9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

L-tone Focus and Word-internal PFC in Kaifeng Mandarin

Lei Wang[1,2]

[1]Tongji University, China
[2]Radboud University Nijmegen, the Netherlands
wl1410512@gmail.com, l.wang@let.ru.nl

Abstract

This investigation examines the effect of prosodic focus on Kaifeng Mandarin L-tone in disyllables through a production experiment involving eight speakers. It is revealed that based on fine-grained acoustic analysis, L-tone focus is characterized by both on-focus pitch range expansion and post-focus compression, along with an increased relative duration and intensity of the host syllable. A pre-L raising effect induced by word-final L focus is also shown to be present. These results may contribute to the cross-linguistic distribution of PFC and provide some insights into our understanding of L-tone focus in general.

Index Terms: L-tone focus, PFC, pre-low raising, Kaifeng Mandarin

1. Introduction

It has been well documented that focus in Mandarin as well as in many other languages is encoded by various prosodic parameters. In Beijing Mandarin, for instance, aside from a greater duration and intensity of the on-focus target, prosodic focus is mainly characterized by increasing the pitch/intensity of the focal item while suppressing the pitch/intensity of the post-focal one [1–3], a phenomenon known as 'post-focus compression or PFC' [4]. PFC, in particular, has been verified to facilitate the identification of focus, hence playing an indispensable function in speech communication [4]. Nevertheless, its absence in a number of languages has also been reported, including Taiwan Mandarin, a closely related language variety with Beijing Mandarin [4]. The first purpose of this current investigation is to provide phonetic evidence that Kaifeng Mandarin, one of the varieties in Zhongyuan Mandarin according to [5], is a PFC language, so as to contribute to the typological distribution of PFC.

Second, different from many previous elicitation paradigms where the whole word is inside the focus and PFC is thus largely examined across phrase boundaries (e.g. [1]), this investigation examines how prosodic focus is encoded within a disyllabic compound. In Kaifeng Mandarin, both syllables within the compound can receive prosodic focus. It should be noted that the prosodic manifestations of focus are dependent on the tone the syllable bears. In general, H-tone tends to have a raised pitch under the focus condition and L-tone a lowered one [1–3,6,7]. Perceptual research indicated that Tone 3 focus is mostly often misidentified, due to its local dissimilatory effect on adjacent tones and the confined capacity of pitch range expansion in the lower range [7]. The local dissimilatory effect includes 'pre-low raising effect' (e.g. [8]) and 'post-low bouncing effect' [9,10] occurring within the same prosodic phrase. The lack of phonetic space at the lower end in itself suggests that L-tone focus may be enhanced by other features. To better address these issues, we investigate the L-tone focus in great detail using new data from Kaifeng Mandarin.

There are four full lexical tones in Kaifeng Mandarin, transcribed impressionistically as [24, 41, 55, 31/312] in five-level Chao digits [11] by [12] and [13]. The phonetic status of these tones was empirically verified and they are phonologically represented as LH, HL, H and L, respectively, by [14]. H/L-tone in sequence triggers tone insertion as an OCP repair strategy with the inserted tone attached to the left-hand syllable, so that H-H and L-L surface as HL-H and LH-L, respectively. The OCP rule is unexceptional and post-lexical in Kaifeng Mandarin.

2. Method

2.1. Stimuli

Six disyllable compounds were chosen in the production experiment. As shown in Table 1, they represent all the logic combination of L and LH, HL, H, excluding underlying L-L which phonologically neutralizes with underlying LH-L. Compounds ending with L have the same morpho-syntactic structure, so are the compounds with an initial L.

Table 1: *Disyllabic compounds.*

Tone	Character	IPA	Gloss
LH-L	拉面	[la mjan]	Stretched noodle
HL-L	捞面	[lau mjan]	Strained noodle
H-L	卤面	[lu mjan]	Stewed noodle
L-LH	豆沫	[tou mwɤʌ]	Soy soup
L-HL	豆苗	[tou mjau]	Bean seedling
L-H	豆奶	[tou nai]	Soy milk

2.2. Subjects

Eight native speakers of Kaifeng Mandarin (mean age: 44.6), four males and four females, participated in the experiment. Half of them received a small amount of fees and the remainder volunteered. None of them had any speaking and listening problems according to their self-report.

2.3. Procedure

Recording was carried out in a quiet room. Prior to the recording, speakers were first given a reading list to familiarize themselves with the words and the three recording sessions. In the first session, speakers were asked to read the target words in isolation where no narrow focus is expected, while in the last two, they were asked to emphasize the initial and the last syllable of each compound, respectively. Syllable prominence is elicited using the corrective focus paradigm, i.e.

10.21437/SpeechProsody.2018-86

不是AB, 是A"C" (Not AB, but A"C"). In the recording, all the target words were presented using Prorec [15], with the triggering context (if present) and six repetitions of a target word separated by the Chinese period printed at the centre of a slide. Quotation marks were put around the emphasized syllable for better visualization. Each slide was presented twice, yielding twelve repetitions per word and per session. Speakers were instructed to read the stimuli in a constant manner and to have a 2–3 second pause between repetitions. All the speakers had no difficulty in following the instruction and if they were not satisfied with their pronunciation, they were allowed to read it again. Speech samples were collected using a head-worn Shure SM10A dynamic microphone connected directly to an ASUS laptop through a Shure X2u pre-amplifier. The sampling frequency is 20000 Hz.

2.4. Analysis

For each subject, ten repetitions per word per focus condition were selected for acoustic analysis, yielding a corpus of 1440 words (6 words × 10 repetitions × 3 focus conditions × 8 speakers). Phonetic annotation and data extraction were carried out in Praat [16], using a script [17]. For each syllable rhyme, eleven equi-distant f0 sampling points, minimum f0, mean intensity, duration and f0 excursion size were collected. Raw f0 values were first converted to semitone with a reference value of 50 Hz and then z-score transformed [18]. Missing values (due to creak or bad sound quality) were replaced by group means.

Growth curve analysis (GCA) [19], was used to access the changes in contour shapes in different focus conditions. Each curve was fitted in a second-order orthogonal polynomial function, where the intercept, the linear and the quadratic terms indicate the mean f0 height, the overall slope and the degree of curvature, respectively. Using orthogonal polynomials instead of the natural ones is to avoid collinearity among different time terms, allowing for independent evaluation of their parameter estimates. The model was structured with fixed effects of focus conditions and random effects of speakers on all time terms.

For duration and mean intensity, the ratio between the two syllable rhymes were calculated. The advantage of using the relative value over the absolute one is that speakers vary in their idiosyncratic speech style, such as speech rate, and within-word focus is most probably based on the relative duration/intensity between the two syllables, as argued by [20] in their case of Mandarin neutral tone.

3. Results

3.1. f0 curves

The f0 contours are presented in Figure 1, where the left panels (a–c) show L-initiated words and the right panels (d–f) L-ended words. Focus conditions are indicated by different line types (black dash = broad focus; black solid = initial focus; grey solid = final focus). Each curve was pooled over across eight subjects. The first sampling point of the first syllable and the last sampling point of the second syllable are omitted to avoid overfitting perception-irrelevant perturbations in that they often show the reverse f0 pattern. Words ending with focal L is not displayed due to interspeaker variation, which will be briefly addressed in 3.1.2. Due to space limitation, L-tones were only statistically compared with those in broad focus.

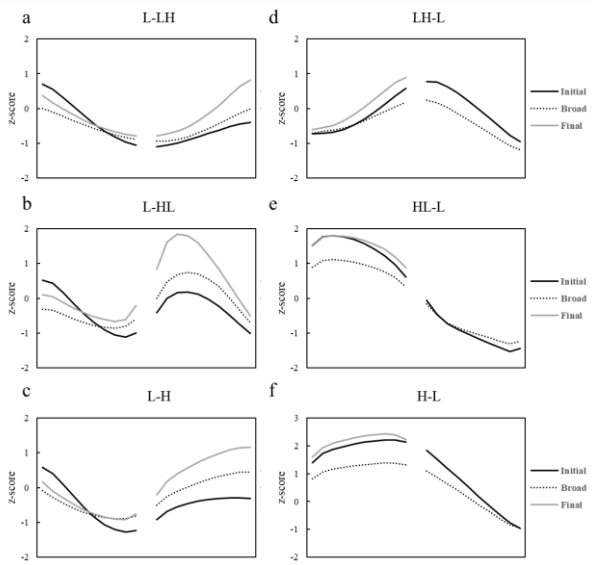

Figure 1: Pitch contours

3.1.1. L in word-initial position

There are three observations based on visual inspection of panels (a–c) in Figure 1. First, word-initial focal L shows a noticeable pitch range expansion, hence a steeper slope, compared with Ls in broad and pre-focus conditions. Results from GCA indicate that the effect of initial focus on all three terms significantly improved the model fit when pooling across the three words. The presence of the effect of initial focus on the intercept (*Estimate* = 0.180, *SE* = 0.008, *p* = 0.000) indicates that the focal L has a greater mean f0 than its non-focal counterpart. The presence of the effect of initial focus on the linear term (*Estimate* = -1.201, *SE* = 0.026, *p* = 0.000) and the quadratic term (*Estimate* = 0.074, *SE* = 0.026, *p* = 0.005) together indicates that the focal L has a steeper f0 slope. This result is further confirmed by comparing the f0 excursion size using linear mixed effect modeling in that the focal L (8.05 ST) shows a much larger excursion size than the non-focal one (3.54 ST) [$\chi^2(1) = 528.84$, $p = 0.000$].

Second, the scaling of the low target can remain unaffected or be lowered under focus, dependent on individual speakers. Specifically, based on inspection of the pronunciation of individual speakers, two of them (one male and one female) fail to show a lowering effect on minimum f0 which is supported by the result of a linear mixed effect modeling, structured with fixed effects of focus and random effects of speakers and words. The effect of focus on the scaling of the L-target turned out to be insignificant for them [$\chi^2(1) = 3.60$, $p = 0.058$]. By contrast, the rest six speakers show a significant lowering of the minimum f0 when the L is focused [$\chi^2(1) = 139.3$, $p = 0.000$].

Third, post-L LH, HL and H show remarkable pitch lowering, i.e. PFC and a leveled-off f0 contour. GCA analyses were conducted independently for each word, comparing post-focal LH, HL and H with those in broad focus. It can be seen that the effect of post-focus context on the intercept significantly improved the model fit for all three words, suggesting that post-focal tones are considerably lowered (LH: *Estimate* =-0.184, *SE* = 0.011, *p* = 0.000; HL: *Estimate* = -0.493, *SE* = 0.017, *p* = 0.000; H: *Estimate* = -0.571, *SE* = 0.012, *p* = 0.000). Furthermore, the effect of focus on the

linear term improved the model fit for L-LH (*Estimate* = -0.242, *SE* = 0.035, *p* = 0.000) and L-H (*Estimate* = -0.390, *SE* = 0.038, *p* = 0.000), not for L-HL (*Estimate* = 0.090, *SE* = 0.053, *p* = 0.093] and the effect of focus on the quadratic term improved the model fit for L-LH (*Estimate* = -0.157, *SE* = 0.350, *p* = 0.000) and L-HL (*Estimate* = 0.248, *SE* = 0.054, *p* = 0.000), not for L-H (*Estimate* = -0.055, *SE* = 0.038, *p* = 0.146), indicating post-focal tones have a flattened f0. Examination of the f0 excursion size in broad and post-focus conditions was carried out. It turned out to be significant for all three tones [LH: $\chi^2(1) = 26.61$, $p = 0.000$; HL: $\chi^2(1) = 16.33$, $p = 0.000$; H: $\chi^2(1) = 31.98$, $p = 0.000$], indicating that post-focal tones undergo pitch range suppression.

3.1.2. L in word-final position

There are three observations concerning L-focus in word-final position. (Only the first point is elaborated due to space limitation.) First, word-final focal L raises the pitch of the preceding tone, i.e. pre-L raising, to the extent that it can be more extreme than the initially focused ones. The pre-L raising effect is quantitatively supported by the result that the effect of focus on the intercept significantly improved the model fit for LH (*Estimate* = 0.324, *SE* = 0.013, *p* = 0.000), HL (*Estimate* = 0.654, *SE* = 0.101, p = 0.000) and H (*Estimate* = 0.863, *SE* = 0.012, *p* = 0.000), as in CGA analysis comparing the tones with those in broad focus. LH and H have an increased slope and a sharper peak which is backed up by GCA analysis. The effect of pre-focal-L context on the linear (LH: *Estimate* = 0.632, *SE* = 0.042, *p* = 0.000; H: *Estimate* = 0.190, *SE* = 0.039, *p* = 0.000) and quadratic terms (LH: *Estimate* = 0.126, *SE* = 0.042, *p* = 0.003; H: *Estimate* = -0.156, *SE* = 0.039, *p* = 0.000) are both significant. For HL, the fixed effect on the linear term failed to improve the model fit (*Estimate* = 0.025, *SE* = 0.047, *p* = 0.593), while the effect on the quadratic term is marginally significant (*Estimate* = -0.102, *SE* = 0.047, *p* = 0.03), indicating that the shape of HL does not undergo drastic change under the influence of pre-L raising. F0 excursion size of pre-focal-L LH [$\chi^2(1) = 153.72$, $p = 0.000$] and H [$\chi^2(1) = 101.36$, $p = 0.000$] is significantly increased, whereas no such effect is reported in HL [$\chi^2(1) = 3.263$, $p = 0.070$], indicating that LH and H undergo pitch range expansion besides a higher f0 scaling under the influence of the L focus.

Second, unlike LH, HL and H, post-focal Ls do not show a noticeable and consistent PFC effect. It seems that only post-LH Ls show undershoot of the L-target, which is absent for post-HL/H Ls.

Third, word-final focal L shows inter-subject variation. In general, enhancement of phrase-final L induces more creak at the point of the L-target, more dipped realization at the second half of the contour, or a lengthened L f0-stretch in the terminal portion of the contour, all of which are optional since there exist speakers who lack any of these enhancement strategies.

3.2. Duration

For the six set of disyllables, the duration ratio between the two syllables under three different focus conditions is shown in Figure 2. The broad focus condition, plotted in between the two narrow focus conditions, serves as the baseline for observing the durational shift induced by focus. Visual inspection reveals that focal L, be it occurring phrase-initially or finally, shows more sizable relative lengthening effect compared with LH, HL and H. Specifically, apart from H, word-initial focal LH, HL and L all show a relative lengthening effect. By contrast, unlike LH, HL and H, word-finally only L is relatively lengthened under corrective focus. These observations were confirmed by the results from a set of linear mixed effect models conducted separately for each word, with duration ratio as the independence variable, focus condition (initial, broad and final) and speaker as fixed and random effects, respectively. It was revealed that for each word, the effect of focus is highly significant [LH-L: $\chi^2(2) = 167.59$, $p = 0.000$; HL-L: $\chi^2(2) = 148.57$, $p = 0.000$; H-L: $\chi^2(2) = 79.74$, $p = 0.000$; L-LH: $\chi^2(2) = 109.35$, $p = 0.000$; L-HL: $\chi^2(2) = 103.01$, $p = 0.000$; L-H: $\chi^2(2) = 40.04$, $p = 0.000$]. Results of multiple comparisons (Tukey) within the focus conditions for each word showed that duration ratio of words containing the focal L is significantly different from either of the two remainder conditions. By contrast, some conditions of particular words failed to reach significance (H-L, broad vs initial; L-LH, broad vs final; L-HL, broad vs final; L-H, broad vs final: though marginally significant, but the duration pattern is counter-intuitive). This indicates that for L-tone focus, lengthening the rhyme is a consistent feature in both phrase-initial and phrase-final positions, unlike the other three tones where no such effect is maintained in both contexts.

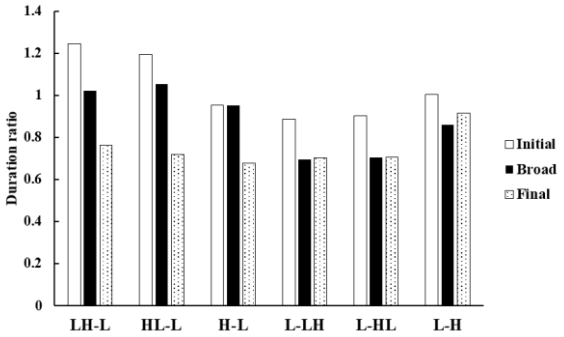

Figure 2: Duration ratio (σ1/σ2)

3.3. Intensity

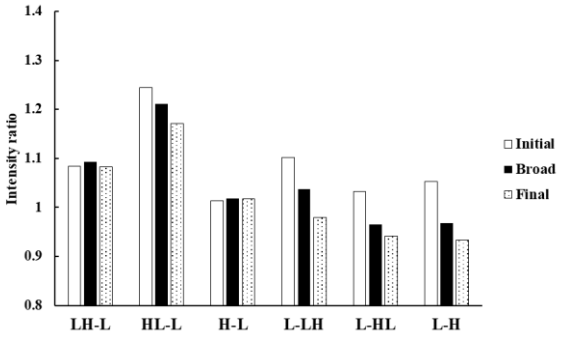

Figure 3: Intensity ratio (σ1/σ2)

The intensity ratio for each word in three different focus conditions is depicted in Figure 3. As shown, L-initiated words have the same intensity pattern. That is, the intensity of a target syllable goes relatively higher when focused. This pattern, however, is only weakly maintained in HL-L, but remains absent in LH-L and H-L. As expected, results from separate linear mixed effect models for each word yielded a significant effect of focus conditions on the intensity ratio for L-LH [$\chi^2(2) = 233.35$, $p = 0.000$], L-HL [$\chi^2(2) = 180.58$, $p =$

0.000], L-H [$\chi^2(2) = 40.04$, $p = 0.000$] and HL-L [$\chi^2(2) = 88.14$, $p = 0.000$]. Tukey Post-hoc test revealed that all three focus types within each of these four words are significantly different from each other. For LH-L [$\chi^2(2) = 4.04$, $p = 0.13$] and H-L [$\chi^2(2) = 1.36$, $p = 0.51$], however, no such effect of focus is reported.

4. Discussion

The current investigation has established that L-tone focus in Kaifeng Mandarin exhibit both cross-linguistic and language-specific properties. Specifically, word-initial L-tone focus is characterized by a pitch range expansion, a greater relative duration and intensity of the syllable carrying the on-focus L-tone. This is in line with the widely attested prosodic ways of encoding focus reported in the literature [1–3,6,7]. The pitch range expansion, in particular, is mainly achieved by raising the f0 onset, and partly by lowering the low-target. A similar picture is provided by [7, pp. 5, Fig. 2] in the case of Tone 3 focus in Standard Mandarin.

The scaling of stressed Tone 3 in Beijing Mandarin has long been noticed. [11, pp. 35], for example, noted that 'when a 3rd Tone is stressed it is dipped lower…'. This impression was empirically verified by the accumulated acoustic evidence [1–4,6,7]. Our current data, however, suggest that lowering of the L-target is speaker-specific. One male and one female failed to show this lowering effect, namely that the pitch range expansion of L focus is solely achieved by a raised f0 onset. It also implies that the absence of a lowered scaling of the L-target in L does not seem to be dependent on the gender of the speaker. Further research with more speakers may spell out this point.

Another debate is concerned with the scaling of L-target in dynamic tones such as HL and LH. [11, pp. 35]'s impressionistic and informal observation suggests that 'when a 4th Tone (HL) is stressed, it starts higher and falls lower'. Our current data, however, seem to suggest that L-scaling in prominent LH and HL tends to remain constant in the case of focal LH (panel d of Figure 1) or slightly raised as expressed by both focal LH and HL (see panels a, b, e in Figure 1). Though lacking statistical support at this current stage, the different performance in the L-target in static and dynamic tones may suggest the invariant property of the low feature for L, hence providing acoustic evidence for its phonological status. The initial falling movement of L, which is expanded in focal condition, serves as a way of auditory enhancement, as noted by [14,21], together with temporal extension and intensity increase of the tone-bearing syllable.

Relative temporal extension of the focal L is consistent in both linguistic contexts, which property is only maintained for LH and HL in word-initial position. One possible explanation is that L-target is not easily lowered, as argued by [22], so that L needs other enhancing features, e.g. lengthening, to stand out from a background and default contextual low pitch. Perceptual research of Tone 3 in Standard Mandarin by [23] substantiated the claim that proportional lengthening of Tone 3 enhances this tone because of an extended 'f0 initial fall' [24] as a result of manipulation which provisionally served as the main perceptual correlates.

Unlike the consistency in relative duration across linguistic contexts, relative intensity is only crucial for L in L-initiated words. This, again, shows the effect of PFC, which induces a greater intensity ratio by suppressing the intensity of the second syllable. Phrase-final Ls, however, lack this intensity increase, most probably due to a combined effect of pre-L raising effect and limited capacity of intensity increase of the L itself. This, in turn, highlights unique and additional enhancement strategies for L, especially in the final position.

Cross-linguistically, pre-L tones tend to have a higher pitch than those followed by non-L tones, a dissimilatory effect, often termed as f0 polarization [25], anticipatory dissimilation [26,27], H-raising [28], anticipatory raising [1] or pre-low raising [8] by a constellation of studies. This raising effect is shown to be magnified when the triggering L is emphasized in Kaifeng Mandarin, similar to the result in Standard Mandarin Tone 3 (e.g. [7]). It also can be observed that quite unexpectedly, the magnitude of this raising effect is even greater than that of the initially focused tones. The underlying mechanism of pre-L raising remains unclear. It could be automatic, as argued by [8], or driven by perceptual mechanisms, i.e. the desire to maximize the perceptual distance between adjacent tones [29,30].

Another main finding of this experiment is that PFC is reported for L. Specifically, when the initial L-tone is focused, post-L LH, HL and H show noticeable pitch range suppression, a reduced f0 height, and an inhibited and fattened f0 contour. The presence (Beijing Mandarin) and absence (Taiwan Mandarin) of PFC in closely related Mandarin families have called for large-scale investigation of PFC both within and across language families [4]. The presence of PFC in Kaifeng Mandarin is in accordance with the hypothesis of the historical origin and development of PFC raised by [4] that as a member of Mandarin family, Kaifeng Mandarin is most likely to show a similar PFC effect as in Beijing Mandarin. This finding clearly extends PFC further south to the middle part of China. Further, PFC in Kaifeng is not likely to be gained via language contact with Beijing Mandarin in that post-L within-word PFC effect is reported to be absent in Standard Mandarin [7], in which, instead of PFC, a post-low bouncing effect was reported. Hence, a reasonable assumption is that PFC in both Kaifeng and Beijing is inherited from a common Proto-language.

At last, L-tones do not seem to serve as a better target for PFC and word-final Ls, when focused, tend to be enhanced by different means, depending on different speakers. Further research is needed to clarify these points.

5. Conclusions

Results of a multi-speaker acoustic study show that L-tone focus in Kaifeng Mandarin exhibits some widely attested prosodic cues of signaling focus, pitch range expansion, temporal extension and intensity increase of the tone-bearing syllable as well as local dissimilatory effects on contiguous tones. Particularly, a PFC effect is found after focal L, which has only been reported to occur after non-L tones in other varieties in the same family, e.g. Beijing Mandarin. It is argued that the presence of PFC in the word level of Kaifeng Mandarin is in support of the hypothetical origin of PFC, and it also highlights the need for L to be enhanced in various prosodic ways.

6. Acknowledgements

This investigation is supported by a CSC scholarship and a travel grant provided by Radboud University Nijmegen. I am grateful to Carlos Gussenhoven for discussions of this topic.

7. References

[1] Y. Xu, "Effects of tone and focus on the formation and alignment of f0 contours," *Journal of Phonetics*, 27(1), 55–105, 1999.

[2] F. Liu and Y. Xu, "Parallel encoding of focus and interrogative meaning in Mandarin intonation," *Phonetica*, 62, 78–87, 2005.

[3] Y. Chen and C. Gussenhoven, "Emphasis and tonal implementation in Standard Chinese," *Journal of Phonetics*, 36, 724–746, 2008.

[4] Y. Xu, S. Chen and B. Wang, "Prosodic focus with and without post-focus compression: a typological divide within the same language family?" *Linguistic Review*, 29, 131–147, 2012.

[5] W. He, "Zhongyuan Guanhua Fenqu(gao) [Classification of Zhongyuan Mandarin (a draft)]," *Fangyan [Journal of Dialect]*, 2, 136–140, 2005.

[6] C. Shih, "Tone and intonation in Mandarin," *Working papers of the Cornell Phonetics Laboratory*, 3, 83–109, 1988.

[7] Y.-C. Lee, T. Wang and M. Liberman, "Production and perception of Tone 3 focus in Mandarin Chinese," *Frontiers in Psychology*, 7(1058), 1–13, 2016.

[8] A. Lee, S. Prom-on and Y. Xu, "Pre-low raising in Japanese pitch accent," *Phonetica*, 74(4), 231–246, 2017.

[9] F. Liu and Y. Xu, "The neutral tone in question intonation in Mandarin," In *Proceedings* of *Interspeech 2007*, Antwerp. August, 630–633, 2007.

[10] S. Prom-on, F. Liu and Y. Xu, "Post-low bouncing in Mandarin Chinese: acoustic analysis and computational modeling," *Journal of the Acoustical Society of America*, 132, 421–432, 2012.

[11] Y. Chao, *A Grammar of Spoken Chinese*. Berkeley: University of California Press, 1968.

[12] Q. Zhang, T, Chen and Y. Cheng, *Henan fangyan yanjiu* [A study of Henan dialects]. Kaifeng: Henan University Press, 1993.

[13] D. Liu, "Kaifeng fangyan jilüe [Kaifeng dialect – a précis]," *Fangyan* [Journal of Dialects], 4, 271–285, 1997.

[14] L. Wang, C. Gussenhoven and J. Liang, "How to pronounce a low tone: lessons from Kaifeng Mandarin," submitted manuscript.

[15] M. Huckvale, Prorec: a program for field workers. [computer program], Version 1.45, 2014, retrieved from http://www.phon.ucl.ac.uk/resource/

[16] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 5.4.08, 2015, retrieved from http://www.praat.org/

[17] Y. Xu, "ProsodyPro – A Tool for Large-scale Systematic Prosody Analysis," In *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (TRASP 2013), Aix-en-Provence, France, 7–10.

[18] P. Rose, "Considerations in the normalization of the fundamental frequency of linguistic tone," *Speech Communication*, 6, 261–282, 1987.

[19] D. Mirman, *Growth curve analysis and visualization using R*. Boca Raton: CRC Press, 2014.

[20] W.-S. Lee and E. Zee, "Prosodic characteristics of the neutral tone in Beijing Mandarin," *Journal of Chinese Linguistics*, 36(1), 1–28, 2008.

[21] C. Gussenhoven, "Intonation," In De Lacy, P. (Eds), *The Cambridge handbook of phonology*. (pp. 253–280) London: Cambridge University Press, 2007.

[22] J. Pierrehumbert, *The phonology and phonetics of English intonation*. PhD thesis, MIT, 1980.

[23] D. Blicher, R. Diehl and L. Cohen, "Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: evidence of auditory enhancement," *Journal of Phonetics*, 18, 37–49, 1990.

[24] X. Shen and M. Lin, "A perceptual study of Mandarin Tone 2 and 3," *Language and speech*, 34(2), 145–156, 1991.

[25] L. Hyman and R. Schuh, "Universals of tone rules: evidence from West Africa," *Linguistic Inquiry*, 5(1), 81–115, 1974.

[26] J. Gandour, S. Potisuk and S. Dechongkit, "Tonal coarticulation in Thai," *Journal of Phonetics*, 22(4), 447–492, 1994.

[27] Y. Xu, "Contextual tonal variations in Mandarin," *Journal of Phonetics*, 25(1), 61–83, 1997.

[28] B. Connell and D. Ladd, "Aspects of pitch realization in Yoruba," *Phonology*, 7(1), 1–29, 1990.

[29] J. Gandour, S. Ponglorpisit, S. Dechongkit, F. Khunadorn, P. Boongird and S. Potisuk, "Anticipatory tonal coarticulation in Thai noun compounds after unilateral brain damage," *Brain and Language*, 45, 1–20, 1993.

[30] S. Potisuk, J. Gandour and M. Harper, "Contextual variations in trisyllabic sequences of Thai tones," *Phonetica*, 54, 22–42, 1997.