

Research on the Effect of Focus on Neutral Tone Words in Chinese: The Variation of Spectral Emphasis

Maolin Wang

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

ABSTRACT - In this study, the spectral emphasis and the effect of focus on it of neutral tone words in Chinese is analyzed. It is shown that, the spectral emphasis of the onset is always greater than that of the rhyme, and the effect of stress and focus on spectral emphasis is great, so spectral emphasis is a more reliable correlate of stress and focus in Chinese. The rhyme of the stressed syllable is the most prominent part of the neutral tone word, so the increase of spectral emphasis on it under focused condition is comparatively great.

Keywords - Focus, spectral emphasis, onset, rhyme

I. INTRODUCTION

This experiment involves the acoustic signaling of two factors, focus and stress. Focus refers to the important part of a speech utterance which expresses the central part of attention, denoting the specific part which the speaker displays as being significant or which the speaker supposes to be more informative to the listener. Focus can be demonstrated acoustically. For instance, it is generally believed that focus is related to pitch and duration. The acoustic representation of focus can be explained in the following way: Firstly, there is generally a great and sudden ascent in pitch on the focused word or expression [1-3]; secondly, an enlargement in duration of the focused syllables [4, 5]; and thirdly, an overall pitch compression in the post-focus parte either through a low plateau, a late while steady decline or a constant fall until the end of the speech utterance [2, 3]. Besides pitch and duration, it is proved that spectral emphasis is also a safe correlate of focal accent. Heldner [6] argues that, compared with intensity, spectral emphasis is a more reliable correlate, as the effect on it of position in the expression, word stress and vowel height was less pronounced and as it proved a better indicator of focal stress in general and for a majority of the speakers.

Literature shows that there exist several measures that would be believed to be the spectral emphasis category. In their influential work by Sluijter & van Heuven [7], a measure called 'spectral balance' was adopted as the intensity in four connecting frequency bands: 0–0.5, 0.5–1, 1–2 and 2–4 kHz. Other authors have also measured spectral emphasis as the difference between the total intensity and the intensity in a low-pass-filtered signal [8]. One of the methods is to compute the difference (in dB) between the general intensity and the intensity in the signal that was low-pass filtered at 1.5 times the f_0 mean for each speech utterance. The reason behind a filter cut-off frequency at 1.5 times f_0 is to 'segregate' the fundamental frequency from the rest of the frequencies and to get a normalized measurement of the energy in the higher frequency bands [6].

Some research work has been done on the realization of pitch and duration of focus in Chinese. It is displayed that focus patterns are realized as pitch range variations super-imposed on different parts of a speech utterance. The pitch range of the tonal contours rightly under focus is dramatically expanded; the pitch range following the focus is greatly suppressed; and the pitch range preceding the focus does not change much from the neutral-focus condition. Thus, there seem to be three apparent focus-related pitch ranges: swelled on non-final focused words, diminished on post-focus words, and neutral on all the other words. It is also indicated that the on-focus force enlarges the rising slope of the rising tone in Chinese, and research work on focus in both English and Chinese has displayed many similarities between the two languages [3, 9].

In regard to the lengthening of focused constituent, it is indicated that when the word is in the utterance medial position, focus brings about robust lengthening. While when a focused domain is multi-syllabic, the distribution of prolonging is non-uniform: there is an obvious tendency of boundary effect with the final syllable lengthened the most. There is still spill-over lengthening on the adjacent syllables outside the focused constituent. The extent of such lengthening is on condition of prosodic boundaries in that word boundaries weaken lengthening more than syllable boundaries [5].

In Chinese, there are tones, and it is not a stress language, therefore, syllables in most Chinese words are roughly equally stressed, except those with neutral tones. Lin et al. [10] studied the maximum intensity value of disyllabic words in Chinese, and found that the maximum intensity value of the first syllable is in most cases greater than that of the second one. In Chinese, there are some words in which one or more syllables have no tones, i.e., they are of neutral tone, which are termed as neutral tone words, and the basic feature of neutral tone syllable is unstressed. The relationship between focus and stress is an interesting topic in phonetic study. Chinese is a tone language, therefore, the relationship between focus and stress in Chinese becomes an even more interesting topic.

The present study will examine the effect of focus and stress on the basis of spectral emphasis of neutral tone words in Chinese. In particular, it is intended to answer the following questions: What are the patterns of spectral emphasis for neutral tone words under unfocused and focused conditions? What is the influence of focus on spectral emphasis of neutral tone words in Chinese?

II. METHODOLOGY

2.1 Speakers and stimuli

The speakers for this experiment are eight native speakers of Standard Chinese, four male and four female, who participated in the recording. The stimuli are 20 neutral tone disyllabic verbs, in the pattern of 'Onset₁ Rhyme₁ Onset₂ Rhyme₂', such as 'xiahu' (to scare) and 'hunong' (to fool). In Chinese, most of the syllables are constituted of two parts, the onset and the rhyme, except the 'zero-onset' syllables. For instance, in the syllable of 'xia', the onset is 'x' and the rhyme is 'ia'. But in a zero-onset syllable like 'ai', there is no onset in it, only the rhyme 'ai'. In the present experiment, only syllables with both onset and rhyme were used, and the spectral emphasis of onset and rhyme will be examined 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.

The 20 verbs are all neutral tone words, with the second syllable as neutral tone ones, i.e. the stress pattern of the key word is 'stressed + unstressed'. They occur at sentence medial position in the carrier form 'Nana VERB Lili', where 'Nana' and 'Lili' are assumed to be two girls' names. The speech utterances were read at two focus conditions: one focusing at the initial word 'Nana', and the other at the VERB. As a result, there result in two focus conditions for the VERB, unfocused and focused. Foci were triggered by questions. In the first case the question is 'Shui VERB Lili? (Who VERB Lili?)', 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

When recording, the orders of the sentences are randomized. The questions for triggering foci are recorded in advance and played from a loudspeaker, and the speakers are instructed to read the answer after the question was played. Each speaker read the sentences on each focus condition once, producing a total of 320 recorded sentences (8 speakers \times 20 sentences \times 2 focus conditions).

Acoustic data were segmented and labeled after the recording, with boundaries of onsets and rhymes of both the stressed and the unstressed syllables of the key verbs marked, and intensity extracted using Praat [11]. The segmentation was at first done by a segmenting program and then manually corrected. For spectral emphasis, the difference (in dB) between the total intensity and the intensity in a signal that was low-pass filtered at 1.5 times the f_0 mean for each speech utterance was calculated. Computation was done by a self-written visual basic program, by which the mean of the spectral emphasis values of the onset and the rhyme of each syllable of the key word were computed. Statistic analysis was done in the software of SPSS.

III. RESULTS

3.1. Main effect

Fig. 1 displays the spectral emphasis of the onset and the rhyme of both the stressed and the unstressed syllable, under both unfocused and focused conditions. Repeated measures ANOVA result shows that, as far as main effect is concerned, the effects of all the three factors are significant, focus condition: $F(1, 159) = 368.1$, $p < 0.001$; stress condition: $F(1, 159) = 181.7$, $p < 0.001$; onset/rhyme condition: $F(1, 159) = 102.5$, $p < 0.001$, with spectral emphasis comparatively large for focused condition, stressed syllable, and the onset. In the following sub-sections, detailed analysis will be presented about them.

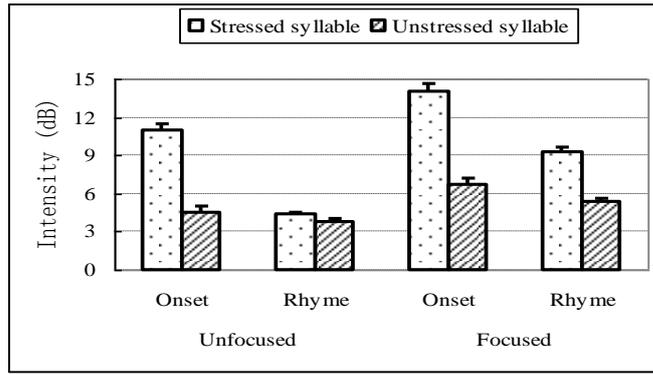


Fig. 1 Spectral emphasis of onset and rhyme at two stress conditions and under two focused conditions

3.2 Onset versus rhyme

It is indicated from repeated measures ANOVA results that there are significant differences between the spectral emphasis of the onset and the rhyme under all of the conditions: unfocused condition, stressed: $F(1, 159) = 171.2, p < 0.001$; unstressed: $F(1, 159) = 4.62, p = 0.033$; focused condition, stressed: $F(1, 159) = 50.3, p < 0.001$; unstressed: $F(1, 159) = 8.55, p = 0.004$, with the spectral emphasis of the onset larger than that of the rhyme.

3.3 Stress

3.3.1 Under unfocused condition

Repeated measures ANOVA result reveals that, under the unfocused condition, there are significant differences between the spectral emphasis of the stressed and unstressed syllables for both the onset and the rhyme. For onset: $F(1, 159) = 120, p < 0.001$; for rhyme: $F(1, 159) = 13.1, p < 0.001$, with the spectral emphasis of the stressed syllable larger than that of the unstressed one.

3.3.2 Under focused condition

It is presented from repeated measures ANOVA result that, under focused condition, for both the onset and the rhyme, spectral emphasis of the stressed syllable is greater than that of the unstressed one, for onset: $F(1, 159) = 101.9, p < 0.001$; for rhyme: $F(1, 159) = 88.8, p < 0.001$.

3.4 Focus

3.4.1 Spectral emphasis

The influence of focus on spectral emphasis is obvious. Repeated measures ANOVA results display that, whether the onset or rhyme, and regardless of the stressed or unstressed syllable, the impact of focus on spectral emphasis is always significant, with that at focused condition much greater than that under unfocused one. For onset, stressed syllable: $F(1, 159) = 124.5, p < 0.001$; unstressed syllable: $F(1, 159) = 68.9, p < 0.001$. For rhyme, stressed syllable: $F(1, 159) = 228.9, p < 0.001$; unstressed syllable: $F(1, 159) = 85.9, p < 0.001$.

3.4.2 Emphasis degree

In the foregoing subsection, it is displayed that the effect of focus on spectral emphasis is obvious. In this subsection, emphasis degree will be examined. Emphasis degree refers to the difference of spectral emphasis between the unfocused condition and the focused condition, as is shown in (1).

$$Dsp = Spe_F - Spe_U \quad (1)$$

In (1), Dsp is emphasis degree, Spe_F stands for spectral emphasis value under focused condition, and Spe_U stands for that under unfocused condition.

Fig 2 shows the emphasis degree for the onset and rhyme for both of the two stress conditions. Repeated measures ANOVA results reveal that, for stress condition, there are significant difference for both the onset and the rhyme, for onset: $F(1, 159) = 6.99, p = 0.009$; for rhyme: $F(1, 159) = 78.8, p < 0.001$, with emphasis degree of stressed syllable greater than that of the unstressed one. As for onset and rhyme, there is interactive effect. For stressed syllable, the emphasis degree of the rhyme is comparatively great: $F(1, 159) = 23.9, p < 0.001$, while for the unstressed syllable, that of the onset is marginally great: $F(1, 159) = 4.56, p = 0.034$.

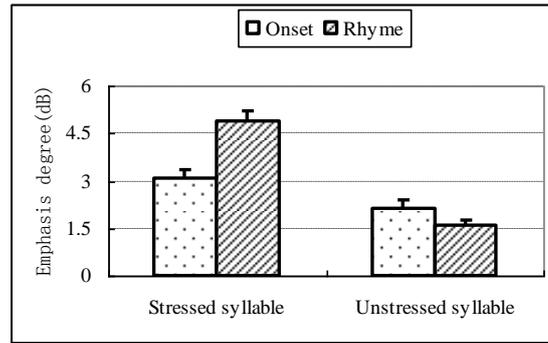
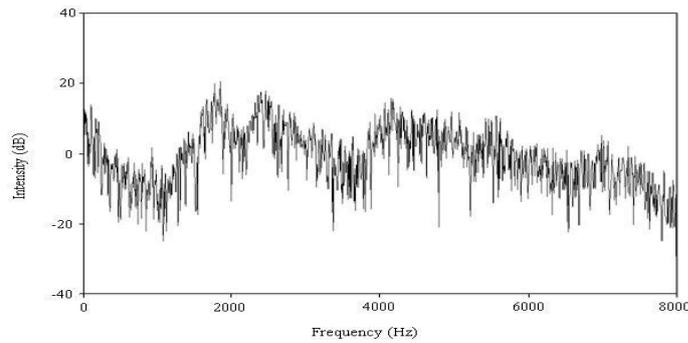


Fig. 2 Emphasis degree for the onset and the rhyme at two stress conditions

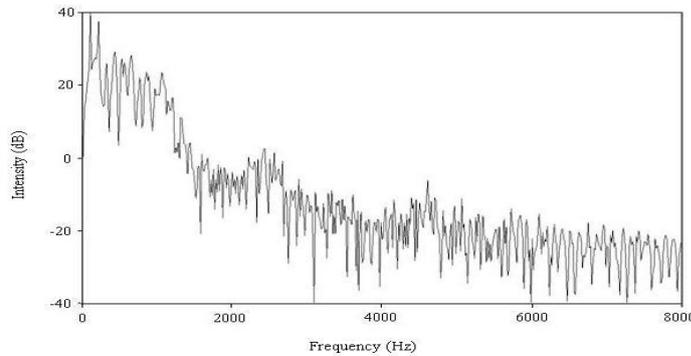
IV. DISCUSSION

Analysis of this experiment indicated that, first of all, the spectral emphasis of the onset of the syllable is much greater than that of the rhyme. It is speculated that the reason for this is as follow: Generally speaking, in most instances, the onset is consonant, while the rhyme is vowel. In Chinese, most of the consonants are voiceless ones, and the vowels are all voiced. For voiced sounds, the energy in the lower frequency bands is high, while for the voiceless sounds, the energy in the higher frequency bands is relatively great. In this study, spectral emphasis is a measurement of the energy in the higher frequency bands of the sound, excluding the fundamental frequency. Therefore, the spectral emphasis of the onset is much greater than that of the rhyme.

Fig. 3 presents the spectrums of the voiceless consonant ‘sh’ (Fig. 3-a) and the voiced vowel ‘ou’ (Fig. 3-b), from which it is shown that, in the lower frequency bands, the energy of the vowel is much great, but in the higher frequency bands, the energy of the vowel reduces to a very low level, while that of the consonant remains at a roughly constant level. This reveals that energy of the higher frequency bands of the voiceless consonant is comparatively great.



(a) The spectrum of consonant ‘sh’



(b) The spectrum of vowel ‘ou’

Fig. 3 The spectrums of (a) consonant ‘sh’ and (b) vowel ‘ou’

Results from the previous section also show that, under both focused and unfocused conditions, and for both of the onset and the rhyme, spectral emphasis of the stressed syllable is greater than that of the unstressed one. Lin and Yan [12] analyzed the acoustic property of unstressed syllables, and found that the major feature of the unstressed syllable is: Firstly, the duration is dramatically reduced; secondly, the unstressed syllables lose their original tones and their pitch is determined by the foregoing stressed syllable; and thirdly, the nuclear vowel of the unstressed syllable is apt to be centralized. Regarding intensity, they investigated 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. From their study, it is shown that maximum intensity is not a reliable cue for stress. In the present study, spectral emphasis is analyzed and it is found that, whether the onset or the rhyme, and regardless of focused or unfocused condition, the spectral emphasis of the stressed syllable is always greater than that in the unstressed one. That is to say, spectral emphasis is a more reliable cue for stress in Chinese.

In this experiment, the effect of focus is studied and it is found that its effect is significant. For both the stressed and the unstressed syllable, and for both the onset and the rhyme, the spectral emphasis under focused condition is always greater than that under unfocused condition. The spectral emphasis of both the onset and the rhyme, of both the stressed and the unstressed syllable, are all increased under focused condition. This result shows that, spectral emphasis is not only a reliable cue for stress, but also a reliable cue for focus.

Emphasis degree for focus is also calculated in this study, and it is found that, for both the onset and the rhyme, emphasis degree of the stressed syllable is greater than that of the unstressed one. We speculate that this is due to the fact that in a neutral tone word, the stressed syllable is the more prominent part, as the spectral emphasis of the stressed syllable is greater than that of the unstressed. No matter in the higher frequency bands or in the lower frequency bands, the energy of the word is mainly carried by the stressed syllable. When the word gets focused, the spectral emphasis will increase, and the location where increase most will fall on the more prominent part. Therefore, the emphasis degree of the stressed syllable will be comparatively great. It contributes more on manifesting focus. In other words, the effect of focus on the spectral emphasis of the stressed syllable is relatively great. As a result, the emphasis degree of the stressed syllable is greater than that of the unstressed one.

As for emphasis degree of the onset and the rhyme, the result is complicated to some extent. For stressed syllable, the emphasis degree of the rhyme is greater than that of the onset. We speculate that the reason for this is as follow. Generally speaking, for neutral tone words, the energy of the stressed syllable is greater than that of the unstressed one. When a word is focused, the emphasis degree of the stressed syllable is greater than that of the unstressed. Emphasis degree refers to the difference of spectral emphasis between the focused condition and the unfocused condition. Under focused condition, the emphasis degree of the rhyme is also greater than that of the onset, as the overall intensity of the rhyme is greater than that of the onset. The rhyme is the more prominent part of the syllable, and it will contribute more on manifesting focus. Because of this, for stressed syllable, the emphasis degree of the rhyme is greater than that of the onset.

However, for the unstressed syllable, the emphasis degree of the onset is comparatively great. This is because of the difference between the unstressed and the stressed syllables. In a neutral tone word, the rhyme of the stressed syllable, whose overall intensity is the greatest, is the important part, and it will be obliged to indicate focus. The rhyme of the unstressed syllable is not as loud as that of the stressed syllable, and will not contribute much in manifesting focus. What is more, most of the onsets in the unstressed syllable will get voiced in the intervocalic position, due to the intervocalic voicing and de-stressing mechanism. When they get voiced, they will obtain some of the properties of the voiced sound, and will have greater increase on spectral emphasis than the voiceless consonant. As a result, in unstressed syllable, the emphasis degree of the onset becomes greater than that of the rhyme.

V. CONCLUSION

The pattern of spectral emphasis, as well as the effect of stress and focus on the spectral emphasis of neutral tone words in Chinese is analyzed in this experiment. It is displayed that, for voiceless consonants, the energy in the higher frequency bands is relatively great, therefore, the spectral emphasis of the onset is larger than that of the rhyme. Spectral emphasis is a reliable correlate of stress in Chinese, as there is always a significant difference between the stressed and the unstressed syllable for spectral emphasis. There is also a significant effect of focus on spectral emphasis. The rhyme of the stressed syllable is the prominent part of the neutral tone word, so there is a great increase of spectral emphasis on it under focused condition.

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] Y. Xu and C. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, 2005, 33, 159–197.
- [3] A. Turk and A. White, "Structural influences on accentual lengthening in English," *Journal of Phonetics*, 1999, 27(2): 171–206.
- [4] Y. Chen, "Durational adjustment under corrective focus in Standard Chinese," *Journal of Phonetics*, 2006, 34: 176–201.
- [5] D. Dahan and J. Bernard, "Inter-speaker variability in emphatic accent production in French," *Language Speech*, 1996, 39(4), 341–374.
- [6] M. Heldner, "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish," *Journal of Phonetics*, 2003, 31: 39–62.
- [7] A. C. Sluijter and V. J. van Heuven, "Spectral balance as an acoustic correlate of linguistic stress," *Journal of the Acoustical Society of America*, 1996, 100(4): 2471–2485.
- [8] H. Traunmuller and A. Eriksson, "Acoustic effects of variation in vocal effort by men, women, and children," *Journal of the Acoustical Society of America*, 2000, 107(6): 3438–3451.
- [9] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, 1999, 27: 55–105.
- [10] M. Lin, J., Yan and G. Sun, "A preliminary experiment on normally stressed disyllabic words in Beijing Chinese," *Fangyan*, 1984 (1): 57–73.
- [11] P. Boersma, "Praat, a system for doing phonetics by computer," *Glott International*, 2001, 5:9/10, pp. 341–345.
- [12] M. Lin and J. Yan, "The acoustic feature of neutral tone in Beijing Chinese", *Fangyan*, 3, pp. 166-178, 1980.