



# Phonetic Realizations of Post-nuclear Accent under Dual-focus Conditions in Standard Chinese

Yi YUAN<sup>1</sup>, Aijun LI<sup>2</sup>, Yuan JIA<sup>2</sup>, Jianhua HU<sup>2</sup>, Balázs Surányi<sup>3</sup>

<sup>1</sup>Institute of Linguistics, Chinese Academy of Science, China

<sup>3</sup>Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary

jessie1111@126.com, liaj@cass.org.cn, summeryuan\_2003@126.com, hujh@cass.org.cn, suranyi@nytud.hu

## Abstract

Previous studies indicated that the rightmost unit is the default position for bearing nuclear accent in multiple-focus condition in Standard Chinese. The present research investigated the accent realization for dual-focus sentences which has a leftmost nuclear accent. The syntactic form of [Subject Verb (Modifier1) Object1 (Modifier2) Object2] was adopted. A corrective focus was always assigned to left-headed Subject while another focus was assigned to the right constituents in various sentence length, and the leftmost Subject was always realized as the nuclear accent. Both perceptual and phonetic analysis were induced to explore the accent realization. We found that i) When nuclear accents are placed on the left-head, multiple foci could be realized by post-nuclear accents. ii) Both F0 expansion and PFC are related cues of nuclear accent, but the expansion amplitude depends on sentence length, longer the sentence structure, smaller the expansion. For non-final post-nuclear accent, F0 is not a reliable cue but PFC always has effect. iii) In this particular syntactic structure, there is an obvious prosodic boundary after Object1, which divides this structure into two prosodic phrases with Subject attracting the accent of intonation phrase. vi) Results of perceptual experiment indicate that the SC exists hierarchical-multiple accents.

**Index Terms:** dual-focus, information structure, phonetic realization, syntactic structure

## 1. Introduction

Focus can be defined as the tendency to accentuate or highlight portions of a sentence for reasons related to information structure [1]. It could also be realized through ways including syntax, morphology, phonology and phonetics [2-4].

Phonetically, a focused word has higher F0, longer duration and greater amplitude compared to the unfocused counterpart. For single focus, Xu [5] has reported that there is an F0 rising on the focus and post-focus compression (PFC). But neither of these cues are consistent in some Chinese dialects, like in Taiyuan dialect the pitch range of on-focus syllables doesn't observed significant expanding [6], while in Taiwanese and Taiwan Mandarin there is a lack of PFC [7]. For double foci, there are some research focusing on English, Dutch and Chinese respectively [8, 9]. In a study on English, Eady et al. [8] studied sentences, such as "*The teacher gave the paper to the student*", and observed that the word following an initial focused item in a double-focus sentence

did not have the low F-value characteristics of words that follow focused items. In a study on Chinese, Kabagema-Bilan et.al [10] studied the multiple (non-sentence-final) foci in Mandarin and reported that each focus may require some prominences. But they assumed that in the multiple-focus condition, prominence is assigned to the rightmost focus: multiple focus in Mandarin shows phonetic focus effects on the second, but not on the first of the two foci. Recently, Wang and Féry [11] did a specific research on dual-focus intonation in Standard Chinese (SC), which studied SVO sentences. Their results showed that the F0 increased for both foci, but no prosodic boundary was inserted after the first focus and the whole sentence was realized as one intonation phrase. Jia [12] took a detailed study on the double and multiple rheme foci conditions in SC. In both conditions, the rightmost positions are the bearing unit of nuclear accent. And the studies stated that there are two accent patterns, nuclear accent and pre-nuclear accent entities, existing in Standard Chinese.

Here we should define two terms, accent and stress which are used in our study. Accent, as Beckman [13] said, is a system of syntagmatic contrasts used to construct prosodic patterns. While for stress, Ladd [14] demonstrated that it is an abstract phonological property of a syllable within a prosodic structure. What we used here, the post-nuclear accent, was firstly noted in Halliday's work as second tonic [15]. According to Halliday [15], there are two 'compound tones' and the second one expresses information in a secondary way. Later, Ladd [14] redefined this term as post-nuclear accent according to the intonation structure of British school [16-19]. In Ladd's [14] definition, it stands for some certain tonal events that are related to prominence in some way after the nuclear accent.

In Hungarian, nuclear accent is left-aligned in Broad Focus (BF) and falls on the pre-verbal position. Narrow Focus (NF) elements are usually moved to this position in the syntax. If an NF element is unable to occupy this position and is further to the left, no rephrasing occurs and post-focal deaccenting is not obligatory [20]. For a cross-language comparison, we will take a further look into the multiple-focus in Standard Chinese in our current study. As mentioned in Jia's study [12], the nuclear accent is always realized on the rightmost position. What if the nuclear accent cannot be placed in the default position as mentioned in the Hungarian's research? We try to answer the following questions: I) How is the dual-focus realized perceptually and phonetically? II) Can there be any post-nuclear accent? And what are their phonetic correlates? III) What is the relation between prosodic structure and accent hierarchy?

## 2. Materials and methods

Information structure can interact with prosodic phrasing, which in turn interacts with syntax. In this study, we concentrated on a complex syntactic structure, double-object structure, which can provide different combinations of foci (single-focus and dual-focus) under different information structures.

### 2.1. Materials

Two main principles were adopted here for selecting the materials: i) the sentence should be long enough to provide various positions for multiple foci. ii) The nuclear accent should be controlled at the onset of the sentence in order to expand the previous research [12]. The syntactic structure selected in the current study was [Subject + Verb + (M<sub>1</sub>) + Object1 + (M<sub>2</sub>) + Object2], the double-object structure. By varying the numbers of modifiers, we obtained four types of structure: SVO<sub>1</sub>O<sub>2</sub>, SVM<sub>1</sub>O<sub>1</sub>O<sub>2</sub>, SVO<sub>1</sub>M<sub>2</sub>O<sub>2</sub>, SVM<sub>1</sub>O<sub>1</sub>M<sub>2</sub>O<sub>2</sub>.

For the information structure categories, every constituent of the sentence could bear a focus according to different stimuli questions. The elicited utterances carried various information structures, and the Subject of the sentence was always assigned with a corrective focus in order to bear nuclear accent in either narrow or dual focus situations. The information structure categories applied here followed Steedman [21]. Examples of the elicited questions and the answer-target sentences are presented in Table 1. There were 80 sentences in total, covering all four lexical tones in Standard Chinese.

Table 1. *Focus stimuli in the study*

<b>Broad focused stimuli</b>	发生了什么事? Fa sheng le shen me shi ? What happened?
<b>Narrow focused stimuli</b>	谁扔给八哥香蕉? Shui reng gei ba ge xiang jiao? Who gave the mynah bananas?
<b>Answer Target Sentence</b>	S V M <sub>1</sub> O <sub>1</sub> M <sub>2</sub> O <sub>2</sub> 张咪扔给[窗边的]八哥[三根]香蕉。 ZangMi gave [near the window] the mynah [three] bananas.
<b>Double Foci-Su-O1</b>	田华扔给谁香蕉? Tian hua reng gei shui xiang jiao? Who did TianHua give the bananas?
<b>Double Foci-Su-O2</b>	田华扔给八哥什么 Tian hua reng gei ba ge shen me What did Tian Hua give the mynah?
<b>Answer Target Sentence</b>	不是田华, 张咪扔给[窗边的]八哥[三根]香蕉。 Not TianHua, ZangMi gave [near the window] the mynah [three] bananas.

### 2.2. Experiment Design

We recruited 10 participants to record the answer-target utterances and recruited another 10 participants to complete

the perception experiments. We take F<sub>0</sub> and duration as the main parameters to examine the acoustic variations in different focus conditions.

#### 2.1.1 Speakers and Recording

Ten native speakers of Standard Chinese from Beijing International Studies University (five females and five males aged from 20 to 22) were recruited in the recording. They were required to read all the 96 question-answer pairs. 880 target utterances were obtained for the following analysis.

#### 2.1.2 Perception Experiment

Another 10 participants from the Capital Normal University in Beijing participated in the perception experiment (five females and five males, all native Standard Chinese speakers aged from 20 to 22 without any hearing problems).

The participants were required to mark the primary prominent words with '[ ]', the secondary prominent words with '\_\_\_' and pauses with '/'. Every annotation marked by the participants scored as 1 point for both the primary and secondary prominent words in the final data. The participants were also instructed that there should be only one primary accented word but multiple choices for the secondary ones and pauses. Only when the Subject was marked by '[ ]', the data of the whole sentence is valid for calculating.

In the annotation system, the most prominent words (i.e., Subjects) correspond to the nuclear accent (i.e., primary accented) or sentence stress and the secondary stressed words correspond to the post-nuclear accent or prosodic phrase stress. The prosodic boundary was related to the pauses marked with '/'. One sample was given as below (the numbers following the syllables represent lexical tone in Chinese, e.g., 'Zhang1' denotes syllable 'Zhang' with Mandarin tone 1):

[Zhang1 Mi1] reng1 gei3/ ba1 ge1 xiang1 jiao1.

[Zhang1 Mi1] gives/ mynah bananas.

### 2.3. Data Annotation and Processing

Then we had these target utterances segmented automatically by a force alignment tool and prosodically annotated by two trained transcribers according to C-TOBi [22] in Praat. F<sub>0</sub> and duration data of each syllable were extracted and F<sub>0</sub> data were normalized in LZ-Score [23]. One-Way ANOVAs were employed with SPSS 10.0 to examine differences of pitch and duration in each syntactic structure between various information categories.

## 3. Analysis of Results

### 3.1. Perception Experiment Data

#### 3.1.1. Perception results of prominence

The materials recorded by the 10-participant production group were used in this perception experiment. Most participants chose Subject as the most prominent word in all conditions, which suggests that nuclear accent was always identified at the Subject position. The secondary prominent words varied across different information structures, which corresponded to the post-nuclear accents.

Figure 1 shows the perception performances of accents measured in participants' marking scores, as a function of the position of accents during perception (S, V, M<sub>1</sub>, O<sub>1</sub>, M<sub>2</sub> and O<sub>2</sub>) and the focus conditions during production (Broad Focus (BF),

focus on Subject (F-SU), Subject & M1 (F-SU-M<sub>1</sub>), Subject & O<sub>1</sub> (F-SU-O<sub>1</sub>), Subject & M<sub>2</sub> (F-SU-M<sub>2</sub>), and Subject & O<sub>2</sub> (F-SU-O<sub>2</sub>). Left panels and right panels are based on the same sets of data with left and right panels respectively grouped by the position of accents during perception and focus conditions during production on the x-axes.

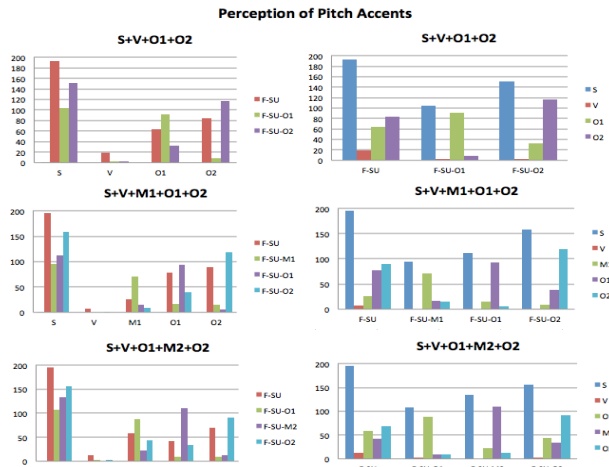


Figure 1: The accent perception results with different syntactic structures

Fig.1 indicated: For the single focus condition (F-SU) where Subject was the nuclear accent, both Objects (O<sub>1</sub> and O<sub>2</sub>) attract stronger prominence than the two modifiers. For the dual-focus conditions, all nuclear accents were realized on the Subject positions, and all other constituents had potential abilities to bear a post-nuclear accent (except for the Verbs).

### 3.1.2. Perceptual Results of Boundary

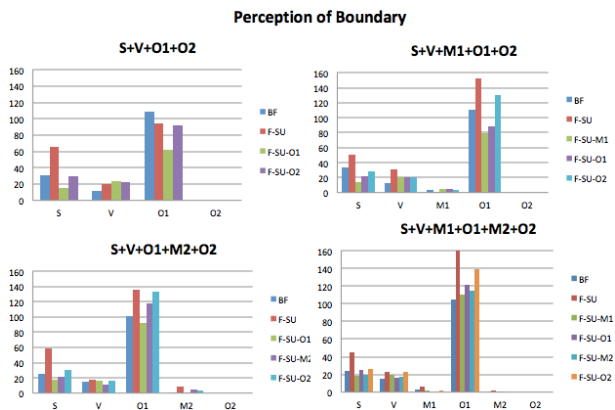


Figure 2: The boundary perception results

The annotation tasks also assess the ability to indicate boundaries, using the prosodic cues. Fig. 2 shows that boundary perception as a function of the position of accents during perception and the position of foci during production. Boundaries could be perceived normally and have various possible positions as the data shown in Fig. 2. We found a consistency that an obvious prosodic boundary after Object1 in all types of syntactic structures, indicating a Prosodic Phrase boundary, while two smaller boundaries were perceived after Subject and Verb indicating Prosodic Word boundary.

## 3.2. Acoustic Experiment Data

### 3.2.1. F<sub>0</sub> patterns

As mentioned in part 2.1, one of the target sentences was “Zhang1 Mi1 Reng1 Gei3 [Chuang1 Bian1 De0] Ba1 Ge1 [San1 Gen1] Xiang1 Jiao1. (Zhang Mi gives [near to window] mynah [three] bananas.)”. We take this H tone group of  $S+V+M_1+O_1+M_2+O_2$  as an example, which can realize all focus conditions including broad-focus, single-focus and dual-focus conditions. The broad focus condition was taken as the baseline for the comparison with other conditions. A One-Way Within-Subject Repeated Measures ANOVA analysis was conducted on the minimum, maximum pitch values and pitch range of each syllable in the sentence, respectively.

*Single focus.* Single focus condition refers to the answer-target sentences elicited by wh-question and had the nuclear accent on Subjects only (F-SU, see Table 1). We found significant rising on the focus ( $p<0.05$ ) and subsequent compression of the F<sub>0</sub> range in post-focus part (post-focus compression, PFC), compared to the broad focus condition.

*Double foci.* We controlled the Subject as a corrective focus, which is also a rheme focus, bearing nuclear accent in dual-focus conditions (see Table 1). In Fig. 3(Fig.3 shows the H tone sentences, the rest experimental utterances of other tone conditions have a similar pattern, which we do not have space to present here), the red rectangular boxes denote the positions of accents in different double foci conditions. There are two dimensions observed, F<sub>0</sub> register and F<sub>0</sub> range. There were no significant differences of F<sub>0</sub> ranges on the two foci between the double-focus conditions while the PFC varying dependently. Longer distance between the two foci, clearer of the PFC effects occur after the nuclear accent. For the post-nuclear accent, the PFC is always a reliable cue to its realization. When second focus is on rheme M<sub>1</sub>, there is a rising on the contour and F<sub>0</sub> followed the focus lowering like the single focus ( $F_{-max}(2,27)=29.017, p<0.001$ ,  $F_{-min}(2,27)=42.775, p<0.001$ ). What we concentrated is the Verb, the pre-focus part, which in this sentence, also the post-focus part of corrective focus. On this constituent, the F<sub>0</sub> register is lowered in a small scale compared with the same level in single focus condition. The results of ANOVA show that the max and min pitch of verb are:  $BF>F-SU-M_1>F-SU$ . And the Bonfroni post hoc test support that, on M<sub>1</sub>, there is no significant between F-SU-M<sub>1</sub> and BF while the contrary with the F-SU ( $p<0.001$ ).

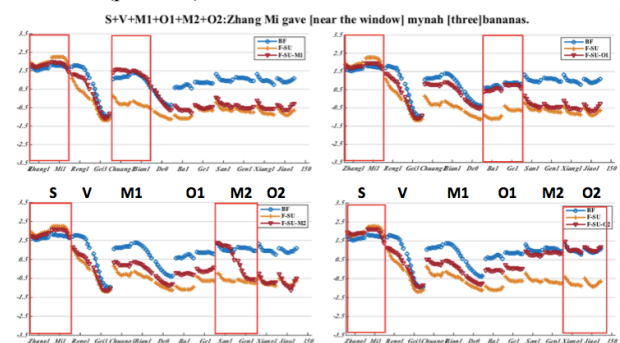


Figure 3: Dual-focus accent realizations for  $S+V+M_1+O_1+M_2+O_2$

The position of second focus moved to the next constituent, the Object1. The raising on focus O<sub>1</sub> is not obvious compared with Broad focus while the PFC contributes

more. And the  $F_0$  of pre-focus part was also barely changed compared with the baseline condition. (Results of the Bonferroni post hoc test shows that the maximum pitch values of M1 is different from Broad focus with  $P < 0.001$  while the range is  $P > 0.001$ ). When come to  $M_2$ , as the second focus, the post-focus part is lowered on pitch register obviously including  $M_1$  and  $O_1$ . What interesting for the on-focus  $M_2$  is, the first syllable, which means a NumP, has the raising of  $F_0$  while the CIP lowered along with the post-focus compression ( $F\text{-max}=58.719$ ,  $p < 0.001$ ,  $F\text{-min}=47.158$ ,  $p < 0.001$ ,  $F\text{-range}=17.690$ ,  $p < 0.001$ ). And the Bonferroni post hoc test support this observation with all the value  $p < 0.001$  compared with BF. For the sentence-final position of second rHEME focus, the post-focus has lowering pitch register on  $M_1$  and  $O_1$  while the  $F_0$  of  $M_2$  was raised to the similar level with the BF condition ( $P\text{-max}=0.331$ ,  $P\text{-min}=0.309$ ).

Then we calculated the pitch ranges of focused words and post-focus words in four double-focus conditions with all tone combinations (H, R, L and F). When Subject and  $M_1$  are both focused, after the first focus, the pitch range was compressed by 0.328. When the distance between two foci increased, the PFC effect became more explicit (in F-SU- $O_1$ , compressed by 1.005). The data of double-focus condition indicates that  $F_0$  expansion on Subject is consistent with BF and F-SU conditions. There is no  $F_0$  effect on non-final post-nuclear accents. PFC after Subject focus shows an anticipatory effect from the following post-nuclear accent.

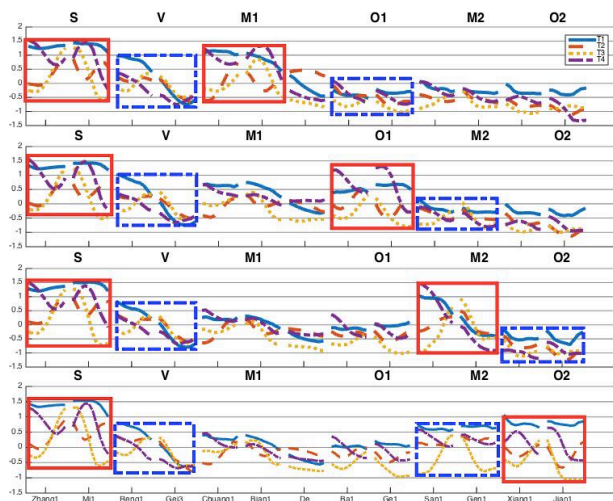


Figure 4: Tone registers of dual-focus accent realizations for  $S+V+M_1+O_1+M_2+O_2$

### 3.2.2. Duration

Within-Subject paired t-tests were conducted on duration of each syllable comparing different focus conditions. Results showed there were no significant differences of syllable duration, including the on-focused syllables, between broad focus condition and other four types of condition F-SU- $O_1$ - $O_2$ , F-SU- $M_1$ - $M_2$ , F-SU- $M_1$ - $O_2$ , F-SU- $O_1$ - $M_2$  (all  $p > 0.05$ ).

## 4. Discussion and Conclusions

In this study, we concentrated on the multiple foci conditions in Standard Chinese (SC). Previous research showed that SC has the phonological entities of nuclear accent and pre-nuclear accent, and in the double and multiple foci

conditions, the rightmost component, as the default position, is the nuclear accent bearing unit [12]. In the study by Kabagema-Bilan et al. [10], the first two foci did not show  $F_0$  effects, and no subsequent  $F_0$  lowering or compression of  $F_0$  range was found after the first focus. When having controlled the corrective or contrastive focus at the start of the sentences (i.e., Subjects), we obtained some new findings related to dual foci realization.

*Focus condition.* In different focus condition, the foci are realized in varied ways. In broad focus condition, two prominences are observed on  $SV(M_1)O_1(M_2)O_2$  structure. Focus scope is not consistent with accent domain and accent assignment is syntactically constrained. In single focus condition, when at the leftmost position, the Subject was realized as nuclear accent, but there could be another post-nuclear accent if a PP boundary exists. For dual-focus conditions, when in the normal condition [10], the rightmost focus is realized as a nuclear accent, and the first one could be realized as a pre-nuclear accent in Standard Chinese. While for this study, the controlled left-headed corrective focus is realized as a nuclear accent on Subject (left headed constituent), the second focus is realized as a post-nuclear accent.

*Phonetic realization.* For nuclear accent, both  $F_0$  expansion and PFC are related cues, but the expansion amplitude is affected by sentence length. Longer the sentence, smaller the expansion. For non-final post-nuclear accent,  $F_0$  is not always a reliable cue but the PFC always shows effects. The lowering effect of PFC shows a clear anticipatory effect from the following post-nuclear accent. There is no obvious difference in various focused regions. We suggest that it is not the case that each [F]-feature [10] directly triggers raising in Focus, but the following lowering and compression is more reliable which has an anticipatory effect from the following post-nuclear accent and prosodic boundary.

*Hierarchical multiple-accents.* For  $SV(M_1)O_1(M_2)O_2$ , there is an obvious prosodic boundary standing after  $O_1$ , which indicates that the sentence has two prosodic phrases. Smaller boundaries can be perceived after Verb as Prosodic Word boundary. The accent is hierarchically organized with prosodic structure [14]: subject bears the Intonation Phrase accent, while the post-nuclear accents are the PP accent or PW accent accordingly in the present corrective focus structure.

To sum up, the phonetic realization of focus for nuclear accent, both  $F_0$  expansion and PFC are related cues, but the expansion amplitude is affected by sentence length, longer the sentence structure, smaller the expansion. For non-final post-nuclear accent,  $F_0$  is not a reliable cue but PFC always has effect. The lowering effect of PFC shows a clear anticipatory effect from the following post-nuclear accent. For duration, there is no obvious difference in different focused regions. Combining the previous research, we suggest that in Standard Chinese, the nuclear accent could be placed both on the left-head and right part of the sentence without changing the prosodic boundary, the nuclear accent is unique and obligatory, pre-nuclear or post-nuclear accents are optional.

## 5. Acknowledgements

This research is supported by joint project between IL, CASS and Research Institute for Linguistics, Hungarian Academy of Sciences, as well as National Program on Key Basic Research Project (973 Program) under Grant 2013 CB329301.



## 6. References

- [1] M. Halliday, "Notes on transitivity and theme in English," *Journal of Linguistics*, vol. 3, pp.199-244, 1967.
- [2] D. Q. Liu, and L. J. Xu, "Jiao dian yu Beijing hua ti ji han yu 'lian'zi ju (Focus and background, topic and lian-sentence in Chinese)," *Chinese Language*, vol. 4, pp. 243-252, 1998.
- [3] J. Xu, *Pu bian Yu fa Yuan ze yu Han yu Yu fa Xian xiang (Grammatical principles and grammatical phenomena)*. Beijing: Beijing University Press, 2001.
- [4] L. J. Xu, "Manifestation of informational focus," *Lingua*, vol. 114, pp. 277-299, 2004.
- [5] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, vol. 27, pp.55-105, 1999.
- [6] W. J. Duan, and Y. Jia, "Contrastive Study of Focus Phonetic Realization between Jinan Dialect and Taiyuan Dialect;" in *O-COCOSDA 2015, October 28-30, Shanghai, China, Proceedings*, pp. 47-52, 2015.
- [7] Y. Xu, SW. Chen, and B. Wang, "Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family?" *Linguistic Review*, vol. 29, pp. 131-147, 2012.
- [8] S. J. Eady, W. E. Cooper, G. V. Klouda, P. R. Mueller, and D. W. Lotts, "Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments," *Language and Speech*, vol. 29, pp. 233-250, 1986.
- [9] H. H. Rump, and R. Collier, "Focus conditions and the prominence of pitch-accented syllables," *Language and Speech*, vol. 39, pp.1-17, 1996.
- [10] E. Kabagema-Bilan, B. Lopez-Jimenez, and H. Truckenbrodt, "Multiple focus in Mandarin Chinese", *Lingua*, vol. 121, pp. 1890-1905, 2011.
- [11] B. Wang, and C. Féry, "Dual-focus Intonation in Standard Chinese," in *O-COCOSDA 2015, October 28-30, Shanghai, China, Proceedings*, pp. 101-106, 2015.
- [12] Y. Jia, *Phonetic and Phonological Analysis of Focus in Standard Chinese*. Beijing: The CASS Press, 2012.
- [13] M. E. Beckman, *Stress and Non-stress Accent*. Foris, Dordrecht, 1986.
- [14] R. Ladd, *Intonational Phonology, Second Edition*. Cambridge: Cambridge University Press, 2008.
- [15] M. A. K. Halliday, *A course in spoken English: intonation*. Oxford: Oxford University Press, 1970.
- [16] H. Palmer, *English Intonation, with Systematic Exercises*. Cambridge: Heffer, 1922.
- [17] J. D. O'Connor, and G. F. Arnold, *Intonation of colloquial English (2<sup>nd</sup> edn)*. London: Longman, 1973.
- [18] R. Kingdon, *The Groundwork of English Intonation*. London: Longman, 1958.
- [19] D. Crystal, *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press, 1969.
- [20] I. Shinichiro, B. Suranyi, and F. Schubö, "Multiple positions of default sentence-level prominence: Meeting Focus Prominence in Hungarian," *Presentation at the 6th international Conference on Tone and Intonation in Europe, Utrecht, 10-12 September 2014, 2014*.
- [21] M. Steedman, "Information structure and the syntax-phonology interface," *Linguistic Inquiry*, vol. 31, pp. 649-689, 2000.
- [22] A. J. Li, "Chinese Prosody and Prosodic Labeling of Spontaneous Speech". In B. Bel and I. Marlin (eds), *Proceedings of the Speech Prosody 2002 Conference. Aix-en-Provence, France, 2002*, pp. 39-46, 2002.
- [23] X. N. Zhu, "F0 Normalization: How to Deal with Between-Speaker Tonal Variations?" *Linguistic Sciences*, vol. 9, pp. 3-19, 2004.