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 2 The sound system of Lili Wu Chinese[7]

2.1 Introduction

Lili Wu Chinese (黎里方言) is a Wu dialect (吴语, ISO 639-3; code: wuu) spoken by approximately 38,000 people who reside in the town of Lili (黎里镇), one of the ten major towns in the Wujiang district (吴江区). The Wujiang district belongs to the prefectural-level municipality of Suzhou city (苏州市) in Jiangsu province (江苏省), the People's Republic of China. It is located at the juncture area of the city of Shanghai (上海市), the city of Suzhou, and the city of Jiaxing (嘉兴市), as shown in Figure 2.1.

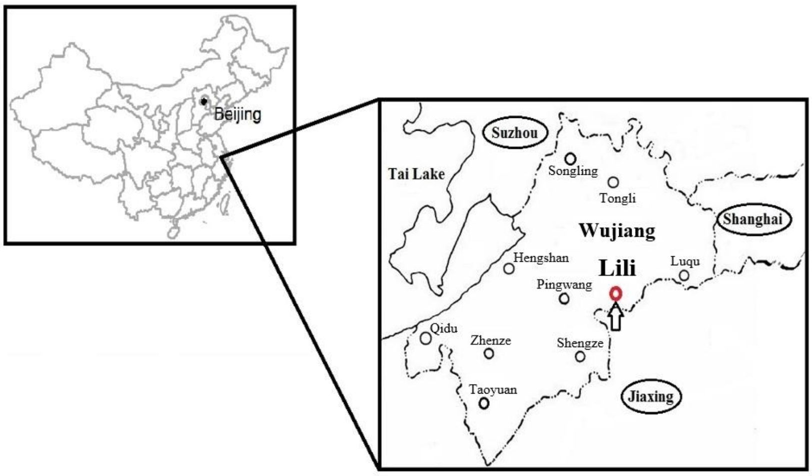


Figure 2.1 Map of the Wujiang dialects (modified based on the map in Ye, 1983).

Lili Wu Chinese is commonly considered to belong to the Suhujia dialect cluster (苏沪嘉小片), which in turn is classified as a member of the Tai Lake subgroup (太湖片) of the Northern Wu dialect group (北部

[7] A version of this chapter has been accepted for publication by *Journal of the International Phonetic Association*. Shi, M., & Chen, Y. (forthcoming). Lili Wu Chinese.

吴语), a Sinitic branch within the Sino-Tibetan family (Wurm et al., 1987). The dialect is famous for the so-called 'aspiration-induced tonal split' phenomenon (ATS), which refers to the lowering of *f0* contours after voiceless aspirated obstruents in certain tonal contexts. Lili Wu Chinese has therefore attracted much attention over the last six decades, which led to a handful of descriptions not only on the dialect but also on its closely-related dialects in the Wujiang area, which appear to have similar aspiration-induced tonal splits (see Section 2.2 for further details on lexical tones and ATS). Perhaps because of this salient tonal-split feature in the dialect, much less attention has been paid to the segmental properties of Lili Wu Chinese in the existing literature.

The description is mainly accompanied by recordings of a sixty-eight-year-old male native speaker, who was born in 1948 and raised in Lili town. All acoustic data presented in this description were elicited from this consultant. The consultant spent most of his life living in Lili and speaking Lili Wu Chinese, except for the three years attending a college in a nearby city. According to his self-report, he can speak accented Standard Chinese and limited Shanghainese when the situation requires him to do so, but he speaks only Lili Wu Chinese at home.

2.2 Lexical tones and aspiration-induced tonal split (ATS)

2.2.1 Lexical tones

There are eight lexical tones in Lili Wu Chinese. Plotted in Figure 2.2 are the *f0* contours of the example morphemes listed in Table 2.1, labeled as T1 to T8, respectively. Generally speaking, lexical tones marked as odd numbers start within a higher *f0* range (above 160 Hz, high register hereafter), while those marked as even numbers start within a lower range (under 160 Hz, low register hereafter). T1 (black solid) has a level *f0* contour within a high register (high–level) while T2 (dark gray solid) is a low-register-rising tone (low–rising). T3 (black round dot) starts within the high register and falls (high–falling). T4 (dark gray round dot) is a low-register-level tone (low–level). T5 (black square dot) has a convex contour which starts at the high register, falls and ends with a slight rise (high–dipping). T6 (dark gray square dot) is realized with a similar *f0* contour to that of T5 but starts at the low register (low–dipping). Both T7 (black dash-dotted) and T8 (dark gray dash-dotted) are associated with syllables that have a much shorter duration than the other tone-bearing syllables. T7 starts within the high register and despite the slight falling contour, sounds like a high-register-level tone (short–high–level). T8 is a low-register-level tone (short–low–level).

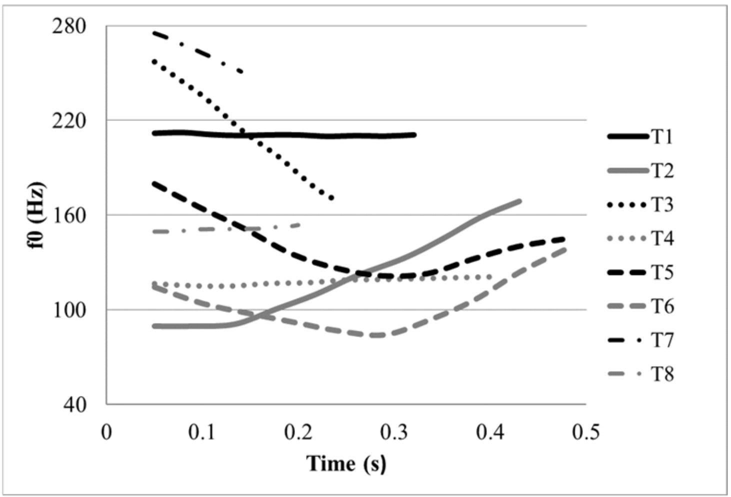


Figure 2.2 *f0* contours of the lexical tones in Lili Wu Chinese.

Table 2.1 Examples of the lexical tones in Lili Wu Chinese.

Lexical tone	Tonal contour	Example	Orthography	Gloss
T1	high–level	/tʊŋ¹/	东	'east'
T2	low–rising	/dʊŋ²/	铜	'copper'
T3	high–falling	/tʊŋ³/	懂	'to understand'
T4	low–level	/dʊŋ⁴/	动	'action'
T5	high–dipping	/tʊŋ⁵/	冻	'to freeze'
T6	low–dipping	/dʊŋ⁶/	洞	'cave'
T7	short–high–level	/tʊʔ⁷/	督	'supervision'
T8	short–low–level	/dʊʔ⁸/	读	'to read'

These eight lexical tones exhibit interesting co-occurrence patterns with both the onset and coda. Lili Wu Chinese features the three-way laryngeal contrast in obstruents, known as voiceless unaspirated, voiceless aspirated and voiced, respectively (see Section 2.3 for more details). Syllables with voiceless unaspirated onsets only allow high-register tones (T1, T3, T5, and T7); while voiced onsets co-occur with low-register tones (T2, T4, T6, and T8). T1 to T6 only co-occur with open syllables or syllables with a nasal coda (舒声) and are therefore also known as smooth/non-checked tones (Middle Chinese [MC] *Ping–Shang–Qu*; 平–上–去; level–rising–departing), while the T7 and T8 only co-occur with closed syllables with a glottal coda /ʔ/ (促声) and are known as abrupt/checked tones (MC *Ru*; 入; entering). MC is a sound system reconstructed based mainly on written records such as rhyme dictionaries (see a comprehensive introduction in Norman, 1988). In the development of the tonal system from MC to the modern Wu dialects, the four-way tonal contour contrast of MC (i.e., *Ping*, *Shang*, *Qu*, and *Ru*) split into a new dual-register, eight-tone system, conditioned by onset consonants, which is evident in modern Wu varieties (Pulleyblank, 1978; Ting, 1984; Norman, 1988).

In the vast majority of Northern Wu dialects such as Shanghainese (Chen & Gussenhoven, 2015), both voiceless unaspirated and aspirated onsets condition high-register tones, leaving voiced onsets to co-occur with low-register tones. What makes Lili Wu Chinese interesting is the effect of obstruent aspiration on lexical tonal realization, as exemplified by /tʰʊŋ^{1}/ 'unblocked', /tʰʊŋ^{4}/ 'to unify', /tʰʊŋ^{6}/ 'ache', and /tʰʊʔ^{8}/ 'baldy'. Their *f0* contours are plotted in Figure 2.3 (labeled as T1-A, T3-A, T5-A, and T7-A where A indicates voiceless aspirated onsets), in comparison to the *f0* contours of the presumably same lexical tones realized after voiceless unaspirated onsets (indicated with U). Except for T1 (i.e., T1-U vs. T1-A), we see a clear *f0*-lowering effect in syllables with voiceless aspirated onsets. This lowering effect, as if a split of the same tone into two as a function of voiceless unaspirated vs. aspirated onsets, is known as ATS.

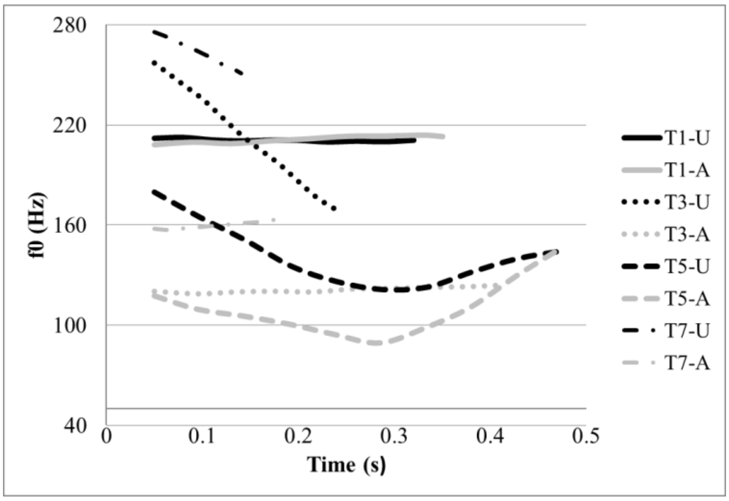


Figure 2.3 *f0* contours of the lexical tones of the words with voiceless unaspirated onsets (U, black) and those with voiceless aspirated (A, light gray) onsets.

2.2.2 New analysis of ATS

Perhaps due to ATS, the tonal inventory in Lili Wu Chinese has been a point of debate in recent decades. To my knowledge, there are at least eleven descriptive works focusing on ATS in Lili Wu Chinese (Chao, 1928;

Ye, 1983; Zhang & Liu, 1983; Shi, 1992; Qian, 1992; Shen, 1994; Wang, 2008, 2010; Xu, 2009; Hirayama, 2010; Yanhong Xu, 2013).

Chao (1928) is the first to note this aspiration induced tonal-split phenomenon. Subsequent representative studies include Shi (1992), the first which reported acoustic data from two native speakers of Lili Wu Chinese; and Wang (2010), which is the first attempt at a comprehensive description of the phonological system of Lili Wu Chinese. However, researchers differ greatly in their treatment/interpretation of ATS. The main debate lies in the question of whether the *f0* contours of lexical tones after aspirated onsets have been merged with those after voiced onsets or emerged as distinct tonal categories independent of the existing eight tonal categories. As can be seen in Table 2.2, the only agreement in the existing literature on ATS is that it doesn't occur (marked as '–' in Table 2.2) in the MC *Ping* category. With respect to the other three MC tonal categories, researchers differ greatly in their treatment/interpretation of the tonal-split phenomenon. Chao (1928) reports an unmerged distribution (marked as '+' in Table 2.2) in the MC *Qu* and *Ru* categories. Represented by Ye (1983), a lot of researchers claim a three-way distinction of lexical tones in the MC *Shang*, *Qu*, and *Ru* categories. However, Shen (1994) argues for a merging of lexical tones following aspirated and voiced onsets (marked as '++' in Table 2.2) in the three MC categories. Wang (2008, 2010) reports a merging in *Shang* and *Qu*, and Yanhong Xu (2013) maintains a merging in *Qu* only.

How should we understand the discrepancies existing among these studies? Given that there is an over-80-year interval between the earliest and the latest studies, it is reasonable to assume that different proposals indeed reflect the actual changes of the lexical tonal system of Lili Wu Chinese. For example, if comparing the study by Chao (1928) to the study by Ye (1983), one can argue that it is not until the 1980s that ATS happened in the *Shang* category. However, it is worth pointing out that not all discrepancies can be explained by such diachronic interpretations. For instance, as we can see from the studies of Qian (1992) and Shen

(1994), both were finished in the early 1990s by recruiting the same consultants. However, Shen (1994) proposes an opposite treatment compared to Qian (1992). It is important to note that these previous studies typically explored ATS based on impressionistic descriptions (e.g., Chao, 1928), or with data from a very limited number of speakers (e.g., Shi, 1992 for one male and one female speaker; Shen, 1994 for two young speakers). The phonetics and phonology of the tonal system of Lili are still seriously under-studied, which motivated a large-scale experimental study with statistical analyses of the *f0* contours to know the general group-level patterns of *f0* contours of the lexical tones in Lili Wu Chinese.

Table 2.2 Treatments of the lexical tones after voiceless aspirated onsets in Lili Wu Chinese. –: no tonal split; +: a three-way distinction; ++: a merger has happened.

Ping	*Shang*	*Qu*	*Ru*	Study
–	–	+	+	Chao (1928)
–	+	+	+	Ye (1983); Zhang & Liu (1983); Shi (1992); Qian (1992); Xu (2009); Hirayama (2010)
–	++	++	++	Shen (1994)[8]
–	++	++	+	Wang (2008, 2010)
–	+	++	+	Yanhong Xu (2013)

The results of a large-scale experimental study across three generations (see Chapter 3) show although *f0* contours after aspirated onsets show a trend of slightly higher *f0*, they can pattern more like those after voiced onsets, resulting in the merger of the *f0* contours of T3-A and T4 (MC *Shang*), T5-A and T6 (MC *Qu*), T7-A and T8 (MC *Ru*), respectively. However, there are comparable *f0* contours after voiceless aspirated (T1-A) and voiceless unaspirated (T1-U) onsets (MC *Ping*), both of which are

[8] According to the acoustic analyses in Shen (1994), *f0* contours after voiceless aspirated onsets in the MC *Ru* category should be transcribed as /33/. However, he finally argues that the tone after the aspirated onset has merged to that after its voiced counterpart and transcribes both as /23/ for the sake of systematicness.

realized within the high *f*0 register. It suggests that the lexical tonal system of Lili Wu Chinese includes two level tones (high–level T1 and low–level T4), one falling tone (high–falling T3), one rising (low–rising T2) and two dipping tones (high–dipping T5 and low–dipping T6). For short syllables with a glottal coda, two level tones are identified (short–high–level T7 and short–low–level T8). The numerical representations of the eight lexical tones and their co-occurrence patterns with onsets are provided in Table 2.3. Here, the tonal transcription system developed by Chao (1930) was adopted. In this system, 5 indicates the highest end of a speaker's pitch range into levels and 1 the lowest. The results confirm the treatment in Shen (1994).

Table 2.3 Numerical representations of the lexical tones in Lili Wu Chinese. A single number refers to cases where the tone-carrying syllables have a short duration and only co-occur with the glottal coda /ʔ/.

Voiceless unaspirated (U)	T1 high–level 44	T3 high–falling 53	T5 high–dipping 423	T7 short–high–level 5
Voiced (D)	T2 low–rising 13	T4 low–level 22	T6 low–dipping 213	T8 short–low–level 3
Voiceless aspirated (A)	T1 high–level 44	T4 low–level 22	T6 low–dipping 213	T8 short–low–level 3

T1 can co-occur with both voiceless onsets (i.e., unaspirated and aspirated). T3, T5, and T7, on the other hand, can only co-occur with voiceless unaspirated onsets. The three low-register tones (T4, T6, and T8) are licensed by both voiceless aspirated and voiced onsets, while T2 is only allowed after voiced onsets. It is important to note that the co-occurrence pattern (i.e., voiceless onsets co-occurring with high-register tones, while voiced onsets with low-register ones), which is commonly observed in most Northern Wu dialects, falls apart in Lili Wu Chinese where voiceless aspirated onsets can co-occur with low-register tones.

In addition, it is worth noting that in Lili Wu Chinese, sonorants (i.e., nasals and liquids) mainly co-occur with low-register tones and share the same tonal pattern with voiced plosives. A set of words initialed with nasals can also co-occur with high-register tones,[9] such as /mu³/ [məʊ³] 'bound morpheme for the literary address of mother'. With respect to fricatives, since there is only a two-way laryngeal distinction (i.e., voiceless vs. voiced), voiceless fricatives co-occur with high-register tones while their voiced counterparts with low-register ones.

2.3 Consonants

Lili Wu Chinese has 28 consonants. Corresponding keywords/bound morphemes are provided below the consonant chart.

[9] In Northern Wu Chinese, there is a group of shared words initialed with nasals (mainly including 妈/母/美/蛮/囡 'bound morpheme for mother/bound morpheme for the literary address of mother/beautiful/very/little darling') that can co-occur with high-register tones. Such a co-occurrence is argued to be relevant to the affective function of high tones (Zhu, 2004a).

	Bilabial			Labio-dental		Alveolar			Alveolo-palatal			Palatal	Velar			Glottal
Plosive	pʰ	p	b			tʰ	t	d					kʰ	k	g	ʔ
Affricate						tsʰ	ts	dz	tɕʰ	tɕ	dʑ					
Nasal			m					n							ŋ	
Fricative				f	v	s		z	ɕ							h
Approximant			w									j				
Lateral approximant								l								

pʰ	pʰε^{1}	攀	‘to climb’	tʰ	tʰε^{1}	瘫	‘paralysis’					kʰ	kʰε^{1}	开	‘to open’
p	pε^{1}	班	‘class’	t	tε^{1}	单	‘single’					k	kε^{1}	该	‘ought to’
b	bε^{2}	赔	‘to compensate’	d	dε^{2}	台	‘platform’					g	gε^{6}	隑	‘to lean against’
								tsʰ	tsʰε^{1}	猜	‘to guess’	tɕʰ	tɕʰi̟^{1}	欺	‘to deceive’
								ts	tsε^{1}	灾	‘disaster’	tɕ	tɕi̟^{1}	鸡	‘chicken’
								dz	dzε^{2}	随	‘at random’	dʑ	dʑi̟^{2}	奇	‘strange’
m	mε^{6}	妹	‘younger sister’	n	nε^{2}	难	‘difficult’					ŋ	ŋε^{2}	颜	‘color’
f	fε^{1}	翻	‘to turn over’	s	sε^{1}	三	‘three’	ɕ	ɕi̟^{1}	稀	‘rare’	h	hε^{1}	虚	‘turgescence’
v	vε^{2}	烦	‘to bother’	z	zε^{2}	馋	‘greedy’					ʔ	kʌʔ^{7}	割	‘to cut’
w	wε^{2}	还	‘to return’	l	lε^{2}	蓝	‘blue’					j	jɘ̟o^{1}	优	‘superior’

Lili Wu Chinese features the three-way laryngeal contrast in obstruents, known as voiceless unaspirated, voiceless aspirated and voiced, respectively. This three-way contrast is a prominent feature of the Northern Wu dialects (Chao, 1967). The three-way laryngeal contrast, however, has different phonetic manifestations in initial vs. medial positions within a word. Generally speaking, in initial position, these obstruents feature a two-way distinction (i.e., short lag vs. long lag) in terms of voice onset time (VOT). As exemplified by the triplet (a) /tɑ^1/ 'knife', (b) /t^{h}ɑ^1/ 'billow', and (c) /dɑ^2/ 'peach' in Figure 2.4, both the voiceless unaspirated (/t/ as in /tɑ^1/, 2.4a) and voiced onsets (/d/ as in /dɑ^2/, 2.4c) have a short-lag VOT (15 ms and 19 ms), in contrast to the aspirated onset (/t^h/ as in /t^{h}ɑ^1/, 2.4b) which is realized with a long-lag VOT (89 ms).

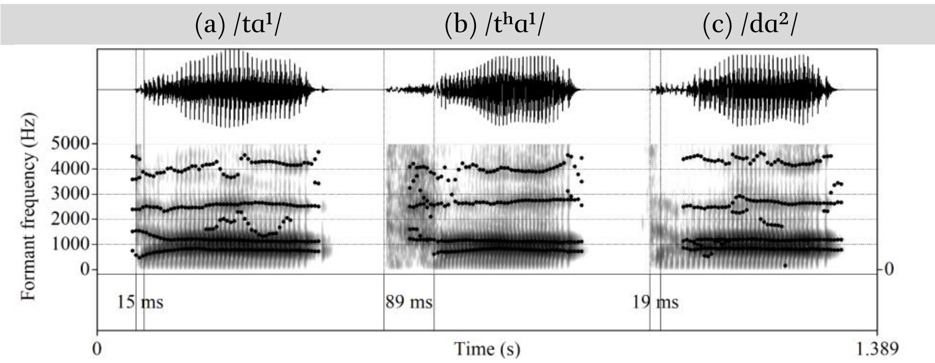


Figure 2.4 Waveforms and spectrograms of (a) /tɑ^1/ 'knife', (b) /t^{h}ɑ^1/ 'billow', and (c) /dɑ^2/ 'peach'. Intervals indicate VOT values.

Table 2.4 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 90 monosyllabic morphemes consisting of 30 sets of triplets. As shown in Table 2.4, irrespective of the place of articulation, the voiceless aspirated plosives have longer VOT values than the other two counterparts. The voiceless unaspirated and voiced plosives have similar short VOT values.

Table 2.4 VOT of unaspirated vs. aspirated vs. voiced plosives in different places of articulation in Lili Wu Chinese, based on 90 monosyllabic morphemes with plosive onsets.

	Bilabial (10 triplets)		Alveolar (10 triplets)		Velar (10 triplets)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Voiceless unaspirated	10 ms	6 ms	12 ms	6 ms	16 ms	7 ms
Voiceless aspirated	94 ms	16 ms	89 ms	18 ms	102 ms	19 ms
Voiced	14 ms	5 ms	17 ms	9 ms	23 ms	7 ms

In medial position, as shown in Figure 2.5, the voiced onset /d/ (2.5c) is realized with noticeable voicing throughout the closure (–61 ms), leading to a three-way laryngeal distinction.

(a) /fi̟^{1} tɑ^1/ (b) /pu^1 tʰɑ^1/ (c) /æŋ^1 dɑ^2/

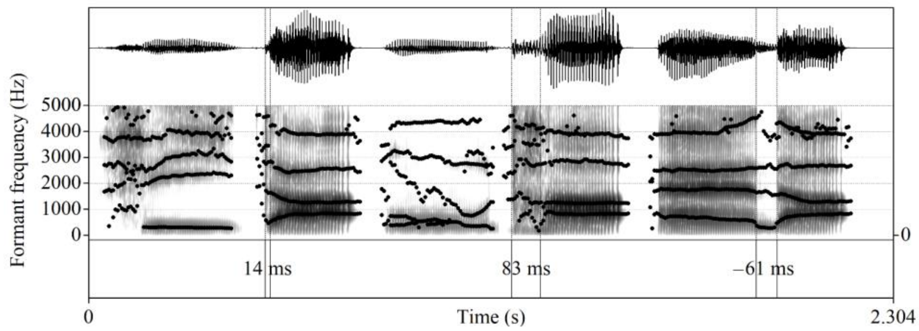


Figure 2.5 Waveforms and spectrograms of (a) /fi̟^{1} tɑ^1/ 'flying knife', (b) /pu^1 tʰɑ^1/ 'great waves', and (c) /æŋ^1 dɑ^2/ 'cherry'. Intervals indicate VOT values.

In addition, in initial position, these obstruents also vary in their phonation from clearly modal (voiceless unaspirated), aspirated (voiceless aspirated), to breathy (voiced) (See Chapter 3 for more information). In Shanghainese, a Northern Wu dialect closely related to Lili Wu Chinese, there are other phonetic properties to signal the three-way laryngeal contrast in both initial and medial positions (e.g., Shen et al., 1987 on closure duration; Ren, 1992 on transillumination data; Gao, 2015 on motion-capture-system data; see also a review in Chen, 2011). Impressionistically speaking, Lili Wu Chinese behaves similarly to Shanghainese. Needless to say, more research is needed to examine if these properties also function in Lili Wu Chinese.

The alveolar consonants /t^h t d/ tend to be denti-alveolar and /s z $ʦ^h$ ʦ ʣ/ apical alveolar, with /s z/ having a contact area slightly further forward than /$ʦ^h$ ʦ ʣ/. The alveolo-palatals /$ʨ^h$ ʨ ʥ ɕ/ are produced via contacting the region between the alveolar ridge and the post-alveolar palatal, with the fricative /ɕ/ having a slightly more frontal contact area.

Fricatives have the voiceless vs. voiced two-way laryngeal contrast. Similar to the plosives and affricates, in initial position, their phonatory states vary from clearly modal in the voiceless ones to slightly breathy in the voiced ones. In medial position, the voiced category is realized with vigorous voicing, leading to a two-way contrast in terms of VOT. It is worth noting that the voicing contrast in fricatives is also signaled via their durational differences, similar to what have been reported for voicing contrast in English fricatives (e.g., Cole & Cooper, 1975), as shown in the following pairs: /f/ (/fu^1/ 'husband') vs. /v/ (/vu^2/ 'to support somebody with one's hand'); /s/ (/$sɛ^1$/ 'three') vs. /z/ (/$zɛ^2$/ 'greedy'). Figure 2.6 illustrates the acoustic realization of /f/ in /fu^1/ (2.6a) and /v/ in /vu^2/ (2.6b). Although neither is realized with regular vocal pulses (i.e., phonetically voiceless), the fricative duration of /f/ (131 ms; 29% of the total duration) is almost 2.4 times longer than that of /v/ (56 ms; 12% of the total duration).

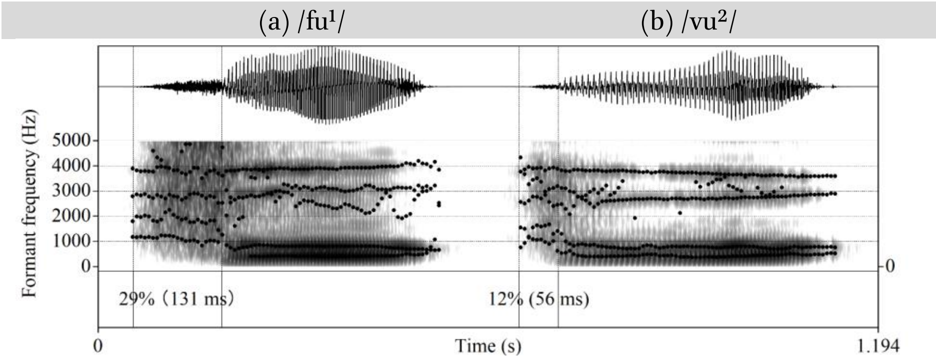


Figure 2.6 Waveforms and spectrograms of (a) /fu^1/ 'husband' and (b) /vu^2/ 'to support somebody with one's hand'. Within a syllable, the percentage of the frication duration with absolute values (ms) between parentheses is indicated.

The durational differences of the voicing contrast in fricatives have not been well investigated in previous studies of Northern Wu Chinese. To further confirm these observations, ten minimal pairs for each minimal set of voicing contrast were elicited. All stimuli were lexemes of relatively high familiarity, as confirmed by the consultant. Both the absolute duration of the frication and the percentage of the frication duration over the whole syllable duration were calculated. The fricative duration was measured from the onset of clear frication noise to the first periodic cycle of the vowel. The results in Table 2.5 show that the percentage of the frication duration of voiceless onsets is significantly greater than that of their voiced counterparts, confirmed by the results of the independent samples *t*-tests (one-tailed) for each pair.

Table 2.5 Average percentage of the frication duration and the independent samples *t*-test results (one-tailed) for each pair of voiceless vs. voiced. The Bonferroni method was adopted for *p*-value adjustments. Parentheses indicate absolute values of the average duration and the standard deviation (ms).

Voiceless vs. voiced	Frication duration	*t*-test
f vs. v	27% (M = 124 ms, SD = 7 ms) vs. 12% (M = 59 ms, SD = 4 ms)	$t(18) = 10.79$, $p < .001$
s vs. z	36% (M = 149 ms, SD = 8 ms) vs. 28% (M = 117 ms, SD = 6 ms)	$t(18) = 2.74$, $p < .01$

Lili Wu Chinese maintains the distinction of affricate /ʣ/ and fricative /z/, as evidenced in the minimal pair /dzɛ^{2}/ ‘at random’ and /zɛ^{2}/ ‘talent’. In both phrase initial and medial positions, there is a clear frication duration difference between the fricative /z/ and the affricate /dz/, as shown in Figure 2.7 (initial position: /dz/ 21% vs. /z/ 30%) and Figure 2.8 (medial position: /dz/ 10% vs. /z/ 23%).

(a) /dzɛ^{2}/ (b) /zɛ^{2}/

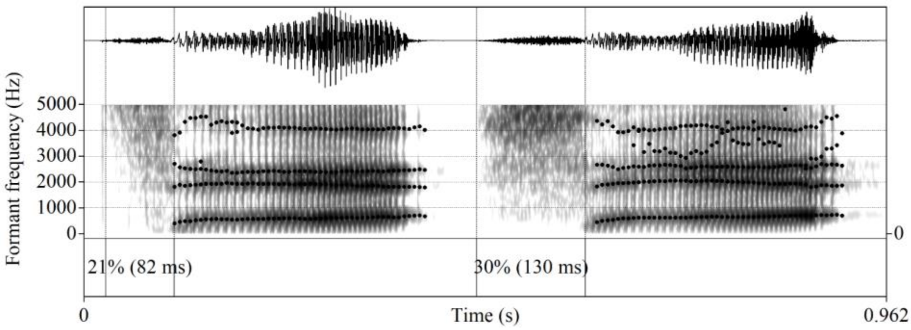


Figure 2.7 Waveforms and spectrograms of (a) /dzɛ^{2}/ ‘at random’ and (b) /zɛ^{2}/ ‘talent’. Within a syllable, the percentage of the frication duration with absolute values between parentheses is indicated.

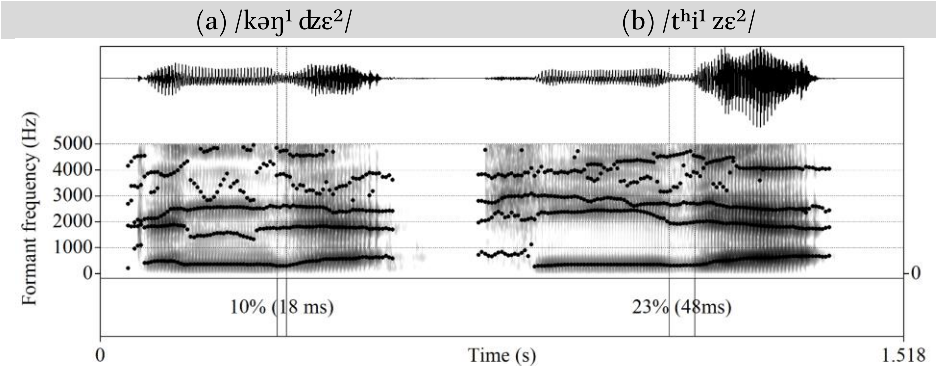


Figure 2.8 Waveforms and spectrograms of (a) /kəŋ^{1} dʑɛ^{2}/ 'to follow' and (b) /t^{h}i^{1} zɛ^{2}/ 'genius'. Within a phrase, the percentage of the frication duration with absolute values (ms) between parentheses is indicated.

In addition, /dz/ and /z/ can sometimes be distinguished according to the degree of formality with which the lexical item is produced. Chao (1928) first observes the variation between /dz/ and /z/ in some morphemes. Very recently, on the basis of recordings of Shanghainese in 1853 by a British Protestant missionary – Joseph Edkins (1823–1905), Chen (2015) argues that the affricate phoneme /dz/ was borrowed from Hangzhou Chinese (i.e., the Hangzhou dialect spoken in the city center of Hangzhou (杭州市区)), and tends to be preserved in literary lexical items. In Lili Wu Chinese, for example, the morpheme /dzɛ^{4}/ 'crime' is pronounced as /dzɛ^{4}/ in the literary lexical item /dzɛ^{4} ɪŋ^{2}/ 'crime', but as /zɛ^{4}/ in /zɛ^{4} ku^{5}/ 'pitiful' – a more colloquial lexical item. However, /dz/ and /z/ are sometimes in free variation for the same lexical item, as exemplified in 'groceries' (/dzaʔ^{8} hu^{5}/ vs. /zaʔ^{8} hu^{5}/). This finding may imply that in Lili Wu Chinese, /dz/ and /z/ are undergoing a merger at the lexical level.

Last but not least, the glottal plosive /ʔ/ only appears in coda position as a phoneme and co-occurs with short syllables as in /paʔ^{7}/ 'hundred'. Phonetically speaking, the [ʔ] segment can also be observed at the beginning of onsetless syllables with the high-register tones (i.e., T1, T3, T5, and T7) (see Section 2.8 for further details on onsetless syllables).

2.4 Sonorants

/n l/ are typical laminal alveolar. The alveolar nasal /n/ is palatalized before high front segments (i.e., /i i̟ y j/), as in /ni^2/ [ɲi^2] 'year' and /njɛ^6/ [ɲjɛ^6] 'to read'. Labial and velar nasals can form syllable nuclei as in /ŋ̍^{4}/ 'five' and /m̩^{4}/ 'parcel of land'. These two syllabic nasals can be found in many Southern Chinese dialects (i.e., Wu, Min, Hakka, Xiang, and Yue) but are relatively rare in dialects belonging to the Mandarin family (Shen, 2006). In addition, /ŋ/ occurs as a nasal coda as well, but its acoustic realization varies according to the preceding vowel. After a front vowel, the nasal coda acquires the anterior feature and sounds like [n] (as in /zɪŋ^2/ [zɪn^2] 'to look for' and /tɕʏŋ^1/ [tɕʏn^1] 'army'), as in contrast to a non-front vowel (as in /dzəŋ^1/ 'deity' and /dʊŋ^2/ 'copper'). Following the treatment of Chen and Gussenhoven (2015) for Shanghainese, an underlying /ŋ/ in coda position is posited.

There are two glides /j/ and /w/ in Lili Wu Chinese. Glides are typically defined as vowel-like segments that function as consonants and belong to the approximant class (Ladefoged & Maddieson, 1996). In Lili Wu Chinese, /j/ and /w/ differ from the corresponding vowels (i.e., /i/ and /u/) in that both tend to be produced with a narrower constriction of the vocal tract indicated via lower F1 values. Following Maddieson and Emmorey (1985), I compared mean F1 of the beginning interval (i.e., 50 ms) of /j/ (/jə̯o^1/ 'surname, Ou') with /i/ (/i^1/ 'smoke') and /w/ (/wɛ^2/ 'to return') with /u/ (/u^2/ 'river'), respectively. The results showed that the F1 values of /j/ (265 Hz) and /w/ (314 Hz) were lower than the corresponding vowels (/i/: 271 Hz; /u/: 354 Hz). Existing descriptions of Lili Wu Chinese such as Wang (2010) have typically posited high vowels /i u/ instead of glides /j w/ in words like /jə̯o^1/ and /wɛ^2/ (/iəu^1/ and /uᴇ^2/ in Wang, 2010),[10] despite the consensus among sinologists that they are glides. The approximants /j

[10] Sinologists frequently use /ᴇ/ to describe the phoneme between /e/ and /ɛ/, which however, does not exist in the IPA.

w/ have been adopted to transcribe the sounds. Needless to say, large-scale and well-designed studies are needed to investigate further the differences in glides vs. vowels in Chinese dialects. Additionally, note that before rounded vowels /o/ and /ø/, /j/ is realized as [ɥ] as in /joʔ^{8}/ [ɥoʔ^{8}] 'bath' and /jø^{2}/ [ɥø^{2}] 'rounded'. Because of the complementary distribution, /ɥ/ is treated as a context-specific (i.e., before /o/ and /ø/) variation of /j/.

A controversial issue is whether it is necessary to posit /j/ after an alveolo-palatal affricate or fricative onset (i.e., /tɕʰ tɕ dʑ ɕ/) in Wu Chinese (see a brief discussion in Chen & Gussenhoven, 2015). Historically, these alveolo-palatal onsets are commonly believed to develop from the velar or glottal onsets (i.e., /kʰ k g h/) due to the palatalization process triggered by the following high front segments (e.g., Wang, 1985: 394). Synchronically, there is no contrast between /tɕʰ tɕ dʑ ɕ/ and /tɕʰj tɕj dʑj ɕj/ in Lili Wu Chinese. More remarkably, the transition from the alveolo-palatal affricate to the following vowel is rather brief. Figure 2.9 illustrates the different transitional characteristics among /tɑ^{1}/ 'knife' (2.9a), where there is no glide, /tjɑ^{1}/ 'marten' (2.9b) and /tsjɑ^{1}/ 'scorched' (2.9c), where there is commonly recognized presence of /j/, and /tɕɑ^{1}/ 'to converge' (2.9d), where I propose absence of /j/. Adapted the method of Chitoran (2002), I marked the beginning of the transition at the start of the sonorant part (i.e., glide or vowel) and the end of the transition as the turning point from a falling F2 to an F2 steady-state, before it falls consistently less than 20 Hz. The F2 values were automatically measured in Praat with a window length of 5 ms. Note that one would have expected a much more stable realization of /j/ with a longer transition from /tɕ/ to /ɑ/ if one assumed the presence of a glide /j/ following /tɕ/. These observations motivated me not to posit an underlying /j/ after alveolo-palatal onsets (following the analysis of Chen & Gussenhoven, 2015 for Shanghainese). But it is worth stressing the importance of further experimental studies to investigate the phonological status and phonetic realization of /j/ after alveolo-palatals in Lili Wu Chinese as well as other Chinese dialects.

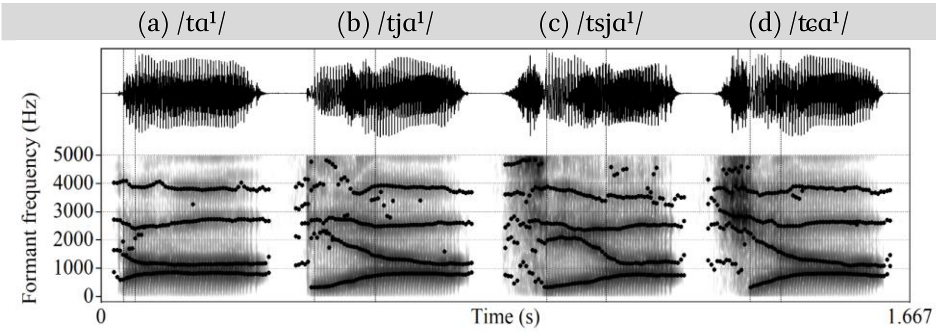


Figure 2.9 Waveforms and spectrograms of (a) /tɑ^{1}/ 'knife', (b) /tjɑ^{1}/ 'marten', (c) /tsjɑ^{1}/ 'scorched', and (d) /tɕɑ^{1}/ 'to converge' respectively. Within each syllable, the transition from the end of the preceding consonant to the time that the F2 converges toward the value of /ɑ/ is indicated.

2.5 Vowels

Vowels in open syllables are plotted in Figure 2.10, and those in closed syllables in Figure 2.11. Except for /u/ (5), ten tokens of each vowel were used. The vowel ellipses were calculated based on the covariance of the tokens by following Pols et al. (1973). Formant values were automatically measured via VoiceSauce (Shue et al., 2011) using the Snack method. Formant values were converted from Hertz to Bark using the built-in formula[11] in Praat but were still indicated in corresponding Hertz values[12] in both plots for ease of comparison to the existing literature.

In open syllables, there are nine monophthongs (2.10a) and one diphthong (2.10b), as plotted in Figure 2.10. These nine monophthongs of Lili Wu Chinese (/i y ɨ̞ ɛ ø u o ɔ ɑ/) constitute a four-way distinction (i.e., close, close-mid, open-mid, and open) in height and a two-way distinction

[11] $\text{Bark} = 7 \cdot \ln(\text{Hertz} / 650 + \sqrt{1 + (\text{Hertz} / 650)^2})$.

[12] $\text{Hertz} = 650 \cdot \sinh(\text{Bark} / 7)$.

(i.e., front and back) in backness. /i y/ contrast in roundness. In addition, there is one diphthong occurring in open syllables, with /ɘ̯o/ gliding towards the back.

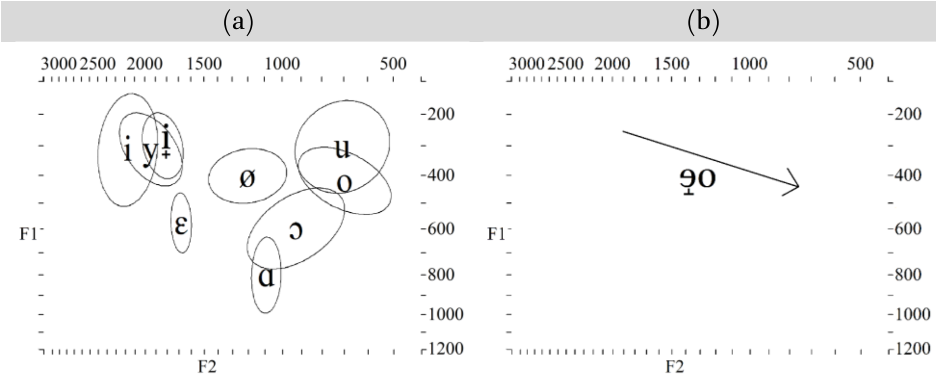


Figure 2.10 Relative F1/F2 formant values of monophthongs and the diphthong in open syllables (a: for monophthongs; b: for the diphthong). The arrow demonstrates the trajectory of the gliding.

Monophthongs in closed syllables are plotted in Figure 2.11, where six (/ɪ ʏ æ ə ʊ ɑ/) occur in syllables closed by a nasal coda (2.11a) and four (/ɪ a ʊ ʌ/) in syllables closed by a glottal coda (2.11b). Compared to the vowels in open syllables, the number of vowels in closed syllables is largely reduced and so is their acoustic vowel space. Generally speaking, vowels in closed syllables are more central and lower than those in open syllables. Following Chen and Gussenhoven (2015), the same set of symbols (i.e., /ɪ/ and /ʊ/) for monophthongs followed by a nasal coda and those by a glottal coda was adopted, although their articulations do differ.

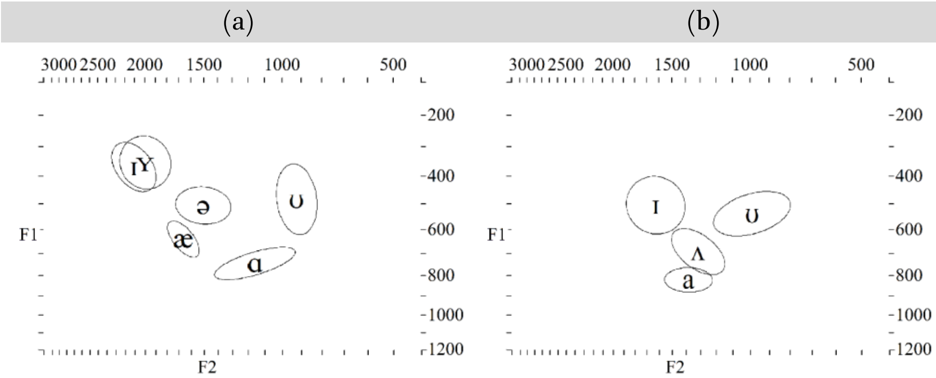


Figure 2.11 Relative F1/F2 formant values of monophthongs in closed syllables (a: for vowels in syllables closed by a nasal; b: for vowels in syllables closed by a glottal coda).

Monophthongs and diphthong in open syllables

i	ti¹	颠	'jolt'	u	u¹	乌	'crow'
y	ɕy¹	虚	'unreal'	ø	kø¹	肝	'liver'
i̞	ti̞¹	低	'low'	o	ko¹	瓜	'melon'
ɛ	kɛ¹	该	'ought to'	ɔ	zɔ²	柴	'firewood'
ɘ̯o	jɘ̯o¹	优	'superior'	ɑ	kɑ¹	高	'high'

Monophthongs in closed syllables

ɪ	tɕɪŋ¹	金	'gold'	ʊ	kʊŋ¹	公	'public'
	tɕɪʔ⁷	急	'hurry'		kʊʔ⁷	郭	'surname, Guo'
ʏ	tɕʏŋ¹	军	'army'	ə	kəŋ¹	跟	'to follow'
æ	kæŋ¹	庚	'age'	ʌ	kʌʔ⁷	革	'to reform'
a	kaʔ⁷	夹	'to clip'	ɑ	kɑŋ¹	刚	'just now'

Lili Wu Chinese presents an interesting case of fricative vowel, as illustrated in Figure 2.12 which plots the spectrograms of the minimal pair /i/ in /ti³/ 'dot' (2.12a) and /i̞/ in /ti̞³/ 'bottom' (2.12b). The F2 of /i/ (2399 Hz) is higher than the F2 of /i̞/ (2009 Hz). Perceptually, a striking difference between /i/ and /i̞/ is the frication present in /i̞/ (Chao, 1928; Wang, 1987; Hu, 2007; Ling, 2007, 2011). Figure 2.13 exhibits narrow band spectrograms of /ti³/ (2.13a) and /ti̞³/ (2.13b). Harmonics can be clearly identified in /ti^3/ but are not in /ti̟^{3}/, especially in the frequency bands above 2 kHz. Furthermore, there is a substantial amount of aperiodic energy in the higher frequency region, particularly above 4 kHz in /ti̟^{3}/, which suggests the presence of strong fricative noise. This observation is further confirmed by the HNR (Harmonics-to-Noise Ratio) results, with /i̟/ in /ti̟^{3}/ (8.1 dB) showing more noise than /i/ in /ti^3/ (9.8 dB).

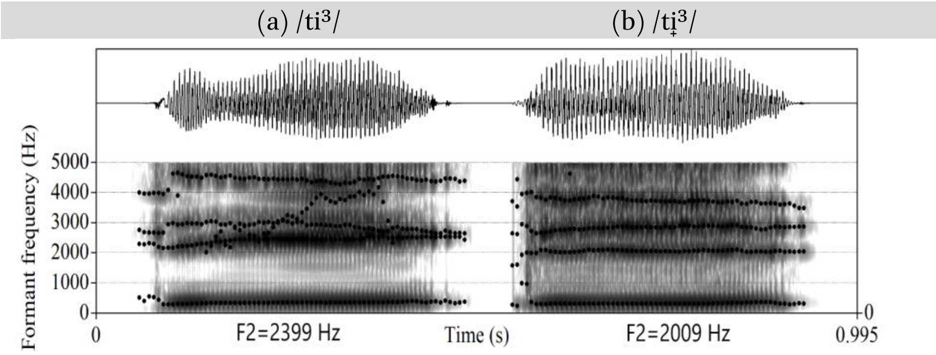


Figure 2.12 Waveforms and spectrograms of (a) /ti^3/ 'dot' and (b) /ti̟^{3}/ 'bottom'. F2 values are indicated.

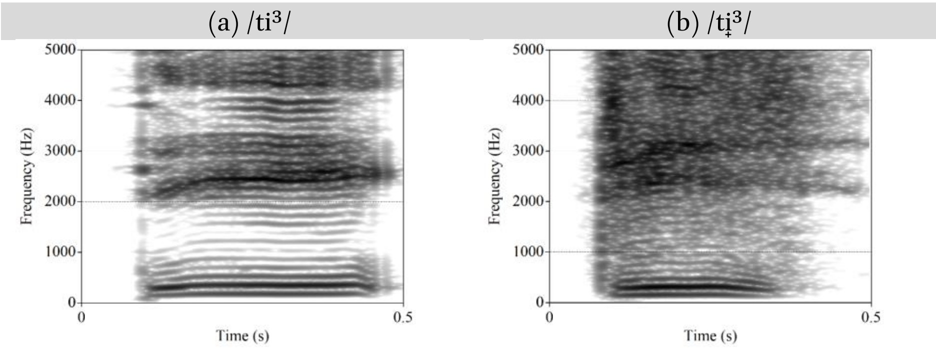


Figure 2.13 Narrow band spectrograms of (a) /ti^3/ 'dot' and (b) /ti̟^{3}/ 'bottom'.

A similar contrast has been reported in Suzhou Wu Chinese. In order to illustrate the frication, Ling (2011) adopts the symbol /i$_z$/ for the phoneme and [ʒ̻̍] (i.e., the syllabic laminal postalveolar voiced fricative) for its phonetic realization. However, this treatment is problematic. First, a subscript /$_z$/ does not meet the convention of diacritic symbols in the IPA. Second, articulatory (i.e., palatographic, linguagraphic, and electromagnetic articulographic (EMA) studies) data of Suzhou Wu Chinese have shown that the constriction of /i̟/ is located at a more anterior position (Ling, 2007, 2011; Hu & Ling, 2019) than /i/. Consequently, the lengthening of the back resonating cavity lowers the F2 of /i̟/ as argued by Ling (2011) following Stevens (1989: 10). Third, it is also worth noting that the frication in Lili Wu Chinese, compared to that in Suzhou Wu Chinese, is not consistently audible for all /i̟/ words produced by the consultant and also not as strong as that in Suzhou Wu Chinese. For instance, there is little frication in /fi̟¹/ [fᵊɨ¹] 'to fly' (which tends to be diphthongal). Given the three reasons, the symbol ̟ has been adopted to highlight the more anterior constriction of /i̟/ and the less friction. Such an articulatory gesture is also accompanied by a remarkable raising of the lower jaw in words such as /i̟¹/ 'clothes', which however, is not observed in words such as /i¹/ 'smoke'. The contrast of high front vowels is an areal phenomenon in many Chinese dialects, especially in the Jianghuai Mandarin family (江淮官话) (R. Shi, 1998; Zhu, 2004b; Zhao, 2007). Similar contrasts have also been reported in some modern African languages, such as Len Mambila (Connell, 2007) and Ring languages (Faytak & Merrill, 2015).

Both /u/ (in /u¹/ 'crow') and /o/ (in /ko¹/ 'melon') are close/close-mid back monophthongs with compressed lip rounding. The lips for /o/ are more protruding but for /u/ they are less rounded and more compressed (similar to the /u o/ contrast in Shanghainese as discussed in Chen & Gussenhoven, 2015). After bilabial and labio-dental, /u/ is produced as [v̩] (i.e., the syllabic labiodental voiced fricative), as exemplified in /pu¹/ [pv̩¹] 'wave'. After alveolar, alveolo-palatal and velar consonants, /u/ is realized with diphthong quality (i.e., [əʊ]), as shown in /ku¹/ [kəʊ¹] 'song'. According to a Suzhounese syllabary named *A syllabary of the Suchow dialect*, recorded by A Committee of the Soochow Literary Association in 1892 for missionaries in acquiring Suzhounese, such a diphthongal realization of /u/ after alveolar, alveolo-palatal and velar consonants can be traced back to the beginning of the 20th Century.

The front vowel /ø/ tends to be produced with a lower F2 such as in /ø^{1}/ [ʔø̈^{1}] ‘in safe’ (1228 Hz) than in /jø^{2}/ [ɥø^{2}] ‘rounded’ (1425 Hz). Both, however, are produced with a lip rounding gesture. Following the treatment in Wang (2010), an underlying /ø/ is posited.

/ɘ̯o/ is a diphthong and only co-occurs with the glide /j/ (e.g., /vjɘ̯o^{2}/ ‘to float’ and /kjɘ̯o^{1}/ ‘to tick off’) or alveolo-palatals (e.g., /dʑɘ̯o^{6}/ ‘used’ and /ɕɘ̯o^{1}/ ‘to rest’). It is transcribed as /əu/ in Wang (2010) but seems not to be supported by the formant trajectory produced by the consultant.

As for vowels in closed syllables, the contrast between /æ/ and /ɑ/ only exists before the nasal coda /ŋ/. As illustrated in /tsʰæŋ^{6}/ [tsʰæ̃^{6}] ‘unimpeded’ and /tsʰɑŋ^{6}/ [tsʰɑ̃^{6}] ‘to sing’, both vowels are consistently nasalized without recognizable velum closure at the end.

Vowels preceding a glottal stop coda show a much shorter duration. When high vowels (i.e., /ɪ/ and /ʊ/) occur before /ʔ/, a general displacement towards an open back position often results in a brief schwa after nuclei, such as /tɕɪʔ^{7}/ [tɕɪᵊʔ^{7}] ‘hurry’ and /kʊʔ^{7}/ [kʊᵊʔ^{7}] ‘surname, Guo’.

2.6 Syllabic approximants

There are two syllabic approximants in Lili Wu Chinse, which are exemplified in /sɹ̩^{1}/ [sɹ̩̺^{1}] ‘silk’ and /sɹ̩̹^{1}/ [sʷɹ̩̹̺^{1}] ‘book’. The syllabic approximant /ɹ̩/ [ɹ̩̺] in Lili Wu Chinese is similar to that in Standard Chinese.

With respect to /ɹ̩̹/, two features are to be further noted. First, the lip rounding gesture of the approximant contributes to the labialization of the preceding alveolar sibilant onset (i.e., /s/ [sʷ] before /ɹ̩̹/). Labialized alveolar sibilants are rare in the world’s languages (but see Lao, a Kra-Dai language reported in Erickson, 2001). The rounding feature is believed to evolve from /u/ or /y/, the two rounded vowels reported to be present instead of /ɹ̩̹/ in other Wu dialects, such as /su^{1}/ in Danyang (丹阳) Wu and /ɕy¹/ in Songjiang (松江) Wu for ‘book’ (Qian, 1992: 88). In addition, /ɹ̹̍/ is articulated more laminally. Laminal consonants have been widely reported to exist in Australian languages (Butcher, 1990; Anderson & Maddieson, 1994). Such an articulatory gesture of /ɹ̹̍/ is reflected in Figure 2.14 as a lowered F4 (3375 Hz, compared to 4221 Hz of /ɹ̩/ in /sɹ̩¹/) and the proximity of F3 and F4. F4 lowering is generally said to be related to articulatory retraction (e.g., Fant, 1960; Stevens & Blumstein, 1975; Vaissière, 2011). The proximity of F3 and F4 is known as a consequence of weakly coupled resonators by forming a relatively larger frontal cavity (Stevens, 1989: 19). For instance, a significant convergence of F3 and F4 is observed in laminal alveolar and postalveolar fricatives in English, as well as in apico-laminal alveolars in French (Dart, 1991: 104). In short, /ɹ̹̍/ is produced with a more laminal articulation combined with a lip rounding gesture than its counterpart /ɹ̩/. Such differences were also noticed by the consultant who offered his native intuition voluntarily with us. Given the impressionistic nature of the description, needless to say, more instrumental studies (e.g., ultrasound) are needed for a precise description of their articulation and acoustic consequences.

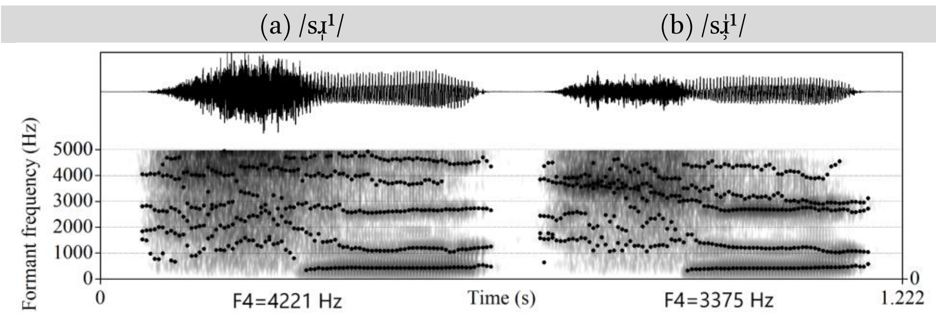


Figure 2.14 Waveforms and spectrograms of (a) /sɹ̩¹/ ‘silk’ and (b) /sɹ̹̍¹/ ‘book’. F4 values are indicated.

It is worth noting that there exist different proposals to transcribe these sounds. For example, among sinologists (after Karlgren, 1915: 294), /ɹ̩/ and /ɹ̹̍/ have often been transcribed as /ɿ/ and /ʮ/, respectively, and are known as ‘apical vowels’. However, as criticized by Lee-Kim (2014: 264), both symbols 'are obsolete in the IPA and not informative for non-Chinese linguists' and are therefore avoided in this description. /ɹ̩/ is usually treated as [z̩] (e.g., Ladefoged & Maddieson, 1996: 314; Wiese, 1997: 239; Duanmu, 2000: 36 for Standard Chinese; Chen & Gussenhoven, 2015 for Shanghainese). Such a treatment, however, has been questioned with ultrasound imaging data (Lee-Kim, 2014; Faytak & Lin, 2015) and acoustic analyses (Howie, 1976: 10). Lee-Kim (2014) further argues that it is more appropriate to describe [z̩] as 'dental approximant [ɹ̪̍]' (see a similar description in Lee & Zee, 2003: 111).[13]

Last but not least, an increasing body of literature has shown that such syllabic approximants are known to affect diachronic changes of high vowels in different languages at different time points, across an overwhelmingly large number of Sino-Tibetan languages (e.g., Baron, 1974; R. Shi, 1998; Zhu, 2004b; Zhao, 2007; Hu & Ling, 2019).

2.7 Syllable structure

Generally speaking, eight syllable combinations can be identified in Lili Wu Chinese. The canonical syllable minimally consists of an obligatory nucleus (V) and a lexical tone as in /u¹/ 'crow' and /ø²/ [ɦø̈²] 'cold'. V stands for vowel or syllabic consonant (i.e., /ɹ̩ ɹ̩̹ m̩ ŋ̍/).[14] It may also contain up to three optional elements in the following linear structure: $(C_1)(G)V(C_2)$,[15] where C_1 can be any consonant in the consonant inventory except for /ʔ/, G is either /j/, as in /kjɘ̞o¹/ 'to tick off', or /w/, as in /kwɛ¹/ 'to close'; C_2 is either /ŋ/ or /ʔ/ as in /kʊŋ¹/ 'public' and /kaʔ⁷/ 'to clip'. All combinations are demonstrated in Table 2.6.

[13] In Lee-Kim (2014: 264), the syllabic diacritic is not used for the sake for simplicity.

[14] In Lili Wu Chinese, /ɹ̩ ɹ̩̹/ are obligatory to contain an onset.

[15] Parentheses indicate optional constituents.

Table 2.6 Syllabic combinations in Lili Wu Chinese.

Combination	Example					
V	/u^{1}/	乌	'crow'	/ø^{2}/	寒	'cold'
GV	/wɔ^{4}/	坏	'broken'	/jø^{2}/	圆	'rounded'
C_1V	/zɔ^{1}/	柴	'firewood'	/kɑ^{1}/	高	'high'
VC_2	/ʊŋ^{1}/	翁	'surname, Weng'	/aʔ^{8}/	盒	'box'
C_1GV	/kjə̠o^{1}/	勾	'to tick off'	/kwɛ^{1}/	关	'to close'
C_1VC_2	/kʊŋ^{1}/	公	'public'	/kaʔ^{7}/	夹	'to clip'
GVC_2	/jʊŋ^{2}/	熊	'bear'	/wʌʔ^{7}/	活	'alive'
C_1GVC_2	/zjæŋ^{1}/	墙	'wall'	/kwʌʔ^{7}/	国	'country'

As illustrated in Table 2.7, co-occurrence constraints on onset and rhyme combinations can be observed.

First, /i ɪŋ ɪʔ i̟/ behave similarly except that /i̟/ can appear after labio-dentals as in /fi̟^{1}/ 'to fly' and /vi̟^{2}/ 'fat'. /i/, on the other hand, is prohibited in this context (i.e., */fi/, */vi/). Second, /y ʏŋ/ are only allowed after alveolar sonorants and alveolo-palatals, or without an onset. Third, before /ø o ɔ ɑ æŋ/, labio-dentals are prohibited (*/fø fo fɔ fɑ fæŋ/) but /ɛ u/ are possible as in /fɛ^{1}/ 'to turn over' and /vu^{2}/ 'to support somebody with one's arm'. Fourth, the two syllabic approximants /ɹ̩ ɹ̹̩/ occur only after alveolar homorganic sibilant onsets /ʦ ʦʰ ʣ s z/.

The distribution of the two glides is summarized in Table 2.8. /j/ is allowed in the majority of cases (e.g., /pjɑ^{1}/ 'watch', /vjə̠o^{2}/ 'to float', /tjɑ^{1}/ 'marten', /kjə̠o^{1}/ 'to tick off', /hjə̠o^{3}/ 'to roar', and /jø^{2}/ 'rounded') except after alveolo-palatals. /w/, however, is more constrained and only allowed after velars (e.g., /kwɛ^{1}/ 'to close'), glottal fricative /h/ (e.g., /hwɛ^{1}/ 'dust'), or serves as a glide onset (e.g., /wɛ^{2}/ 'to be back').

Table 2.7 Observed onset-rhyme combinations in Lili Wu Chinese. +: Yes; –: No.

Onset	i ɪŋ ɪʔ	i̞	y ʏŋ	ɛ u ɑŋ ʊŋ əŋ aʔ ʊʔ ʌʔ	ø o ɔ ɑ æŋ	ɹ̩ ɻ̩	ə̞o
Bilabial (pʰ p b m)	+	+	–	+	+	–	–
Labio-dental (f v)	–	+	–	+	–	–	–
Alveolar plosives (tʰ t d)	+	+	–	+	+	–	–
Alveolar sonorants (n l)	+	+	+[16]	+	+	–	–
Alveolar sibilants (tsʰ ts dz s z)	+	+	–	+	+	+	–
Alveolo-palatal (tɕʰ tɕ dʑ ɕ)	+	+	+	+[17]	+[18]	–	+
Velar and Glottal (kʰ k g h ŋ)	–	–	–	+	+	–	–
j	–	–	–	+[19]	+[20]	–	+
w	–	–	–	+	+[21]	–	–
Zero onset	+	+	+	+	+	–	–

[16] /nʏŋ/ and /lʏŋ/ cannot be observed.
[17] /tɕʰ tɕ dʑ ɕ/ + /ɛ u əŋ/ cannot be observed.
[18] /tɕʰ tɕ dʑ ɕ/ + /o/ cannot be observed.
[19] /jɛ/, /ju/, and /jəŋ/ cannot be observed.
[20] /jo/ cannot be observed.
[21] /wu/, /wø/ and /wɑ/ cannot be observed.

Table 2.8 Observed onset-glide combinations in Lili Wu Chinese. +: Yes; –: No.

Onset	j	w
Bilabial	+	–
Labio-dental	+	–
Alveolar	+	–
Alveolo-palatal	–	–
Velar and Glottal	+	+
Glottal	+	+
Glide onset	+	+

2.8 Onsetless syllables

In onsetless syllables with high-register tones (i.e., T1, T3, T5, and T7), the phonetic segment [ʔ] can be observed at the onset of the tone-bearing syllable, as in /ø¹/ [ʔö¹] 'in safe' and /si̟¹ ø¹/ [si̟⁴⁴ ʔö⁴²] 'a city, Xi'an'. With respect to onsetless syllables with low-register tones (i.e., T2, T4, T6, and T8), phonetic realization of [ɦ] before a non-high vowel (e.g., /ø²/ [ɦö²] 'cold', /ɔ²/ [ɦɔ²] 'shoes', and /aʔ⁸/ [ɦaʔ⁸] 'box') is observed, in contrast to other cases when there is a high vowel or glide (e.g., /i²/ [ʝi²] 'salt', /jø²/ [ɥø²] 'rounded', /u²/ 'river', and /wʌʔ⁸/ 'alive'). [ɦ] disappears in non-initial position within a prosodic word, e.g., /tʰɑ⁴ ɔ²/ 'galoshes'. The general pattern is therefore similar to Shanghainese (Chen & Gussenhoven, 2015).

In Lili Wu Chinese, syllables with low-register tones show relatively stronger breathiness than those with high-register counterparts. As indicated by Figure 2.15, the Fast Fourier transform (FFT) spectrum of /ø¹/ [ʔö¹] 'in safe' (dark) and /ø²/ [ɦö²] 'cold' (light) shows the phonatory contrast in the vowel /ø/, taken within an interval of approximately 30 ms from the first regular vocal pulse of the vowel. As shown by the measurements on H1–H2 (i.e., amplitude difference between the first and second harmonics), there is a phonatory difference between the two vowels with /ø²/ (4.5 dB) showing more breathiness than /ø¹/ (2 dB). This contrast has also been observed in some other Northern Wu dialects (Cao & Maddieson, 1992).

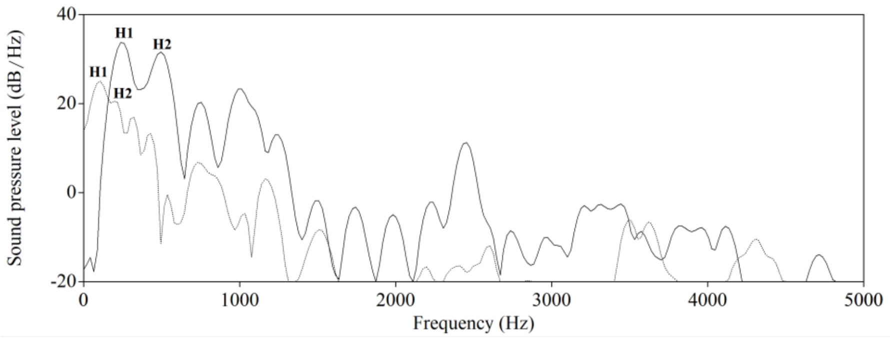


Figure 2.15 FFT spectrum of /ø/ in /ø¹/ 'in safe' (dark) and /ø²/ 'cold' (light), taken over an interval of approximately 30 ms from the first regular vocal pulse of the vowel. The first two harmonics (H1 and H2) of each syllable are indicated.

2.9 Tone sandhi

Lexical tones over monosyllabic morphemes undergo changes when they are combined into compounds or phrases. In this description, I offer some preliminary observations concerning tone sandhi variations in Lili Wu Chinese over disyllabic compounds (hereafter called the tone unit). Tonal realization is contingent upon the lexical tone of the second syllable (σ_2). Two general patterns are observed.

First, when σ_2 carries an abrupt tone (i.e., T7 and T8 over a glottal-coda syllable), regardless of the syllable structure of the first syllable (σ_1), only level *f0* contours surface, and the specific *f0* height is dependent on the lexical tone of σ_1. After a high tone, a low tone appears; while after a low tone, a high tone appears. Both patterns are illustrated in Figure 2.16, which shows T1 (high–level) + T7/T8 (2.16a, /tsʰəŋ¹ tsɪʔ⁷/ 'the Spring Festival'; 2.16b, /tɕɪŋ¹ dʑɪʔ⁸/ 'Peking Opera') and T6 (low–level) + T7/T8 (2.16c, /tʰɔ⁶ kwʌʔ⁷/ 'Thailand'; 2.16d, /tɕʰi̧⁶ dɪʔ⁸/ 'steam whistle').

Second, when σ_2 carries a non-abrupt tone (i.e., T1 to T6 over an open syllable or a syllable with a nasal coda), the lexical tonal contour of

σ_1 remains and affects the pitch realization of σ_2. The specific *fo* contour of σ_2 hinges upon the lexical tonal register of σ_1. When σ_1 is produced with a high-register tone (i.e., T1, T3, T5, and T7), σ_2 is typically realized with a falling *fo* contour, as shown in Figure 2.17 (e.g., 2.17a, /sɪŋ^{1} zəŋ^{4}/ 'new kidney'; 2.17b, /kɛ^{3} zɑ^{4}/ 'to remold'; 2.17c, /tɕɔ^{5} zɑ^{4}/ 'introduction'; 2.17d, /kʊʔ^{7} tʰu^{4}/ 'territory'). However, other patterns have also been observed. For example, in the combination of T7 + σ_2, when σ_2 bears T1 (e.g., /kʊʔ^{7} kʰu^{1}/ 'orthopedics'), the monosyllabic citation form of T1 in /kʰu^{1}/ (high–level) is preserved, instead of a predictable falling contour like /tʰu^{4}/ in /kʊʔ^{7} tʰu^{4}/ 'territory'.

When σ1 is pronounced with a low-register tone (i.e., T2, T4, T6, and T8), the sandhi pattern tends to be more complicated. The tonal contour of σ2 seems to also exert influence on the overall tonal realization. For example, Figure 2.18 shows the contrast of /pʰɔ^{6} tɕʰi^{4}/ 'to dispatch' (2.18a) vs. /ʨʰi̟^{6} pʰɑ^{6}/ 'bubble' (2.18b). Here, T4 in /tɕʰi^{4}/ completely loses its monosyllabic citation form (low–level) and is realized with a high-falling contour, similar to Shanghainese (Chen & Gussenhoven, 2015). However, the lexical tone of the preceding tone T6 in /pʰɑ^{6}/ (low–dipping) is only preserved to a certain extent. The same tone (i.e., T6) is realized with an audible pitch level difference: T6 in /pʰɔ^{6} tɕʰi^{4}/ is overall lower than that in /ʨʰi̟^{6} pʰɑ^{6}/. Further research is needed to investigate the extent to which listeners are sensitive to the differences.

Before closing off this section, it is worth pointing out that syllables with aspirated onsets show two different patterns of changes. They pattern either with syllables that have unaspirated onsets and carry T1, or with syllables that have voiced onsets and carry T4, T6, or T8. For example, the sandhi change of /tʰʊŋ^{1} fʊŋ^{1}/ 'to ventilate' patterns with that of /tʊŋ^{1} fʊŋ^{1}/ 'east wind'; while /ʦʰʌ^{8} djɵ̝o^{2}/ 'to stand out' patterns with /zʌ^{8} djɵ̝o^{2}/ 'tongue'.

It is important to conclude here that even within the arguably simplest construction beyond a monosyllabic morpheme (i.e., disyllabic compounds), Lili Wu Chinese already exhibits different patterns of tonal realization from its neighboring Northern Wu dialects such as Shanghainese (Chen & Gussenhoven, 2015). It is not only subject to the influence of the preceding tone on tonal realization but also seems sensitive to tonal properties of the second syllable. In this illustration, I have just presented a preliminary glimpse into the pitch contours of disyllabic compounds in Lili Wu Chinese. Needless to say, more data and further research are needed.

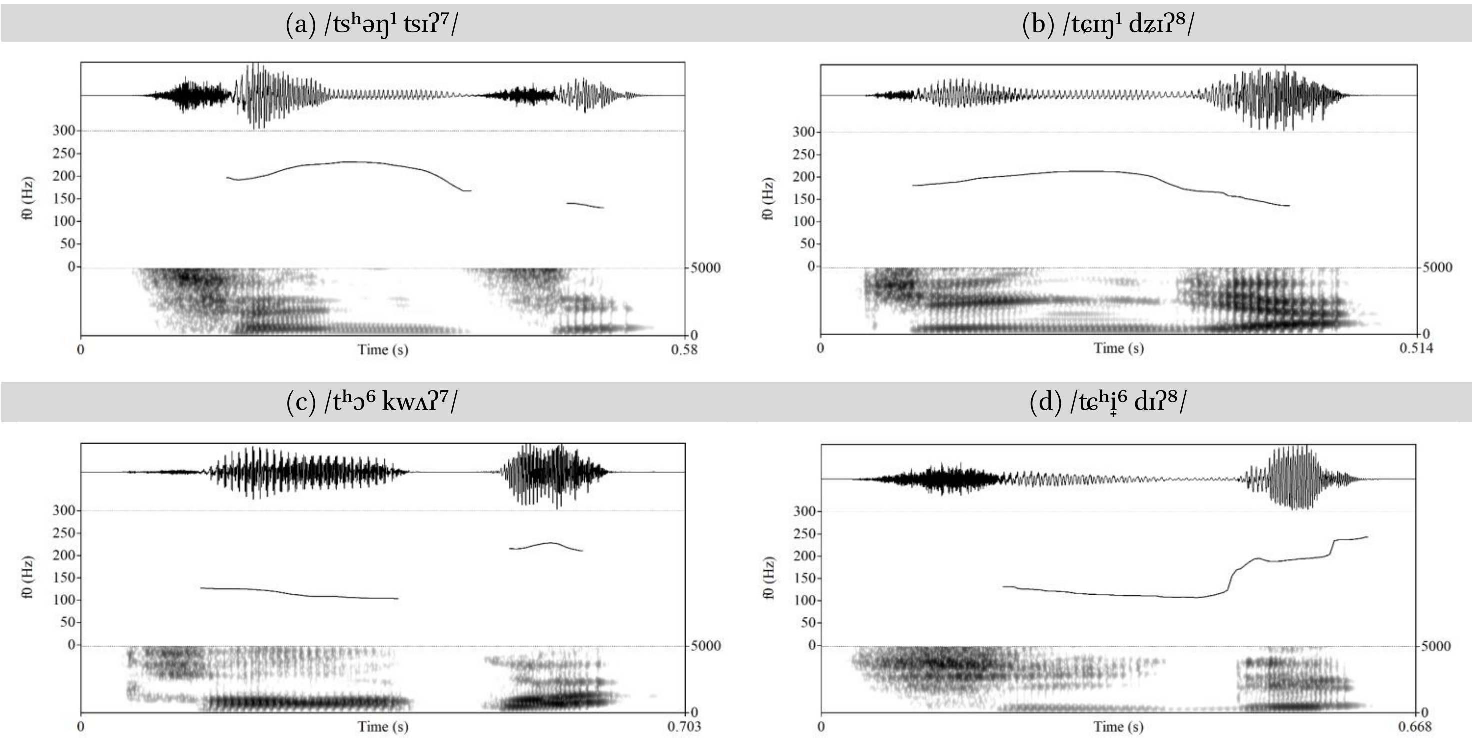


Figure 2.16 Waveforms, *f*0 tracks, and spectrograms of (a) /tsʰəŋ^{1} tsɪʔ^{7}/ 'the Spring Festival', (b) /tɕɪŋ^{1} dʑɪʔ^{8}/ 'Peking Opera', (c) /tʰɔ^{6} kwʌʔ^{7}/ 'Thailand', and (d) /tɕʰi̠^{6} dɪʔ^{8}/ 'steam whistle'.

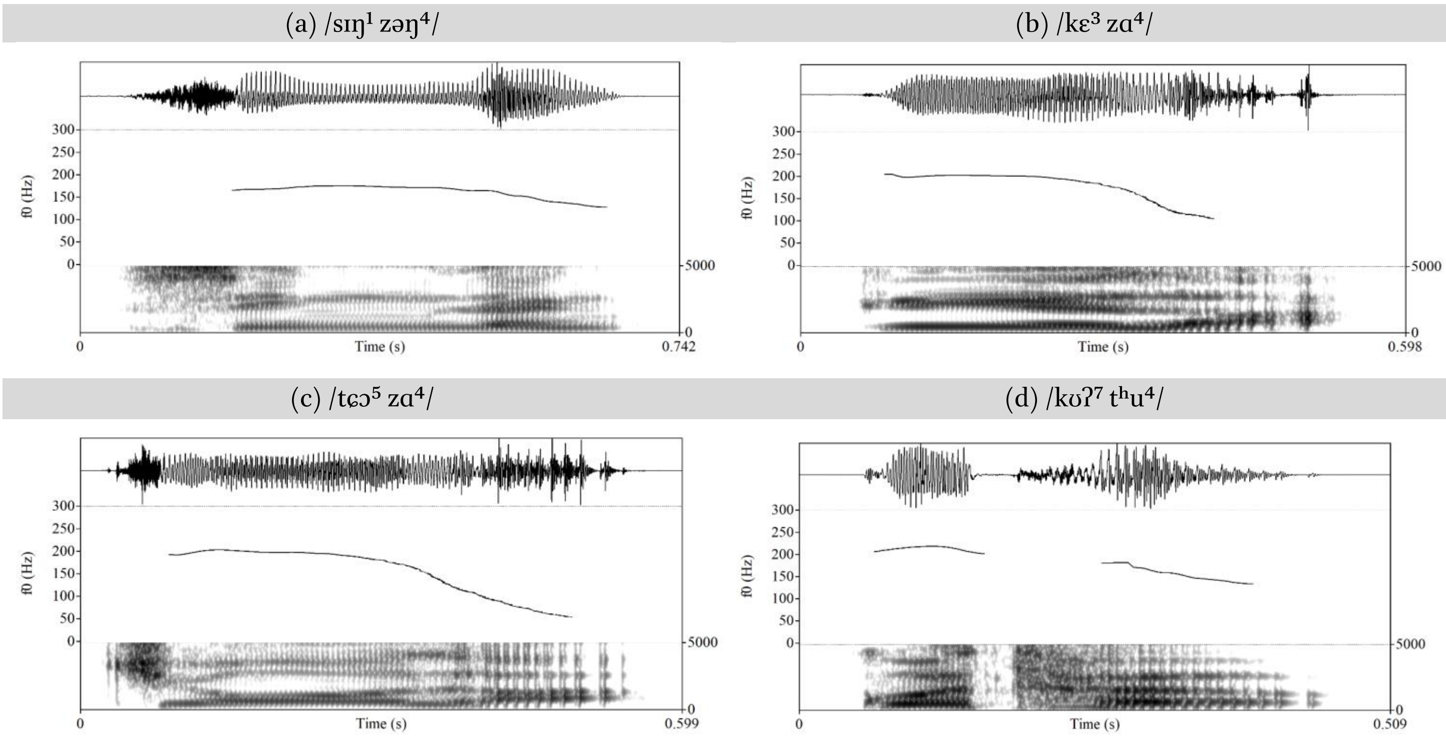
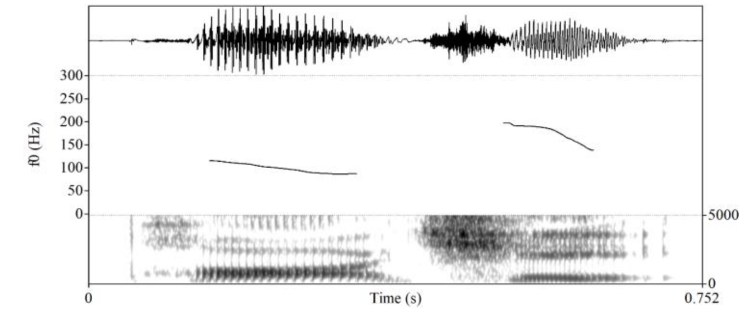


Figure 2.17 Waveforms, *f0* tracks, and spectrograms of (a) /sɪŋ^{1} zəŋ^{4}/ 'new kidney', (b) /kɛ^{3} zɑ^{4}/ 'to remold', (c) /tɕɔ^{5} zɑ^{4}/ 'introduction', and (d) /kʊʔ^{7} tʰu^{4}/ 'territory'.

(a) /p^{h}ɔ^{6} tɕ^{h}i^{4}/



(b) /tɕ^{h}i̯^{6} p^{h}ɑ^{6}/

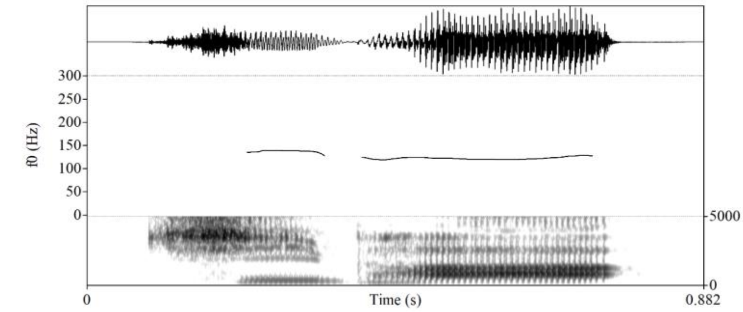


Figure 2.18 Waveforms, *f0* tracks, and spectrograms of (a) /p^{h}ɔ^{6} tɕ^{h}i^{4}/ 'to dispatch' and (b) /tɕ^{h}i̯^{6} p^{h}ɑ^{6}/ 'bubble'.