INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

Pitch Declination and Final Lowering in Northeastern Mandarin

Ping Cui[1], *Jianjing Kuang*[2]

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

Abstract

Northeastern Mandarin has a similar lexical tone system as Beijing Mandarin. However, the two dialects significantly diverge at higher prosodic structures. T1 in Northeastern Mandarin always changes to a falling tone in domain-final positions. Previous studies have analyzed this variation as a type of tone sandhi, but we propose it is related to more global prosodic processes such as final lowering. We addressed this issue by conducting both production and perception experiments with native bidialectal speakers of Northeastern Mandarin and Beijing Mandarin. Our findings suggest that T1 variation is essentially a domain-final lowering effect. Other tones also show some kind of final lowering effects. Compared to Beijing Mandarin, Northeastern Mandarin generally has greater global pitch declination and greater final lowering effects. Our perception experiment further showed that both prosodic effects play important roles in identifying the Northeastern Mandarin accent, and final lowering cues are more perceptually salient than the global declination cues. These findings support the notion that pitch declination and final lowering effects are linguistically controlled, not just a by-product of the physiological mechanisms.

Index Terms: pitch declination, final lowering, Northeastern Mandarin, Beijing Mandarin

1. Introduction

Northeastern Mandarin (abbreviated to NEM) is a major Mandarin variety spoken in the Northeastern part of China. Its distribution includes Heilongjiang Province, Jilin Province, most of Liaoning Province, and the eastern part of Inner Mongolia Autonomous Region ([1], [2] and [3]).

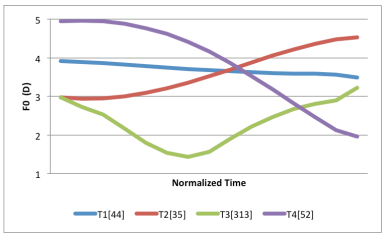

Figure 1: *F_0 results of the four tones in monosyllables on a normalized 1–5 numerical scale in Northeastern Mandarin* [4].

As shown in Figure 1, the citation tone patterns in Northeastern Mandarin are very similar to those in Beijing Mandarin (abbreviated to STM). Northeastern Mandarin has four lexical tones, namely, level (T1), rising (T2), low-dipping (T3) and high-falling (T4). The phonetic realizations of the four tones are also very similar to those in Beijing Mandarin, except that T1 is slightly lower than the T1 in Beijing Mandarin. However, despite the similarity of the citation tones between the two Mandarin dialects, the tone sandhi patterns in Northeastern Mandarin are significantly different from Beijing Mandarin. In Beijing Mandarin, there is only one sandhi rule associated with T3, but in Northeastern Mandarin, T3 becomes a rising tone not only when it precedes T3, but also when it is before T1 ([5], [6] and [7]). Moreover, as the interest of this paper, while T1 in Beijing Mandarin has no phonological variation, T1 in Northeastern Mandarin always changes to a mid-falling tone when it occurs at the second syllable of a disyllabic word ([7] and [8]), regardless of the preceding tones. To our best knowledge, the nature of "T1 sandhi" has not been systematically investigated. Based on our observation, unlike T3 sandhi, which is triggered by following tones, the so-called "T1 sandhi" is clearly tied to domain-final positions. Then the question is, what type of domain final would trigger this tone variation?

Our preliminary observation suggests that T1 especially clearly changes to a falling tone at the end of a large phrase such as intonational phrases. As demonstrated in Fig. 2, in a sentence with all T1 syllables, in addition to an overall falling trend of T1 due to pitch declination, there is a remarkably steep falling at the end of the sentence. We thus hypothesize that the T1 variation is related to the final-lowering effect of large prosodic boundaries.

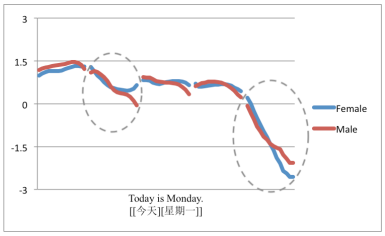

Figure 2: *F_0 trajectories of an all T1 sentence (jin1 tian1 xing1 qi1 yi1 "today is Monday") from two speakers. Two falling T1 syllables were circled.*

Pitch declination, known as the global downward trend throughout an utterance, has been widely reported among languages (e.g., [9], [10], [11] and [12]). Pitch declination is driven by physiological factors as well as language-specific linguistic structures. Final lowering, defined as additional abrupt lowering confined to phrase and utterance ends, is found to be a mechanism independent from global pitch declination [18], and tends to be more language specific (e.g., [11], [13]).

Pitch declination in Beijing Mandarin has been extensively discussed in a number of studies (e.g., [14], [15], [16], [17],

http://dx.doi.org/10.21437/Interspeech.2020-1987

[18], [19] and [20]). All T1 sentences have been a good test ground to exhibiting these effects. The general consensus is that Beijing Mandarin has a clear pitch declination effect, which is significantly correlated with the utterance length. Moreover, Beijing Mandarin also exhibits a language-specific declination slope [20]: Compared to American English, Beijing Mandarin has an overall steeper declination slope. The discussion on the final-lowering effect in Beijing Mandarin is however less conclusive. Beijing Mandarin at best has a small final-lowering effect that is independent from pitch declination or tone influence [19, 20, 21, 22], if not totally lacks such an effect [18].

In light of the studies of Beijing Mandarin, it is particularly intriguing to investigate the nature of T1 falling, and how much this tone variation in Northeastern Mandarin is related to more global prosodic organization. Given that Northeastern Mandarin is an extremely closely related language variety to Beijing Mandarin, studying this case would provide important insights for the understanding of the universal vs. language-specific nature in pitch declination and final lowering. Specifically, in this study we intend to investigate (1) whether the slope of falling T1 is sensitive to the strength of the prosodic boundaries; (2) whether all the tones in Northeastern Mandarin have the same pitch declination and final lowering effects as T1? To address these questions, we conducted a production experiment with bidialectal speakers to compare the pitch declination and final lowering effects in both Beijing Mandarin and Northeastern Mandarin.

Moreover, although no experimental studies have been done for this Mandarin variety, it has been widely acknowledged among Chinese people that Northeastern Mandarin has a special intonational accent. We speculate that the different declination and final lowering effects might contribute to the recognition of Northeastern Mandarin. Therefore, in a follow-up perception experiment, we further test whether these prosodic features are perceptually salient to the native speakers.

2. Study 1: Production experiment

2.1. Methods

2.1.1. Stimuli

16 like-tone sentences (4 tones * 4 breaks) were designed for the production experiment. That is, all the syllables in the sentences have the same tone; and for each tone, there are 4 sentences varying in length from one sentence-medial break to four sentence-medial prosodic breaks. These prosodic breaks roughly are the boundaries of prosodic words or small phrases. Examples of T1 sentences are shown in Table 1.

2.1.2. Participants and procedure

Twelve participants (7 females and 5 males) who were native bidialectal speakers of Northeastern Mandarin and Beijing Mandarin took part in the study. Participants were asked to produce all sentences in Northeastern Mandarin first, and then in Beijing Mandarin. A total of 16 sentences * 2 dialects * 2 repetitions * 12 speakers =768 utterances are recorded. All sentences were arranged in a random order. The recordings were collected online through Audacity and WeChat because of coronavirus pandemic. Since this study is based on within-speaker comparison of F_0, the differences of individual microphones are unlikely to influence the results.

Table 1: *Examples of T1 like-tone sentences.*

Type	Stimuli
one-break	今天|星期一。(Today is Monday.)
two-breaks	今天|修|收音机。(Today I repair the radio.)
three-breaks	高多斌|今天|修|收音机。(Today Gao Duobin repaired the radio.)
four-breaks	山东|高多斌|今天|修|收音机。(Today Gao Duobin from Shandong repaired the radio.)

As shown in Table 1, they are all T1 like-tone sentences with different lengths. Same design goes to T2, T3, and T4 as well.

2.1.3. Measurements

Praat was used for the extraction of fundamental frequency. We first used the forced aligned developed by Penn Phonetics Lab to do the segmentation, and then manually annotated the tonal nuclei by defining the stable points of the rhyme (Fig. 3).

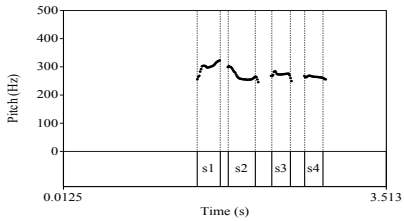

Figure 3: *segmentation of tonal sequence s (x-axis: Hz; y-axis : s)*

The tonal nuclei were analyzed using Takahiro's script [23]. 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.

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

In addition, to quantitatively measure the slope of pitch declination, linear regression lines were fitted to F_0 contours of the entire sentences as well as individual syllables, using the least-squares method.

2.2. Results

Here we report our results and analyses of 16 like-tone sentences for 12 speakers. Fig. 4 shows the pitch trajectories of T1 sentences. It can be seen that the overall declination slope in Northeastern Mandarin is significantly steeper than that in Beijing Mandarin, as confirmed by a paired t-test averaged across all sentences (t=-3.977, p=0.028). This is partially driven by the final lowering for T1 in Northeastern Mandarin. There is a dramatic falling contour at the end of each break, especially at the end of the sentences. It seems that the slope of the falling is correlated with the strength of the

prosodic boundaries. Bigger boundaries have greater falling slopes.

Fig. 5 shows that there is no significant difference in T2 between Northeastern Mandarin and Beijing Mandarin in terms of the global pitch declination, as confirmed by a paired t-test (t=-1.174, p=0.325). However, importantly, T2 in Northeastern Mandarin also has a final lowering effect, which usually exhibits as an additional falling tail after the rising trend of the final syllables of the sentences.

Fig. 6 illustrates the F_0 trajectories of the T3 sentences. Similar to T2 sentences, T3 sentences in Northeastern Mandarin and Beijing Mandarin both have a similar slope of pitch declination. There is generally no significant difference between the two dialects of the global pitch declination averaged across all sentences (t=1.583, p=0.212), although the slope of T3 in Beijing Mandarin appears to be slightly lower than that in Northeastern Mandarin, especially the sentence with three breaks (Fig. 6c). Besides, T3 also undergoes tone sandhi in non-final position, changing to a rising tone.

According to Fig. 7, T4 sentences are similar to T2 and T3. Both of Northeastern Mandarin and Beijing Mandarin have a clear pitch declination. But the global declination slope in Beijing Mandarin is marginally steeper than that in Northeastern Mandarin (t=2.127, p=0.123).

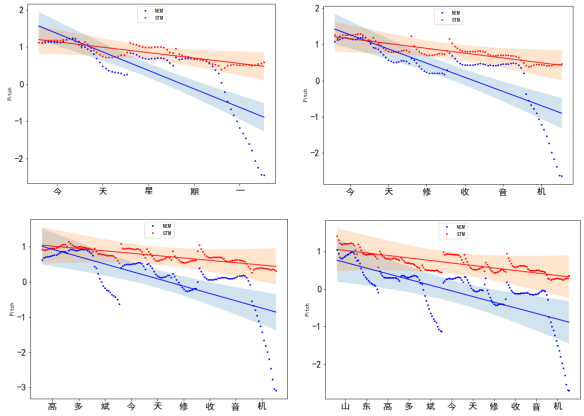

Figure 4: *F_0 patterns and slopes of T1 like-tone sentences both in Northeastern Mandarin and Beijing Mandarin. Four pitch plots respectively represent four sentences with different prosodic breaks (one break: top left, two breaks: top right, three breaks: bottom left, and four breaks: bottom right). The blue pitch track represents the acoustic realization in Northeastern Mandarin, while the red one represents Beijing Mandarin. Same as below.*

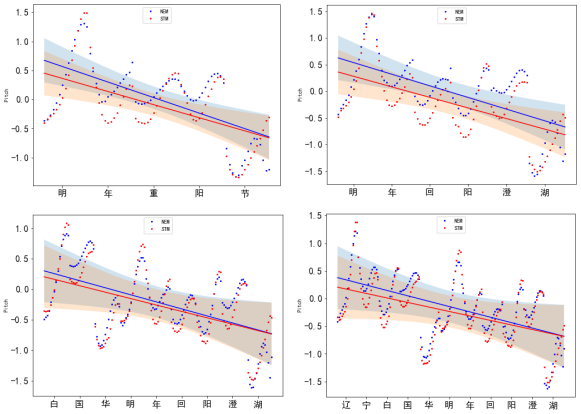

Figure 5: *F_0 patterns and slopes of T2 like-tone sentences both in Northeastern Mandarin and Beijing Mandarin.*

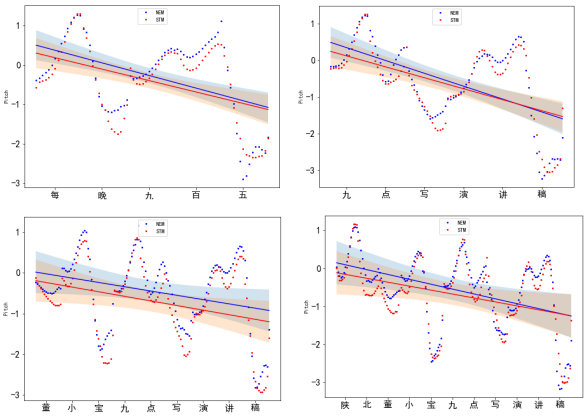

Figure 6: *F_0 patterns and slopes of T3 like-tone sentences both in Northeastern Mandarin and Beijing Mandarin.*

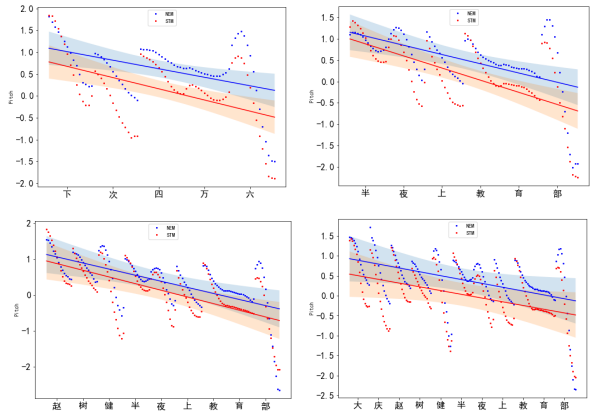

Figure 7: *F_0 patterns and slopes of T4 like-tone sentences both in Northeastern Mandarin and Beijing Mandarin.*

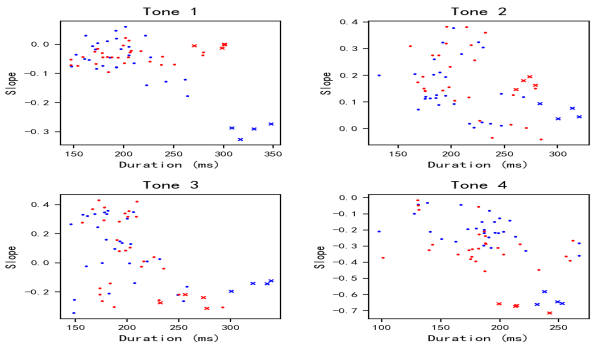

Figure 8: *Correlation analysis between the slope and duration of individual syllables for four tones both in Northeastern Mandarin and Beijing Mandarin. The blue pitch track represents Northeastern Mandarin, while the red one represents Beijing Mandarin. The cross-shaped symbol indicates the last syllable of each sentence. Four plots respectively represent correlation of four tones (T1: top left, T2: top right, T3: bottom left, and T4: bottom right).*

Correlation analysis was further carried out between the slope and duration for individual syllables. As shown in Fig 8a, for T1 syllables, only Northeastern Mandarin exhibits a significant negative correlation between syllable duration and declination slope (r=-0.83, p<0.001). That is, the longer the syllable, the steeper the falling slope. This result further confirmed the observation from Fig. 4. Other tones show similar correlations between syllable slope and duration for both dialects. Moreover, as expected, regardless of the tones,

syllables at the end of the sentences (indicated by cross-shaped symbols in Fig. 8) generally have a substantially longer duration and a steeper slope. However, the duration of sentence-final syllables is significantly longer in Northeastern Mandarin (p=0.02). The two dialects also differ in the effect size of final lowering. As shown in Figure 8a, T1 in particular has a much steeper falling at the end of the sentences, on top of generally steeper global pitch declination.

3. Study 2: Perception experiment

Study 1 suggests that T1 in Northeastern Mandarin has significantly greater final lowering and pitch declination compared with Beijing Mandarin. Do these features of production are also perceptual salient in Northeastern Mandarin? Are native speakers sensitive to these two effects in Northeastern Mandarin? To verify these issues, we further conduct a perception experiment.

3.1. Methods

3.1.1. Stimuli

The perception stimuli came from the materials of production (see Table 1). A female speaker recorded the stimuli using Beijing Mandarin in a sound-treated booth in the Phonetics Lab at the University of Pennsylvania. The recordings were made with a SHURE WH320 Condenser microphone at a sampling rate of 32 bit /44.1 kHz.

Firstly, to test the final lowering effect, the slope of sentence-final T1 syllables was manipulated to create a level-to-falling continuum. 7 steps of stimuli were made for each sentence, starting from the original value and then lowering the endpoint by 1.5st every step through Pitch Synchronous Overlap and Add(PSOLA) in Praat. A total of 28 sentences (4 breaks * 7 steps of the last syllable) were created. Secondly, to test the global declination effect, the overall slope of the entire sentence was manipulated to create a less declination to more declination continuum. 9 steps of stimuli were made for each sentence, starting from the original value and lowering by 2st every time. A total of 18 sentences (2 breaks * 9 steps of the overall sentence) were created. All the stimuli (46 sentences) were arranged in random order.

3.1.2. Participants and procedure

Thirty-one participants (16 females and 15 males) who were native speakers of Northeastern Mandarin and can also speak Beijing Mandarin took part in the study, including 12 speakers in the production study.

Listeners were instructed to identify whether the sentences they heard were in Northeastern Mandarin or Beijing Mandarin through Qualtrics. They were allowed to hear the sounds as many times as they wanted, and were instructed to respond by clicking on the button with the corresponding dialect on the screen. A practice session was conducted prior to the actual experiment.

3.2. Results

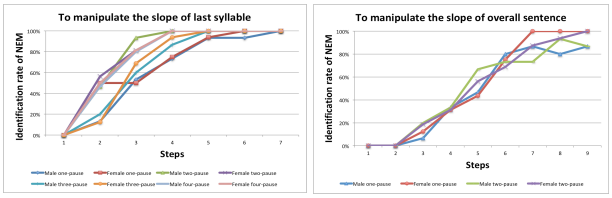

Figure 9: *Identification rate of Northeastern Mandarin. (Left: To manipulate the slope of the last syllable, Right: To manipulate the slope of the overall sentence.)*

As indicated by Fig. 9, overall native listeners of Northeastern Mandarin are very sensitive to both the slope of the sentence-final syllables (Fig. 9 left) as well as the overall slope of the sentences (Fig. 9 right), but final lowering cues appear to be more salient. When manipulating final lowering cues, most subjects accept step3 as Northeastern Mandarin (Fig.9 left), however, the crossing point is not until step 5 for the manipulation of global declination (Fig. 9 right), even though the step-interval is bigger for the global slope condition.

In addition, we noticed that gender and sentence-length may also have an effect on the identification rate. The overall identification rate of the female is relatively higher than that of the male. Also, the sentence with two-breaks has the highest identification rate, while the identification rate of the one-break sentence is lowest. Due to limited space, we do not discuss these two effects in detail here.

4. Discussions and Conclusions

By conducting both production and perception experiments with native bidialectal speakers, this study provides the first systematic investigation of the nature of T1 variation in Northeastern Mandarin, and how much this variation is associated with global prosodic features. In the production experiment, we found that T1 like-tone sentences in Northeastern Mandarin have significantly greater global pitch declination as well as additional dramatic falling at the end of the large prosodic boundaries. Moreover, the falling slope of final T1 is correlated with the syllable duration and the position of the syllable: Larger boundaries (especially sentence final) introduce a longer duration and much steeper falling. This finding suggests that T1 variation is essentially a domain-final lowering effect, and it is more appropriate to treat the T1 variation as a boundary tone rather than a tone sandhi at the word level. This conclusion was further supported by the analyses of other tonal categories. We found a clear falling tail after the rising contour for final T2 syllables. The final lowering effects for T3 and T4 are mostly cued by longer duration, as T3 and T4 already have a low pitch target. Taken together, unlike the small final lowering effect found in Beijing Mandarin, the final lowering effect in Northeastern Mandarin is much more salient.

Furthermore, consistent with the production experiment, although both global pitch declination and final lowering effects are perceptually salient in the identification of Northeastern Mandarin, native Northeastern speakers are more sensitive to final lowering cues than the global declination cues.

5. References

[1] X. Song et al, “The phonemic structure of Liaoning dialect/辽宁语音说略”, *Journal of Chinese Linguistics*, vol.2, pp.104-114, 1963.

[2] W. He, “Distribution of Northeastern Mandarin/东北官话的分区”, *Fangyan(Dialect)*, vol.3, pp. 172-181, 1986.

[3] S.A. Wurm, R. Li and T. Baumann, *Language atlas of China*. Australian Acad. of the Humanities; Longman Group (Far East), 1987.

[4] P. Cui and Y.J. Wang, “Acoustic and perceptual research on tone sandhi of T3T1 in fushun dialect/抚顺话上声+阴平变调的声学和知觉研究”, *Journal of Chinese Phonetics*, vol.11, pp. 88-99, 2019.

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

[6] C.H. Zhao and H.B. Chen, “Speech characteristic of Anshan dialect/鞍山方言的语音特征”, *Journal of University of Science and Technology Liaoning*, vol.2, pp. 184-187, 2014.

[7] P. Cui, J.J. Kuang, and Y.J. Wang, “The effect of focus and prosodic boundary on the two T3 sandhi in Northeastern Mandarin” in *Speech Prosody 2020 – 10th Annual Conference of the International Speech Communication Association, May 25-28, Tokyo, Japan, Proceedings*, 2020.

[8] L. Ye (2013). “The speech characteristics showed by disyllabic tone sandhi in Harbin dialect/哈尔滨方言两字组连读变调中的方音特色”, Master's thesis, Naikai University, 2013.

[9] D.R. Ladd, “Declination: a review and some hypotheses”, *Phonology*, vol.1, pp. 53–74, 1984.

[10] M. Liberman and J. Pierrehumbert, “Intonational invariance under changes in pitch range and length” in Aronoff, M., Oerhle, R. (Eds.), *Language Sound Structure*. MIT Press, pp. 157–233, 1984.

[11] B. Connell and D.R. Ladd, “Aspects of pitch realization in Yoruba”, *Phonology*, vol.7, pp. 1-29, 1990.

[12] B. Connell, “Downdrift, downstep, and declination,” in Proceedings of Typology of African Prosodic Systems Workshop, Bielefeld University, Germany, 2001.

[13] D.R. Ladd, “Commentary: tone languages and laryngeal precision”, *Journal of Language Evolution*, vol.1, no.1, pp. 70–72, 2016.

[14] J. Shen, “Pitch range and intonation of the tones of Beijing Mandarin/北京话声调的音域和语调”, In *Acoustic Studies of Beijing Mandarin*. Beijing University Press, pp. 73–130, 1985.

[15] Y. Xu and Q.E. Wang, “What can tone studies tell us about intonation?” in *Intonation: Theory, models and applications*, 1997.

[16] Y. Xu, “Effects of tone and focus on the formation and alignment of F0 contours”, *Journal of Phonetics*, vol.27, no.1, pp. 55–105, 1999.

[17] P. Wang, L. Shi, and F. Shi, “Declination and downstep effect in declarative sentences of Chinese Mandarin/普通话陈述句中的音高下倾和降阶”, *Journal of Chinese Phonetics*, vol.3, pp.54-60, 2012.

[18] C. Shih, “A declination model of Mandarin Chinese”, in *Intonation*, Springer, pp. 243-268, 2000.

[19] J.H. Yuan, C. Shih, and G. Kochanski, “Comparison of declarative and interrogative intonation in Chinese”, in *Speech Prosody 2002, International Conference, April 11-13, Aix-en-Provence, France, Proceedings*, 711-714, 2002.

[20] J.H. Yuan and M. Liberman, “F0 declination in English and Mandarin Broadcast News Speech”, *Speech Communication*, vol.65, pp. 67–74, 2014.

[21] W. Lai, Y. Li, H. Che, S. Liu, J. Tao, X. Xu et al., “Final lowering effect in questions and statements of Chinese Mandarin based on a large-scale natural dialogue corpus analysis”, in *Speech Prosody 2014 – 7th Annual Conference of the International Speech Communication Association, May 20-23, Dublin, Ireland, Proceedings*, pp. 653–657, 2014.

[22] W. Lai, X. Xu, Y. Li, H. Che, S. Liu, and J. Tao, “Phonological influences on the realization of final lowering evidence from dialogue chinese mandarin,” in *2014 17th Oriental Chapter of the International Committee for the Co-ordination and Standardization of Speech Databases and Assessment Techniques (COCOSDA)*. IEEE, pp. 1-6, 2014.

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