

## FOCUS AND INTONATIONAL PHRASE BOUNDARY IN STANDARD CHINESE

Yiya Chen (*yiya@ling.ed.ac.uk*)

Theoretical and Applied Linguistics Department, University of Edinburgh

### ABSTRACT

This paper reports results of an experiment investigating the relation of focus and prosodic boundary. We tested the hypothesis that focus in Standard Chinese introduces an intonational phrase (IP) boundary before a focused constituent by examining the durational adjustment of syllables in different prosodic positions (i.e. IP initial vs. IP medial) and focus conditions (i.e. focused vs. unfocused). Results show that under both focus conditions, IP initial onset was significantly longer than IP medial onset but little difference was observed in rhyme duration. Focus, however, tended to induce lengthening more consistently in rhyme than in onset in both prosodic positions. Furthermore, the magnitudes of lengthening on onset and rhyme tended to be comparable in terms of their percentage of lengthening. This suggests that the effect of prosodic position on segment duration is localized and restricted to onset while the effect of focus is relatively more global and spans over the whole focused constituent. We also found that an IP initial unfocused syllable differed significantly from an IP medial focused syllable in both onset and rhyme duration. We thus conclude that there is no durational evidence that focus inserts an IP boundary to the left edge of a focused constituent in Standard Chinese.

### 1. INTRODUCTION

This paper reports results of an experiment investigating the relation of focus and prosodic organization. The specific type of focus elicited here is corrective focus, defined here as the linguistic mechanism employed to signal contrast or correction in discourse [1].

It is by now widely accepted that utterances are phrased into constituents based mainly on their morpho-syntactic and information structure; these constituents are hierarchically structured into various domains at different levels of the prosodic organization ([2, 3], among others; see [4] for review). Such a view of prosody emphasizes prosodic constituency and in particular, the effect of domain and domain edges on speech production.

Within this framework of prosody, focus<sup>1</sup> has been viewed as to either insert a new prosodic boundary before the focused constituent or after the focused constituent. Selkirk [5], for example, proposed that focus in English introduces an intonational phrase boundary to the right edge of a focused constituent. In Standard Chinese, Shih

[6] proposes that focus inserts an intonational phrase boundary to the left edge of a focused constituent. (See also works on the relation of focus and phrasing in [7] for Hungarian, [8] for Shanghai, [9] for Korean, among others).

Although the proposal that focus introduces an intonational phrase boundary provides a nice account for the 3<sup>rd</sup> low tone sandhi phenomena in Shih [6], this paper aims to further test the hypothesis by examining patterns of durational adjustment correlated with focus and IP boundary.

The rationale behind this goal is based on results of a series of recent studies which have shown convincingly that segments at different prosodic positions correlate with their strength of articulation ([10], and references therein). Domain initial segments are articulated with more strength than domain medial ones; furthermore, the higher level the domain is at the prosodic hierarchy, the stronger the articulation of the domain-initial segment will be. Different strength of articulation correlates with different durational patterns of a segment.

Following Shih [6], we assume that when a syllable in IP medial position is focused, a new IP boundary would be inserted to the left edge of the syllable. Therefore a focused IP medial syllable would be promoted to be prosodically comparable to an IP initial syllable. To test this hypothesis, we will first examine the durational pattern of both IP initial and IP medial syllables when neither is focused (IP initial<sub>uf</sub> vs. IP medial<sub>uf</sub>). If the same syllable is indeed articulated differently in different prosodic positions and therefore realized with different durations, we will then examine the durational patterns of an IP medial syllable under focus and compare it to an IP initial syllable without focus (IP medial<sub>f</sub> vs. IP initial<sub>uf</sub>). If focus introduces only a new IP boundary, a focused IP medial syllable is predicted to be comparable to an unfocused IP initial syllable. However, if focus not only introduces an IP boundary, but also lengthening of the focused segment(s) (as predicted by a large body of work such as [11] for English, [12] for Standard Chinese, we then predict that a focused IP medial syllable should be more comparable to a focused IP initial syllable (IP medial<sub>f</sub> vs. IP initial<sub>f</sub>). To summarize, we will examine the durational pattern of the same syllable in the following three pairs of prosodic and focus conditions (1).

- (1) a. IP initial<sub>uf</sub> vs. IP medial<sub>uf</sub>  
 b. IP initial<sub>uf</sub> vs. IP medial<sub>f</sub>  
 c. IP initial<sub>f</sub> vs. IP medial<sub>f</sub>

## 2. METHOD

### 2.1. Test Materials

The test materials in this experiment included two sets of data with comparable segmental and tonal compositions which are in two different prosodic positions: One in intonational phrase initial position and the other in intonational phrase medial position (illustrated in table I).

Table I. Schematization of utterances with two different prosodic positions. (Target syllables are underlined and bolded; Square brackets indicate IP boundaries.)

Prosodic Positions	Schematized templates
Intonational Phrase Initial	$\sigma \sigma \sigma \sigma,_{ip}[\underline{\sigma} \sigma \sigma \sigma$
Intonational Phrase Medial	$\sigma \sigma \sigma,_{ip}[\sigma \underline{\sigma} \sigma \sigma \sigma$

Exemplified in (2a) is from the first set where the target syllable is in intonational phrase initial position and that in (2b) is from the second set where the target syllable is in intonational phrase medial position.

- (2) a. Intonational Phrase initial  
 zhōu bīn zhāng shuō, miāo nán hěn duō.  
 ‘zhōu bīn zhāng said that miāo is much more difficult.’  
 b. Intonational Phrase medial  
 zhōu bīn shuō, shuō miāo nán hěn duō.  
 ‘zhōu bīn said that it is much more difficult to say miāo.’

Here the underlined is the target syllable, which varied in all four lexical tones. To keep the data set manageable for analysis within a reasonable amount of time, only one syllable structure was included for the target syllable (CGVG *miao*). The syllable that precedes the target syllable varied between *shuo* with high-level tone (H), meaning “to say”, and *xie* with low-falling tone (L), meaning “to write”. The syllable that follows the target syllable varied between *man* with a falling tone (F), meaning “slow”, and *nan* with a rising tone (R), meaning “difficult”. Thus the target syllable was preceded by tones that end high or low and followed by tones that start high or low. The specific choices of the X, Y, and Z were made based upon four factors: semantic meanings of the syllables, possible tonal combinations (since not all syllables carry four different lexical tones), easy segmentation, and availability of the desired syllable structures.

### 2.2. Subjects and recording

Data reported here were from 2 male and 1 female native speakers of Standard Chinese. Two subjects were born and grew up in Beijing. One was not born in Beijing

but has spent many years in Beijing and speaks standard Chinese without any detectable accent. All sentences with different focus conditions were automatically randomized with a computer program. They were asked to repeat the task three times; each time with different randomized order. Recording was done with Sony Digital Mega Bass MZ-R55 at the sampling rate of 16000. All recording sessions were carried out in the sound booth of the Phonetics lab at the State University of New York at Stony Brook. All subjects were told that the purpose of the recording was for a study of emphasis in Standard Chinese, but they were naïve of what exactly was examined. During the recording, whenever subjects responded without emphasis on the target syllable, they were asked by the monitor to repeat the utterance.

### 2.3. Elicitation of focus

Focus was elicited by providing subjects with relevant contextual information. Subjects saw the sentence in Chinese characters (shown in pinyin in (3)) on a computer screen. They were also shown the question in (4). A typical answer from the subject is shown in (5) with emphasis on *ma* (underlined).

- (3) a. zhōu bīn shuō, shuō mā nán hěn duō.  
 ‘Zhoubin said that it is more difficult to say *ma*.’  
 b. zhōu bīn shuō, shuō bí nán hěn duō.  
 ‘zhōubīn said that it is more difficult to say *bi*.’  
 (4) Question:  
 Suppose he said sentence (a), and I said he said sentence (b), how would you utter (a) to correct me?  
 (5) Response:  
 zhōu bīn shuō, shuō mā nán hěn duō.

### 2.4. Measurements and criteria

Individual sentences were first extracted and saved as separate ESPS sound files in the ESPS/waves+ signal processing software package (Entropic Inc.). They were then segmented in Praat. Segmentation was based on both spectrogram and zoomed waveforms, with reference to audio judgement. We measured the duration of bilabial nasal onset (*m*) and the duration of the rest of the syllable (*iao*), which we will refer to as rhyme though no commitment was made to the syllabic role of the onglide (*j*, spelled as *i*) and offglide (*w*, spelled as *o*). In the sentence medial position, the onset of nasal constriction duration was measured from a clear formant discontinuity from the preceding vowels and the following glides. In the sentence initial position, the onset of nasal was measured from the start of voicing.

## 3. RESULTS AND DISCUSSION

Figure 1 illustrates the patterns of durational adjustment of the target syllable, in terms of onset and

rhyme durations, in different focus and prosodic conditions.

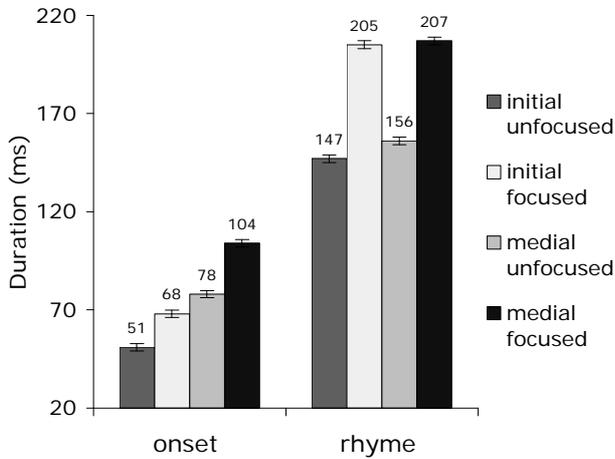


Figure 1. Mean durations (and standard errors) of onset and rhyme of *miao* in different focus conditions (focused vs. unfocused) and prosodic positions (intonational phrase initial vs. intonational phrase medial).

Results of a Repeated Measures Analysis by Subjects showed a significant effect of position on onset duration [ $F(1, 2) = 74.12, p < .025$ ], but not rhyme duration. As a contrast, the effect of focus was near significant on rhyme duration [ $F(1, 2) = 14.89, p < .10$ ] but not on onset duration. No significant interaction of focus and position was found on onset or rhyme duration. We further conducted separate Repeated Measures of Analyses by Subjects examining first the effect of position in different focus conditions and then the effect of focus in different prosodic positions.

When unfocused, an IP initial syllable differed from an IP medial syllable significantly in onset duration [ $F(1, 2) = 22.37, p < .05$ ] but not rhyme duration. The mean duration of the onset of an IP initial syllable (51 ms) was shorter than the mean duration of an IP medial syllable (78 ms). The mean duration of the rhyme of an IP initial syllable (147 ms) was also shorter than that of an IP medial syllable (156 ms), though not significantly. The magnitude of the mean difference for onset ( $\Delta D_{\text{onset}} = 27$  ms), however, was much bigger than that for rhyme ( $\Delta D_{\text{rhyme}} = 9$  ms). In terms of percentage, an IP medial onset was 51% longer than an IP initial onset (calculated as  $\Delta D_{\text{onset}} / D_{\text{initial onset}}$ ); while an IP medial rhyme was only 6% longer than an IP initial rhyme (calculated as  $\Delta D_{\text{rhyme}} / D_{\text{initial rhyme}}$ ).

When under focus, a similar positional effect on the durational pattern of onset and rhyme was observed. Specifically, there was again a significant effect of position on onset duration [ $F(1, 2) = 42.81, p < .025$ ] but not on rhyme duration. The mean duration of the onset of

an IP initial syllable (68 ms) was shorter than that of an IP medial syllable (103 ms). The mean duration of the rhyme of an IP initial syllable (205 ms) was also shorter than that of an IP medial syllable (207 ms). The magnitude of the mean difference for onset ( $\Delta D_{\text{onset}} = 36$  ms) was again much bigger than that for rhyme ( $\Delta D_{\text{rhyme}} = 2$  ms). In terms of percentage, an IP medial onset was 52% longer than an IP initial onset (calculated as  $\Delta D_{\text{onset}} / D_{\text{initial onset}}$ ); while an IP initial rhyme was only 1% longer than that of an IP medial rhyme (calculated as  $\Delta D_{\text{rhyme}} / D_{\text{initial rhyme}}$ ).

Thus we see evidence from durational patterns that segments were realized differently due to their different prosodic positions and such a difference held under both focus conditions. Given that onset was in general shorter than rhyme and that the durational difference due to prosodic position was much greater and consistent for onset than for rhyme, we may conclude that the observed durational differences induced by different prosodic positions was to a large extent localized. In other words, it was the onset, the left-most segment, aligned with a prosodic edge, which exhibited robust durational difference. The localization of prosodic domain initial effect has also been observed in Korean where prosodic position was reported to have a significant effect on C but not V in CV sequence (Cho & Keating 2001).

The effect of focus on onset and rhyme duration was further examined by separating data into two prosodic positions: intonational phrase initial vs. medial. When in initial position, no significant effect of focus was found on onset duration but there was a near significant effect of focus on rhyme duration [ $F(1, 2) = 10.8, p < .10$ ]. On average, a focused onset was 17 ms (25%) longer than an unfocused onset. A focused rhyme was 58 ms (28%) longer than an unfocused rhyme. When in medial position, focus had a near significant effect on onset duration [ $F(1, 2) = 16.74, p < .10$ ] and also a significant effect on rhyme duration [ $F(1, 2) = 22.91, p < .05$ ]. On average, a focused onset was 26 ms (28%) longer than an unfocused onset; a focused rhyme was 52 ms (25%) longer than an unfocused rhyme. Given that there were robust differences in the observed means for both onset and rhyme under different focus conditions but the statistical results were not consistently significant, we further examined the durational adjustment of individual subjects separately and confirmed that despite the variations in the magnitude of durational changes within and across subjects, all three subjects exhibited consistent patterns of durational adjustment reported so far in terms of their individual means.

It is thus safe to conclude that the durational adjustment induced by focus condition was different from that induced by prosodic position. While the prosodic positional effect on segment duration was localized and restricted to onset position only, focus tended to induce

durational adjustment on both onset and rhyme. Lengthening on rhyme tended to be more consistent than that on onset although their magnitudes of lengthening in percentage tended to be comparable.

Repeated Measures Analysis by Subjects was also conducted to further confirm that an unfocused IP initial syllable is not comparable to an IP medial focused syllable. Results showed that the duration of an IP initial onset without focus (51 ms) differed significantly from the duration of an IP medial onset under focus (102 ms) [F (1,2) = 83.19, p < .025]. The duration of an IP initial rhyme without focus (146 ms) differed significantly from the duration of an IP medial rhyme under focus (207 ms) [F (1,2) = 27.27, p < .05].

Table II summarized the results of this experiment. We observed that durational adjustment induced by prosodic positions was maintained under different focus conditions; on the other hand, durational adjustment induced by focus conditions was maintained in different prosodic positions. These two patterns conjointly argue for the view that the effects of prosodic boundary and focus on segment durational adjustment are of different natures. Furthermore, that an unfocused IP initial syllable was significantly different from a focused IP medial syllable in both onset and rhyme durations further argues against the possibility that focus inserts an intonational phrase boundary, and “promotes” an intonational phrase medial constituent to become intonational phrase initial.

Table II. Durational differences in onset and rhyme, induced by different prosodic positions and focus conditions (shaded ones are not statistically significant).

Onset	Rhyme
IP initial <sub>uf</sub> < IP medial <sub>uf</sub>	IP initial <sub>uf</sub> » IP medial <sub>uf</sub>
IP initial <sub>uf</sub> < <b>IP medial<sub>f</sub></b>	IP initial <sub>uf</sub> < <b>IP medial<sub>f</sub></b>
<b>IP initial<sub>f</sub> &lt; IP medial<sub>f</sub></b>	<b>IP initial<sub>f</sub> » IP medial<sub>f</sub></b>

#### 4. CONCLUSION

In a nutshell, this study tested the hypothesis that focus introduces an intonational phrase boundary to the left edge of a focused constituent by examining the patterns of durational adjustment correlated with focus and IP boundary. We found that 1) there was a significant difference between an IP initial onset and an IP medial onset when neither was focused (IP initial<sub>uf</sub> vs. IP medial<sub>uf</sub>) but no significant difference was found in rhyme duration; 2) a focused IP medial syllable differed significantly from a focused IP initial syllable (IP medial<sub>f</sub> vs. IP initial<sub>f</sub>) in onset duration but not in rhyme duration; 3) when an IP medial syllable was focused, both its onset and rhyme duration were significantly different from an IP initial syllable without focus (IP medial<sub>f</sub> vs. IP initial<sub>uf</sub>). This led us to conclude that in Standard Chinese, IP

boundary induced durational adjustment and focus induced durational adjustment should be viewed as independent of each other. There was no durational evidence to support the hypothesis that focus inserts an intonational phrase boundary before the focused constituent.

Note that there might exist two types of intonational phrase boundaries: One with a perceptually clear pause between IPs and the other without pause. When there is pause, only the duration of the nasal closure opening is measurable acoustically, which leads to a shorter consonant duration at IP initial position, as observed in this experiment. The IP boundary that blocks tonal sandhi may be equivalent to one without a pause, as pointed out by Chilin Shih (personal communication). Experiments are thus needed to further investigate this issue.

#### 5. REFERENCES

- [1] Gussenhove, Carlos. “Types of Focus in English”. In Daniel Buring, Matthew Gordon & Chungmin Lee (eds.) *Topic and Focus: Intonation and Meaning. Theoretical and Crosslinguistic Perspectives*. Dordrecht: Kluwer. 2004.
- [2] Selkirk, E. “On Derived Domains in Sentence Phonology.” *Phonology Yearbook* 3: 371-405. 1986
- [3] Nespor, M. & I. Vogel. *Prosodic Phonology*. Dordrecht, Foris. 1986.
- [4] Shattuck-Hufnagel, S. & A. Turk. “A Prosody Tutorial for Investigators of Auditory Sentence Processing.” *Journal of psycholinguistic research* 25(2): 193. 1996.
- [5] Selkirk, Lisa. “Contrastive FOCUS vs. Presentational Focus.” *Proceedings of Speech Prosody 2002*. 2002.
- [6] Shih, Chlin. Mandarin Third Tone Sandhi and Prosodic Structure. In J. Wang and N. Smith (eds.). *Studies in Chinese Phonology*, 81-124. Dordrecht: Foris. 1997
- [7] Vogel, I. and I. Kenesei. “The Interface between Phonology and Other Components of Grammar: The Case of Hungarian.” *Phonology Yearbook* 4: 243-263. 1987.
- [8] Selkirk, E. & T. Sheng. “Prosodic domains in Shanghai Chinese.” In S. Inkelas & D. Zec (eds.) *The phonology-syntax*. Chicago: University of Chicago Press. 1990
- [9] Jun, S.-A. *The Phonetics and Phonology of Korean Prosody: Intonational Phonology and Prosodic Structure*, Garland, New York. 1996.
- [10] Fougeron, C. and P. Keating. “Articulatory strengthening at edges of prosodic domains.” *Journal of the Acoustical Society of America* 101: 3728-3740. 1997.
- [11] Turk, A. and J. Sawusch. “The domain of accentual lengthening in American English.” *Journal of Phonetics* 25: 25-41. 1997.
- [12] Chen, Y. “Durational Adjustment under Corrective Focus in Standard Chinese”. *Journal of Phonetics*. (To Appear)
- [13] Cho, T. and P. A. Keating. “Articulatory and Acoustic Studies on Domain-Initial Strengthening in Korean.” *Journal of Phonetics* 29 (2): 155-90. 2001

<sup>1</sup> Here I will use the cover term *focus* to refer to the different types of focus that the researchers might have elicited when using different methods to obtain the data.