DISYLLABIC TONE SANDHI PATTERNS IN SHAOXING WU CHINESE

Xinyi Wen, Yiya Chen, Lisa Lai-Shen Cheng

Leiden University Centre for Linguistics, Leiden University, the Netherlands
x.wen@hum.leidenuniv.nl, yiya.chen@hum.leidenuniv.nl, l.l.cheng@hum.leidenuniv.nl

ABSTRACT

We report acoustic data on disyllabic tone sandhi patterns in Shaoxing Wu Chinese. With quantitative analyses, we confirmed the impressionistic observations that 1) both tonal register and contour condition disyllabic tone sandhi forms in Shaoxing Wu Chinese; and 2) not only the initial but also the non-initial tones play a role. These findings present novel experimental evidence on the independence of register and contour for lexical tone representations and sandhi processes. Furthermore, they provide a more nuanced picture of the left- vs. right-dominant tone sandhi patterns in Northern Wu dialects; both can be present in the same dialect. No durational correlates of the left- and right-dominant sandhi patterns, however, were found, raising questions about the phonetic substance of metrical prominence structure based on tone sandhi dominance patterns.

Keywords: Tone Sandhi, Northern Wu Chinese, Tonal Register, Contour

1. INTRODUCTION

Lexical tones can contrast in both pitch register and contour [1]. The structural relationship between the two has been a long-debated issue [2], with different models proposed on tonal representations (e.g., [3], [4]). These models differ in their predictions of whether tonal pitch register and contour are independent or interdependent. While some data suggest that they can spread separately and therefore are independent (e.g., [5], [6]), these observations have been drawn mainly upon introspective impressionistic descriptions. Experimental results are therefore needed to confirm the tonal patterns and investigate further how exactly tonal register and contour interact. One goal of this paper is to report data from the *f0* realizations of disyllabic tone sandhi patterns in Shaoxing Wu Chinese, aiming to shed light on this issue and provide insights on the way register and contour give rise to the surface *f0* contours of contextual tonal realization.

Shaoxing Wu Chinese is a Northern Wu dialect, known to contrast in both register and contour. There are eight lexical tones, with four different contours in both high and low pitch registers (e.g., [7], [8]), as illustrated in Figure 1. Impressionistic studies have shown that tone sandhi patterns of disyllabic words in Shaoxing Wu are conditioned by both tonal register and contour of initial syllables. In addition, the contours of final syllables are preserved when following an initial falling tone, though neutralized following a rising tone [9]. Such complex tone sandhi patterns present an interesting test case on how tonal register and contour condition contextual tonal realization. Yet, their effects on *f0* trajectories have not been thoroughly examined with controlled experiments.

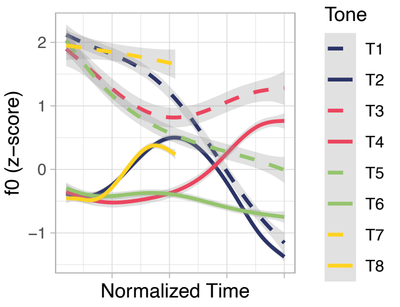

Figure 1: Lexical tones in Shaoxing Wu.

The impressionistically observed disyllabic tone sandhi patterns, if borne out by experimental data, will shed light on the issue of left- vs. right-dominant sandhi patterns in Wu Chinese. Words in Northern and Southern Wu are known to differ in their tone sandhi patterns (e.g., [6], [10]). The former tends to undergo left-dominant sandhi, with the initial tone determining the sandhi form and the non-initial tone often neutralized. The latter undergoes right-dominant sandhi, with the non-initial tone preserving the citation form and the initial one neutralized. However, there are also Northern Wu dialects with the non-initial tone conditioning tone sandhi (e.g., [11], [12]), and the effects of register and contour vary across dialects [13]. Given the potential role tones in the final position play in Shaoxing Wu, the second goal of this paper is to report on data which enable us to have a detailed look at how tones in the final position condition sandhi patterns, especially the effects of register and contour separately. This would further our understanding of the right-dominant patterns of tone sandhi in Northern Wu dialects, in which the prevalent sandhi pattern is left-dominant.

Related to the left vs. right-dominant sandhi is the proposal that such a tone sandhi difference is due to a typological difference in metrical prominence across

Chinese dialects. In the existing literature, it has been argued that only prominent syllables can preserve underlying tones (e.g., [6], [14]). Yet, few of them provided evidence other than tonal patterns, leading to a circular argument for the relationship between sandhi-dominant patterns and metrical structures. While evidence from cross-dialect comparisons suggests the need to delink metrical prominence from durational correlates [15], it is nevertheless worth investigating whether there is a durational difference between words undergoing left- vs. right-dominant tone sandhi within the same dialect. As mentioned earlier, in Shaoxing Wu Chinese, words starting with falling tones resemble right-dominant tone sandhi patterns, in contrast to words starting with rising tones that show left-dominant tone sandhi patterns. Corresponding differences in syllable duration for words with falling vs. rising tones in the initial syllable can elucidate the potential link between tone sandhi-derived metrical prominence structure and its possible durational correlates.

In sum, this study aimed to address the following research questions: (1) How do tonal registers and contours of initial (σ1) and final syllables (σ2) condition the *f0* trajectories of disyllabic words in Shaoxing Wu? (2) Do words starting with falling and rising tones differ in their duration patterns? In doing so, we will better understand the interplay of register and contour in contextual tonal realizations as well as the potential link between sandhi dominance and metrical prominence.

2. METHOD

2.1. Participants

16 middle-aged native speakers of Shaoxing Wu were recruited as participants of the current study (10 female and 6 male, mean = 50.7, SD = 3.6). They were all born and raised in the urban area of Shaoxing. They use Shaoxing Wu frequently for daily communication. None of them reported difficulties in speaking or hearing.

2.2. Stimuli

A list of disyllabic words was constructed as the stimuli. To examine whether and how registers and contours of σ1 and σ2 condition sandhi contours, the tone of both syllables was controlled as follows: for *registers*, both High and Low register tones were included; for *contours*, Rising and Falling tones were included. Two different items were included for each tonal combination. In total, 512 tokens were collected (2 σ1 tonal registers × 2 σ1 tonal contours × 2 σ2 tonal registers × 2 σ2 tonal contours × 2 items × 16 speakers).

2.3. Procedures

Recordings were conducted in a quiet room in Shaoxing, using a laptop equipped with an external sound card (Cakewalk UA-G1) and a headset (Sennheiser PC150). Stimuli were displayed on the screen with a randomized order for each participant using FieldPhon [16]. Participants read the stimuli aloud at a normal speech rate, which were recorded with 44 kHz sampling rate and 16-bit rate.

2.4. Data Preparation

Data from one male speaker were excluded from analysis due to creakiness in most of the tokens produced. Another 52 tokens were excluded due to disfluency or mistakes in production, as judged by a native speaker of Shaoxing Wu. As a result, 428 tokens from 15 speakers were included for further analysis. The sound files were manually segmented using Praat. *f0* values in Hz were extracted by taking ten equidistant points in the rime of each syllable and normalized to within-speaker z-scores. Rime duration of each syllable was also extracted.

2.5. Data Analysis

To analyse the *f0* trajectories, generalized additive mixed models (GAMM; [17]) were applied with the *mgcv* package in R [18]. The response of the model is the normalized *f0* value of disyllabic words. To understand whether and how the underlying registers and contours of both syllables condition disyllabic tone sandhi patterns, it is important to formally assess their effects on the height or shape of *f0* trajectories separately. We first created a variable as the predictor to present the interaction between underlying tonal registers and contours of σ1 and σ2, and then re-specified the models with ordered-factor difference smooths (following [19]). Factor smooths modelling non-linear differences over time for each speaker and each item were included as random effects.

For the analysis of duration, linear mixed-effects models were built with the *lme4* package in R [18]. Tonal context (start with falling vs. rising tone), syllable position (σ1 vs. σ2), and their interaction were entered into the models as fixed factors in a stepwise fashion. By-speaker and by-item intercepts and slopes for the effect of syllable position were included as random effects.

3. RESULTS

3.1. *f0* trajectories

We first investigated the effect of σ1 register and contour on the *f0* trajectories of disyllabic words.

Fitted *f0* trajectories of words ending with High Falling, Low Falling, High Rising, and Low Rising tones (T1-T4) were plotted in Figure 2A-D. As shown in each panel of Figure 2, the contrasts of register and contour are preserved on σ1 (before the vertical dashed lines). Underlyingly high register tones are realized higher than low register tones, and underlyingly falling and rising tones differ in their surface contours: falling tones surface with high falling and low level pitch contours, while rising tones surface with high level and low rising contours. Zooming in on the *f0* realizations of σ2 (after the vertical dashed lines), they are also conditioned by the underlying contour of σ1. For words ending with rising tones (Figure 2C and 2D), their σ2 surface with falling tones after an initial rising tone, and surface with level-ish contours after a falling tone.

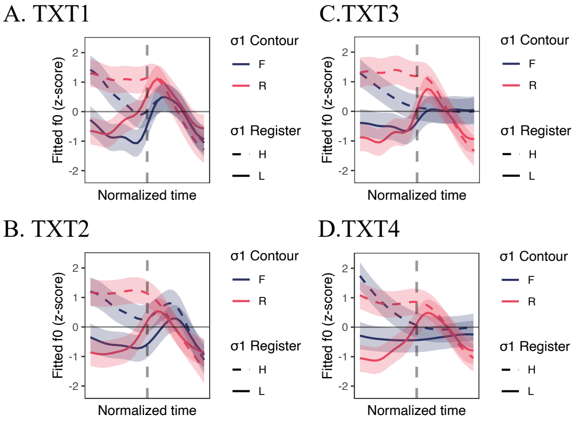
A. TXT1 C.TXT3 B. TXT2 D.TXT4

Figure 2: Fitted *f0* trajectories of disyllabic words ending with the same tone. Line types represent the register, colours the contour of σ1, and vertical dashed line syllable boundary.

To illustrate the effect of σ2 register and contour, fitted *f0* trajectories were re-plotted by presenting words starting with the same tone in the same panel of Figure 3.

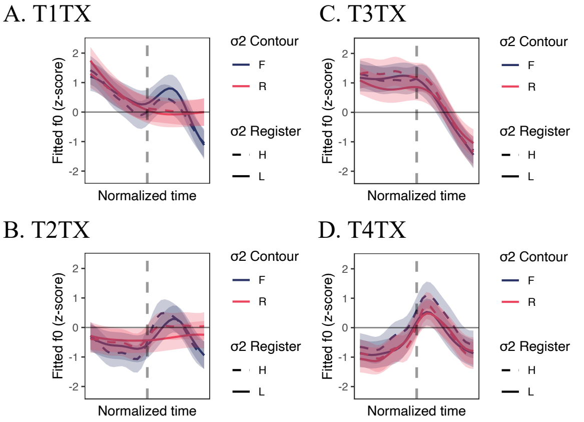
A. T1TX C. T3TX B. T2TX D. T4TX

Figure 3: Fitted *f0* trajectories of disyllabic words starting with the same tone. Line types and colours represent the register and contour of σ2 respectively.

Figure 3 illustrates that the *f0* realizations of σ1 are neither conditioned by the underlying register nor the contour of σ2: T1-T4 surface with high falling, low level, high level, and low rising pitch contours respectively, regardless of the tonal identities of their following syllables. As for the *f0* realizations of σ2, the underlying registers of σ2 seem to be neutralized: in each panel, underlyingly high register and low register tones surface alike. Regarding the effect of σ2 contour, however, there is a sharp contrast between words starting with underlyingly falling (T1 and T2, Figure 3A-B) and rising tones (T3 and T4, Figure 3C-D). Following a falling tone, the contours of σ2 are preserved to some extent. After a rising tone, σ2 are neutralized to a falling contour.

The statistical results in Table 1-4 further confirmed the above observations. Briefly speaking, both the register and contour of σ1 showed a significant effect, while only the contour of σ2 had a significant effect in words starting with falling tones.

σ1 × σ2		*β*	SE	*t*	*p*
F × H.F	H − L	0.678	0.254	2.670	0.015 *
R × H.F	H − L	0.597	0.248	2.404	0.032 *
F × L.F	H − L	0.705	0.236	2.978	0.006 **
R × L.F	H − L	0.913	0.256	3.567	< 0.001 ***
F × H.R	H − L	0.630	0.247	2.548	0.022 *
R × H.R	H − L	1.065	0.264	4.022	< 0.001 ***
F × L.R	H − L	0.773	0.250	3.087	0.004 **
R × L.R	H − L	0.902	0.233	3.871	< 0.001 ***

Table 1: Parametric terms modelling the constant difference between disyllabic words starting with High and Low register tones.

σ1 × σ2		edf	*F*	*p*
H × H.F	F − R	4.039	6.942	< 0.001 ***
L × H.F	F − R	3.137	2.439	0.068
H × L.F	F − R	5.279	10.118	< 0.001 ***
L × L.F	F − R	5.255	8.431	< 0.001 ***
H × H.R	F − R	1.000	58.777	< 0.001 ***
L × H.R	F − R	1.000	34.672	< 0.001 ***
H × L.R	F − R	1.000	49.57	< 0.001 ***
L × L.R	F − R	3.637	7.734	< 0.001 ***

Table 2: Smooth terms modelling the non-linear difference between disyllabic words starting with Falling and Rising tones.

σ1 × σ2		*β*	SE	*t*	*p*
H.F × F	H − L	-0.179	0.287	-0.622	1
H.R × F	H − L	0.193	0.272	0.711	0.955
L.F × F	H − L	0.190	0.277	0.685	0.987
L.R × F	H − L	0.451	0.273	1.649	0.198
H.F × R	H − L	0.082	0.273	0.301	1
H.R × R	H − L	0.044	0.288	0.153	1
L.F × R	H − L	0.430	0.273	1.576	0.230
L.R × R	H − L	0.095	0.273	0.349	1

Table 3: Parametric terms modelling the constant difference between disyllabic words ending with High and Low register tones.

σ1 × σ2		edf	*F*	*p*
H.F × H	F − R	4.927	9.765	< 0.001 ***
H.R × H	F − R	1.000	0.197	1
L.F × H	F − R	4.796	9.131	< 0.001 ***
L.R × H	F − R	1.623	0.267	1
H.F × L	F − R	5.520	12.697	< 0.001 ***
H.R × L	F − R	1.000	3.488	0.124
L.F × L	F − R	5.444	10.445	< 0.001 ***
L.R × L	F − R	1.002	0.498	0.956

Table 4: Smooth terms modelling the non-linear difference between disyllabic words ending with Falling and Rising tones.

3.2. Duration

Figure 4 shows that there is a difference in the duration of σ1 and σ2 in words starting with falling and rising tones, respectively. The effect of tonal context was significant, with the mean duration of words starting with falling tones being longer than those starting with rising tones ($\chi2(1) = 18.62$, $p < 0.001$). Additionally, the mean duration of σ2 was significantly longer than that of σ1 ($\chi2(1) = 15.64$, $p < 0.001$). The interaction between tonal context and syllable position was significant ($\chi2(1) = 11.86$, $p < 0.001$): the mean duration of σ2 was significantly longer than that of σ1 in both words starting with falling ($\chi2(1) = 19.42$, $p < 0.001$) and rising tones ($\chi2(1) = 5.66$, $p = 0.017$), and the interaction is mainly due to different magnitudes of lengthening in σ2.

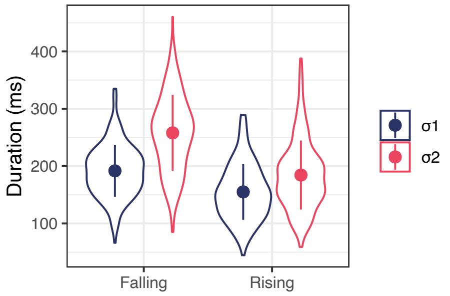

Figure 4: Duration of σ1 (blue) and σ2 (red) in words starting with falling and rising tones. The error bars represent the standard deviation, and the outlines of the violin plot indicate the kernel probability density.

4. DISCUSSION AND CONCLUSION

Our results demonstrated that in Shaoxing Wu, both the tonal register and contour of σ1 have significant effects on the *f0* trajectories of a disyllabic word. Specifically, the underlying register of σ1 is preserved; the contour contrast is also preserved on σ1, via conditioning the surface forms which are not identical to the underlying contours. The tone of σ2 also plays a role: after an initial falling tone, a falling tone preserves its contour, while a rising tone surfaces with a level-ish contour. Taken together, the results confirm impressionistic observations on Shaoxing Wu (e.g., [9]) with acoustic data.

These findings confirm that in Shaoxing Wu, registers can be preserved when contours undergo changes (e.g., σ1 in Figure 2A), and contours can be maintained despite register neutralization (e.g., σ2 in Figure 3A). These patterns contradict the idea that tonal register dominates tonal contour [3], and provide experimental evidence for the independence of tonal register and contour in sandhi processes of Wu dialects. They lend support to models of tonal representation proposing independent relationship between register and contour (e.g., [4]), which draw evidence only from impressionistically observed register or contour spreading.

Our findings also offer a more nuanced picture of the left- vs. right-dominant sandhi typology in Wu dialects. We have shown that σ2 can maintain their contours despite the overall left-dominant tendency, indicating that both patterns are present in the same dialect. This supports previous research on the complex tone sandhi patterns in Northern Wu dialects (e.g., [11], [12]). In addition, the results demonstrate that the effects of register and contour can differ between σ1 and σ2. [13] proposed that in Northern Wu, σ1 tends to preserve registers, and σ2 preserves contours. Our data align with their findings, showing that σ2 does preserve contours to some extent, and σ1 preserves both register and contour.

The findings on syllable duration provide further insights into the relationship between sandhi dominance and metrical prominence. We found that σ2 is longer than σ1, be it in words starting with falling or rising tones, although the syllable-position-based durational contrast is greater when σ1 bears a falling tone (Figure 4). As confirmed by the results on *f0*, words starting with falling (Figure 3A-B) and rising tones (Figure 3C-D) undergo right- and left-dominant sandhi respectively. If this distinction is due to different metrical structures [14], we would expect to see different durational patterns between the syllables in each type of metrical structures; σ2 should be longer than σ1 in a right-dominant structure, and vice versa in a left-dominant structure. However, no such durational differences were observed, raising questions about cues of metrical prominence motivated by tone sandhi dominance patterns.

To conclude, with controlled experimental data, this study provides evidence that tonal register and contour behave independently in sandhi processes, and offers a more nuanced understanding of the complex tone sandhi patterns in Northern Wu dialects. In addition, our findings suggest the lack of a straightforward relationship between tone-sandhi based metrical prominence and syllable durational contrasts.

5. ACKNOWLEDGEMENTS

This research is supported by CSC scholarship to Xinyi Wen and NWO VICI grant (No. vi.c.181.040) to Yiya Chen.

6. REFERENCES

[1] K. Pike, *Tone languages*. Ann Arbor: University of Michigan Press, 1948.

[2] M. Yip, *Tone*. Cambridge University Press, 2002.

[3] M. Yip, “Contour Tones,” *Phonology*, vol. 6, no. 1, pp. 149–174, 1989.

[4] Z. Bao, *On the Nature of Tone*. Oxford University Press, 1999.

[5] Z. Bao, “Tonal Contour and Register Harmony in Chaozhou,” *Linguist. Inq.*, vol. 30, no. 3, pp. 485–493, 1999.

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

[7] F. Wang, “Shaoxinghua Jiyin (Sounds of Shaoxing),” in *Yuyanxue Luncong (Issues of Linguistics)*, vol. 3, Shanghai: Shanghai Educational Publishing House, 1959.

[8] J. Zhang, *The phonology of Shaoxing Chinese*. Utrecht: LOT, 2006.

[9] F. Wang, *Shaoxing Fangyan Yanjiu (Studies on Shaoxing Dialect)*. Beijing: Language and Culture Press, 2015.

[10] J. Zhang, “A directional asymmetry in Chinese tone sandhi systems,” *J. East Asian Linguist.*, vol. 16, no. 4, pp. 259–302, Nov. 2007, doi: 10.1007/s10831-007-9016-2.

[11] M. Y. Chen and H. Zhang, “Lexical and postlexical tone sandhi in Chongming,” in *Studies in Chinese Phonology*, J. Wang and N. Smith, Eds. Berlin, Boston: De Gruyter, 1997. doi: 10.1515/9783110822014.13.

[12] M. K. M. Chan, “Contour-tone spreading and tone sandhi in Danyang Chinese,” *Phonology*, vol. 8, no. 2, pp. 237–259, Aug. 1991, doi: 10.1017/S0952675700001391.

[13] J. Kuang, J. Tian, and Y. Zhou, “The common word prosody in Northern Wu,” in *Tonal Aspects of Languages 2018*, Jun. 2018, pp. 7–11. doi: 10.21437/TAL.2018-2.

[14] S. Duanmu, “Metrical and Tonal Phonology of Compounds in Two Chinese Dialects,” *Language*, vol. 71, no. 2, p. 225, Jun. 1995, doi: 10.2307/416163.

[15] C. Guo and F. Chen, “Phonetic Realizations of Metrical Structure in Tone Languages: Evidence From Chinese Dialects,” *Front. Psychol.*, vol. 13, pp. 1–17, Jul. 2022.

[16] W. Pan, L. Li, and X. Han, “Field Phon (Version F2.0.2).” 2015.

[17] S. N. Wood, *Generalized Additive Models: An Introduction with R*, 2nd ed. New York: Chapman and Hall/CRC, 2017. doi: 10.1201/9781315370279.

[18] R Core Team, “R: A language and environment for statistical computing.” 2021.

[19] M. Wieling, “Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English,” *J. Phon.*, vol. 70, pp. 86–116, Sep. 2018, doi: 10.1016/j.wocn.2018.03.002.