

## **Effect of Focus on the Intensity of Disyllabic Words in Chinese**

Qi Gong

*College of Foreign Studies, Jinan University, Guangzhou, China*

---

**Abstract:-** In this experiment, the intensity pattern and the effect of focus on it for disyllabic words in Chinese is analyzed. Results show that, in most cases, the average intensity of rhyme in the first syllable is comparatively great, while for the onset, due to intervocalic voicing, that in the second syllable is great. Focus has great effect on intensity. Under focused condition, the overall intensity will greatly increase. The results from this study will be useful in speech engineering.

**Keywords:-** Onset, rhyme, intensity, focus, syllable

---

### **I. INTRODUCTION**

The study of focus is of great importance in speech engineering. Focus refers certain part of an utterance which is the centre of attention. It indicates what the speaker presents as being significant or what the speaker supposes to be more revealing to the listener. Focus can be presented acoustically. For example, it is generally believed that focus is intimately related to pitch and durations. The acoustic realization of focus can be explained in the following way: Firstly, there is usually a dramatic rise in pitch on the focused phrase [1-3]; secondly, an extension in duration of the focused syllables [4, 5]; and thirdly, a general pitch reduction in the post-focus sequence [2, 3].

Besides pitch and duration, it is exhibited that intensity is also a dependable correlate of focus. Heldner [6] argues that intensity is a dependable correlate of focal accent in the sense that there are statistically significant differences between focal and non-focal words, in all positions and for all speakers in the analyzed materials, as well as in the sense that they are helpful for automatic detection of focus. In addition, he argues that the results in his research work might prove useful in modeling for speech synthesis. Plag et al. [7] examined the acoustic correlates of primary and secondary stress in North American English, and it is found that the differences in pitch and intensity between the two stressed syllables are great, especially in accented left-prominent words.

In Chinese, much research work has been conducted on the representation of pitch and duration of focus. It is displayed that focus arrangements are realized as pitch range variations imposed on different parts of an utterance. The pitch range of tonal contours directly under focus is usually expanded; the pitch range after the focus is greatly repressed; and the pitch range before the focus does not change much from the neutral-focus condition. Thus, there seem to be three focus-related pitch ranges: expanded in non-final focused words, reduced in post-focus words, and neutral in all other words. It is also indicated that the on-focus force increases the rising slope of the rising tone in Chinese, and research on focus in both English and Chinese has displayed many similarities between the two languages [3, 8].

Regarding the duration lengthening of focused constituent, it is indicated that when the word is in utterance medial position, focus triggers great lengthening. When a focused domain is multi-syllabic, the pattern of lengthening is non-uniform: there is a strong trend of edge effect, with the last syllable lengthened the most. There is also spill-over lengthening on the adjacent syllables outside the focused domain. The extent of such lengthening is conditioned by prosodic boundaries where word boundaries decrease lengthening more than syllable boundaries [5]. Chinese is a tone language, so syllables in most Chinese words are of roughly equal stress, except those with neutral tones. Lin et al. [9] investigated the maximum intensity of disyllabic words in Chinese, and found that in most cases the maximum intensity of the first syllable is larger than that of the second one.

The present study will examine the effect of focus on the intensity of disyllabic words in Chinese. In particular, it will try to answer the following questions. What are the intensity patterns of disyllabic words under unfocused and focused condition? What is the effect of focus on intensity of disyllabic words in Chinese?

### **II. METHODOLOGY**

#### **A. Speakers and stimuli**

The subjects for this experiment are eight native speakers of Standard Chinese, four male and four female. The stimuli are twenty disyllabic verbs, in the form of 'Onset<sub>1</sub> Rhyme<sub>1</sub> Onset<sub>2</sub> Rhyme<sub>2</sub>', like 'Shanghai' (hurt) and 'Xinshang' (appreciate). In Chinese, most of the syllables consist of two parts, the onset and the

---

rhyme, except the ‘zero-onset’ syllables. For instance, in the syllable of ‘shang’, the onset is ‘sh’ and the rhyme is ‘ang’. However, in a zero-onset syllable like ‘ai’, there is no onset, only the rhyme ‘ai’. In experiment 1, only syllables with both onset and rhyme were used, and the intensity of onset and rhyme will be examined separately. For the twenty stimuli, the onsets include fricatives like ‘x’, ‘sh’, etc, and nasals like ‘n’, ‘m’, and the rhymes include monophthongs like ‘i’, ‘u’, etc, diphthongs like ‘ai’, ‘ao’, etc, triphthongs like ‘iou’, and VN combinations like ‘in’, ‘ang’, etc.

All the twenty verbs are normally stressed, without neutral tones. They occur in sentence medial position, in the carrier structure ‘Nana VERB Lili’, in which ‘Nana’ and ‘Lili’ are assumed to be two girls’ names. The sentences were read under three focus conditions, one focusing on the initial noun ‘Nana’, one on the VERB, and still another on the final noun ‘Lili’. As a result, there yielded three focus conditions for the VERB, post-focused, focused, and pre-focused. Foci were triggered by questions. In the first case the question is ‘Shui VERB Lili? (Who VERB Lili?)’, in the second case, ‘Nana zenme Lili? (What did Nana do to Lili? or How does Nana like Lili?)’, and in the last case it is ‘Nana VERB shui? (Who does Nana VERB?)’.

## B. Procedure and measurements

The orders of the sentences are randomized when recording. The questions for triggering foci are recorded beforehand and played from a loudspeaker, and the subjects read the answer after the question was played. Each speaker read the sentences on each focus condition once, yielding a total of 480 recorded sentences (8 speakers  $\times$  20 sentences  $\times$  3 focus conditions).

After the recording, acoustic data were segmented and labeled, with onsets and rhymes of both the first and the second syllables of the key words marked, and intensity extracted using Praat [10]. The segmentation was first done by a segmenting program and then manually corrected. The time step for extracting intensity is 10 milliseconds. Analysis was done by a self-written visual basic program, by which the average of the intensity values within the onset and the rhyme of each syllable of the key word were computed. For example, if the duration of the rhyme of a syllable is 200 ms, the extraction will get 20 intensity values within it, and the average intensity is the mean of the 20 intensity values. Statistic analysis was done in SPSS.

## III. RESULTS

### A. Post-focused vs. focused

Fig. 1 displays the average intensity values of the onset and the rhyme of both the first and the second syllable, under post-focused and focused conditions. In the following sub-sections, detailed analysis will be illustrated about them.

#### 1) Onset and rhyme

It is displayed from repeated measures ANOVA results that the main effect of onset versus rhyme is significant, i.e. there are significant difference between the average intensity values of onset and rhyme:  $F(1, 159) = 352, p < 0.001$ , with the intensity value of the rhyme much greater than that of the onset.

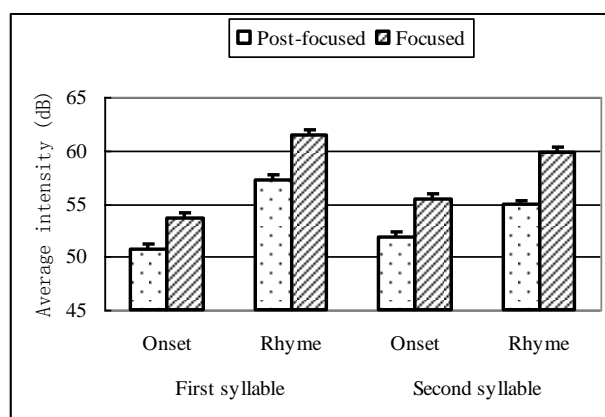


Fig. 1. Average intensity values of onset and rhyme at two syllable positions and under two focused conditions

#### 2) The first and the second syllable

A. Under post-focused condition: Repeated measures ANOVA results show that, under post-focused condition, significant difference also exists between the average intensity values of the first and second syllables for both the onset and the rhyme. But there is interactive effect. For onset,  $F(1, 159) = 4.82, p = 0.03$ , with the

average intensity of the second syllable a little greater than that of the first. For rhyme,  $F(1, 159) = 39.23$ ,  $p < 0.001$ , with the average intensity of the first syllable much greater than that of the second.

B. Under focused condition: The intensity pattern under focused condition is quite similar to that under post-focused one. Repeated measures ANOVA results show that, under focused condition, there are also significant difference and interactive effect. For onset,  $F(1, 159) = 10.13$ ,  $p = 0.002$ , with the average intensity of the second syllable greater than that of the first syllable. For rhyme,  $F(1, 159) = 10.52$ ,  $p = 0.001$ , with the average intensity of the first syllable greater than that of the second.

### 3) Focus

Results show that the effect of focus on intensity is great. It is revealed from repeated measures ANOVA results that, whether the onset or the rhyme, and regardless of the first or the second syllable, the effect of focus on average intensity is always significant, with the average intensity value under focused condition much greater than that under post-focused one. For onset, first syllable:  $F(1, 159) = 104.9$ ,  $p < 0.001$ ; second syllable:  $F(1, 159) = 132.8$ ,  $p < 0.001$ . For rhyme, first syllable:  $F(1, 159) = 204.2$ ,  $p < 0.001$ ; second syllable:  $F(1, 159) = 221.7$ ,  $p < 0.001$ .

Fig. 2 presents the waveform and intensity contours of the sentence ‘Nana xinshang Lili’ (Nana likes Lili), with the key word ‘xinshang’ at the medial position, which is (a) under focused condition and (b) under post-focused condition. From Fig. 2 it is shown that for both the onset and the rhyme in the first and the second syllable, the intensity of those under focused condition are greater than those under post-focused condition.

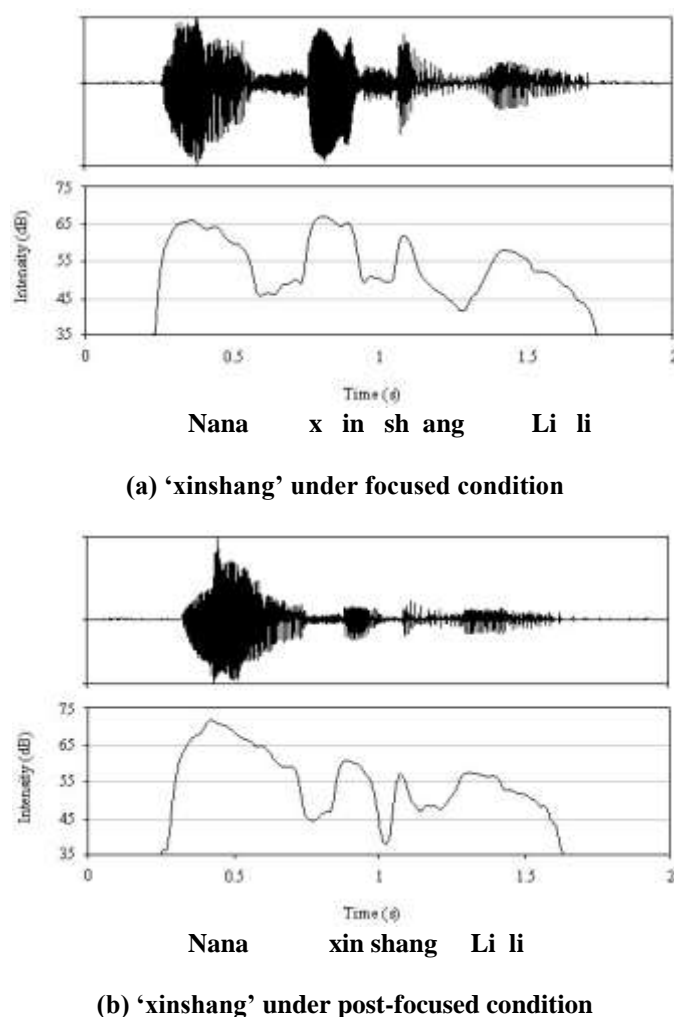
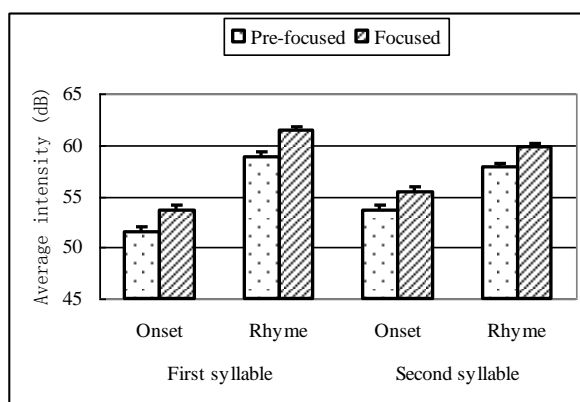


Fig. 2. The waveform and intensity contour of ‘Nana xinshang Lili’, with ‘xinshang’ under (a) focused and (b) unfocused condition

## B. Pre-focused vs. focused

Fig. 3 presents the average intensity values of the onset and the rhyme of both the first and the second syllable, under pre-focused and focused conditions. Detailed analysis will be presented about them in the following sub-sections.



**Fig. 3. Average intensity values broken down by syllable position, onset and rhyme and focus condition**

### 1) Onset and rhyme

It is shown from repeated measures ANOVA results that the main effect of onset and rhyme is significant, i.e. there are significant difference between the average intensity values of onset and rhyme:  $F(1, 159) = 301.9$ ,  $p < 0.001$ , with the average intensity of the rhyme significantly greater than that of the onset.

### 2) The first vs. the second syllable under pre-focused condition

Repeated measures ANOVA results display that, under pre-focused condition, significant difference exists between the mean intensity values of the first and second syllables for the onset:  $F(1, 159) = 10.9$ ,  $p = 0.001$ , with mean intensity of the second syllable greater than that of the first. However, as for the rhyme, there is no significant difference:  $F(1, 159) = 2.46$ ,  $p = 0.119$ . The first and the second syllable are of similar average intensity.

### 3) Focus

When the key words occur at the pre-focused position, the effect of focus on intensity is also great. Repeated measures ANOVA results indicate that, whether the onset or the rhyme, and regardless of the first or the second syllable, the effect of focus on average intensity is consistently significant, with the average intensity under focused condition much greater than that under pre-focused one. For onset, first syllable:  $F(1, 159) = 63.4$ ,  $p < 0.001$ ; second syllable:  $F(1, 159) = 38.4$ ,  $p < 0.001$ . For rhyme, first syllable:  $F(1, 159) = 31.4$ ,  $p < 0.001$ ; second syllable:  $F(1, 159) = 44.1$ ,  $p < 0.001$ .

## IV. DISCUSSION

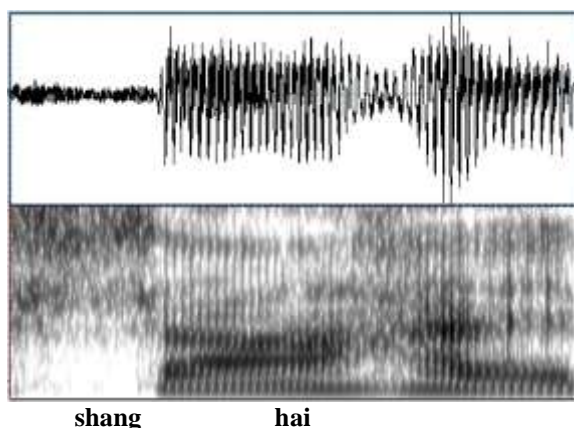
Results from this experiment showed that, first of all, the intensity of the rhyme is much greater than that of the onset. Generally speaking, the onset is optional and the rhyme obligatory in a syllable. If the onset occurs in a syllable, it must be a consonant or consonant cluster. The rhyme is usually a vowel or a combination of vowel and consonant. The vowel is usually called the ‘nuclear’ and the consonant in the rhyme is called the ‘coda’. The vowel is perceptually louder than the consonant, i.e. the intensity value of the vowel is greater than that of the consonant, and it is we believed that this is true in any language. Therefore, the intensity of the rhyme is greater than that of the onset.

It is also indicated from the previous section that the intensity value of the rhyme in the first syllable is much greater than that in the second one. We conjecture that the reason for this is similar to that for pitch. It is found that in an utterance, the pitch is usually high at initial position and low at final position [11]. It is also the case for intensity. In an utterance, the energy of the initial word is greater than the latter ones. Similarly, in a word, the energy of the first syllable is usually greater than the subsequent ones. In a syllable, the vowel of the rhyme is the nuclear, so the rhyme contributes more to the energy than the onset. Therefore, the intensity of the rhyme in the first syllable is greater than that in the second.

However, when the key word is under pre-focused condition, it is not the case. There is no significant difference between the average intensity values for the rhymes of the first and the second syllables. The first reason for this is that, under pre-focused condition, the intensity values of both the first and the second syllables

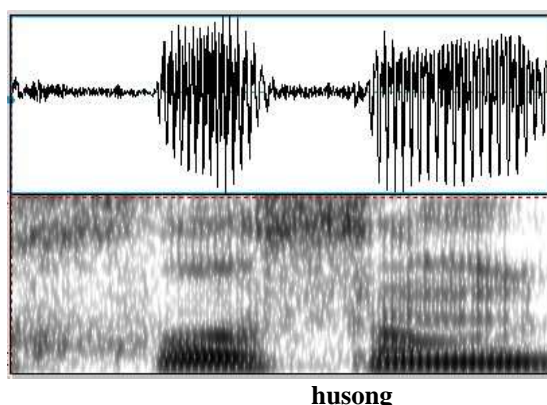
are restrained, not as large as those under focused condition. When they are restrained, the difference between the first and the second syllable will not be as obvious as that under focused condition. The second reason is that, key word at position, with the focus following it. The intensity of the focused word is great, and what immediately precedes the focused word is the rhyme of the second syllable. Due to the anticipatory effect of the following focused word, the intensity of the rhyme of the second syllable will increase. Owing to the dual effect of restraint and the anticipatory influence, the difference between the intensity values of the rhymes in the first and the second syllables disappears.

Regarding the onset, however, it is not the case: The intensity of onset in the second syllable is a little larger than that in the first syllable. As is mentioned above, the onset contributes little to the energy of the syllable, so the intensity mode of the onset is not in line with that of the rhyme. We assume that the reason for greater onset intensity in the second syllable is intervocalic voicing. A consonant occurring at intervocalic position will sometimes get voiced. When a voiceless consonant get voiced, its intensity will get greater. Fig. 4 presents the waveform and spectrogram of ‘shanghai’ (hurt) under unfocused condition, from which it can be seen that ‘sh’ in the first syllable remains voiceless, while ‘h’ in the second syllable gets voiced.



**Fig. 4 The wave form and spectrogram of ‘shanghai’ (hurt) under pre-focused condition**

Fig. 4 displays an example of intervocalic voicing. However, intervocalic voicing is not quite common in Chinese. When the data are examined closely, it is found that in most cases the intervocalic consonants keep unchanged in Chinese disyllabic words. Fig. 5 is one of the case, where the intervocalic consonant ‘s’ does not get voiced.



**Fig. 5. The wave form and spectrogram of ‘husong’ (escort) under unfocused condition**

In this experiment, it is also found that the effect of focus on intensity value is great. No matter the first or the second syllable, and regardless of the onset or the rhyme, the intensity value under focused condition is consistently greater than that under unfocused condition. The intensity values of both the onset and the rhyme, of both the first and the second syllable, are all increased. Intensity is a dependable correlate of focus.

## V. CONCLUSION

In this study, the intensity pattern as well as the effect of focus on the intensity of disyllabic words in Chinese is analyzed. Results show that, energy is mainly carried by the rhyme of the syllable. In disyllabic

words, in most cases, the intensity of the rhyme in the first syllable is great, while under pre-focused condition, there is no significant difference between the average intensity values for the rhymes of the first and the second syllables. In regard to the onset, due to intervocalic voicing, the intensity of the second syllable is comparatively great. Focus has great effect on intensity. Under focused condition, the overall intensity will increase. These results will be applicable in engineering like speech synthesis and speech recognition.

#### REFERENCES

- [1] Mary Beckman, *Stress and non-stress accent*, Dordrecht: Foris Publications, 1986.
- [2] Delphine Dahan and Jean-Marc Bernard, "Inter-speaker variability in emphatic accent production in French," *Language Speech*, vol. 39, no. 4, pp. 341–374, 1996.
- [3] Yi Xu and Ching X. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, vol. 33, pp. 159–197, 2005.
- [4] Alice Turk and Laurence White, "Structural influences on accentual lengthening in English," *Journal of Phonetics*, vol. 27, no. 2, pp. 171–206, 1999.
- [5] Yiya Chen, "Durational adjustment under corrective focus in Standard Chinese," *Journal of Phonetics*, vol. 34, pp. 176–201, 2006.
- [6] Mattias Heldner, "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish," *Journal of Phonetics*, vol. 31, pp. 39–62, 2003.
- [7] Ingo Plag, Gero Kunter and Mareile Schramm, "Acoustic correlates of primary and secondary stress in North American English," *Journal of Phonetics*, vol. 39, no. 3, pp. 362–374, 2011.
- [8] Yi Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, vol. 27, pp. 55–105, 1999.
- [9] Maocan Lin, Jingzhu Yan and Guohua Sun, "A preliminary experiment on normally stressed disyllabic words in Beijing Chinese," *Fangyan*, no. 1, pp. 57–73, 1984.
- [10] Paul Boersma, "Praat, a system for doing phonetics by computer," *Glott International*, 5:9/10, pp. 341–345, 2001.
- [11] Yetunde O. Laniran and G. N. Clements, "Downstep and high raising: interacting factors in Yoruba tone production," *Journal of Phonetics*, vol. 31, pp. 203–250, 2003.