that not all voiced onsets are realized with periodic voicing during the closure (i.e., lead VOT, see Section 4.4), the *f*0 values in Hz were measured at twenty equidistant points over the duration of the following vowel for all syllables using VoiceSauce in the Straight method automatically. For obstruents, vowel onset was defined as the onset of the first periodic pattern following the burst in the acoustic waveform. For sonorants (i.e., nasal and lateral), vowel onset was defined as the point with an abrupt change of F2 in the spectrogram. Vowel offset was defined as the cessation of periodicity in the waveform. As we can see in Figure 4.14, the trajectory after voiceless unaspirated onsets has an overall higher *f*0 contour than that after voiced onsets, which confirms the impressionistic observation. The contours after both voiceless aspirated and voiced onsets, however, show quite some similarity. The only difference seems to be that voiced onsets tend to trigger a slightly lower beginning than the voiceless aspirated counterparts. Very interestingly, the trajectory after sonorants is much closer to that after voiceless unaspirated counterparts, rather than after voiced counterparts in terms of the overall pitch height. All four *f0* contours, although starting with different pitch heights, converge towards a rising contour over the last one-third of the syllable.

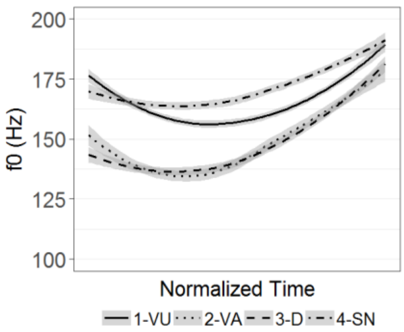

Figure 4.14 Mean *f0* of T2 after different onsets. Each mean *f0* contour (light gray areas indicate ± *SE*) was averaged over 50 samples, 10 of which formed minimal quartets of onsets. VU: voiceless unaspirated; VA: voiceless aspirated; D: voiced; SN: sonorants.

A distinctive pitch contour characteristic of the lexical tone is consistently observed in other lexical tonal categories. Figure 4.15 shows onset perturbation effects on *f0* in T1 (4.15a), T3 (4.15b), T4 (4.15c), and T5 (4.15d), respectively. Each contour was extracted by averaging across different amounts of samples contrasting in onsets (i.e., 76 for T1; 43 for T3; 47 for T4; 61 for T5). Wherever possible, voiced, voiceless aspirated, voiceless unaspirated, and sonorant onsets have been included.

As observed in Figure 4.15(a–c), voiceless unaspirated onsets (solid line) consistently introduce a higher *f0* onset than its aspirated counterparts (dotted line). However, such an effect vanishes gradually and both *f0* contours tend to merge towards the end. Visual inspection of the contours of both T3 and T4 suggests that the perturbation effect of voiceless aspirated onsets is quite salient relative to that of voiceless unaspirated onsets, which ranges from 25 Hz in T3 to 50 Hz in T4.

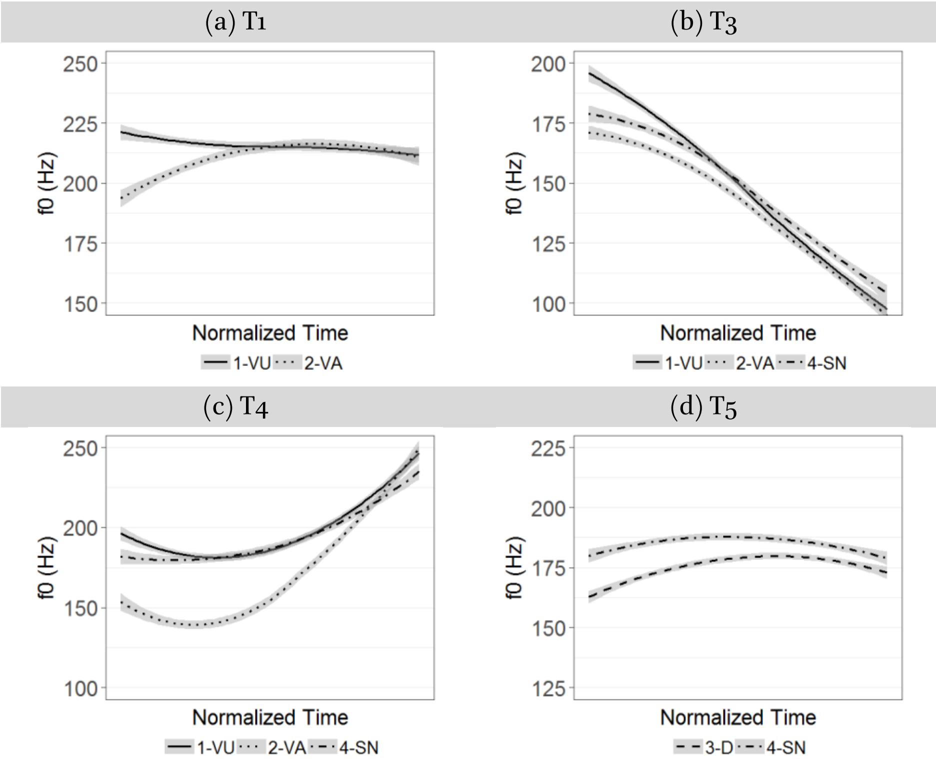

Figure 4.15 Mean *f0* of (a) T1, (b) T3, (c) T4, and (d) T5 after different onsets. Light gray areas indicate ± *SE*. VU: voiceless unaspirated; VA: voiceless aspirated; D: voiced; SN: sonorants.

For the perturbation effect of sonorant onsets, as shown in Figure 4.15(b–c), the *f0* contours after sonorant onsets are similar to those after voiceless unaspirated counterparts for both T3 and T4. Furthermore, sonorant onsets tend to introduce a much higher *f0* contour than voiceless aspirated onsets. As mentioned earlier (Section 4.2), T5 does not co-occur with voiceless onsets. Figure 4.15(d) shows the *f0* realizations of T5 after voiced obstruent and sonorant onsets. Interestingly, the *f0* trajectory after sonorants shows a higher overall pitch than that after voiced onsets, similar to what has been observed in T2.

These observations suggest that in Shuangfeng Xiang Chinese, voiceless unaspirated onsets share more similar perturbation effects with sonorant counterparts, while voiced onsets resemble voiceless aspirated ones. Following Hanson's motivation of using nasals as a reference (Hanson, 2009), C*F*0 in Shuangfeng Xiang Chinese might be summarized that *f*0 is lowered relative to the sonorant baseline after voiced and voiceless aspirated onsets but is unaffected (or minimally affected) after voiceless unaspirated onsets. This trend, however, is opposite to the results reported in Kirby and Ladd (2016), which found that in 'true voicing' languages (i.e., French and Italian in their study), compared to *f*0 after a sonorant, voiceless onsets introduce a salient *f*0 raising but no distinguishable *f*0 change following the release of a voiced consonant. Needless to say, more research with a larger group of participants is needed to further verify the differences.

4.10 Tone sandhi

Lexical tones over monosyllabic morphemes may undergo changes when they come into contact with each other in connected speech, due to a range of factors (Chen, 2012). Given a disyllabic constitution, the most relevant contextual tonal variation is tone sandhi. Chinese dialects, especially those spoken in southern and eastern areas, such as Wu and Min dialects, are famous for their manifold sandhi variations. Interested readers are referred to the seminal work of Chen (2000), which gathers a constellation of descriptions on tone sandhi across Chinese dialects. (Tone sandhi is often difficult to be distinguished from tonal coarticulation; see Li et al., 2019 for further discussion with data from Tianjin Mandarin.)

A preliminary observation over all possible 25 disyllabic compounds (5 × 5) indicates that in Shuangfeng Xiang Chinese, no obvious sandhi change can be straightforwardly identified. Quite commonly, in a disyllabic compound, irrespective of the syllable position, both lexical tonal contours are maintained. For example, Figure 4.16 and 4.17 plot minimal pairs of /sja^{1} pen^{1}/ 'three sides' (4.16a) vs. /sja^{1} pen^{4}/

'three times' (4.16b), and /ɕan^{1} tʃɻ̩^{3}/ 'a new sheet of paper' (4.17a) vs. /ɕan^{4} tʃɻ̩^{3}/ 'letter paper' (4.17b), respectively. All tonal contours produced in the isolated form are held. This observation accords with the report on tonal variations by Xiang (1960). But tonal coarticulation does seem to exist, which requires quantitative data with more tokens produced by multiple speakers.

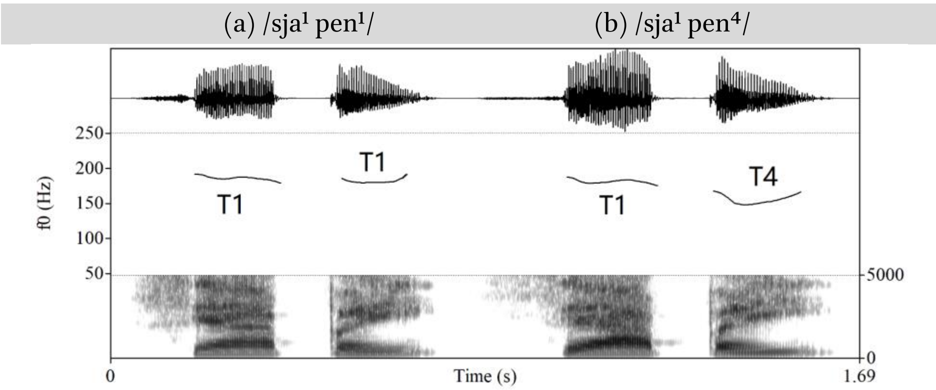

Figure 4.16 Waveforms, *f0* tracks, and spectrograms of (a) /sja^{1} pen^{1}/ 'three sides' and (b) /sja^{1} pen^{4}/ 'three times'.

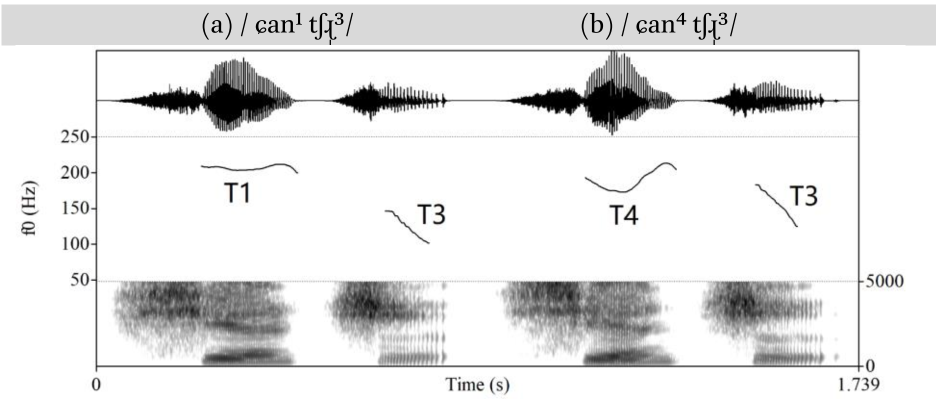

Figure 4.17 Waveforms, *f0* tracks, and spectrograms of (a) /ɕan^{1} tʃɻ̩^{3}/ 'a new sheet of paper' and (b) /ɕan^{4} tʃɻ̩^{3}/ 'letter paper'.

Worth noting is that in Shuangfeng Xiang Chinese, there are some grammatical morphemes which are phonetically realized as neutral tone (Chao, 1968; Chen & Xu, 2006). Typical examples include the plural suffix

/li^{N}/ such as in /t^{h}o^{1} li^{N}/ ‘they’ and the nominal suffix /tsɹ̩^{N}/ in /tja^{1} tsɹ̩^{N}/ ‘list’ (also see Bao, 2015 for more examples). Here, the superscript ^{N} is used to mark syllables with a neutral tone.

Various tone sandhi changes have been reported in Xiangxiang Xiang Chinese. It is argued that the tonal contours of the second syllable are reduced to three level tones, depending on tonal identity of the syllable (Zeng, 2019). These changes, however, have not been observed in Shuangfeng Xiang Chinese. It is important to conclude here that even within the same subgroup of Xiang Chinese (i.e., Old Xiang Chinese), Shuangfeng Xiang Chinese already exhibits different sandhi patterns from its neighboring dialects such as Xiangxiang Xiang Chinese. It seems that tone sandhi patterns, at least within Old Xiang Chinese, can be highly dialect-specific. Needless to say, more data and further research are needed.