Journal of Computational Science & Engineering 5 (2013) 226-231

Journal of Computational Science & Engineering

Available at http://www.asocse.org ISSN 1710-4068

Spectral Emphasis and the Effect of Focus on it for Disyllabic Words in Chinese

Maolin Wang[a]*, Xuejiao Huang[a], Wei Xiong[a] and Ziyu Xiong[b]

[a] *College of Chinese Language and Culture, Jinan University, Guangzhou, China*
[b] *Institute of Linguistics, Social Academy of Social Sciences, Beijing, China*

Article Information

Article history:
Received 10 February 2013
Revised 17 February 2013
Accepted 20 February 2013
Available online 10 March 2013

Keywords:
Focus
Spectral emphasis
Syllable

Abstract

In this paper, the spectral emphasis and the effect of focus on it for disyllabic words in Chinese is analyzed. It is found that, the spectral emphasis of the onset is greater than that of the rhyme. Due to intervocalic voicing, under unfocused condition, the spectral emphasis of the onset of the first syllable is greater than that of the second. Under focused condition, the emphasis degree of the first syllable is greater than that of the second, and that of the rhyme is greater than that of the onset.

1. Introduction

This study deals with the acoustic realization of focus. Focus refers to some part of an utterance which expresses the centre of attention. It denotes the part which the speaker presents as being important or which the speaker assumes to be more informative for 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, a increase in duration of 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

* Corresponding author: wmljd@126.com

2013 © The American Computational Science Society. All rights reserved.

constant fall until the end of the utterance [2, 3].

Besides pitch and duration, it is shown that spectral emphasis is also a reliable correlate of focal accent. Heldner [6] argues that, compared to intensity, spectral emphasis is more reliable a correlate, as the influence on it of position in the phrase, word accent and vowel height was less pronounced and as it proved a better predictor of focal accents in general and for a majority of the speakers.

There exist several measures that would fall into the spectral emphasis category. In the influential work by Sluijter & van Heuven [7], a measure called 'spectral balance' was defined as the intensity in four contiguous frequency bands: 0–0.5, 0.5–1, 1–2, 2–4 kHz. Some authors have also measured spectral emphasis as the difference between the overall intensity and the intensity in a low-pass-filtered signal [8]. One of the methods is to calculate the difference (in dB) between the overall intensity and the intensity in a signal that was low-pass filtered at 1.5 times the f0 mean for each utterance. The rationale behind a filter cut-off frequency at 1.5 times f0 is to 'separate' the fundamental from the rest of the harmonics and to obtain a normalized measure of the energy in the higher frequency bands [6].

Much research work has been done on the representation 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, 9].

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, except those with neutral tones. Lin et al. [10] analyzed the maximum intensity of disyllabic words in Chinese, and found that in most cases the maximum intensity of the first syllable is greater than that of the second one. They, however, did not compare the intensity of focused and unfocused words.

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

2. 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 'Shanghai' (hurt) and 'Xinshang' (appreciate). In

2013 © The American Computational Science Society. All rights reserved.

Chinese, most of the syllables are composed of two parts, the onset and the rhyme, except the 'zero-onset' syllables. For example, in the syllable of 'shang', the onset is 'sh' and the rhyme is 'ang'. 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 thyme were used, and the spectral emphasis of onset and rhyme will be investigated separately. For the 20 stimuli, the onsets include fricatives like 'x', 'sh', etc, and nasals like 'n', 'm'. 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 20 verbs are normally stressed, with no neutral tones. 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 initial word 'Nana', 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 '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.

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 first and the second syllables of the key words marked, and intensity extracted using Praat [11]. The segmentation was first done by a segmenting program and then manually corrected. For spectral emphasis, the difference (in dB) between the overall intensity and the intensity in a signal that was low-pass filtered at 1.5 times the f0 mean for each utterance was calculated. Analysis was done by a self-written visual basic program, by which the average of the spectral emphasis values within the onset and the rhyme of each syllable of the key word were calculated. Statistic analysis was done in SPSS.

3. Results

3.1 Onset versus rhyme.

It is shown from repeated measures ANOVA results that the main effect of onset versus rhyme is significant, i.e. there are significant difference between the spectral emphasis of onset and rhyme: $F(1, 159) = 110.8$, $p < 0.001$, with the spectral emphasis of the onset much greater than that of the rhyme.

3.2 The first versus the second syllable.

Under unfocused condition. Repeated measures ANOVA result shows that, under unfocused condition, there are significant differences between the spectral emphasis of the first and second syllables for both the onset and the rhyme. But there is interactive effect. For onset, $F (1, 159)=5.95$, $p=0.016$, with the spectral emphasis of the first syllable greater than that of the second. For rhyme, $F (1, 159)=23.6$, $p<0.001$, with the spectral emphasis of the second syllable much greater than that of the first.

Under focused condition. It is shown from repeated measures ANOVA result that, under focused condition, there is no effect of syllable position on the spectral emphasis of the onset: F(1,

2013 © The American Computational Science Society. All rights reserved.

159)=1.65, p=0.2. However, the effect on that of the rhyme is significant: F (1, 159)=7.48, p=0.007, with the spectral emphasis of the first syllable greater than that of the second.

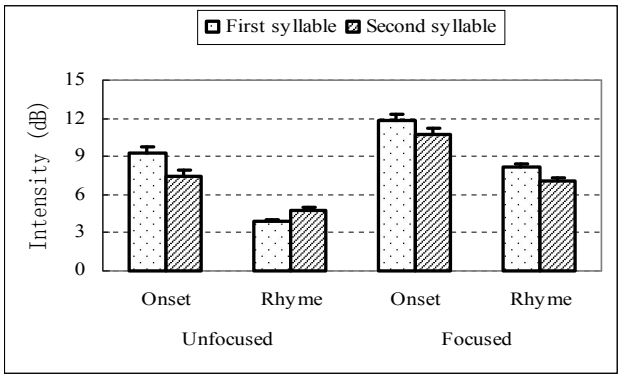


Fig. 1 Spectral emphasis of onset and rhyme at two syllable positions and under two focused conditions.

3.3 Focus.

Spectral emphasis. The effect of focus on spectral emphasis is great. Repeated measures ANOVA results show that, whether the onset or the rhyme, and whether the first or second syllable, the effect of focus on spectral emphasis is always significant, with that under focused condition much greater than that under unfocused one. For onset, first syllable: F(1, 159) 106.2, p<0.001; second syllable: F(1, 159)=120.9, p<0.001. For rhyme, first syllable: F(1, 159)=186, p<0.001; second syllable: F(1, 159)=93.1, p<0.001.

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

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

In (1), *Dsp* stands for emphasis degree, *SpeF* for spectral emphasis value under focused condition, and *SpeU* for that under unfocused condition.

Fig 2 presents the emphasis degree for onset and rhyme at two syllable positions. Repeated measures ANOVA results show that there is no significant main effect between the emphasis degrees of onset and rhyme: F(1, 159) = 2.57, p = 0.111, as there is a there is a significant syllable position × onset/rhyme interaction: F(1, 159) = 32.8, p < 0.001. The effect of syllable position is significant: F(1, 159) = 5.24, p = 0.023, with emphasis degree of the first syllable greater than that of the second. Further analysis shows that, regarding the first syllable, emphasis degree of the rhyme is greater than that of the onset: F(1, 159) = 22.7, p < 0.001.

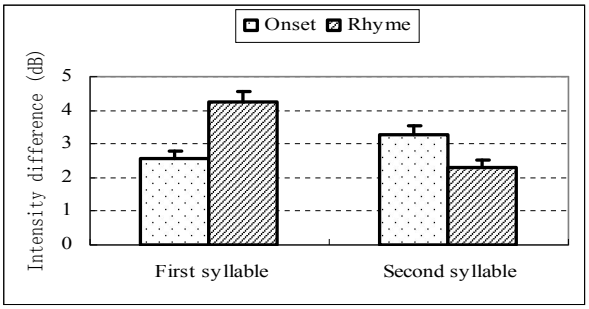


Fig. 2 Emphasis degree for the onset and the rhyme at two syllable positions

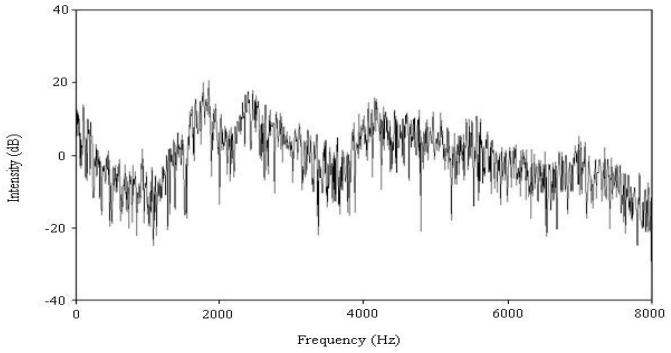
4. Discussion

Results of this experiment showed that, first of all, the spectral emphasis of the onset is much greater than that of the rhyme. We speculate that the reason for this is as follow. Generally speaking, in most cases, the onset is consonant and the rhyme is vowel. In Chinese, most of the consonants are voiceless, and vowels are always voiced. For voiced sounds, the energy in the lower frequency bands is great, but for voiceless sounds, the energy in the higher frequency bands is comparatively great. In this study, spectral emphasis is a measure of the energy in the higher frequency bands, excluding the fundamental. Therefore, the spectral

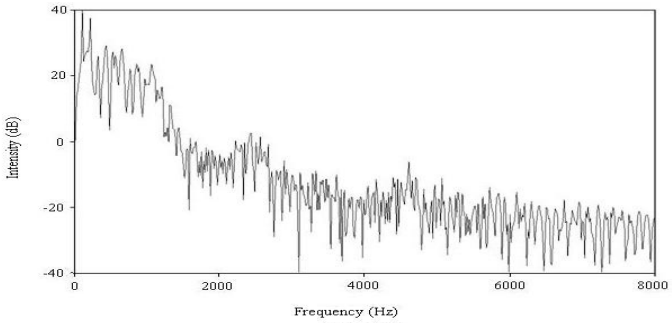
2013 © The American Computational Science Society. All rights reserved.

emphasis of the onset is greater than that of the rhyme.

Figure 3 shows the spectrums of consonant 'sh' (Figure 3-a) and vowel 'ou' (Figure 3-b), from which it can be seen that, in the lower frequency bands, the energy of the vowel is much great, but in the higher bands, the energy of the vowel drops to a very low level, while the that of the consonant remains at a medium level. As a result, energy of the higher frequency bands of the consonant is great.



(a) The spectrum of vowel of 'sh'



(b) The spectrum of vowel 'ou'

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

It is also shown from the previous section that, under unfocused condition, as far as the onset is concerned, the spectral emphasis of the first syllable is greater than that of the second. We suppose that the reason for smaller spectral emphasis of onset in the second syllable is intervocalic voicing. A consonant occurring at intervocalic position tends to become voiced. As is mentioned above, for voiced sounds, the energy in the lower frequency bands is great, but that in the higher frequency bands is comparatively small. When the consonant in the second syllable gets voiced, the energy in the higher bands gets reduced, that is, the spectral emphasis gets reduced. Therefore, the spectral emphasis of onset in the first syllable is greater than that of the second.

However, for the rhyme, it is just the opposite. For rhyme, the spectral emphasis of the second syllable is greater than that of the first. Generally speaking, the energy of the vowel is great and that of the consonant is small. However, for spectral emphasis, it is the opposite. The spectral emphasis of the consonant is great and that of the vowel is small. For disyllabic words, the energy of the rhyme in the first syllable is greater than that in the second syllable. Similar to the case of consonant and vowel, for rhyme, the spectral emphasis of the second syllable is greater than that of the first.

The effect of focus is studied in this experiment, and it is found that the effect is significant. For the rhyme, when the key word is under unfocused condition, the spectral emphasis of the second syllable is greater than that of the first. However, under focused condition, it is just the opposite. The spectral emphasis of the first syllable is greater than that of the second. We speculate that the reason for this is as follow. Generally speaking, for disyllabic words, the energy of the first syllable is greater than that of the second. When the word is focused, the emphasis degree of the first syllable is greater than that of the second. Emphasis degree refers to the difference of spectral emphasis between the focused condition and the unfocused condition. What is more, 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, and the rhyme contributes more on manifesting focus. Under these dual effects, under focused condition, the spectral emphasis of the first syllable gets

2013 © The American Computational Science Society. All rights reserved.

greater than that of the second.

Coming to the onset, when the key word is under unfocused condition, the spectral emphasis of the first syllable is greater than that of the second. However, under focused condition, there is no effect of syllable position on the spectral emphasis. When the key word is under focused condition, the spectral emphasis of both of voiced and voiceless sounds will increase. Comparatively, the voiced sounds will have greater increase than the voiceless sounds, as the voiced sounds contribute more on manifesting focus. As is mentioned above, some of the onsets in the second syllable will get voiced in the intervocalic position. When they get voiced, they will have greater increase on spectral emphasis than the onset in the first syllable, and as a result, the difference between them disappears. Therefore, there becomes no effect of syllable position on the spectral emphasis.

In this study, emphasis degree for focus is calculated, and it is found that emphasis degree of the first syllable is greater than that of the second. It has been found that in disyllabic word, the energy of the first syllable is greater than the second. When the word is focused, the first syllable will have greater increase on spectral emphasis than the second syllable, as it contributes more on manifesting focus. Therefore, the emphasis degree of the first syllable is greater than that of the second.

It is also found that for the first syllable, the emphasis degree of the rhyme is greater than that of the onset. The reason for this is similar to that mentioned above. The energy of the rhyme is greater than that of the onset. Under focused condition, the emphasis degree of the rhyme is comparatively great, as it contributes more on manifesting focus. Therefore, the emphasis degree of the rhyme is greater than that of the onset.

5. Conclusion

In this experiment, the pattern of spectral emphasis, as well as the effect of focus on disyllabic words in Chinese is analyzed. It is found that, for voiceless sounds, the energy in the higher frequency bands is comparatively great, so the spectral emphasis of the onset is greater than that of the rhyme. Due to intervocalic voicing, under unfocused condition, the spectral emphasis of the onset of the first syllable is greater than that of the second. However, for the rhyme, the spectral emphasis of the second syllable is greater than that of the first. Under focused condition, there is no effect of syllable position on the spectral emphasis of the onset. For the rhyme, the spectral emphasis of the first syllable is greater than that of the second. As the first syllable and the rhyme contribute more on manifesting focus, the emphasis degree of the first syllable is greater than that of the second, and that of the rhyme is greater than that of the onset.

References

[1] M. Beckman, “Stress and non-stress accent,” Dordrecht: Foris Publications, 1986.

[2] Y. Xu, C. Xu, Journal of Phonetics, 33 (2005) 159-197.

[3] A. Turk, A. White, Journal of Phonetics, 2 (1999) 171-206.

[4] Y. Chen, Journal of Phonetics, 34 (2006) 176-201.

[5] D. Dahan, J. Bernard, Language Speech, 4 (1996) 341-374.

[6] M. Heldner, Journal of Phonetics, 31 (2003) 39-62.

[7] A.C. Sluijter, V.J. van Heuven, Journal of the Acoustical Society of America, 4 (1996) 2471-2485.

[8] H. Traunmuller, A. Eriksson, Journal of the Acoustical Society of America, 6 (2000) 3438-3451.

[9] Y. Xu, Journal of Phonetics, 27 (1999) 55-105.

[10] M. Lin, J. Yan, G. Sun, Fangyan, 1 (1984) 57-73.

[11] P. Boersma, Glot International, 9 (2001) 341-345.

2013 © The American Computational Science Society. All rights reserved.