10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

The Effect of Focus and Prosodic Boundary on the two T3 sandhi in Northeastern Mandarin

Ping Cui[1], *Jianjing Kuang*[2], *Yunjia Wang*[3]

[1 3]Peking University
[2]University of Pennsylvania
cuiping@pku.edu.cn, kuangj@ling.upenn.edu, wangyunjia@pku.edu.cn

Abstract

Northeastern Mandarin has a similar tonal system as Beijing Mandarin. However, their disyllabic tone sandhi patterns are different. T3 becomes a mid-rising tone like T2 when it is followed by T1 or T3 in Northeastern Mandarin. Recently it has been found that in Beijing Mandarin syllable stress and prosodic boundary have an effect on T3 sandhi. In this study, we ask if the prominence of T3 bearing syllable and prosodic boundary affect the acoustic realization of the two types of T3 sandhi in Northeastern Mandarin. Do T3-sandhi before T1 and T3-sandhi before T3 have the same phonetic realization? Overall results show that the underlying citation form of T3 is better retained in sandhi contexts when i) the syllable that bears T3 is focused, and ii) strong prosodic boundary is present within T3T3 and T3T1. In addition, there is also a difference between T3T3 and T3T1. T3-sandhi before T3 could cross the larger prosodic boundary, though it is still conditioned by the stress of the syllable. However, T3-sandhi before T1 cannot span the strong prosodic boundary, regardless of its syllable stress. The results also imply that the acoustic realization of T3 sandhi is gradient and continuous.
Index Terms: Tone sandhi, contrastive focus, prosodic boundary, Northeastern Mandarin

1. Introduction

Arguably the most well-known case of tone sandhi that takes place in Beijing Mandarin is T3 sandhi (often referred to as Standard Mandarin). The T3 sandhi rule in Beijing Mandarin has been well established since [15]: when two T3 tones are in a row, the first T3 changes into a rising tone.

Northeastern Mandarin is similar to Beijing Mandarin in the tone type and the pitch value except that T1 is slightly lower than Beijing Mandarin. Many scholars describe it as /44/ or /33/. As shown in Table 1 (also see Fig. 1).

Table 1: *Various transcriptions of citation tone in Northeastern Mandarin*

	T1	T2	T3	T4
Song (1963)	[44]/[33]	[35]/[24]	[213]	[53]/[52]/[42]
He (1986)	[44]/[33]	[35]/[24]	[213]/[212]	[53]/[42]
Jia et al (2009)	[33]	[35]	[213]	[41]
Ye (2013)	[44]	[24]	[313]	[51]
Cui & Wang (2019)	[44]	[35]	[313]	[52]

The transcriptions are given in 'Chao numbers', whereby a speaker's tonal range from low to high is represented by a numerical scale from '1' to '5' [1], [6].

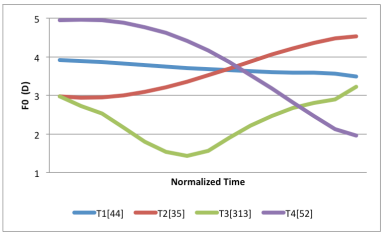


Figure 1: *F_0 results of the four tones in monosyllables on a normalized 1–5 numerical scale in Northeastern Mandarin* [28].

[28] find that although the citation tone patterns of Northeastern Mandarin are very similar to those of Beijing Mandarin, their disyllabic tone sandhi patterns are significantly different. T1 always changes to a mid-falling tone in the domain-final position (Fig. 2 left), while T3 becomes a rising tone when it precedes T1 or T3 (Fig. 2 right). Meanwhile, the experimental results show that native speakers cannot distinguish the minimal pair of "T3+T1" and "T2+T1".

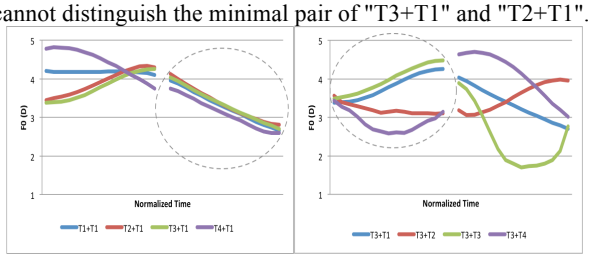


Figure 2: *F_0 results of T1 sandhi and T3 sandhi in disyllables in Northeastern Mandarin.*

In Beijing Mandarin, although the results of the perception experiment prove that native speakers can't distinguish the T3 sandhi and T2 ([3], [5]), whether the T3 sandhi is the same as T2 in phonetics, the opinions of the academic community are not consistent. [2] believes that T3 sandhi is the same as T2 in the absence of stress, but when the stress is carried, the pitch value of T3 sandhi is /24/ instead of /35/. A series of acoustic analyses by [9], [11-12] and others have confirmed that there is a slight difference between T3 sandhi and T2: the average f_0 and the low point on the pitch are lower than those of T2. [22-23] further validated their opinions in the comparison of large-scale natural corpora and true and novel words.

On this basis, many scholars have attempted to conduct further research and analysis of T3 sandhi in more complex

10.21437/SpeechProsody.2020-52

speech contexts, such as different prosodic boundaries and focus conditions. [7] believes that the form of T3 sandhi in the non-final position depends on whether the following T3 belongs to the same foot as it. [10] also believes that the occurrence of tone sandhi is related to syntactic and rhythmic tightness, but in extreme cases (such as faster speech rate), the two may not work. [13] proposed the concept of minimal rhythmic unit in the framework of optimization theory. If there are continuous T3 combinations in one MRU, the former T3 will change; if two adjacent T3 belong to two MRUs, there are two cases based on the tone of the latter syllable of the second MRU.

The above studies are mainly purely theoretical and phonological. [17] used empirical methods to observe whether the T3 changes and the specific acoustic realization of the pitch change under different prosody conditions in Beijing Mandarin. They find that the T3 sandhi could apply cross prosodic boundaries, but the distance between two T3 syllables is responsible for the likelihood of the application of tone sandhi. That is, it's more likely to occur on smaller boundaries.

[26] further added the condition of stress to compare the acoustic realizations of T3 sandhi and T2 in Beijing Mandarin. They observed the difference between T3 sandhi and T2 in two stress conditions and on three rhythmic boundaries respectively. The results show that the acoustic T3 sandhi-T2 differences under all the prosodic conditions are significant, and the magnitude of the difference under all the prosodic conditions varies along the order of pause-lengthening > across foot > distinctive stress within foot > undistinctive stress within foot. However, they only examined the effect of stress on the pitch in the same foot and didn't compare the effect of stress on the other two prosodic boundaries. Therefore, the different realizations of T3 sandhi in their results may be caused by different information focus of the sentence produced by different speakers.

[27] compared the pitches of T3 sandhi and T2 in a more complicated speech context, and observed the regressive effect of the T3 sandhi and T2 on the preceding syllable. It is found that when the focus syllable and the post-focus syllable are across foot rhythmic boundaries, the regressive effect of T3 sandhi in the focus position on the preceding tone is greater than that of T2. While they are on the same foot, T3 sandhi and T2 in the focus position have the same regressive effect on the preceding tone.

Is the T3 sandhi in Northeastern Mandarin similar to that of Beijing Mandarin? That is, whether the stress of the syllable that bears T3-sandhi and the prosodic boundary between the adjacent tones affect the realization of T3-sandhi? Do T3-sandhi before T1 and T3-sandhi before T3 have the same phonetic realization? These issues haven't been specifically studied.

Given the above reasons, this paper intends to study the following specific issues: (1) Whether the focus will affect the T3 sandhi so that it retains more underlying features of the citation tone, thereby increasing the difference between T3 sandhi and T2; (2) As the prosodic boundary between syllables becomes looser, will the sufficiency of the realization of the T3 sandhi rule become weaker; (3) If the underlying features of citation tone retained by the T3 sandhi is a "function" of the focus and prosodic boundary, how to explain the cause of the T3 sandhi? Is the acoustic variation of the T3 sandhi continuous or categorical? In all, this paper is to investigate if the effects of these intonation levels can highlight the pitch characteristics at the phonological level. It supplements and expands its research on the T3 sandhi in Beijing Mandarin.

2. Methods

2.1. Data collection

Recordings were made on a field trip to the northeastern part of China in 2019. Twelve females and twelve males (age 25 to 69 at the time of recording) who were native speakers of Northeastern Mandarin participated in the study. They all grew up in this region and have no long-term (more than one year) experience living outside. The recording was done in a quiet office using xRecorder [25], with a Terra Tec DMX 6-Fire sound card and isk RM10 micro-phone, at a sampling rate of 44.1 kHz.

In the material for production, we put 16(4*4) disyllabic words further into different carrying sentences, so that they are in different focus positions (pre-focus, focus, and post-focus) and in different prosodic boundaries (within rhythmic foot, across the foot and pause-lengthening). We used the contrastive focus in the carrying sentences. Only the syllables preceding the target words were different in the two conditions. The first had a high tone(T1), while the second had a low tone (T3). Each word is presented twice, and a total of 16*12*2=384 pronunciation items are obtained. All words are arranged in random order. Some examples of pronunciation items are shown in Table 2.

Table 2: *Examples(T3+T3) of pronunciation items*

Type	Focus-PostFocus	PreFocus-Focus
Pause-lengthening(subject-predicate construction)	高**[醒]**早上就走了，不是高飞(Gao Xing left in the morning, not Gao Fei).	高醒**[早]**上就走了，不是晚上(Gao Xing left in the morning, not at night).
Across the foot(modifier-head construction)	我找高波导游，不找高**[醒]**导游(I am looking for Guide Gao Bo, not looking for Guide Gao Xing).	我找高醒董事，不找高醒**[理]**事(I am looking for Executive Director Gao Xing, not looking for Director Gao Xing).
Within rhythmic foot	我是说张姐，不是说**[李]**姐(I mean Ms. Zhang, not Ms. Li).	我是说李哥，不是说**李**[姐](I mean Mr. Li, not Ms. Li).

For example, the prosodic boundary between the two syllables of the target words("醒" and "早") in the first row is pause lengthening(subject-predicate construction), and the focused syllable in the first column is "xing(醒)", while the focus of the second column is "morning(早)".

2.2. Acoustic measures

The software used for the extraction of fundamental frequency is Praat. We manually annotated the tonal nuclei by defining the feature points of the tone, as shown in Fig. 3.

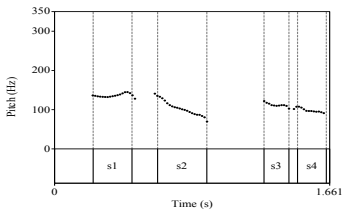


Figure 3: *Annotation of tone nuclei for tonal sequences (x-axis: Hz; y-axis: s)*

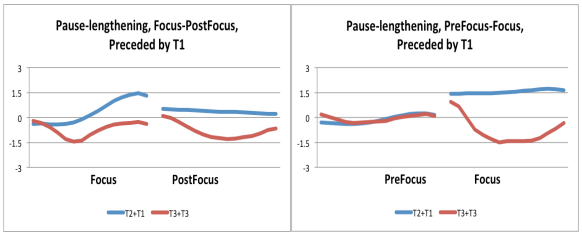
The tonal nuclei were analyzed using Takahiro's script [27]. 15 f_0 points on the contour were evenly extracted from each tone nucleus. Within the speaker, normalization was done for f_0 values using z-score by equation (1), where x stands for specific f_0 values, μ and σ respectively stand for the mean and deviation of all f_0 values from a certain participant.

$$z=(x-\mu)/\sigma \quad (1)$$

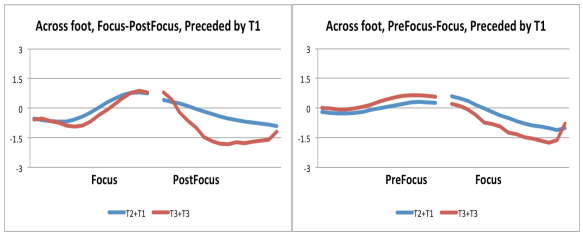
The special note here is the pitch measurement of the syllable carrying T3. Because T3 often has a creaky voice, the periodic information of T3 in the PointProcess file automatically generated by Praat is often incorrect. At this time, we mainly solve this problem by adjusting the parameters of the Praat software, observing the waveform and manually correcting the PointProcess file.

3. Results

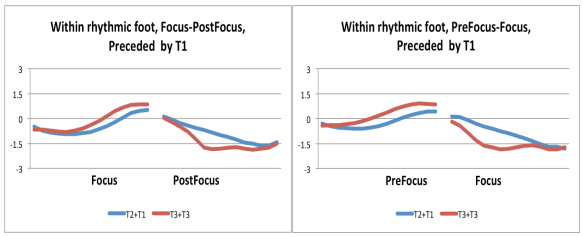
Because we are examining the effects of focus and prosodic boundary on the tone, we will mainly analyze the pitch contour of the target words for the time being, regardless of the tone of other syllables in the carrying sentences. We focus on the analysis of T3+T3 and T3+T1 in different focal positions and prosodic boundaries, and also compare the two T3 sandhi with T2.



(a)



(b)



(c)

Figure 4: *F_0 patterns of T3+T3 vs. T2+T1. a, b, c respectively represent the prosodic boundary of the target word as pause-lengthening, across the foot, and within rhythmic foot. The red pitch track represents the acoustic realization of T3+T3, while the blue one represents T2+T1. The preceding tone of the target words is T1.*

Fig. 4 shows the pitch contours of T3+T3 posit in different focal positions and cross different prosodic boundaries. As shown in Fig. 4(a), we may observe that when the prosodic boundary between the two syllables of the target words is pause lengthening, the focused syllable carrying T3 retains more underlying features of the citation tone. While the unfocused tone loses its concave character and becomes a flat or slightly rising tone. In Fig. 4(b), when the prosodic boundary between the two syllables of the target words is across the foot, the focused T3 still retains some underlying features, while the unfocused T3 in the non-final position changes to a rising tone similar to T2. However, the rising slope is relatively small. In Fig. 4(c), when the prosodic boundary between the two syllables of the target words is within rhythmic foot, T3 becomes a rising tone similar to T2 before T1, whether it is focused or unfocused. That's to say, the focus syllable usually retains more features of the underlying tone. Besides, the pitch curve has an obviously rising or falling trend and the pitch range is much more broad in the focal position. The pitch range of T3 of the first syllable in different focal positions and prosodic boundaries is shown in Table 3.

Table 3: *The average pitch range of T3 of the first syllable in the two tonal combinations(unit: z-score)*

T3+T3	Pause-lengthening	Across foot	Within rhythmic foot
focus	1.01	1.81	1.84
non-focus	0.54	0.71	1.34
T3+T1	**Pause-lengthening**	**Across foot**	**Within rhythmic foot**
focus	1.59	1.27	1.51
non-focus	1.56	0.83	1.41

So the effect of prosodic boundary on T3+T3 in Northeastern Mandarin and Beijing Mandarin has both similarities and differences. In Beijing Mandarin, the T3 sandhi could cross the foot and even larger prosodic boundary, and the looser the prosodic boundary between the two tones, the more the monophonic features of the T3 sandhi are retained acoustically. However, T3 sandhi cannot span the pause-lengthening boundary in Northeastern Mandarin. Besides, tone sandhi is also related to focus, that is, the focused syllable carrying T3 retains more underlying features of the citation tone.

Let's then compare T3+T1 and T2+T1 in different focal positions and prosodic boundaries.

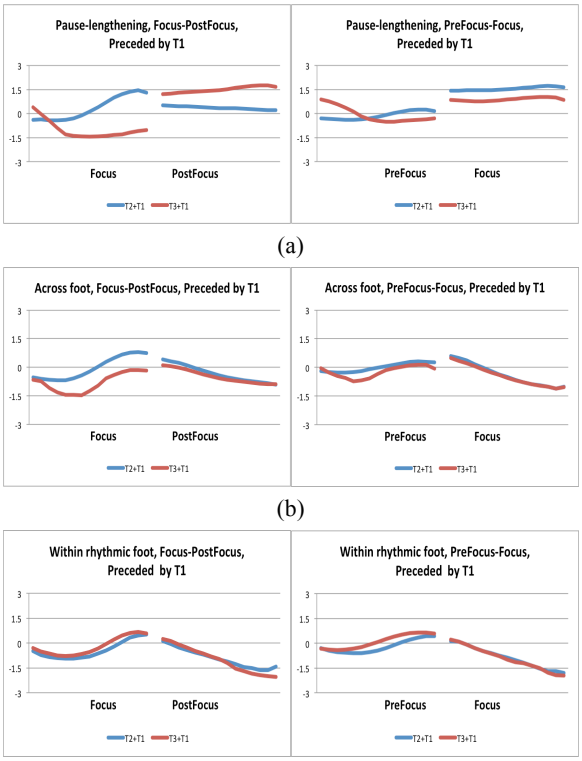


Figure 5: *F_0 patterns of T3+T1 and T2+T1.*

Fig. 5 shows the pitch contours of T3+T1 in different focal positions and prosodic boundaries. As demonstrated in Fig. 5(a), we may observe that when the prosodic boundary between the two syllables of the target words is pause lengthening, T3 in the non-final position retains more underlying features of the citation tone, whether it is focused or unfocused. It's a little different from the T3+T3 combination, in which the focus has an effect on T3 of the first syllable. We further compare the first T3 in T3+T3 and T3+T1 through paired sample T-test. The result is shown in Table 4.

Table 4: *T-test result of pitch difference of T3 of the first syllable (unit: z-score)*

Item	Highest point	Lowest point	Pitch range	Pitch contour mean
T3+T3	0.48	-0.06	0.54	0.21
T3+T1	-0.29	-1.3	1.01	-0.87
p	0.046*	0.2	0.429	0.64

From the pitch values(As shown in Table 4, Fig. 4(a) and 5(a)), the pitch curves of the two sets are obviously different. Under the condition of pause-lengthening and unfocus, the first T3 of T3+T3 is closer to T3 sandhi similar to T2, while the T3 of T3+T1 retains more features of the underlying tone, such as [+low], [+rising]. However, the statistical results show that there is only a significant difference between the two at the highest point.

In Fig. 5(b), when the prosodic boundary between the two syllables of the target words is across the foot, the focused T3 still retains some underlying features, while the unfocused T3 in the non-final position changes to a rising tone similar to T3 sandhi in disyllabic combinations. In Fig. 5(c), when the prosodic boundary between the two syllables of the target words is within rhythmic foot, T3 becomes a rising tone similar to T2 before T1, whether it is focused or unfocused. Besides, T1 changes to a falling tone in the final position under the condition of cross-foot and within rhythmic foot.

Like T3+T3, the focus syllable usually retains more features of the underlying tone in T3+T1. Besides, the pitch curve has an obviously rising or falling trend and the pitch range is much more broad in the focal position. The pitch range of T3 of the first syllable in different focal positions and prosodic boundaries is shown in Table 4.

4. Discussions and Conclusions

The present study investigates how do the prosodic boundaries and focus conditions work on the two T3 sandhi in Northeastern Mandarin. We found that as the prosodic boundaries between syllables become looser, T3 sandhi rules become less likely to apply accordingly, which is similar to Beijing Mandarin [17]. However, compared to T3 sandhi in Beijing Mandarin, T3 sandhi in Northeastern Mandarin restrains to smaller prosodic boundaries. T3 sandhi cannot apply across phrase boundaries in Northeastern Mandarin. Meanwhile, tone sandhi in Northeastern Mandarin is also sensitive to focus conditions. The focused syllable carrying T3 in the non-final position retains more features of the underlying tone, thereby increasing the difference between T3 sandhi and T2.

Importantly, there is a crucial difference between T3+T3 and T3+T1. When the prosodic boundary of the target words is pause-lengthening, a relatively larger boundary, the T3 of T3+T1 bears more features of the underlying tone, such as [+low], [+rising]. However, the first unfocused T3 of T3+T3 loses its concave character and only keeps the low feature. This means that the rules of T3+T3 sandhi are stronger than those of T3+T1 sandhi. T3+T3 sandhi can cross the larger prosodic boundaries than T3 + T1. The above results also imply that the acoustic realization of T3 sandhi is gradient. At least at the acoustic level, the realization of the sandhi rules may be continuous. It seems easier to account for the continuous variation of T3 sandhi if the prototype of T3 is considered to have both the target of [+low] and [+rising]. Based on this hypothesis, we agree that the rising form of T3 sandhi is caused by the anticipatory raising effect of its following T3 or T1([26], [28]). This regressive effect is weakened when the first syllable of the sandhi domain is focused, or the prosodic boundary between the two tones is loose and involving phrase-final lengthening.

Moreover, we also find that T1 sandhi is more complicated, which is subject to the following tones and the phrasal boundary, leading to tonal split. The change of T3 before T1 and the change of T1 itself may involve factors at the diachronic level. Due to space limitations, we will not discuss it here.

In sum, both the prominence of the syllable and the prosodic boundary between the adjacent tones affect the T3 sandhi in Northeastern Mandarin. The underlying features of citation tone retained by the T3 sandhi is a "function" of the focus and prosodic boundary.

5. References

[1] Y.R. Chao, *Mandarin primer: an intensive course in spoken Chinese*, Cambridge: Harvard University Press, 1948.

[2] S.E. Martin, "Problems of hierarchy and indeterminacy in Mandarin phonology", *Institute of History and Philology*, vol. 29, pp. 209-229, 1957.

[3] W.S-Y. Wang, K.P. Li, and R. L. Brotzman, "Research on Mandarin phonology", *Project on Linguistic Analysis*, Ohio State University Research Foundation 6:1-63, 1963.

[4] X. Song et al, 辽宁语音说略["The phonemic structure of Liaoning dialect"], 《中国语文》[*Journal of Chinese Linguistics*], vol.2, pp.104-114, 1963.

[5] W.S-Y. Wang and K.P. Li, "Tone 3 in Pekinese", *Journal of Speech and Hearing Research*, vol.10, pp. 629-63,1967.

[6] Y.R. Chao, *A grammar of Spoken Chinese*, Berkeley: University of Berkeley Press, 1968.

[7] C.L. Shih, *The prosodic domain of tone sandhi in Chinese*, PhD dissertation, University of California at San Diego, 1986.

[8] W. He, 东北官话的分区["Distribution of Northeastern Mandarin"],《方言》[*Fangyan(Dialect)*], vol.3, pp. 172-181, 1986.

[9] X. Shen, Tonal coarticulation in Mandarin, *Journal of Phonetics*, vol.18, pp. 281-295, 1990.

[10] J.Shen, 北京话上声连读的调型组合和节奏形式["The Combination of Tone Patterns and Rhythmic Forms of T3 sandhi in Beijing Mandarin"], 《中国语文》[*Journal of Chinese Linguistics*], vol.4, pp. 274-281, 1994.

[11] Y. Xu, Contextual tonal variations in Mandarin. *Journal of Phonetics*, vol.25, pp. 61-83, 1997.

[12] Y. Xu, Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, vol.27, pp.55–105, 1999.

[13] M.Y. Chen, *Tone sandhi: patterns across Chinese dialects*, Cambridge: Cambridge University Press, 2000.

[14] Y. Xu, Understanding tone from the perspective of production and perception, *Language and Linguistics*, vol.5, pp. 757-797, 2004.

[15] Y.R. Chao, 北平语调的研究["Study of Peiping Intonation"], 《赵元任语言学论文集》[*Yuenren Chao 's Essays on Linguistics*], Beijing: The Commercial Press, pp. 253-272, 2002.

[16] Z.M. Zhang, 东北官话的分区["Distribution of Northeastern Mandarin"],《方言》[*Fangyan(Dialect)*], vol.2, pp. 141-148, 2005.

[17] J.J. Kuang and H.J. Wang, 连上变调在不同韵律层级上的声学表["T3 sandhi at the boundaries of different prosodic hierarchies"],《中国语音学报》[Journal of Chinese Phonetics], vol.1, pp. 125-131, 2006.

[18] J. Kou, 抚顺地区人群语音的识别方法及研究["People's speech recognition method and research in Fushun"], 辽宁警专学报[*Journal of Liaoning Police College*], vol.2, pp. 62-64, 2009.

[19] J. Jia et al, 汉语普通话与沈阳方言转换的研究["Conversion from Chinese Mandarin to the Shenyang dialect"], 《清华大学学报》[*Journal of Tsinghua University*], vol.S1: 1309-1315, 2009.

[20] J. Zhang and Y. Lai, Testing the role of phonetic knowledge in Mandarin tone sandhi, *Phonology*, vol.27, pp. 153–201, 2010.

[21] T. Lin and L.J. Wang, A course in Phonetics(revised version), Beijing: Peking University Press.

[22] C. Zhang and G. Peng, "Productivity of Mandarin Third Tone Sandhi: A Wug Test," In F. Shi F. & Peng G. (eds.), *East Flows the Great River: Festschrift in Honor of Prof. William S-Y. Wang's 80th Birthday*, Hong Kong: City University of Hong Kong Press, pp.255-282, 2013.

[23] J.H. Yuan and Y.Y. Chen, 3rd tone sandhi in Standard Chinese: A corpus approach, *Journal of Chinese Linguistics*, vol.42, pp. 218-237, 2014.

[24] C.H. Zhao and H.B. Chen, 鞍山方言的语音特征["Speech characteristic of Anshan dialect"], *Journal of University of Science and Technology Liaoning*[辽宁科技大学学报], vol.2, pp. 184-187, 2014.

[25] Z.Y. Xiong, xRecorder.exe in TCL/TK(V03.04.03), Chinese Academy of Social Science.

[26] Y.J. Wang, M.X. Yu, and Q. Wu, 重音和韵律对北京话连上变调的作用["The Influences of stress and rhythmic relation on the implementation of third-tone-sandhi rule in Beijing Mandarin."], 《中国语音学报》[*Journal of Chinese Phonetics*], vol.10, pp. 72-82, 2018.

[27] H. Takahiro, 普通话语句中相邻声调间的相互作用——焦点与韵律的制约[Interaction between adjacent tones in the sentences of Mandarin Chinese—constraints of focus and rhythm], PhD dissertation, Peking University, 2019.

[28] P. Cui and Y.J. Wang, 抚顺话"上声+阴平"变调的声学和知觉研["Acoustic and perceptual research on tone sandhi of "T3+T1" in Fushun Dialect "],《中国语音学报》[*Journal of Chinese Phonetics*], vol.11, pp. 88-99, 2019.