



## Prosodic Focus with Post-focus Compression in Lan-Yin Mandarin

Chen Shen<sup>1</sup>, Yi Xu<sup>2</sup>

<sup>1</sup> School of English & Department of Human Communication Sciences, University of Sheffield, UK

<sup>2</sup> Department of Speech, Hearing and Phonetic Sciences, University College London, UK

cshen1@sheffield.ac.uk, yi.xu@ucl.ac.uk

### Abstract

Post-focus compression (PFC), the lowering of pitch range and intensity of the post prosodic focus components, is a phenomenon that has been found in various languages worldwide. The interesting findings of the presence and absence of PFC in two closely-related Mandarin Chinese languages, Beijing Mandarin and Taiwan Mandarin respectively, have brought several discussions on the historical origin and spreading of PFC. This study examined Jincheng subgroup of Lan-Yin Mandarin, a group of Mandarin Chinese mainly spoken in the North-western region of China. Acoustic analyses and statistics showed that all speakers raised their pitch and intensity of focused words, and lowered pitch and intensity of post-focused words except in one condition where post-low bouncing was present. We therefore conclude that Lan-Yin Mandarin exhibits PFC in a very similar way as in Beijing Mandarin, and this further provides support for a hypothetical major typological division among the Chinese languages.

**Index Terms:** Lan-Yin Mandarin, prosodic focus, post-focus compression, post-low bouncing

### 1. Introduction

Prosodic focus is a communicative function that is used to highlight a particular unit of an utterance relative to the rest of the words [3]. In many languages, focus is phonetically realized with the on-focus component displaying expanded pitch range, and post-focus component showing reduced pitch range (also known as post-focus compression (PFC)) [1], [2], [3], [4], [5]. Among these, PFC is one of the most important acoustic cues for perceiving focus [2], [6], [7], [8], and has been found in a number of Indo-European and Altaic languages, such as English [5], German [9], Japanese [10], and Korean [11]. Interestingly, PFC is also found in Mandarin Chinese, a tonal language that belongs to the Sino-Tibetan family [1], [2], [4]. However, PFC is found to be absent in many other languages spoken in China, such as Cantonese [12], Taiwanese [7], and Yi, Wa, and Deang [13]. More surprisingly, PFC is also absent in Taiwan Mandarin, a language that is typologically closely related, and phonetically very similar to Beijing Mandarin [2]. These findings suggest that the distribution of PFC in various languages is rather complicated. Studies such as [2], [14], [15] have demonstrated that PFC is a feature that is hard to spread through language contact. Therefore, questions regarding the historical origin and the possible cross-linguistic transmission of PFC have been raised [16].

Bearing this in mind, the current study is designed to examine another northern dialect of Mandarin Chinese: the

Lan-Yin Mandarin spoken in the Jincheng area (Lanzhou City, Gansu Province, China). Typologically, this Jincheng subgroup of Mandarin Chinese belongs to the *Shan Gan Ning* group of Mandarin [17]. Although comprehensive and systematic studies of the phonetic and prosodic inventory of Lan-Yin Mandarin are rare, there have been some descriptions of the general linguistic and tone features of this dialect. Similar to standard Mandarin, Lan-Yin Mandarin also has four lexical tones, but they differ from those in Standard Mandarin in phonetic tone values [18], [19].

The present study is an investigation of the prosodic encoding of focus in Lan-Yin Mandarin. A specific objective is to find out if PFC occurs in this dialect, as would be predicted based on previous studies, given its northern location in China and its close relation to Beijing Mandarin [2], [16].

## 2. Experiment

### 2.1. Methods

Following previous studies [2], [3], [15], the design of the present study is to use the question-answer paradigm to elicit prosodic focus, and to identify the acoustic correlates of focus through systematic acoustic analysis.

#### 2.1.1. Subjects

Nine native speakers of the Jincheng subgroup of Lan-Yin Mandarin (6 female; 3 male; mean age, 32 year-old) were recruited through personal contact via email, and were recorded remotely over Skype. They had no self-reported speech and hearing disorders.

#### 2.1.2. Stimuli

Two declarative sentences were used as targets (Table 1). Four different focus conditions: neutral (non-focus), initial (focus on word 1), medial (focus on word 2), and final (focus on word 3) in both sentences were elicited by two different sets of pictures together with four precursor questions (Table 2). The four lexical tones of Lan-Yin Mandarin displayed below are transcribed with a standard 5-point scale of representing tone contours, in which 1 means the lowest relative pitch level and 5 the highest [27]. The tonal transcriptions of the two sentences in Beijing Mandarin are also included as reference.

Table 1. Lan-Yin Mandarin target sentences.

	Word 1	Word 2	Word 3
Characters & gloss	老妈 'mother'	摸 'stroke'	狸猫 'civet'
Lan-Yin Mandarin	[lɔ̃35 ma55]	[mɔ̃213]	[li41 mɔ̃55]

Standard Mandarin	[lao214 ma55]	[mo55]	[li35 ma055]
Characters & gloss	老爹 ‘father’	给老妈 ‘mother’	买花 ‘buys flower’
Lan-Yin Mandarin	[lɔ̃35 dɛ55]	[gə̃21 lɔ̃35 ma55]	[me35 hɔ̃a55]
Standard Mandarin	[lao214 dɛ55]	[gei24 lao214 ma55]	[mai214 hɔ̃a55]

Table 2. Precursor questions for eliciting four focuses.

Focus	Precursor Question	Lan-Yin Mandarin Transcription	English Translation
None	图中你看着撒了?	[tu41 ṽñ55 nɪ35 kʰan21 dzɔ̃ə41 sa213 lɔ̃ɔ̃55]	What do you see in the picture?
Initial	谁在摸狸猫?	[fei41 dzɛ213 mɪ213 li41 mɔ̃ɔ̃55]	Who is stroking the civet?
	谁给老妈买花?	[fei41 gə̃21 lɔ̃35 ma55 me35 hɔ̃a55]	Who buys mom flower?
Medial	老妈摸狸猫着呢?	[lɔ̃35 ma55 dzɔ̃a213 li41 mɔ̃ɔ̃55 ɻ43 nə3]	What is mom doing to the civet?
	老爹给谁买花?	[lɔ̃35 dɛ55 gə̃21 fei41 me35 hɔ̃a55]	Who does dad buy flower to?
Final	老妈摸撒着来?	[lɔ̃35 ma55 mɪ213 sa213 dzə3 lɛ41]	What is mom stroking?
	老爹给老妈买的撒?	[lɔ̃35 dɛ55 gə̃21 lɔ̃35 ma55 me35 dɪ 42 sa213]	What does dad buy for mom?

### 2.1.3. Recording

All 9 subjects were recorded remotely using Skype via a piece of third party software: Evaer. It allows the users to record Skype calls with a decent quality for up to 5 minutes without charge. Although this is not the most ideal way of recruiting and recording participants, under the limitations of available resources, this has been tested to be the best possible way to collect data. Moreover, Skype transmission was tested to influence  $F_0$  only minimally [20]. To avoid possible connection issues during high internet traffic period, all the Skype calls were made at around 8 am or 10 pm China Standard Time at the participants’ best convenience. In order to minimize the potential influence that different devices could have on the quality of recordings, each participant was asked to use a MacBook, iPad or iPhone when answering the Skype call. Participants were also encouraged to reduce unnecessary movements so that the distance between their mouth and microphone could remain constant. Upon receiving the Skype call, participants were asked to switch off their video camera, so that high-quality WAV audio file(s) of the call could be automatically generated and saved onto the computer hard disk by Evaer at 44.1 kHz sampling rate, and 16-bit bit depth after the caller clicked the finish button. Several trials were also recorded prior to the actual task to familiarize the subject with the experiment procedure, and for the experimenter to check the quality of the recording. When unavoidable transmission issues occurred during recording process, extra repetition(s) were carried out. The instruction language used by the experimenter throughout the process was the Jincheng variety of Lan-Yin Mandarin.

### 2.1.4. $F_0$ extraction

$F_0$  extraction was done using ProsodyPro, an interactive script [22] running under Praat [21], which allows manual rectification of the vocal pulse marking. The script then generated locally smoothed  $F_0$  contours for each target sentence, and computed mean  $F_0$ , max  $F_0$ , duration, and mean intensity for the voiced part of each syllable of the two sentences.

## 2.2. Analyses and Results

### 2.2.1. Analysis

Figure 1 presents time-normalized mean  $F_0$  contours of the two Lan-Yin Mandarin sentences produced by all 9 subjects, together with cross-subject standard errors. Each curve is composed of 40 sentence repetitions. It can be seen from the figure that, except for the medial focus in the shorter sentence (upper panel), there is a clear increase of the average  $F_0$  of the on-focus words, and substantially lowered  $F_0$  of the post-focus words compared to their neutral counterparts in both sentences.

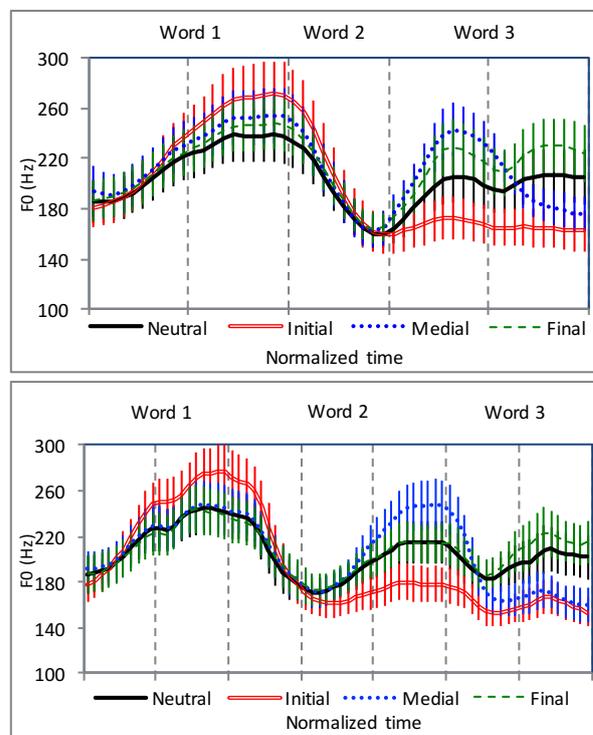


Figure 1: Time-normalized mean  $F_0$  contours of the two Lan-Yin Mandarin sentences, each curve represents an average of 40 sentence repetitions produced by 9 subjects.

Interestingly, the  $F_0$  of the medial focus in the shorter sentence goes higher than not only the neutral focus, but also the final focus initially, and then drops below the neutral level in the later half of the final syllable. This is likely due to an articulatory effect found in Mandarin Chinese known as ‘post-low bouncing’, where the lowering of  $F_0$  in the low tone causes  $F_0$  of the first post-low syllable to rise and then gradually drop [23, 24]. Moreover, the effect of post-low bouncing is found to be stronger when the low tone is under

prosodic focus [24]. Because the on-focused medial word ‘摸’ in the shorter sentence is realized as low tone in Lan-Yin Mandarin, the  $F_0$  of the following syllable is raised at first, and then lowered at the end of the sentence. This also contributes to the less-salient PFC effect on word 3 in the shorter sentence (Figure 2).

### 2.2.2. Results

To verify the observations from the  $F_0$  plots in Figure 1, 2-way repeated measures ANOVAs were performed, with sentence (short vs. long) and focus condition (neutral, pre-, on-, and post-focus) as independent factors. The dependent variables were max  $F_0$ , mean  $F_0$ , duration and intensity. The ANOVA results are presented in Table 3 below.

Table 3. Results of repeated measures ANOVAs. The degrees of freedom for both Initial and Final are 1, 8 and 2, 16, and those for Medial are 1, 8 and 3, 24.

Significant results are shown in boldface (abbreviations: WL – Word Location; Sent – Sentence; Foc – focus).

W L	Factor	Mean $F_0$		Max $F_0$		Duration		Intensity	
		F	p	F	p	F	p	F	p
I	Sent	0.047;	0.834	0.911;	0.3677	0.165;	0.6951	4.724;	0.0615
	Foc	21.791;	<b>&lt;.0001</b>	17.453;	<b>&lt;.0001</b>	8.448;	<b>0.0031</b>	0.149;	0.8631
M	Sent	3.822;	0.0863	27.333;	<b>0.0008</b>	1.099;	0.3251	1.204;	0.3045
	Foc	15.499;	<b>&lt;.0001</b>	14.546;	<b>&lt;.0001</b>	23.005;	<b>&lt;.0001</b>	16.279;	<b>&lt;.0001</b>
F	Sent	6.287;	<b>0.0365</b>	0.511;	0.4951	12.776;	<b>0.0072</b>	75.111;	<b>&lt;.0001</b>
	Foc	19.406;	<b>&lt;.0001</b>	16.956;	<b>0.0001</b>	5.66;	<b>0.0138</b>	30.593;	<b>&lt;.0001</b>

Table 3 shows that sentence has significant effects only on max  $F_0$  in medial word, and on mean  $F_0$ , duration, and intensity in final word. However, focus has significant effects on all four acoustic measurements except intensity in the initial word. In order to find out which of the three focus conditions actually produce significant effect, a Bonferroni/Dunn post-hoc test was conducted, and the results are shown in Table 4.

Table 4. Results of Bonferroni/Dunn post-hoc test. Arrows indicate direction of change relative to neutral focus. (Note that it is not possible for Word 1 to be post-focus or Word 3 to be pre-focus).

WL	Condition	Mean $F_0$	Max $F_0$	Duration	Intensity
I	On-focus	<b>&lt;.0001</b> ↑	<b>&lt;.0001</b> ↑	0.0333 ↑	0.7463 ↑
	Pre-focus	0.3084 ↑	0.1498 ↑	0.0959 ↓	0.596 ↑
M	Post-focus	0.0708 ↓	0.7945 ↓	0.3024 ↓	0.0133 ↓
	On-focus	<b>&lt;.0001</b> ↑	<b>&lt;.0001</b> ↑	<b>&lt;.0001</b> ↑	<b>0.001</b> ↑
F	Pre-focus	0.5266 ↑	0.3178 ↑	0.466 ↑	0.0171 ↑
	Post-focus	<b>0.0007</b> ↓	<b>0.0047</b> ↓	0.1009 ↓	<b>0.0008</b> ↓
	On-focus	0.0699 ↑	0.0227 ↑	0.124 ↑	<b>0.0021</b> ↑

Table 4 shows that mean and max  $F_0$  increased significantly when initial and medial words were on-focus relative to the neutral focus condition. There are also increases of on-focus mean and max  $F_0$  in final word, but they are not significant. This indicates that in Lan-Yin Mandarin, on-focus expansion is present. Although on-focus duration and intensity of all three words increased, only the change in duration and

intensity in medial word, and intensity in final word were significant.

As for the post-focus measurements, reduced mean, max  $F_0$ , intensity, and shortened duration can be seen in both medial and final words, with the lowering of mean and max  $F_0$  as well as intensity in final word being statistically significant. However, despite the consistent on-focus expansion of medial word across all four acoustic factors, post-focus lowering was not significant in this word under any of the four factors. This is likely due to the fact that the tone of this position in the short sentence is realized as a low, and as such it has reached the pitch floor even in the neutral-focus condition, as can be seen in Figure 1. In contrast, the medial word in the long sentence did not have this problem. So we performed additional ANOVAs on the two sentences separately. A significant effect of focus was found in the longer sentence for both mean  $F_0$  ( $F(3,24) = 27.29, p < 0.0001$ ) and max  $F_0$  ( $F(3,24) = 26.26, p < 0.0001$ ). Post-hoc test shows significantly lower post-focus  $F_0$  with both measures relative to the neutral focus condition. For the short sentence, the focus effect was also significant on both mean  $F_0$  ( $F(3,24) = 3.53, p = 0.030$ ) and max  $F_0$  ( $F(3,24) = 7.15, p = 0.014$ ), but a post-hoc test showed *higher* post-focus values relative to the neutral focus condition (Figure 2). As can be seen in Figure 1, this higher  $F_0$  occurs at the beginning of word 2. This is due to the fact that the on-focus  $F_0$  raising in word 1 had a carryover effect, and the fact that the  $F_0$  later in the syllable cannot be lowered due to the pitch-floor effect mentioned above.

Finally, no distinct changes between pre-focus and neutral-focus were found in any of the four acoustic measurements.

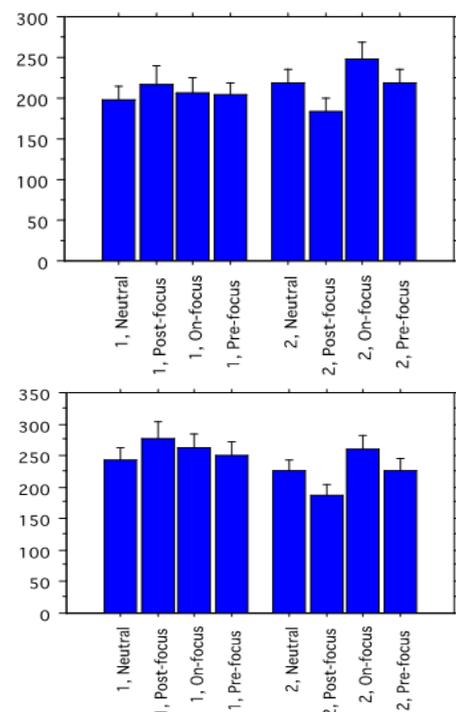


Figure 2: Mean  $F_0$  (upper) and max  $F_0$  (lower) change (semitone) in word 2 of the two Lan-Yin Mandarin sentences produced by 9 subjects (shorter sentence marked as ‘1’ and longer sentence marked as ‘2’ in the graphs) under four focus conditions.

### 3. Discussion

In line with previous studies on the linguistic distribution and realization of prosodic focus in various languages, the results of our analyses of spoken Lan-Yin Mandarin by 9 native Jincheng variety speakers show both on-focus expansion and post-focus compression in their speech production.  $F_0$  and intensity are tested to be the two major acoustic correlates of focus in this Mandarin variety. Duration, on the other hand, is less critical in terms of focus realization. These findings made Lan-Yin Mandarin similar to the Beijing variety of Mandarin, but dissimilar to Taiwan Mandarin as well as Taiwanese and Cantonese [2], [7], [14].

The presence and absence of PFC in various languages have been reported extensively in several studies. The finding of differential existence of PFC in these very closely-related languages is consistent with the hypothetical historical origin of PFC [2], [7], [16]. Following this hypothesis, being a northern variety of Mandarin, Lan-Yin Mandarin should realize focus in a similar way as Beijing Mandarin, which is confirmed by the present results. A complicating factor, however, is that since 1952, the standardized *Putonghua* or Mandarin Chinese has been overwhelmingly used throughout China [25]. Under the influence of certain language policy and some possible sociolinguistic factors, spoken language in Lanzhou has moved dramatically from the local dialect - *Lanzhouhua* - to the more standard Beijing Mandarin. This change has potentially caused the decreasing number of speakers of *Lanzhouhua*, since 4 of the younger participants in the present study claimed that they did not know how to speak proper *Lanzhouhua*. Chen et al. [14] have demonstrated the potential possibility of L2 Mandarin speakers adopting the PFC in their Beijing Mandarin due to frequent use. They did not find, however, the same speakers had adopted PFC in their native Cantonese. So, although we cannot fully rule out the possibility that the Lan-Yin Mandarin speakers in the current study have acquired PFC through close contact with Beijing Mandarin, the likelihood is low based on previous findings [14], [15]. Nevertheless, a comparable study of older speakers of *Lanzhouhua* would offer some insights into whether PFC is a historically existed feature in *Lanzhouhua* or a contemporarily acquired one through pervasive language contact with Beijing Mandarin.

PFC has also been shown to be a hard-to-gain but easy-to-lose feature through language contact by various studies [2], [7], [13], [14], [15]. If this is true, how it is possible that PFC was gained through language contact needs to be reconsidered. According to Xu [16], another possible explanation of the presence of PFC in Lan-Yin Mandarin is that it is inherited. That is, both *Lanzhouhua* and Beijing Mandarin are descendants of an Altaic language, which in turn, is a descendant of an ancient proto-language of a hypothetical macro-language family. This hypothesis further predicts that there is a major division within the Chinese language group [2]. That is, languages spoken in Northern China tend to have PFC, as they are more likely to be genetically related to Beijing Mandarin, whereas those spoken in Southern China tend to have no PFC. The latter prediction has been supported by the absence of PFC in Southern Min (also known as Taiwanese for the variety spoken in Taiwan) [7], [14], and Cantonese [12], [15]. But direct evidence for the former prediction has been scarce until recently [16]. The findings of PFC in Lan-Yin Mandarin in the present study provide further support for the inheritance hypothesis. However, much more

new research on other Chinese languages, as well as non-Chinese languages spoken in China is needed.

### 4. Conclusion

Based on the results of the present experiment, we can conclude that PFC, an effective means of conveying focus through prosody, exists in Lan-Yin Mandarin. This finding provides support for a hypothetical major typological division among the Chinese languages.

### 5. References

- [1] 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.
- [2] Y. Xu, S. Chen and B. Wang, "Prosodic focus with and without post-focus compression: A typological divide within the same language family?", *The Linguistic Review*, vol. 29, no. 1, pp. 131-147, 2012.
- [3] W. Cooper, "Acoustical aspects of contrastive stress in question-answer contexts", *Journal of the Acoustical Society of America*, vol. 77, no. 6, pp. 2142-2155, 1985.
- [4] F. Liu and Y. Xu, "Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation", *Phonetica*, vol. 62, no. 2-4, pp. 70-87, 2005.
- [5] Y. Xu and C. Xu, "Phonetic realization of focus in English declarative intonation", *Journal of Phonetics*, vol. 33, no. 2, pp. 159-197, 2005.
- [6] Y. Xu, C. X., Xu, and X. Sun, "On the Temporal Domain of Focus," *In Proceedings of International Conference on Speech Prosody*, Nara, pp. 81-84, 2004.
- [7] S.-w. Chen, B. Wang, and Y. Xu, "Closely related languages, different ways of realizing focus," *In Proceedings of Interspeech*, Brighton, pp. 1007 - 1010, 2009.
- [8] L. Wang, B. Wang, and Y. Xu, "Prosodic encoding and perception of focus in Tibetan (Anduo Dialect)," *In Proceedings of Speech Prosody*, Shanghai, pp. 286-289, 2012.
- [9] C. Féry and F. Kügler, "Pitch accent scaling on given, new and focused constituents in German", *Journal of Phonetics*, vol. 36, no. 4, pp. 680-703, 2008.
- [10] S. Ishihara, "Syntax-phonology interface of wh-constructions in Japanese," *Tokyo Conference on Psycholinguistics*, Tokyo, pp. 165-189, 2002.
- [11] Y. C. Lee, and Y. Xu, "Phonetic Realization of Contrastive Focus in Korean," *In Proceedings of Speech Prosody*, Chicago, pp. 1003:1-4, 2010.
- [12] W. L. Wu, and Y. Xu, "Prosodic Focus in Hong Kong Cantonese without Post-focus Compression," *In Proceedings of Speech Prosody*, Chicago, pp. 1-4, 2010.
- [13] B. Wang, L. Wang, and T. Kadir, "Prosodic encoding of focus in six languages in China," *In Proceedings of The 17th International Congress of Phonetic Sciences*, Hong Kong, pp. 144 - 147, 2011.
- [14] Y. Chen, S. Guion-Anderson, and Y. Xu, "Post-Focus Compression in Second Language Mandarin," *In Proceedings of Speech Prosody*, Shanghai, pp. 410-413, 2012.
- [15] W. L. Wu, and L. Chung, "Post-focus compression in English-Cantonese bilingual speakers," *In Proceedings of The 17th International Congress of Phonetic Sciences*, Hong Kong, pp. 148 - 151, 2011.
- [16] Y. Xu, "Post-focus compression: Cross-linguistic distribution and historical origin," *In Proceedings of The 17th International Congress of Phonetic Sciences*, Hong Kong, pp., 2011.
- [17] Australian Academy of the Humanities and the Chinese Academy of Social Sciences. *Language Atlas of China. Pacific Linguistics, Series C*, No. 102. Hong Kong: Longman Group (Far East) Ltd., 1998.
- [18] 雒鹏 (L. Peng) 甘肃汉语方言声韵调及特点。西北师大学报(社会科学版), vol. 11, pp. 122-123, 2001.

- [19] 李嘉慧 (J.H. Li) 兰州方言声调特点。语文学刊, vol. 12, pp. 77-78, 2011.
- [20] J. Bulgin, P. De Decker, and J. Nycz, "Reliability of formant measurements from lossy compressed audio," *British Association of Academic Phoneticians Colloquium*, London, 2010.
- [21] Boersma, P., Weenink, D. *Praat*. [Online] available at: <http://www.fon.hum.uva.nl/praat/>.
- [22] 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, pp. 7-10, 2012.
- [23] S. Prom-on, F. Liu and Y. Xu, "Post-low bouncing in Mandarin Chinese: Acoustic analysis and computational modelling", *J. Acoust. Soc. Am.*, vol. 132, no. 1, pp. 421-432, 2012.
- [24] Y. Chen, and Y. Xu, "Production of weak elements in speech—evidence from F0 patterns of neutral tone in Standard Chinese," *Phonetica*, vol. 63, pp. 47–75, 2006. 100–104.
- [25] Wikipedia. "Standard Chinese", Wikipedia.org. [Online]. Available [http://en.wikipedia.org/wiki/Standard\\_Chinese](http://en.wikipedia.org/wiki/Standard_Chinese). [Last Modified: 20 March 2015, 22:42].
- [26] Chao, Y. R., *A Grammar of Spoken Chinese*, Berkeley, CA: University of California Press, 1968.
- [27] Duan, W., and Jia, Y., "The typology of focus realization of Northern Mandarin," in *Proceedings of the 9th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, 492-496, 2014.