Complex Patterns of Tonal Realization in Taifeng Chinese

Xiaoyan Zhang[1], Aijun Li[2], Zhiqiang Li[3]

[1]Beijing International Studies University, China

[2]Center of Corpus and Computational Linguistics, Institute of Linguistics, Chinese Academy of Social Sciences, China

[3]Department of Modern and Classical Languages, University of San Francisco, USA

1366068074@qq.com; liaj@cass.org.cn; zqli@usfca.edu

Abstract

Taifeng Chinese is a Wu dialect that has the smallest inventory of tones while still preserving checked tones and the voicing contrast in syllable onsets. A previous acoustic study identified four surface tones in isolation as a result of tone split and the merger of tonal categories derived from the Middle Chinese tonal system. Notably, a long Yang Shang tone merged with short checked tones and the Yang Ping tone was realized as two surface tones, subject to regional and age-graded variations. Surface realization of tones in Taifeng is further examined in an acoustic investigation of disyllabic tone sandhi in verb-object (VO) combinations. Analyses of pitch contours and tonal duration from multiple speakers reveal complex patterns of tonal realization in tone sandhi. The tone sandhi in VO combinations is best characterized as the right-dominant pattern, in which the second tone has consistently longer duration and retains its citation form while the first tone is realized in a reduced pitch range with much shorter duration. Tonal realization is also governed by register alternation: the two tones are realized in opposite register specifications. In general, tonal realization in tone sandhi exhibits considerable complexities not observed in monosyllables.

Index Terms: tonal realization, tone sandhi, register alternation, Taifeng Chinese, tonal duration

1. Introduction

Lexical tones in Chinese dialects take certain phonetic shapes in isolation. Tone 1 in Mandarin Chinese is high level (transcribed as 55) and tone 2 is high rising (35). When tones come into contact with one another in connected speech, their phonetic shapes are often altered to such a degree that tonal identity is changed. This process of tone alternation is referred to as tone sandhi [1].

Some dialects have extensive tone sandhi. In her survey of 83 Chinese dialects, Yue-Hashimoto [2] classified tone sandhi into two major types. The first type is called "first-syllable dominant", in which the first syllable in the tone sandhi domain maintains its lexically associated tone while the following syllables in the domain assume tonal values different from their citation forms. Most Wu dialects, especially the northern varieties belong to this type, such as Suzhou, Shanghai, Wuxi and Ningbo. The second type is called "last-syllable dominant", in which the last syllable in the domain keeps its underlying tone while the preceding syllables assume sandhi forms. Most Min dialects, southern Wu dialects and Mandarin dialects display this type of tone sandhi [3]. Yue-Hashimoto also noted that the phenomenon of dominance is conditioned by stress.

Tonal realization involving change of tonal identify in Wu dialects can be quite complicated for two major reasons. First, Wu dialects maintain a neat correspondence between onset voicing and tone register [4]. The four tonal categories of the Middle Chinese (MC), referred to as Ping (level), Shang (rising), Qu (departing) and Ru (entering) in the philological tradition, end up in the two registers of Yin and Yang when they go with voiceless and voiced onsets respectively. As these tonal categories are realized in the phonology, tone merger would reduce the number of lexical tones. Second, Wu dialects are notable among Chinese varieties for having complex patterns of tone sandhi. Not only lexical compounds and phrasal structures tend to follow different patterns of tone sandhi, but morphosyntactic structures also comes into play too [3]. For examples, different rules of tone sandhi could apply for modifier-noun structures and verb-object combinations in some dialects [e.g. 5].

In this paper we analyze patterns of disyllabic tone sandhi in Taifeng Chinese, a Wu dialect in the Xuanzhou area of Anhui province, and compare them with the way tones are realized in isolation, as reported in our previous study [6]. As a point of departure, we start with a synopsis of the lexical tone system of Taifeng and then lay out the blueprint for the current study.

As seen in Table 1 [6], Taifeng contrasts four surface tones in isolation: three long tones and one checked tone, represented as M, R, L and ʔH. The numbers in the last column indicate pitch values on a five-point scale, also known as "tone letters" [7]. Several observations can be made of how tones are realized in isolation: (1) Taifeng has the smallest tonal inventory in Wu dialects with four contrasting surface tones in citation forms. (2) The Yang Ping tone is split into two subcategories, T2 and T3, one with voiced obstruent onsets and one with sonorant onsets. Most Wu dialects would not make this distinction in tone mapping. (3) The mapping of T2 shows regional and age-graded variations in three different groups of speakers. (4) The

Yang Shang tone, T5, merges with two Ru tones, T8 and T9. Ru tones are short checked tones, ending in a glottal stop. This kind of merger is rare in Wu dialects.

Table 1: *MC tonal categories and surface tones*

MC tonal categories and onset voicing				Surface tones	
Ping	Yin	–voi		T1	M-22, 33
	Yang	+voi	obs	T2	R-24
				T2	L-213 R-13
				T2	L-213
			son	T3	R-24
Shang	Yin	–voi		T4	L-212, 213
	Yang	+voi		T5	ʔH-5
Qu	Yin	–voi		T6	R-24
	Yang	+voi		T7	R-24
Ru	Yin	–voi		T8	ʔH-5
	Yang	+voi		T9	ʔH-5

Reduction from nine tonal categories to four surface tones results from extensive merger in the current tonal system. The process is made even more complicated with regional variations. Our present study is concerned with disyllabic tone sandhi in verb-object (VO) combinations in Taifeng. Other combinations show different tone sandhi patterns and will be reported elsewhere. Here we ask how lexical tones in Taifeng are realized in the first and second syllables, how they compare with the citation forms and what factors determine tonal realization in VO combinations. Acoustic analyses will be provided for pitch contours and durations of the disyllabic words used in the experiment. A phonological analysis will be attempted too.

2. Materials

2.1. Recording materials

In order to determine whether the MC tonal categories are realized in tone sandhi the same way as in isolation, we created a list of disyllabic words in which each syllable is in one of the nine tonal categories, resulting in 81 combinations. We ended up with recorded data for 79 combinations, with T3-T5 and T4-T9 missing due to lack of appropriate words, as we focus on words in the VO structure. Since no difference was found in tone sandhi between lexical compounds and phrases in Taifeng, we did not deliberately distinguish them in the list (cf. [8]). A partial list of words used in the experiment is given in Table 2.

Table 2: *A partial word list.*

T1-T1: 搬家 [pɶ̃ ka]	T2-T6: 还债 [uɶ̃ tsai]
T3-T1: 磨刀 [mo tɔ]	T4-T7: 煮饭 [tʃʮ uã]
T5-T8: 犯法 [uã xuaʔ]	T6-T7: 看病 [kʰuɶ̃ ɸɦin]
T7-T5: 卖米 [mai miʔ]	T6-T8: 送客 [soŋ kʰɛʔ]
T8-T2: 刷牙 [saʔ ŋa]	T9-T9: 核实 [hɛʔ sɿʔ]

2.2. Recording and data processing

12 participants were recruited for the experiment. They all participated in the previous study [6]. Born and raised in the Taifeng township, they are fluent in the dialect and use it in their daily communication. None of them reported any hearing or speaking difficulties. Data collected from ten speakers (6 females and 4 males) were presented in the analysis. Their average age is 54 (SD = 8.47).

The recording was carried out in a sound-proof room. X-recorder [9] was used to collect data on the computer with a Lexicon-IO22 audio adaptor and a desktop microphone. The sampling rate and quantizing resolution were set to be 44 KHz and 16 bits respectively. Each speaker was asked to read all 79 disyllabic sequences without a carrier sentence two times.

Praat 6.1 [10] was employed to provide segmental annotations for syllabic and subsyllabic elements (initials and finals). Pitch contour was extracted for each syllable, with spurious pitch cycles manually modified. 11 pitch values were extracted at equal intervals on each syllable final (tone bearing unit). Durations of syllable initials and finals were also obtained.

In order to eliminate individual fluctuations, the formula in (1) was used to calculate the z-score f_0 normalization values (z_{lf0i}) for each speaker, based on the mean (m_{lf0}) and standard deviation (s_{lf0}) of the logarithmic f_0 values. The normalized z-scores were later converted into pitch values on the five-point scale.

$$Z_{lf0_i} = \frac{lf0_i - m_{lf0}}{s_{lf0}} = \frac{\lg f0_i - \frac{1}{n}\sum_{j=1}^{n} \lg f0_j}{\sqrt{\frac{1}{n-1}\sum_{j=1}^{n} (\lg f0_j - \frac{1}{n}\sum_{k=1}^{n} \lg f0_k)^2}} \quad (1)$$

3. Results and Discussion

Recall that T2 is realized differently in isolation in three groups. Group 1 consists of 4 speakers (1 male and 3 females) living outside Zhengtan village with an average age of 51.8. In their speech, T2 merges with T3. Together they also merge with T6 and T7. These four tones are realized as a rising tone, R. Groups 2 and 3 include speakers living in Zhengtan village. Two older male speakers (with an average age of 67.5) produce an alternating pattern for T2, merging with T3 in some words and with T4 in others. Other 4 speakers (3 males and 1 female) form Group 3. They merge T2 with T4 in their speech. See Table 3 for details. The regional and age-related variations of T2 are also reflected in the tone sandhi data. Due to space limitations, our discussion will focus on output patterns in tone sandhi in Group 1.

Table 3: *Mapping of tones in three groups.*

Group 1: **M**-T1; **R**-T2, T3, T6, T7; **L**-T4; **ʔH**-T5, T8, T9
Group 2: **M**-T1; **R**-T2, T3, T6, T7; **L**-T2, T4; **ʔH**-T5, T8, T9
Group 3: **M**-T1; **R**-T3, T6, T7; **L**-T2, T4; **ʔH**-T5, T8, T9

3.1. Pitch patterns in disyllabic tone sandhi (Group 1)

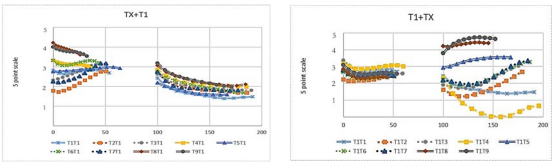


Figure 1: *Realizations of T1 in disyllabic combinations.*

Pitch data for each disyllabic combination were normalized and then converted into the five-point scale so as to compare patterns of tonal realization across different tonal contexts. Given the large number of disyllabic combinations, we present the pitch data in four sets, corresponding to the four contrasting surface tones in isolation. Two graphs were generated for a given tonal category (Ti) as it is combined with all 9 tonal categories (Tx) in Taifeng: when it is final (Tx+Ti) and when it is initial (Ti+Tx) in disyllabic combinations.

The two graphs in Figure 1 show f_0 patterns of tonal combinations when T1 is final (left) and initial (right). In both, the final tones are closer to their citation forms than the tones in the first syllable, which are realized in a narrow pitch range. Specifically, T1 is realized as a mid-level tone (M) consistently in all three groups in its citation form [6], but in the disyllabic context, its phonetic shape is modified differently in initial and final positions. In the final position (left), the f_0 contour of T1 is approaching the tonal target of M at the end of its host syllable while the beginning coincides with where the preceding tones end, probably due to the effect of coarticulation. In the initial position (right), T1 is shorter in duration than the following tones and also T1 in the final position. The f_0 contour does not show much movement. The four contrasting surface tones are clearly discernible in the final syllable. This pattern of tonal realization points to the "last-syllable dominant" type of tone sandhi, which underlies the initial-final asymmetry [11].

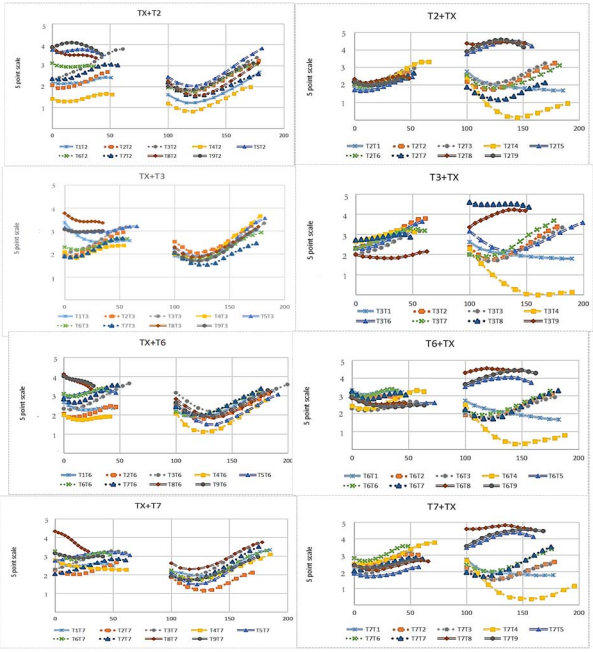


Figure 2: *Realizations of T2,3,6,7 in disyllabic combinations.*

The eight graphs in Figure 2 show f_0 patterns of tonal combinations when T2, 3, 6 and 7 are final (left) and initial (right). These four tonal categories merge into one tone and are realized as R in isolation. Here the same initial-final asymmetry is manifested; ergo, the final syllable is more prominent with respect to tonal realization. The four graphs on the left clearly show a somewhat delayed rising f_0 contour, the canonical shape of a fully realized R in Chinese dialects [12]. This means the merger of these four tonal categories is also attested in tone sandhi. It is worth nothing that at least in some cases the first and the second tones seem to be realized in opposite register specifications of tones. In other words, if the first tone occurs in a higher register of pitch, the second will be realized in a lower register. Take T3-T2 (in the top left graph) as an example. The shorter T3 is realized as a much higher R than the more prominent T2 in the final position. The similar register alternation is also observed in T1-T1 in Figure 1. The initial M is realized higher than the more fully realized M in the final position. In the four graphs on the right, the four contrasting contours are identified in the final syllable despite subtle differences in the way each tone is implemented. For example, T7 is mapped to R in T2+T7 (in the top right graph), but it is actually realized as a lower R than T2, T3 and T6. The coarticulatory effect is also visible. The f_0 contour of T1 is slightly elevated at the beginning before reaching the M target. In the initial position, the tone is usually realized as a weakened form of R, a short H or M.

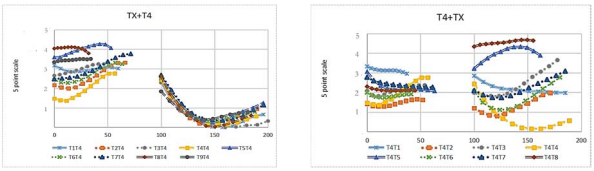


Figure 3: *Realizations of T4 in disyllabic combinations.*

The two graphs in Figure 3 show f_0 patterns of tonal combinations when T4 is final (left) and initial (right). T4 in Taifeng is a low dipping tone similar to the Mandarin third tone, transcribed as 213. The long version of T4 only appears in the final position (left), which is supposed to be a metrically prominent position which can license a long contour tone [11]. In the initial position (right), T4 is always truncated and realized a small R or a short M. In terms of register alternation, all tones seem to be realized in a higher register than T4, which operates in the lower register of tone (left).

The six graphs in Figure 4 (next page) show f_0 patterns of tonal combinations when T5, 8, and 9 are final (left) and initial (right). T8 and T9 are historically checked tones in syllables with glottal endings. T5 is a Yang Shang tone, but it merges with checked tones in Taifeng. They are realized as a short H, notated as ʔH. The preservation of checked tones in Taifeng is noteworthy since there are only four contrasting tones in the lexical tone system. Merger of a long tone with short checked tones is rare among Wu dialects. In disyllabic contexts, T5, T8 and T9 exhibit strikingly similar patterns of tonal realization. They all surface as a short H when they are final (left). In the initial position (right), they are realized as a downstepped H, lower than the three checked tones in the final syllable, but higher than other tones. The four contrasting contours are well maintained when they occur in the final syllable with some variations, probably due to the effect of coarticulation. The long tones in the initial syllable (left) appear to be contained in a narrow range in a much lower register of pitch than the following checked tones.

A succinct summary of disyllabic tone sandhi patterns for VO combinations in Group 1 is provided in Table 4. despite extensive contextual variations, four general patterns seem to emerge, based on the system of surface tones in Table 3. In all four patterns, the contrasting tonal contours in the lexical tone system are preserved in the final position where tones in the first position undergoes a process of neutralization. The initial-final asymmetry in tonal realization is accompanied by variations in duration patterns, to be analyzed in 3.2. The register shift also happens in some combinations, as we will see in 3.3.

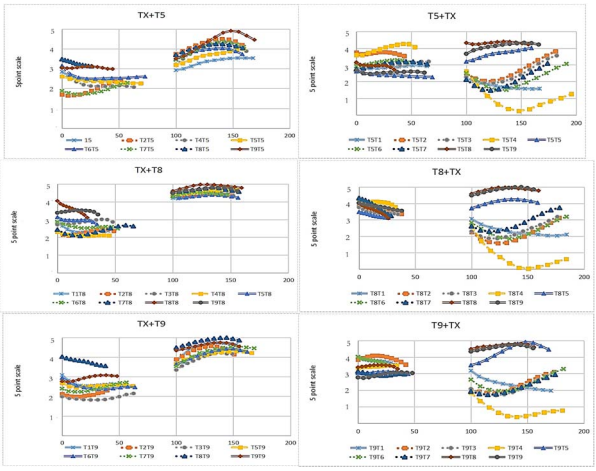


Figure 4: *Realizations of T5,8,9 in disyllabic combinations.*

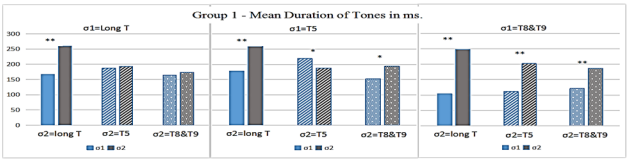
Table 4: *Disyllabic tone sandhi patterns in Group 1.*

A: Tx + T1: M + M	T1 + T2: M + R
B: Tx + T2,3,6,7: M + R	T2 + T6: M + R
C: Tx + T4: M + L	T4 + T1: M + M
D: Tx + T5,8,9: M + ʔH	T5 + T6: ʔM + R

3.2. Duration patterns in disyllabic VO combinations

Duration was measured for the syllable final, which is taken to be equivalent to the duration of tones. Mean duration of tones from all three groups is given in Figure 5. The discussion below is provided only for Group 1, but the other two groups exhibit consistent duration patterns.

In the long-long combinations of tones (i.e. no checked tones involved), the initial syllables are on average about 60% as long as the final syllables. The duration distribution is consistent with the postulate that the disyllabic tone sandhi in Taifeng VO combinations is metrically right-headed. In the long-short combinations (i.e. the three checked tones T5, T8 and T9 are in the final position), there does not seem to be significant differences in duration. In the short-long combinations, the final syllable is often twice as long as the initial syllable in checked tones. In this scenario, syllables in T5 are longer than syllables in T8 and T9. In fact, the duration of T5 is closer to the duration of long tones in the initial position than that of other checked tones. In many cases, the glottal ending is still present to differentiate T5 from long tones. Further study with segmental composition better controlled for will be needed to explore duration patterns of the three checked tones in more detail. In the short-short combinations, the positional effect manifests itself in the final tone being longer than the initial tone.



(*, p<0.05; **, p<0.01)

Figure 5: *Mean duration of tones in ms.*

3.3. Right-prominent tone sandhi and register alternation

Our analysis of the pitch and patterns in the disyllabic VO combinations indicates that the disyllabic tone sandhi in Taifeng is a right-prominent system in which the final syllable is the metrically prominent position which licenses the lexically associated tone of the syllable. The longer duration of the final syllable also allows lexically contrasting surface tones to maintain their distinctive phonetic shapes. In Taifeng, they are M, R, L and ʔH. In the non-prominent position, tones are implemented in a much narrow range and have often lost their distinctive contours.

Tonal realization in disyllabic tone sandhi is also governed by register alternation, assuming that a tone has a contour and a register specification in the phonological representation [1]. What we have discovered is that in general the tone in the prominent position retains its register specification and the tone in the non-prominent position (i.e. initial syllable) takes on the opposite register. There are exceptions: when both tones are the Yin tones, as determined by onset voicing, the register profile will be high and low; when both tones are the Yang tones, the register profile will be low and high. Checked tones in the final position are always in the high register.

4. Conclusions

Taifeng Chinese presents an interesting case in understanding complex patterns of tonal realization in isolation and disyllabic tone sandhi. Analyses of pitch contours and vowel duration data from multiple speakers reveal complex patterns of tonal realization in tone sandhi. The tone sandhi in VO combinations is best characterized as the right-dominant pattern, in which the second tone has consistently longer duration and retains its citation form while the first tone is realized in a reduced pitch range with much shorter duration. Tonal realization is also governed by register alternation: the two tones are realized in opposite register specifications. Similar patterns of register alternation seem to exist in the closely related Yanchi dialect which distinguishes five tones, four long tones and one checked tone [13]. In general, tonal realizations in disyllabic tone sandhi in Taifeng exhibits considerable complexities, which can be traced back to the lexical tone system in monosyllables and the metrical prominence in disyllabic combinations.

5. Acknowledgements

This work is supported by the National Social Science Fund of China Key Project grant (No. 15ZDB103) and the “Four Batches” Talent Project grant awarded to Aijun Li.

6. References

[1] M. Yip, Tone. New York: Cambridge University Press, 2002.

[2] A. O. Yue-Hashimoto, “Tone sandhi across Chinese dialects,” Wang Li Memorial Volumes, HK: Joint Publishing Co, 1987, pp. 445-474.

[3] M. Chen, *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge University Press, 2000.

[4] Y. R. Chao, *A Study of Modern Wu Dialect*. Beijing: Science Press, 1928.

[5] S. Lü, “Tone system of the Danyang dialect,” *Dialect*, no. 2, pp. 85-122, 1980.

[6] A. Li, X. Zhang and Z. Li, “Acoustic and phonological analyses of Taifeng Chinese,” *in Speech Prosody 2020 - 10th Biennial Conference of the International Speech Communication Association, May 25-28, Tokyo, Japan, Proceedings,* 2020, pp. 100–104.

[7] Y. R. Chao, “A system of ‘tone-letters’”, Le Maître Phonétique, 30, pp. 24–27, 1930.

[8] B. Xu, Z. Tang and N. Qian, “Tone Sandhi in New Shanghai Dialect,” *Dialect*, no.2, pp. 145-155, 1981.

[9] Z. Xiong, “X-recorder: recording tool for speech corpus,” https://mp.weixin.qq.com/s/ewBC2hDg4seuV1YWoPVi-A.

[10] P. Boersma and D. Weenink, “Praat: doing phonetics by computer,” http://www.praat.org/, 2020.

[11] J. Zhang, “A directional asymmetry in Chinese tone sandhi systems,” *Journal of East Asian Linguistics,* vol. 16, no. 4, pp. 259-302, 2007.

[12] Y. Xu, “Fundamental frequency peak delay in Mandarin,” Phonetica, vol. 58, pp. 26-52, 2001.

[13] M. Shen, “Disyllabic tone sandhi in the Yanchi dialect in the Xuancheng city of Anhui,” *Fangyan*, no. 2, pp. 144-157, 2016.

[This paper was published at 12th International Symposium on Chinese Spoken Language Processing (ISCSLP) 2021]