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

Tonal variations in Honggu Chinese

Chunyu Ge[1], *Aijun Li*[2]

[1]Graduate School of Chinese Academy of Social Sciences, China
[2]Chinese Academy of Social Sciences, China
gechunyu92@hotmail.com, liaj@cass.org.cn

Abstract

This paper investigates the tonal variations in Honggu Chinese. The tones in isolation are investigated, as well as in different focus conditions and sentence positions. There are two tones in Honggu Chinese. T1 is a rising tone, and T2 a high tone. It is optional to pronounce T2 as high level, and it is usually convex when embedded in sentences. Growth Curve Analysis is employed to analyse the effects of focus and sentence position on the tones. Focus can influence both the pitch height and pitch contour of the tones. The pitch height of the tones can also be affected by sentence position. One of the interesting findings is that the post-focus condition combined with sentence-final position results in the incomplete neutralization of two tones. Honggu Chinese provides a good example of how tone can vary due to different factors.

Index Terms: Honggu Chinese, tonal variation, focus, sentence position, Growth Curve Analysis

1. Introduction

Tone languages are languages where there is an indication of pitch in the lexical realization of at least some morphemes [1]. This indication of pitch can be influenced by various factors, and its phonetic realization is by no means invariant. It can be influenced by the tonal context in which the tone is located, which is usually referred to as tone sandhi [2]. There is another contextual effect called tonal conarticulation, which is the tonal variations stemming from articulatory constraints [3]. Intonation, which also employs pitch as the primary acoustic parameter, can also influence its phonetic realization [4, 5].

Chinese is well-known as a tone language, and there is a large number of dialects in Chinese, whose tone systems vary greatly. Different sources of tonal variations have been identified in Chinese. The most extensively found tonal variation in Chinese dialects is tone sandhi [2]. Far less is known as to how the phonetic realization of tone is influenced by intonation. To our knowledge, there are only two well-established phenomena in Chinese which exemplifies the influence of intonation on the realization of tone. A canonical study on focus intonation in Mandarin Chinese reveals that the assignment of focus influences the pitch range within which the lexical tone is realized [6]. This influence occurs with reference to the position of the focus, in that pitch range is expanded for syllables on focus, and compressed for syllables after focus. Another interesting phenomenon in Cantonese exemplifies the influence of intonation on the pitch contour of lexical tone. The high boundary tone in yes-no questions in Cantonese can cause the pitch contour of the final syllable to become rising, regardless of its lexical tone. Both perception studies [7, 8] and acoustic studies [9, 10] suggest that some of the tonal contrasts are neutralized.

This paper will consider the tonal variations in Honggu Chinese, a Chinese dialect in north-western China [11]. Chinese dialects in north-western China share the property of less tonal contrast than dialects in other areas. Some attribute this to the close contact with Altaic languages, which do not employ pitch to form lexical contrast [12]. Honggu Chinese stands out in Chinese dialects in that it contrasts only two tones while most other dialects have tone systems of three or more tones. The two tones are traditionally referred to as *Pingqu sheng* and *Shangsheng*, named after the tone categories in Middle Chinese with which they used to belong. Table 1 presents the names and labels used in the paper, along with the transcriptions of these two tones using Chao's system [13]. It is generally accepted that T1 is a rising tone, and T2 a high tone. Whether T2 is a level tone is contentious, due to the fact that apart from high level, almost all authors have reported that there is another high falling variant, which can be transcribed as /53/. One of the authors found that different speakers can pronounce the same tone differently, and even for the same speaker, he can pronounce the tone with high level at one occasion, and high falling at another. This is simply to say that high level and high falling are free variants of the same tonal category.

Table 1: *Tone system of Honggu Chinese and the symbols used in the paper*

Tone	Tonal value	Label
Pingqu Sheng	35	1
Shang Sheng	55	2

Given the relatively simple tone system in Honggu Chinese, tone sandhi in this dialect is somehow complex [14, 15, 16]. Apart from tonal contexts, the grammar structure of the compound can lead to different tone sandhi patterns. The tone sandhi rules for modifier-noun compounds and for verb-object compounds are different. There is another tone sandhi rule for duplicate structures. Previous studies of Honngu dialect are mostly impressionistic descriptions, and there is only one acoustic study [17]. This study confirmed that there are only two tones in Honggu Chinese by clustering and discrimination of the pitch contours of different tone categories in Middle Chinese.

To investigate the tonal variations in Honggu Chinese, the present study manipulates the focus condition and the sentence position of the test words. Since Honggu Chinese is an under-described dialect, this paper will also present its tone system in isolation.

2. Methodology

To investigate tonal variations in Honggu Chinese, we made recordings of native Honggu speakers. The recording materials consist of two parts, monosyllabic words in both isolation and focus conditions. For monosyllabic words in isolation, five

10.21437/SpeechProsody.2020-50

words are selected for each tone category in Middle Chinese. The syllable structures of the monosyllabic words are all CV, where C is preferably simple stops as /p, t, k/, and V preferably cardinal vowels as /i, u, a/.

To investigate the influence of focus on tone, a subset of the test words are selected and embedded in focus sentences. For each tone in Honggu Chinese, seven test words are selected for each tone. The focus conditions and the positions of the test words in the sentence are manipulated. Each test word is placed at sentence-initial, sentence-medial, and sentence-final positions, and its focus condition varies between narrow focus and non-focus. Therefore, the total number of the sentences is $2(tones) \times 7(words) \times 3(sentencepositions) \times 2(focusconditions) \times 6(speakers) = 504$. Contrastive focus is used here because it is most prominent among different kinds of focus. In an under-described dialect as Honggu Chinese, it is better to look first into the most likely case of focus prosody. The test sentences are all simple SVO sentences, with a leading sentence as background. For focus conditions at sentence-medial positions, a modifier is added before the test word. For sentence-final focus sentences, the test words are objects. Due to the space limit, we include only a sentence-initial sentence as illustration, as in (1). The test word is in italic. In (1a), the test word *monkey* is in the focus condition, contrasting with the subject in the leading sentence, dog. The test word *monkey* in (1b) is in non-focus condition as it appears in the leading sentence.

(1) a. kəu pu ɕihuan tʃʰi tʰautsi. *həu* ɕihuan tʃʰi tʰautsi.
dog not love eat peach. *monkey* love eat peach

Dogs do not love eating peaches. *Monkeys* love eating peaches.

b. həu ɕihuan tʃʰi ɕiaŋtɕiɔ. *həu* hai ɕihuan tʃʰi tʰautsi.
monkey love eat banana. *monkey* also love eat peach

Monkeys love eating bananas. *Monkeys* also love eating peaches.

The recordings were conducted in the summer of 2019, on a field trip to Honggu, Lanzhou. Unfortunately, it is when the crops and vegetables were harvested, and only six speakers were recruited. Among the six speakers, three are male and three are female. Most of them are in their 50s ($mean = 50.5, sd = 9.2$). In advance of the recording, we collaborated with a native speaker and adjusted the test words and sentences to be grammatical in the dialect. If the test words and test sentences are unnatural to native speakers or hard to produce, some modifications were made and confirmed again with the native speaker. The recordings took place in a quiet room in the local primary school with curtains drawn. The sound was recorded into a laptop using a `SHURE SM86` microphone and a `Lexicon I-O 22` sound card. The sampling rate is 44100 Hz and the sound was quantized in 16 bit. The recordings were segmented and pitch were extracted in the voiced parts of the syllables only. Ten pitch points were extracted equidistantly for each voiced component. Data analyses were carried out in R [18]. Growth Curve Analysis (GCA) was employed to analyse the effects of focus and sentence position on the tones [19]. GCA is a kind of linear mixed effects model that can capture different aspects of contours, by modeling contours as high-order polynomials.

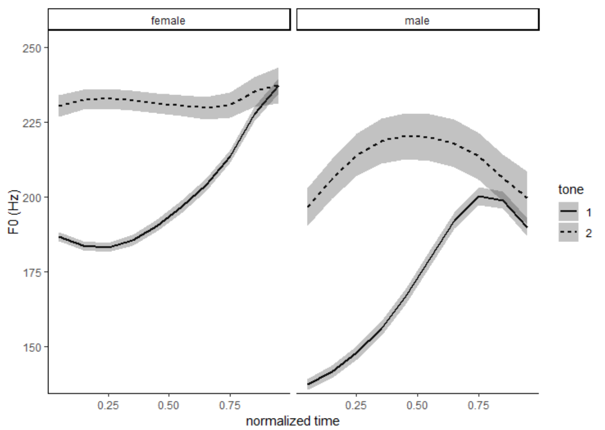

Figure 1: *Averaged pitch contours of two tones in Honggu Chinese (left: female; right: male).*

It is used previously on the tone sandhi and tonal coarticulation in Tianjin Mandarin [20]. GCA was conducted using the `lmerTest` package [21].

3. Results

In this section, we first present the tones of Honggu Chinese in isolation. After that, we concentrate on the focus prosody in Honggu, to see what influence focus has on these tones. Monosyllabic words at different sentence positions are investigated, under both focus and non-focus conditions.

3.1. Tones in isolation

Impression of the first author during his fieldwork confirms the previous report that there are only two tones in Honggu Chinese. Further inspection of the recordings of monosyllabic words in isolation corroborates the impression. For the eight tone categories of Middle Chinese, all but the third tone (Yin Shang) are rising, while the third tone is high. There is only a slight variability for the eldest speaker (M3, the third male speaker) in this experiment. [1] All the first tone and some of the second tone in Middle Chinese are pronounced as low without rising. The patterns of all other speakers are consistent.

Figure 1 shows the averaged pitch contours of the two tones across speakers in Honggu Chinese, with the anomalies in M3 removed. The ribbons surrounding the lines indicate the standard errors. The left panel shows the pitch contours of female speakers, and the right panel male speakers. The pitch contour of T1 is rising with a small dip towards the end for male. The excursion of T1 is large, about 60 Hz for both genders. The end point of T1 nearly approximates the level of T2. For T2, it is high level for female and high convex for male. It is reminiscent of the report of the dialectologists that T2 is sometimes high falling. Given the convex shape of T2 in male speakers, we suspect that the target of T2 is high, whether implement it as high level is optional. It is possible that different strategies are employed in implementing the high target. The broad ribbon of T2 for male indicates that there is more variations of T2 for male speakers.

[1] Only the data of the anomalies in citational tones of M3 are removed, and the data of M3 in sentences are consistent with other speakers and are included in the analyses.

3.2. Focus at sentence-initial positions

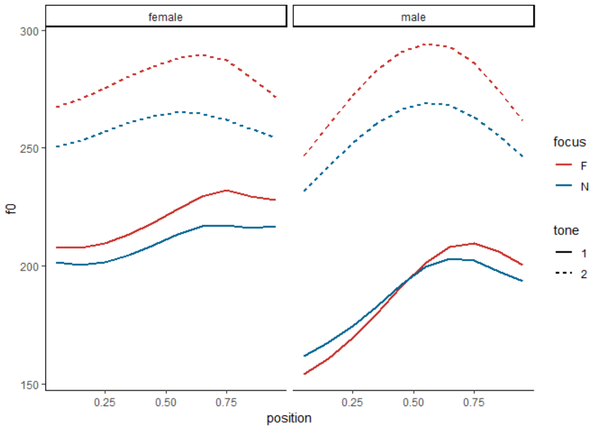

Figure 2: *Averaged pitch contours of tones at sentence-initial positions under different focus conditions (left: female; right: male).*

Figure 2 presents the averaged pitch contours of the two tones under different focus conditions at sentence-initial positions. Red contours are the pitch contours under focus, and blue contours under non-focus conditions. The solid line stands for T1, while the dashed line stands for T2. Unlike that in isolation, T2 is not high level, but high convex for both genders. It is higher under focus than non-focus. T1 is still rising, and it is higher under focus for female, but the contours of two focus condtions cross for male. To analyze the pitch contours of these two tones, two separate models are constructed using GCA. A linear model is constructed for T1, and a quadratic model for T2. Speaker and item are included in the model as random intercepts. The effects of focus condition on different terms are included in the model in a stepwise manner. The models for each tone are compared in terms of log-likelihoood. The same procedure is applied to the following analyses for focus at sentence-medial and sentence-final positions. For T1, adding the effect of focus condition on the intercept improves model fit significantly ($\chi^2(1) = 9.26, p = 0.002$), and adding the effect on the slope improves the model only slightly significantly ($\chi^2(1) = 5.03, p = 0.025$). T1 is estimated to be 5.38 Hz higher under focus than non-focus. For T2, the effects of focus condition on both the linear and quadratic terms do not improve model fit, only the effect on the intercept does ($\chi^2(1) = 146.59, p < 10^{-15}$). That is to say, the main influence focus condition has is on the pitch heights of T1 and T2, while that on pitch contours is weak, at sentence-initial positions.

3.3. Focus at sentence-medial positions

The averaged pitch contours of focus and non-focus conditions at sentence-medial positions are shown in Figure 3. As in Figure 2, the colours indicate focus condition and the line types tone category. T2 is again in convex shape, and higher under focus than non-focus. The case of T1 is slightly different. For both genders, T1 under focus starts lower under focus than non-focus but rises more rapidly under focus, and its ending point under non-focus condition is lower than under focus. Two models are also fitted to the two tones as in sentence-initial positions. The effects of focus condition on both the intercept and the slope improve model fit significantly ($\chi^2(1) = 49.61, p < 10^{-11}$ for

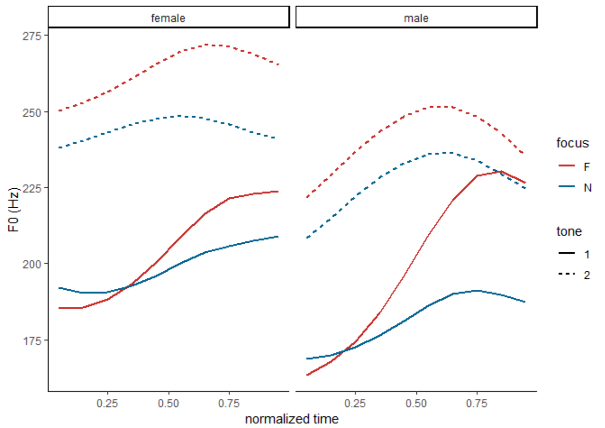

Figure 3: *Averaged pitch contours of tones at sentence-medial positions under different focus conditions (left: female; right: male).*

the intercept, $\chi^2(1) = 54.13, p < 10^{-12}$ for the slope) for T1. The intercept of T1 under non-focus condition is estimated to be 9.36 Hz higher than under focus, but its slope is far flatter. For T2, it is also the case that neither the effects of focus condition on the linear term, nor on the quadratic term improves model fit, and only the effect on the intercept does ($\chi^2(1) = 34.27, p < 10^{-8}$). At sentence-medial positions, both pitch height and pitch contour of T1 are influenced by focus condition, and only the pitch height of T2 is influenced.

3.4. Focus at sentence-final positions

The patterns of focus at sentence-final positions are dramatically different than those at sentence-initial and sentence-medial positions. The averaged pitch contours are presented in Figure 4. The colours and the line types indicate focus condition and tone category, respectively. The shapes of both tones are generally preserved under focus, but are both falling under non-focus conditions. T2 is high convex for male and high level for female, as in isolation, and T1 is rising for both genders under focus. Both tones become falling under non-focus conditions. It seems that focus condition influences both pitch height and pitch contour at sentence-final positions. T1 and T2 are neutralized, and their contrast in pitch height is preserved to some extent. To test if the two tones are completely neutralized, a linear GCA model is fitted to the data of non-focus conditions. Speaker and item are treated as random effects, and the effects of tone on the intercept and the slope are added into the model in turn. The effect of tone on the intercept improves model fit ($\chi^2(1) = 42.58, p < 10^{-10}$), while the effect on the slope does not ($\chi^2(1) = 0.54, p = 0.46$). T2 is estimated to be 8.45 Hz higher than T1 under non-focus conditions. This is small compared to their difference in isolation and in other sentence positions.

3.5. The influence of sentence position

To investigate the influence of sentence position on the tones, the focus condition and non-focus condition are treated separately. For focus conditions, two models with linear and quadratic polynomials are fitted to the two tones separately. The effect of sentence position is both significant on the intercept ($\chi^2(2) = 426.47, p < 10^{-15}$) and on the the linear

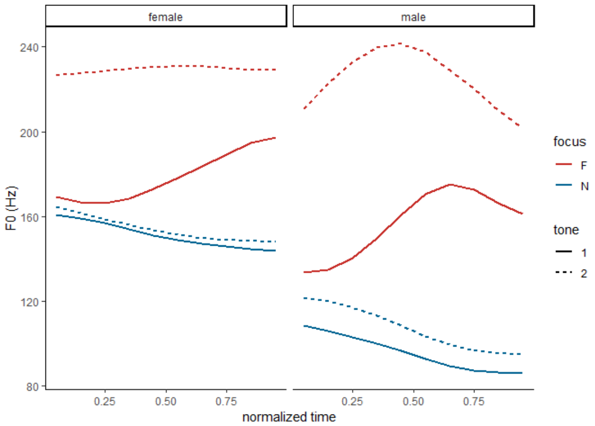

Figure 4: *Averaged pitch contours of tones at sentence-final positions under different focus conditions (left: female; right: male).*

term ($\chi^2(2) = 21.83, p < 10^{-4}$) for T1. For T2, the effect is also significant on the intercept ($\chi^2(2) = 353.29, p < 10^{-15}$) and the linear term ($\chi^2(2) = 11.21, p < 0.00$), but not on the quadratic term ($\chi^2(2) = 3.53, p = 0.17$). Both tones are highest at sentence-initial positions, and lowest at sentence-final positions. The shapes of both tones are steepest at sentence-medial positions, and flattest at sentence-final positions.

The influence of sentence position on tone is obvious under non-focus conditions. Tones at sentence-initial positions preserve their pitch contour as well as their pitch height. The pitch contour of T1 is influenced at sentence-medial positions. At sentence-final positions, the two tones neutralize to low falling, and their contrast in pitch height is preserved to a lesser degree. Under non-focus conditions, sentence position affects both pitch height and pitch contour of the tones in Honggu Chinese.

4. Discussions

It is pointed out at the outset that T2 of Hongggu Chinese has a high level variant and another high falling. The present study finds that the shape of T2 is usually high convex, except for in isolation and under focus at sentence-final positions, and for female speakers only. It is consistent with the impression of the first author and the first dialectologist who reported the tone system of Honggu Chinese (personal communication). We propose that the target of T2 is simply high, without specifying its exact shape. Speakers can implement it as high level or high convex at will. It seems high convex is more usual. When the peak is early relative to the syllable, it can sound like high falling. This accounts for the two variants previously reported.

The influences of focus on tones in Honggu Chinese share properties as Mandarin Chinese. The present design considers only the contrast between focus and non-focus. In fact, the non-focus conditions can be divided into two groups. The non-focus condition at sentence-initial position is pre-focus, and the non-focus conditions at sentence-medial and sentence-final positions are indeed post-focus. The broad focus condition has not been taken into account, which is a flaw of the design. Thus, even if visual inspections show that tones at sentence-initial and sentence-medial positions are higher than tones in isolation, we cannot rush into the conclusion that the pitch ranges of syllables on-focus are broadened due to focus. What is clear in this study is the influence of focus on post-focus components. The pitch range of pre-focus components is only weakly influenced by focus, as shown in sentence-initial positions. The pitch ranges of post-focus components do compress, as in sentence-medial and sentence-final positions. What is more surprising is that tones at sentence-final positions under post-focus conditions all become low falling. This suggests that the tones are neutralized. Since their pitch heights are slightly preserved, this neutralization is incomplete. This finding is of typological significance. In Chinese, the only reported post-lexical tonal neutralization is in Cantonese, due to the high boundary tone in yes-no questions. Focus is found to have influence only on pitch range in Mandarin Chinese, with the pitch contour of tones intact [6]. Honggu Chinese provides a case of post-lexical tonal neutralization due to focus, although combined with the effect of sentence-position.

Another factor that can influence the realisation of tones in Honggu Chinese is sentence position. Tones at sentence-initial positions are higher than at sentence-medial positions, which are in turn higher than sentence-final. This is good indication of declination. Sentence position can also influence the pitch contour of tones. The sentence-final position combined with post-focus condition can lead to tonal neutralization in Honggu Chinese.

Honggu Chinese is notorious for its tonal variation. By manipulating focus condition and sentence position, several factors are identified as the sources of tonal variations. The variation of T2 between high level and high convex is due to the different implementation strategies. Focus can compress the pitch range of post-focus components. Sentence position has an effect on the pitch height of the tones. Sentence-final position, together with post-focus condition, can even lead to the neutralization of the tones in Honggu Chinese. The abundant sources of tonal variation in Honggu Chinese illustrate how tone can vary due to different factors.

5. Conclusions

This paper investigates the tones of Honggu Chinese in different conditions. Tones of different focus conditions, and at different sentence-positions, as well as in isolation are investigated. Several factors that can induce tonal variation are identified, such as focus condition and sentence position. It is proposed that T2 is a high tone, and the variation between high level and high convex is sub-phonemic. Sentence position can influence the pitch height of the tones. Focus can induce tonal variation both in pitch height and in pitch contour. The post-focus condition at sentence-final positions can even induce tonal neutralization. Further study will be conducted on the tonal variations of bi-syllabic words induced by tone sandhi, and the patterns will be compared to that found in monosyllabic words in this paper.

6. Acknowledgements

This research is supported by the Key NSSFC (No. 15ZDB103) and National Key R&D Program of China (2017YFE0111900, 2013CB329301). Thanks also to Prof. Peng Luo for the introduction to the site and the assistance of the fieldwork.

7. References

[1] L. M. Hyman, "Word-prosodic typology," *Phonology*, vol. 23, no. 2, pp. 225–257, 2006.

[2] M. Y. Chen, *Tone Sandhi: Patterns across Chinese Dialects*, ser. Cambridge Studies in Linguistics. Cambridge, UK ; New York: Cambridge University Press, 2000.

[3] Y. Xu and Q. Emily Wang, "Pitch targets and their realization: Evidence from Mandarin Chinese," *Speech Communication*, vol. 33, no. 4, pp. 319–337, 2001.

[4] Y. R. Chao, "Tone and Intonation in Chinese," *Bulletin of the Institute of History and Philology Academia Sinica,*, vol. 4, no. 2, pp. 121–134, Jan. 1933.

[5] L. M. Hyman and K. C. Monaka, "Tonal and Non-Tonal Intonation in Shekgalagari," in *Prosodic Categories: Production, Perception and Comprehension*, ser. Studies in Natural Language and Linguistic Theory, S. Frota, G. Elordieta, and P. Prieto, Eds. Dordrecht: Springer, 2011, pp. 267–289.

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

[7] Y. Mai, "The effect of sentence final rising to the lexical tones in Cantonese [in Chinese]," in *Proceedings of the 7th International Conference on Yue Dialect*, 2000, pp. 163–171.

[8] J. K.-Y. Ma, V. Ciocca, and T. L. Whitehill, "The perception of intonation questions and statements in Cantonese," *The Journal of the Acoustical Society of America*, vol. 129, no. 2, pp. 1012–1023, 2011.

[9] W.-S. Lee, "The effect of intonation on the citation tones in Cantonese," in *Proceedings of TAL*, Beijing, China, 2004, pp. 107–110.

[10] C. Ge and A. Li, "Intonation of Cantonese Interrogative Sentences with and without Sentence Final Particle," in *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 2019, pp. 2455–2459.

[11] P. Luo, "A Chinese dialect with only two tones: The phonological system of Honggu dialect in Lanzhou [in Chinese]," *Journal of Northwest Normal University (Social Science)*, no. 6, pp. 74–77+100, 1999.

[12] M. Hashimoto and X. Wang, "The Altaicization of tones in Chinese," *Journal of Jinzhong University*, no. 02, pp. 112–115+120, 1986.

[13] Y. R. Chao, "A system of Tone-letters," *Le Maître Phonétique*, no. 30, pp. 24–27, 1930.

[14] W. Zhang and W. Deng, "Phonological properties of two-tone language Honggu dialect [in Chinese]," *Studies in Language and Linguistics*, no. 4, pp. 84–88, 2010.

[15] C. Mo and F. Zhu, "Tone sandhi in two tone language Honggu dialect [in Chinese]," *Journal of Gansu Normal Colleges*, vol. 19, no. 01, pp. 43–46, 2014.

[16] L. Li, "Tones and tone sandhi in Honggu dialect in Gansu Province: On tone sandhi of Chinese dialects in Gansu [in Chinese]," *Linguistic Research*, no. 1, pp. 43–47, 2018.

[17] Q. Ran, H. Tian, and B. Qi, "An acoustic analysis of tones in Honggu dialect [in Chinese]," *Chinese Journal of Phonetics*, vol. 4, pp. 82–92, 2013.

[18] R. C. Team, "R: A Language and Environment for Statistical Computing," 2018.

[19] D. Mirman, *Growth Curve Analysis and Visualization Using R*. Boca Raton: CRC Press, 2014.

[20] Q. Li and Y. Chen, "An acoustic study of contextual tonal variation in Tianjin Mandarin," *Journal of Phonetics*, vol. 54, pp. 123–150, Jan. 2016.

[21] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *Journal of Statistical Software*, vol. 82, no. 13, 2017.