

Research on the Effect of Focus and Stress on the Realization of Intensity in Chinese

Maolin Wang

College of Chinese Language and Culture / Overseas Chinese Education Research Institute, Jinan University, Guangzhou, China

ABSTRACT

In this paper, the average intensity and the effect of focus and stress on it for disyllabic words in Chinese is analyzed. The key words in the experiment are 20 verbs, occurring in sentence medial position, and the subjects are eight native speakers of standard Chinese. Average intensity is analyzed, and results show that the average intensity of the rhyme is greater than that of the onset. Due to consonant voicing, for the onset, under unfocused condition, the intensity of unstressed syllable is greater than that of the stressed one, while for the rhyme, under focused condition, that of stressed syllable is greater than the unstressed one.

Keywords - Intensity, focus, stress, syllable

Date of Submission: 02 June 2014



Date of Publication: 15 June 2014

I. INTRODUCTION

This paper reports an investigation into the effect of focus on the realization of intensity of stressed and unstressed syllables in Chinese. Focus refers to syllables or words that are emphasized by the speaker to indicate that they are new or otherwise informative to the listener. Focus can be signaled acoustically. For example, it is generally agreed that focus is closely related to pitch and durations. The acoustic realization of focus can be stated in the following way: Firstly, there is usually a great and sudden rise in pitch on the focused phrase [1-3]; secondly, an increase in duration on the focused syllables [4, 5]; and thirdly, a global pitch compression in the post-focus sequence either through a low plateau, a late but steady fall or a constant fall until the end of the utterance [2, 3]. At the same time, it is reported that intensity is also a reliable correlate of focal accent. Heldner [6] argues that intensity is a reliable correlate of focus in the sense that there are statistically significant differences between focal and non-focal words for all words, in all positions and for all speakers in the analyzed materials, as well as in the sense that they are useful for automatic detection. Furthermore, he argues that the results in his study might prove important in modeling for speech synthesis.

Plag et al. [7] investigated the acoustic correlates of primary and secondary stress in North American English, and found that the differences in pitch and intensity between the two types of stressed syllables are large, especially in accented left-prominent words. Much research work has been done on the realization of pitch and duration of focus in Chinese. It is shown that focus patterns are implemented as pitch range variations imposed on different regions of an utterance. The pitch range of tonal contours directly under focus is substantially expanded; the pitch range after the focus is severely suppressed; and the pitch range before the focus does not deviate much from the neutral-focus condition.

Thus, there seem to be three distinct focus-related pitch ranges: expanded in non-final focused words, suppressed in post-focus words, and neutral in all other words. It is also shown 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 shown many similarities between the two languages [3, 8]. As for the lengthening of focused constituent, it is shown that when the word is in utterance medial position, focus induces robust lengthening. When a focused domain is multi-syllabic, the distribution of lengthening is non-uniform: there is a strong tendency of edge effect with the last syllable lengthened the most. There is also spill-over lengthening on the neighboring syllables outside the focused constituent. The magnitude of such lengthening is conditioned by prosodic boundaries in that word boundaries attenuate lengthening more than syllable boundaries [5].

Chinese is not a stress language, so syllables in most Chinese words are of roughly equal stress. Lin et al. [9] analyzed the maximum intensity of disyllabic normal-stressed words in Chinese, and found that in most cases the maximum intensity of the first syllable is greater than that of the second one. In Chinese, there are some words in which one or more syllables are of neutral tone. These words are called neutral tone words. According to Lin and Yan [10], the nature of neutral tone syllable is unstressed. The relationship of focus and stress is an interesting topic in phonetic study. Chinese is not a stress language, so the relationship of focus and stress in Chinese becomes an even more interesting topic.

The present study will investigate the effect of focus and stress 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 neutral tone words under unfocused and focused condition? What is the effect of focus on the intensity of this kind of word in Chinese?

II. METHODOLOGY

2.1 Speakers and stimuli

Eight native speakers of Standard Chinese, four male and four female, participated in the recording. The stimuli are 20 disyllabic verbs, in the form of 'Onset₁ Rhyme₁ Onset₂ Rhyme₂', such as 'xiahu' (to scare) and 'hunong' (to fool). In Chinese, most of the syllables are composed of two parts, the onset and the rhyme, except for the 'zero-onset' syllables. For example, in syllable like 'xia', the onset is 'x' and the rhyme is 'ia'. But in zero-onset syllable like 'ai', there is no onset, only the rhyme 'ai'. In the present study, only syllables will both onset and rhyme were used, and the intensity of onset and rhyme will be investigated separately. For the 20 stimuli, the onsets include fricatives like 'x', 'h', etc, and nasals like 'n', 'm'. The rhymes include monophthongs like 'i', 'u', etc, diphthongs like 'ia', 'ao', etc, triphthongs like 'iou', and VN combinations like 'in', 'ong', etc.

All the 20 verbs are neutral tone words, with the second syllable being neutral tones, i.e. the stress pattern of the key word is 'stressed + unstressed'. They occur in sentence medial position in the carrier structure 'Nana VERB Lili', where 'Nana' and 'Lili' are supposed to be two girls' names. The sentences were read under two focus conditions, one focusing on the final word 'Lili', and the other on the VERB. As a result, there yielded two focus conditions for the VERB, unfocused and focused. Foci were elicited by questions. In the first case the question is 'Nana VERB shui? (Who does Lili VERB?)', and in the second case it is 'Nana zenme Lili? (What did Nana do to Lili? or How does Nana like Lili?)'.

2.2 Procedure and measurements

The orders of the sentences are randomized when recording. The questions for eliciting foci are recorded beforehand and played from a loudspeaker, and the speakers read the answer after the question was played. Each speaker read the sentences on each focus condition once, yielding a total of 320 recorded sentences (8 speakers × 20 sentences × 2 focus conditions).

After the recording, acoustic data were segmented and labeled, with onsets and rhymes of both the stressed and the unstressed syllable of the key words marked, and intensity extracted using Praat [11]. 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 calculated. 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

3.1 Main effect

Fig. 1 presents the average intensity of the key words, broken down by focus condition, stress condition and onset/rhyme condition. Repeated measures ANOVA result shows that, as far as main effect is concerned, the effects of focus and onset/rhyme condition are significant, focus condition: $F(1, 159) = 37.5, p < 0.001$; onset/rhyme condition: $F(1, 159) = 567.7, p < 0.001$, with average intensity comparatively large for focus condition and the rhyme. The effect of stress condition is not significant: $F(1, 159) = 0.014, p = 0.907$, with no significant difference between the average intensities of stressed and unstressed syllables. Detailed analysis will be done in the following subsections.

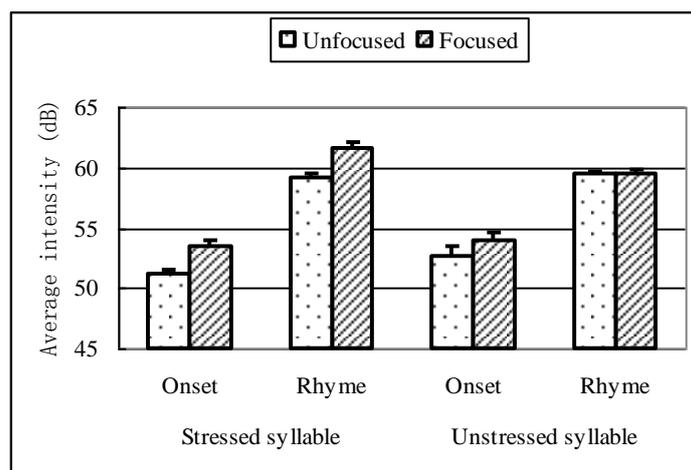


Fig. 1. Average intensity of the key words, broken down by focus, stress and onset\rhyme conditions

3.2 Onset and rhyme

It is shown from repeated measures ANOVA results that there are significant differences between the average intensities of the onset and the rhyme under all of the conditions: unfocused condition, stressed: $F(1, 159) = 277.1, p < 0.001$; unstressed: $F(1, 159) = 99.8, p < 0.001$; focused condition, stressed: $F(1, 159) = 231.9, p < 0.001$; unstressed: $F(1, 159) = 93.9, p < 0.001$, with the average intensity of the rhyme larger than that of the onset.

3.3 Stress

3.3.1. Onset

Repeated measures ANOVA results show that, for the onset, under unfocused condition, there is marginally significant difference between the average intensity values of the stressed and the unstressed syllables: $F(1, 159) = 4.04, p = 0.046$, with that of unstressed syllable greater than the stressed one, while under focused conditions, there is no significant difference: $F(1, 159) = 0.721, p = 0.397$.

3.3.2. Rhyme

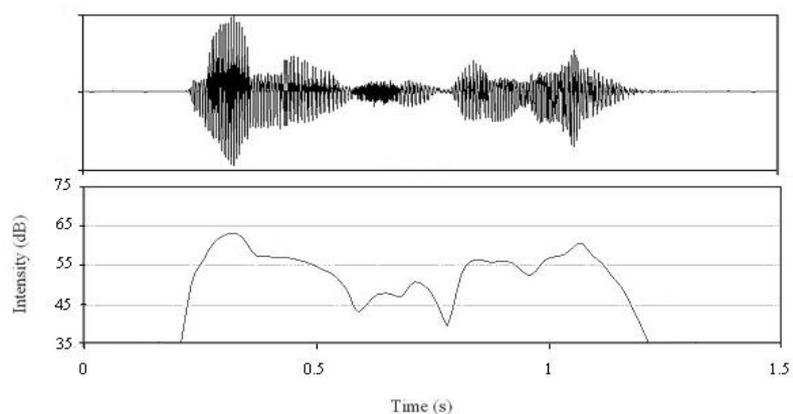
The effect of stress on the intensity of the rhyme is different from that of the onset. It is shown from repeated measures ANOVA results that, for the rhyme, there is significant difference between the average intensity values of the stressed and the unstressed syllables under focused condition: $F(1, 159) = 21.6, p < 0.001$, with that of stressed syllable greater than the unstressed one. Under unfocused condition, there is no significant difference: $F(1, 159) = 127, p = 0.262$.

3.4 Focus

3.4.1. Average intensity

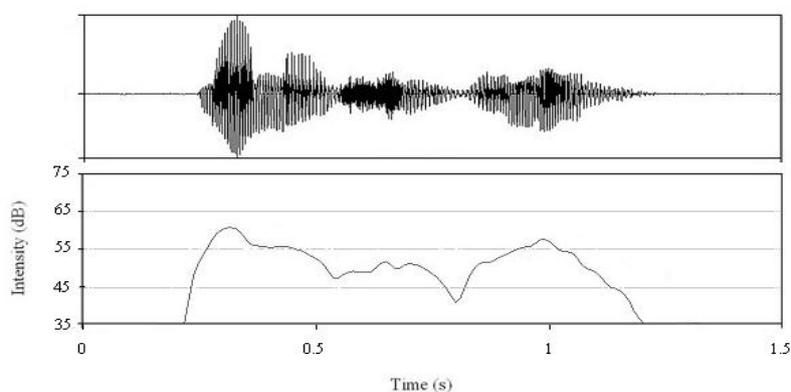
It is shown from repeated measures ANOVA results show that, the effect of focus on intensity is great in most cases. For onset, the effect of focus is significantly great for both stressed and unstressed syllables, stressed syllable: $F(1, 159) = 72.5, p < 0.001$; unstressed syllable: $F(1, 159) = 5.46, p = 0.021$, with the average intensity under focused condition greater than that under unfocused one. For rhyme, there is significant difference for stressed syllable: $F(1, 159) = 80.9, p < 0.001$, with that under focused condition greater than under unfocused one, while for unstressed syllable, there is no significant difference: $F(1, 159) = 0.11, p = 0.741$.

Fig. 2 shows the waveform and intensity contour of the sentence 'Nana xihuan Lili' (Nana likes Lili). The key word is 'xihuan' at the medial position, with (a) under unfocused condition and (b) under focused condition. From Fig. 2 it can be seen that for stressed syllable, the intensity of that under focused condition is greater than that under unfocused one.



Na na xi huan Li li

(a) 'xihuan' under unfocused condition



Na na xi huan Li li

(b) 'xihuan' under focused condition

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

3.4.2. Intensity difference

In the forgoing subsection, it is shown that there is effect of focus on average intensity in most cases. In this subsection, intensity difference will be analyzed. Intensity difference refers to the difference of average intensity values between the focused condition and the unfocused condition, as is shown in (1).

$$\Delta In = In_F - In_U \quad (1)$$

In (1), ΔIn stands for intensity difference, In_F for intensity under focused condition, and In_U for that under unfocused condition.

a) *Onset and rhyme*: Fig 3 presents the intensity difference for onset and rhyme of both the stressed and the unstressed syllables. It is shown from repeated measures ANOVA results that, for unstressed syllable, there is significant difference between the onset and the rhyme: $F(1, 159) = 6.06, p = 0.015$, with the intensity difference of the onset greater than that of the rhyme. However, for the stressed, there is no significant difference: $F(1, 159) = 0.53, p = 0.466$.

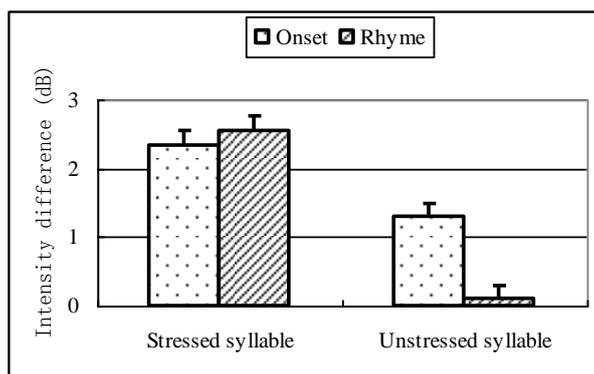


Fig. 3. Intensity difference for onset and rhyme in both the stressed and unstressed syllables

b) *Stress*: In regard to the stressed and the unstressed syllables, repeated measures ANOVA results show that, for onset, there is no significant difference between them: $F(1, 159) = 3.367, p = 0.068$. However, for the rhyme, the difference between the stressed and the unstressed is significant: $F(1, 159) = 54.6, p < 0.001$, with the intensity difference of the stressed syllable greater than that of the unstressed one.

IV. DISCUSSION

Results of this experiment showed that the main effects of focus and onset\rhyme condition on average intensity are significant. For onset and rhyme, no matter the stressed or the unstressed syllable, and regardless of the focused or the unfocused condition, the intensity of the rhyme is always much greater than that of the onset. Generally speaking, in a syllable, the onset is optional and the rhyme obligatory. If 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 nasal. The vowel is louder than the consonant, i.e. the intensity of the vowel is greater than that of the consonant, and we believe this is true in any language. Therefore, the intensity of the rhyme is greater than that of the onset.

It is also shown from the previous section that, under unfocused condition, for the onset, there is marginally significant difference between the average intensity values of the stressed and the unstressed syllables, with that of unstressed syllable greater than the stressed one. This is due to the voicing of the consonant in the unstressed syllable. It is shown from former studies that, in the unstressed syllable, the voiceless consonant tends to get voiced. For example, in ‘gege’ (elder brother), whose second syllable is unstressed, the onset of the second syllable is originally unvoiced, but in an unstressed syllable, it gets voiced. When a consonant gets voiced, its intensity will be increased. Therefore, the average intensity of the unstressed syllable is greater than that of the stressed one.

However, under focused condition, there is no significant difference between the average intensity values of the stressed and the unstressed syllables. This is because that the effect of focus is greater on stressed syllable than on unstressed one. Under unfocused condition, the average intensity of the onset of the unstressed syllable is comparatively great. When the word is focused, the intensity of the onset of the stressed syllable increased more, while there is less increase for that of the unstressed one. As a result, there is no significant difference between the intensity of the onsets of the stressed and the unstressed syllables under focused condition.

In regard to the rhyme, however, it is not the case: under focused condition, the average intensity of the rhyme of the stressed syllable is much greater than that of the unstressed one. Lin and Yan [10] analyzed the acoustic feature of unstressed syllables, and found that the prominent feature of the unstressed syllable is: Firstly, the duration is greatly reduced; secondly, the unstressed syllable loses its original tone and its pitch is determined by the preceding stressed syllable; and thirdly, the vowel of the unstressed syllable tends to be centralized. Regarding intensity, they analyzed the maximum intensity values of the unstressed syllables and found that they are reduced in most cases. However, they also found that in some cases the maximum intensity values of the unstressed syllable roughly equal to those of the stressed ones, and in other cases they are not reduced, but increased. Lin and Yan [10] have not investigated the average intensity values of the unstressed syllables. In the present study, the average intensity values are analyzed and it is found that, under focused condition, the average intensity of the rhyme in the stressed syllable is much greater than that in the unstressed one. This is in line with former studies on stress in English and Dutch [12-15].

Under unfocused condition, however, there is no significant difference between the intensity values of the rhyme of the stressed and the unstressed syllables. This is because of the effect of focus on the word following the unstressed syllable. In this study, when the key word is unfocused, it is at the pre-focused position. If the following word is under focused condition, its intensity will increase. The rhyme of unstressed syllable of the key word is the immediate neighbor of the focus, and there is a transition to the increased intensity on it. Therefore, its intensity will increase to some extent. As a result, there is no difference between the intensity values of the rhyme of the stressed and the unstressed syllables.

It is found in this experiment that the effect of focus on intensity is great in most cases. For onset, the effect of focus is significantly great for both stressed and unstressed syllables, and for rhyme, there is significant difference for stressed syllable, with the average intensity under focused condition greater than that under unfocused one. Further analysis shows that, under focused condition, the increase of intensity on the rhyme of the stressed syllable is much greater than of the unstressed one.

In the two syllables of the key word, the effect of focus on the rhyme of the stressed syllable is great. This is because, in a neutral tone word, the rhyme of the stressed syllable is the most prominent part. The energy of the word is mainly carried by this part. When the word is focused, the overall intensity will increase, and the location where intensity increase most will fall on the most prominent part. Therefore, the intensity increase of the rhyme of the stressed syllable will be the greatest. It contributes most to cue focus. In other words, the effect of focus on the intensity of the rhyme of the stressed syllable is the greatest.

The onset is not as loud as the rhyme, but there is still significant effect of focus on it. For a stressed syllable, there is no significant difference between the intensity differences of the onsets of the rhyme. However, for an unstressed syllable, the intensity difference of the onset is significantly greater than that of the rhyme, which is due to the fact that, for unstressed syllable, there is no significant effect of focus on the intensity of the rhyme.

V. CONCLUSION

In this experiment, the intensity pattern as well as the effect of focus and stress on the it for disyllabic words in Chinese is analyzed. It is found that, energy is mainly carried by the rhyme of the syllable, and the intensity of the rhyme is greater than that of the onset. Due to consonant voicing mechanism, for onset, under unfocused condition, the intensity of unstressed syllable is greater than that of the stressed one, while for rhyme, under focused condition, that of stressed syllable is greater than the unstressed one. The effect of focus on the intensity is great in most cases, and the effect of focus on the intensity of the rhyme of the stressed syllable is the greatest.

ACKNOWLEDGEMENTS

This work was supported in part by the Social Science Research Project of Guangdong Province, Grant No. GD11CWW04, as well as the Humanity and Social Science Research Project for Colleges in Guangdong Province, Grand No. 11WYXM012.

REFERENCES

- [1] M. Beckman, *Stress and non-stress accent*, Dordrecht: Foris Publications, 1986.
- [2] D. Dahan and J. Bernard, "Inter-speaker variability in emphatic accent production in French," *Language and Speech*, 39(4), pp. 341–374, 1996.
- [3] Y. Xu and C. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, 33, pp. 159–197, 2005.
- [4] A. Turk and A. White, "Structural influences on accentual lengthening in English," *Journal of Phonetics*, 27(2), pp. 171–206, 1999.
- [5] Y. Chen, "Durational adjustment under corrective focus in Standard Chinese," *Journal of Phonetics*, 34, pp. 176–201, 2006.
- [6] M. Heldner, "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish," *Journal of Phonetics*, 31, pp. 39–62, 2003.
- [7] I. Plag, G. Kunter and M. Schramm, "Acoustic correlates of primary and secondary stress in North American English," *Journal of Phonetics*, 39(3), pp. 362–374, 2011.
- [8] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, 27, pp. 55–105, 1999.
- [9] M. Lin, J. Yan and G. Sun, "A preliminary experiment on normally stressed disyllabic words in Beijing Chinese," *Fangyan*, 1, pp. 57–73, 1984.
- [10] M. Lin and J. Yan, "The acoustic feature of neutral tone in Beijing Chinese", *Fangyan*, 3, pp. 166-178, 1980.
- [11] P. Boersma, "Praat, a system for doing phonetics by computer," *Glott International*, 5:9/10, pp. 341–345, 2001.
- [12] D. Fry, "Duration intensity as physical correlates of linguistic stress", *Journal of the Acoustical Society of America*, 32, pp. 765–769, 1955.
- [13] D. Fry, "Experiments in the perception of stress", *Language and Speech*, 1(2), pp. 126–152, 1958.
- [14] A. M. Sluijter and V. J. van Heuven, "Spectral balance as an acoustic correlate of linguistic stress", *Journal of the Acoustical Society of America*, 100(4), pp. 2471–2485, 1996.
- [15] A. Okobi, *Acoustic correlates of word stress in American English*. Ph.D. thesis, MIT, 2006.