AN ACOUSTIC STUDY OF TONE SANDHI IN LISHUI WU

Wupeng Lan[1], Xiaocong Chen[1], Caicai Zhang[1]

Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR, China
21046493g@connect.polyu.hk, xiaocong.chen@polyu.edu.hk, caicai.zhang@polyu.edu.hk

ABSTRACT

This study investigated the tone sandhi in the Lishui Wu dialect, an understudied variety of Southern Wu dialects in Zhejiang Province, China. Previous studies showed that the tone sandhi in the Lishui dialect is mostly right-dominant, but there seem to be no uniform patterns across items and speakers. We recruited 4 elderly and 4 young native speakers of Lishui Wu and conducted acoustic analyses of the f0 of the first syllable in both disyllabic real and pseudo words. Overall, our results did not reveal consistent sandhi patterns for both real and pseudo words, but the first syllable tended to be mostly produced as falling tones (with different slopes), with the pitch onset height conditioned by the yin and yang tones. Further analyses also found that young speakers showed a larger similarity between the sandhi and citation tones than elderly speakers, indicating some generational changes in Lishui tone sandhi.

Keywords: Tone sandhi, Lishui Wu, age, lexicality

1. INTRODUCTION

It has been shown that Wu Chinese exhibits complex tone sandhi patterns, but previous acoustic studies mostly focused on the major varieties of Wu dialects such as Shanghai [1][2] and Wuxi dialect [3]. In this study, we investigated the tone sandhi pattern of an understudied variety of Southern Wu, the Lishui (丽水) Wu dialect. The Lishui dialect belongs to the Chuqu (处衢) subgroup of Wu dialects and is spoken mostly in the Liandu district, the main urban area of Lishui city in Zhejiang province, China [4].

The f0 patterns of the citation tones in Lishui Wu are displayed in Fig 1. Due to the historical evolution, the T4 (yangshang) category has merged with other tonal categories, with those derived from the voiced obstruent initials in Middle Chinese (T4a) merging with T2 (yangping), and those derived from the voiced sonorant initials (T4b) merging with T3 (yinshang) [5]. Thus, there are only seven distinctive citation tones [5]-[9], which consist of six smooth tones (T1: yinping 34; T2: yangping 32; T3: yinshang 43; T5: yinqu 51/41; T6: yangqu 341), as well as two checked tones (T7: yinru 45; T8: yangru 23).

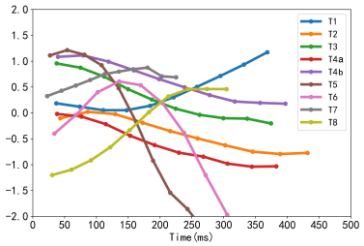


Figure 1. Pitch contour of citation tones in Lishui Wu (y-axis stands for normalized z scores of log-transformed f0)

Previous studies showed that the disyllabic tone sandhi pattern in Lishui dialect is mostly right-dominant, whereby the final syllable in a disyllabic word preserves its citation tone in most cases, while the first syllable tended to change to another tone [5]-[9]. However, previous impressionist phonetic transcriptions and acoustic studies did not consistently reveal well-formed patterns for the sandhi tones of the first syllable [5]-[9], though it was shown that most of the sandhi tones were level tones or falling tones [5][8]. Moreover, as shown in [8], the tone sandhi patterns seemed to vary among three speakers of different ages, indicating that the tone sandhi may undergo some changes across generations. It remains unclear whether old and young speakers show any difference in their tone sandhi patterns. Besides, it also remains a myth whether native speakers have any implicit knowledge of tone sandhi, as no consistent tone sandhi patterns were revealed in previous studies, and less is known about whether tone sandhi in the Lishui dialect can be productive in novel words.

Thus, this study aims to investigate how age and lexicality influence the tone sandhi patterns of the first syllable in disyllabic words of the Lishui dialect. We focused on addressing two questions: (a) What are the tone sandhi patterns of the first syllable in disyllabic real words and pseudowords in the Lishui dialect? (b) Are there any differences in the tone sandhi patterns between young and elderly speakers?

2. METHODS

2.1. Subjects

We recruited 8 native speakers of the Lishui dialect, including four elderly speakers (2 males and 2 females, aged between 72 and 78) and four young

speakers (2 males and 2 females, aged between 23 and 35), all of whom were born and raised in the main urban area of Lishui and had not lived outside Lishui for a long time before 18. They could speak the Lishui dialect fluently and used it frequently in their daily life. All the participants' parents were also born and raised in the main urban area of Lishui and spoke the Lishui dialect. They reported no language or hearing deficits and could read Chinese characters using the Lishui dialect.

2.2. Materials and procedure

Given the practical consideration for the experimental duration, we only examined the tone sandhi patterns of the smooth tones as the initial syllables. A real-word and a pseudo-word list were created. The real words were selected from common disyllabic words in the Lishui dialect, and the pseudowords were constructed by combining two existing morphemes to form nonsense items albeit with some syntactic structures (see below). All the real words and pseudowords were from 48 tonal combinations (6 tonal categories for the first syllable including T1, T2, T3, T5, T6 and the merged T4; 8 tonal categories for the second syllable including T1, T2, T3, T5, T6, T7, T8 and the merged T4), with four items for each of the tonal combinations in both real words and pseudowords respectively. For the four items of each tonal combination for both real and pseudowords, two had the Modifier+Noun (MN) structure and the other two had the Verb+Object (VO) structure (also see [10]). There were 384 items (48×4×2) in total. Four additional Lishui native speakers confirmed the selected real word words as frequent words and the constructed pseudowords as non-existing items. A single-character list was also created for the measurement of citation tones.

The recording took place in a quiet room. The stimuli were presented one by one by DMDX on a laptop computer screen and participants' pronunciations were recorded by a condenser microphone SAMSON C03U sampled at 22050 Hz with 16-bit quantification. Participants were asked to pronounce the single-character list first, and then completed the real-word and the pseudoword list, with the order of the two lists counterbalanced across participants. Before completing the pseudoword list, the experimenter also told participants to pronounce two characters together as if they formed a real disyllabic word and provided four additional pseudowords for practice. The stimuli within each list were presented randomly. Participants were asked to pronounce each item twice. The experimenter listened to each participant's pronunciation and make sure that they did not make a mistake.

2.3. Data preprocessing and coding

All stimulus recordings were imported into Praat for manual annotation and segmentation. 10 equidistant f0 points in the annotated rime portion were automatically extracted using the STRAIGHT algorithm in VoiceSauce [11]. We removed those tokens with inaccurate f0 tracking due to creaky voicing and obtained 2536 tokens for real words and 2528 tokens for pseudowords. We computed the log-z (LZ) values by log-transforming f0 values and performing the z-score normalization to remove any speaker/gender differences [12]. We also averaged the LZ values of the tokens for each item for each participant and transcribed the tones on a 5-level scale. We did not analyse the items when the T4 was the first syllable due to its merging with other tonal categories. We thus only focused on the sandhi tonal realizations for T1, T2, T3, T5 and T6 as the first syllable in the analysis.

3. RESULTS

3.1. Tone sandhi patterns in real and pseudo words

In general, we found that for each tonal combination, the sandhi tones of the first syllables from the same tonal category were realized as more than one type of pitch contour in most cases, and the sandhi tone realizations also varied from speaker to speaker. Thus, no consistent patterns were clearly detected. However, we found that, the pitch onset height of the sandhi tones was mostly conditioned by the yin and yang tones, a phenomenon similarly observed in other Wu dialects [5]. Specifically, the sandhi tones of the yin tones (T1, T3, T5) started at a higher f0 onset (above LZ score 0) whereas those of the yang tones (T2, T6) started at a lower f0 onset (below LZ score 0) (see Fig 2).

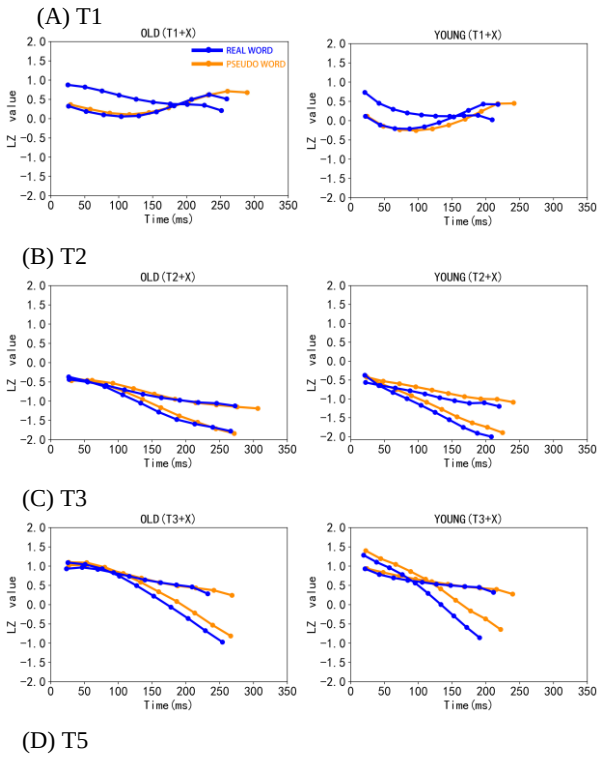
Given the various realizations of the sandhi tones, we further categorized the sandhi tones into the following types: Falling, Level, Rising, and Convex tones, based on the pitch contour of the sandhi tones.[1] Moreover, for falling tone, yin tones (T1, T3 and T5) mainly produced to a high falling tone, with its f0 onset over LZ score 0, whereas yang tones (T2, T6) mainly produced to a low falling tone, with its f0 onset below LZ score 0. Additionally, different slopes of falling tones were also detected. We subdivided the falling tones into four subcategories: high-onset steeper-slope falling tone (FSP+), high-onset shallower-slope falling tone (FSL+), low-onset steeper-slope falling tone (FSP-), and low-onset shallower-slope falling tone (FSL-). We then calculated the proportion of each tonal type for each tonal category and displayed those dominant types of pitch contour of the sandhi tones from each tonal

category in both real and pseudo words for the elderly and young age groups separately in Fig 2. In general, we found that the sandhi tones of the same tonal category exhibit different realizations in both real and pseudo words. Moreover, we also found that the realizations of the sandhi tones were not clearly conditioned by the following tone.

As for T1, in real words, all participants tended to produce it as a high-onset shallower-slope falling tone /43/ (old:29% vs. young:43%), or a rising tone /34/ (old:45% vs. young:41%) that was very similar to the citation tone of T1 /34/. In contrast, for pseudowords, all participants tended to produce it as the rising tone /34/ (old:71% vs. young:83%), indicating that they may choose to preserve the citation tone of T1 in novel word contexts.

For T2, in real words, all participants tended to produce it as low-onset falling tones with a shallower slope /32/ (old:31% vs. young:37%) or a steeper slope /31/ (old:25% vs. young:29%). The same pattern was also observed in pseudowords (/32/: 34% for old people and 39% for young people; /31/: 20% for old people and 23% for young people). Note that these two falling tones are also similar to the citation tone of T2 /32/.

For T3, in real words, all participants tended to produce it as high-onset falling tones with a shallower slope (/43/ (old:50%; young:51%) or a steeper-slope falling /42/ (old:48%; young:30%). A similar pattern was also detected in pseudowords (/43/: 48% for old people and 59% for young people; /42/: 50% for old people and 29% for young people). Again, the sandhi tones are very similar to the citation tone of T3 /43/.



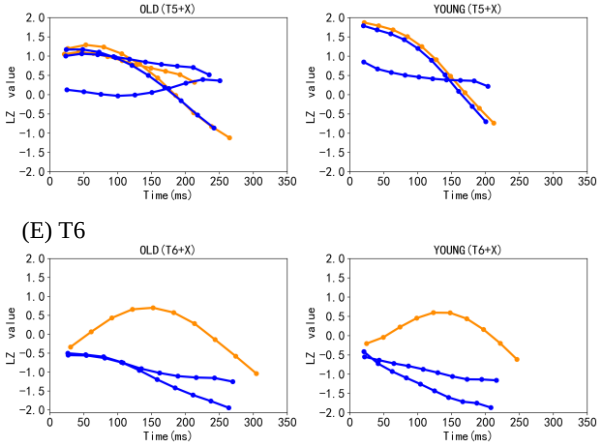


Figure 2. The pitch contour of the dominant types of sandhi tones from T1, T2, T3, T5 and T6 as the first syllable in real and pseudo words for the elderly and young age groups (—real word, —pseudoword).

For T5, in real words, old speakers tended to produce it as a high-onset steeper-slope falling tone /42/ (43%), a high-onset shallower-slope falling tone /43/ (28%) or a rising tone /34/ (20%). In contrast, young speakers tended to produce it as a high-onset steeper-slope falling tone /52/ (44%) or a high-onset shallower-slope falling tone /43/ (24%). In pseudowords, old speakers tended to produce it as a high-onset steeper-slope falling tone /42/ (71%) and a high-onset shallower-slope falling tone (43) (25%) while young speakers tended to dominantly produce it as a high-onset steeper-slope falling tone /52/ (73%), which seemed to more resemble the citation tone of T5 /51/.

For T6, in real words, all participants tended to produce it as low-onset falling tones with a shallower slope /32/ (old: 27% vs. young: 31%) or a steeper slope /31/ (old:21% vs. young:24%). In contrast, for pseudowords, all participants tended to produce it as the convex tone /342/ (old:44% vs. young:42%), which was very similar to the citation tone of T6 /341/. This seemed to indicate that they tended to preserve the citation tone of T6 contexts.

In sum, although we did not detect consistent patterns for each tonal category, we found that the dominant types of sandhi tones were mostly falling tones (with different slopes), and the pitch onset height was conditioned by the yin and yang tones. Moreover, it seemed that the tone sandhi patterns of old and young speakers mostly converged except for T5. In addition, it seemed that people tended to preserve the citation tones for some tonal categories in pseudowords.

3.2 Analysis of the similarity between citation tones and sandhi tones

As shown in the previous section, it seemed that some of the dominant types of the sandhi tones were similar

to the citation tones of the same tonal category, indicating that participants may choose to preserve the citation tone in some cases. Thus, to quantify the degree of tone sandhi, we followed [13] and measured the pitch similarity between the sandhi tones and citation tones by computing the Euclidean distance between the 10 f0 points of the derived tones from each token of the initial syllable and the 10 averaged f0 points of the citation tone from the same corresponding tonal category for each participant respectively. The Euclidean distance was calculated by the following equation:

$$d(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_{10}-y_{10})^2}$$

To further investigate how age and lexicality influenced the degree of tone sandhi, we conducted the linear mixed-effect model analysis on the Euclidean distance using JASP [14], with age and lexicality included as the fixed-effect predictors, and the structure as the control covariate into the model. Due to the model convergence problem, we only included the by-participant and by-item intercept. The model results were summarized in Table 1.

Table 1. Fixed effect estimates for Linear Mixed Models results of Euclidean distance.

Effect	df	F	p
Structure	1, 320.50	0.207	.650
Age	1, 6.00	5.970	.050
Lexicality	1, 320.5	0.109	.741
Structure*Age	1, 4736.40	0.063	.801
Age * Lexicality	1, 4736.42	11.885	< .001
Structure * Lexicality	1, 320.50	<.001	.996
Structure*Age*Lexicality	1, 4736.42	0.066	.797

As shown in Table 1, we found that there was a significant main effect of age and a significant interaction effect between age and lexicality on the Euclidean distance. Further analyses revealed that the mean Euclidean distance was smaller for young speakers than old speakers in both real words (β=88.793, *SE*=34.09, *z*=2.605, *p*=.028) and pseudo words (β=77.586, *SE*=34.09, *z*=2.276, *p*=.046). This showed that the sandhi and citation tones were more similar for young than old speakers, indicating that young speakers tended to preserve the citation tones. Moreover, there was also a significant difference in the mean Euclidean distance between the real and pseudo words among old speakers (β=-7.588, *SE*=2.506, *z*=-3.028, *p*=.010), as old speakers showed more similarity and tended to preserve the citation tones more in pseudowords than real words. In contrast, no lexicality effect was found among young speakers (β=3.618, *SE*=2.501, *z*=1.447, *p*=.148), possibly because young speakers tended to preserve the citation tones in both real and pseudowords.

4. DISCUSSION AND CONCLUSION

In this study, we found that in Lishui disyllabic words, the dominant sandhi tone types of first syllables were mostly falling tones (with different slopes). This is slightly different from previous results [5]-[9], which found that the sandhi tones were predominantly level tones and falling tones. This may be due to the slight difference between the perceptual and subjective judgment for the classification of level tone. Some of the tones classified as shallower-slope falling tones in this study may be perceptually classified as level tones in previous studies. Moreover, we also found that the pitch onset height of the sandhi tones was conditioned by the yin and yang tones, consistent with similar previous observations in Wu dialects [5]. However, this experiment revealed that the tone sandhi patterns of the Lishui disyllabic words were extremely complex and did not exhibit clear patterns for each tonal combination (also see similar observations in other Wu dialects [5]). The reasons for this perplexing complexity are still unknown, which need further examination.

Our further analyses revealed that young speakers tended to show a larger similarity between the sandhi and citation tones in real words. This seems to indicate that the sandhi tones in the Lishui dialect have a trend of preserving the original citation tones from the corresponding tonal category instead of changing to another different tone in younger generations. Moreover, for pseudowords, old speakers also showed a larger distance between the sandhi and citation tones than young speakers, which indicated that old speakers showed a higher degree of tone sandhi and could apply some knowledge of tone sandhi even in pseudowords. However, we also found that old speakers had a greater degree of tone sandhi in real words than pseudowords. That is to say, the realizations of tone sandhi in real words did not fully apply to pseudowords. This seemed to indicate that the sandhi tones were probably stored with the specific lexical items in native speakers' mental lexicon [15]. On the whole, the degree of tone sandhi in old speakers is greater than young speakers. This shows that Lishui dialect has a trend of becoming less tone sandhi as time changes. However, note that our study did not manipulate the semantic transparency [3], which could further affect the tone sandhi productivity. Thus, further research is required.

Moreover, the Euclidean distance used in this study only took the mean LZ value into consideration and may not fully reflect the similarity in terms of the pitch contour. Other similarity measures taking the tonal shape into account need to be explored in future studies.

5. ACKOWLEDGEMENTS

We thank Wanxiang Subdistrict Offices, Liandu District, Lishui City. We thank Tu Huimin, Chong Shufen, Zhao Lijun and Tong Chang, who gave us selfless help in the pre-participant recruitment and the recording place. We thank the eight participants for taking the time to participate in our experiment. Additionally, we would like to thank Gao Ziqi, for his help in writing the code.

6. REFERENCES

[1] Zhu, X. 2006. *A Grammar of Shanghai Wu*. Lincom Europa.
[2] Zhang, J., Meng, Y. 2016. Structure-dependent tone sandhi in real and nonce disyllables in Shanghai Wu. *Journal of Phonetics*, *54*, 169-201.
[3] Yan, H. 2017. *The nature of variation in tone sandhi patterns of Shanghai and Wuxi Wu*. Springer.
[4] Xiong, Z.H., Zhang, Z.X., Huang, X., Dao, B., & Zou, J.Y. 2013. *Language Atlas of China*. The Commercial Press.
[5] You, R.J., & Yang, J.Q. 2001. *Wuyu shengdiao de shiyan yanjiu*[An Experimental Study of Wu tone sandhi]. Fudan University Press.
[6] Ji, X.L. 1988. *Zhongguo Dabaike quanshu:Yuyan Wenzi*[Encyclopedia of China: language and character]. Encyclopedia of China Publishing House.
[7] Wang, W.S. 2002. Zhejiang Lishui fangyan de yuyin tedian[Phonetic features of Lishui dialect of Zhejiang]. *Lishui shifan zhuanke xuexiao xuebao*, *24*(1), 46-49.
[8] Steed, W.J. 2010.*Lishui Wu tone and tone sandhi: an acoustically-based description*. Dissertation, ANU.
[9] Lei, Y.P. 2019. *Zhejiang fangyan ziyuan diancang: Lishui*[Zhejiang dialect resource collection: Lishui]. Zhejiang University Press.
[10] Yan, H. 2013. *The productivity of tone sandhi patterns in Wuxi Chinese.* Master thesis, University of Kansas
[11] Shue, Y.-L., P. Keating , C. Vicenik, K. Yu. 2011. VoiceSauce: A program for voice analysis. *Proceedings of the ICPhS XVII*, 1846-1849.
[12] Zhu, X. 2010. *Yuyin xue*[Phonetics]. The Commercial Press.
[13] Rhee, Nari., Kuang, J.J., & Chen, Aoju. 2022. The effect of musicality on the development of Mandarin prosody. *Speech Prosody*, 704-707.
[14] JASP Team. 2022. JASP (Version 0.16.3)[Computer software].
[15] Hsieh, H. I. 1976. On the unreality of some phonological rules. *Lingua*, *38*(1), 1-19.

[1] Concave tones and rising tones were classified into the same category as the number of concave tones was small and concave tones did not form phonological contrast with rising tones in the Lishui dialect.