

PROSODIC VARIATION OF FOCUSED SYLLABLES OF DISYLLABIC WORD IN MANDARIN CHINESE

Zhenglai GU*, Hiroki MORI†, and Hideki KASUYA†

*Graduate School of Engineering, Utsunomiya University

†Faculty of Engineering, Utsunomiya University

ABSTRACT

This paper describes a study on how disyllabic on-focus words under a certain context condition affects the variations of prosodic correlates of syllables in terms of different focal positions in Mandarin Chinese. The study shows that: (1) there is no direct link between semantic structure of disyllabic on-focus words and the prosodic correlates, but an indirect link through the mediation of focus positions; (2) the prosodic correlates of the second syllable are more variable than those of the first syllable. Specifically, there is a significant asymmetry of vowel duration and fundamental frequency (F0) range between the pre-focus and post-focus syllables.

1. INTRODUCTION

In Chinese, words are composed of characters that are monosyllabic. Unlike the Roman alphabet, which gives relatively precise pronunciation information, but less information on meaning, Chinese monosyllabic characters, however, give less precise pronunciation information, but do give information on meaning. These monosyllabic characters are put together to create disyllabic characters. Thus the meaning of most of these disyllabic words is derived from the characters of the words, e.g. “gang1 cai2 (steel material)”, “tou2 tong4 (head-ache)”, etc. If a disyllabic word is to be focused on, one from three types of focus positions is likely to be selected by Chinese speakers. These focus positions are defined as follows: the first syllable (Fp1), second syllable (Fp2) and both syllables (Fp12). The former two are called narrow focus and the latter is called broad focus [1]. Recent reports have revealed: (1) There is a radical asymmetry of F0 range around the focused word [2,3,4], and (2) The duration of on-focus words remarkably increases, while the duration of the neighboring syllables slightly changes [4], which we call symmetry of duration. It is obvious, however, that these reports simply take a look at the overall prosodic characteristics of the target word under a broad focus condition. The variations, which can occur under narrow focus conditions, are little concerned. In addition, there is doubt that the scope of the focus effect might depend on the tonal combinations [5].

The present study therefore serves as a starting point for overcoming the deficiency of the previous studies by providing insight into interaction between prosodic correlates of the syllable in terms of different focus positions. In particular, the study tries to answer the following two questions: (1) To what extent do the F0 contours of the words with different semantic structure vary while under the same focus position and tonal context conditions (experiment 1)? (2) How do these three focus positions affect the prosodic correlates (experiment 2)?

2. CLASSIFICATION OF WORDS

We classify disyllabic words of Chinese into the following four categories:

Category 1 (C1): words in which the first syllable's contribution to the meaning of the whole word is more important.

Category 2 (C2): words in which the second syllable's contribution to the meaning of the whole word is more important.

Category 3 (C3): words in which each syllable's contribution to the meaning of the whole word is the same. C3 is further classified into three sub-categories, C3₁, C3₂ and C3₃ according to their semantic structures: no contribution, the same contribution by two antonyms and the same contribution by two synonyms, respectively.

Category 4 (C4): words, which can belong to C1, C2 or C3 depending upon the contextual conditions.

In a study, we have found that Chinese speakers commonly retain focus positions consistently on the sense-contributing syllables of the words (Table 1).

Table 1 indicates that the character meaning of the word plays an important role in the assignment of the focus positions. The issue is therefore focused on whether or not the character meaning directly affects the prosodic correlates.

Table 1: Focus positions for each category.

Cat.	Sample	Sense	Focus Positions		
			Fp1	Fp2	Fp12
C1	Zhuan1Yi1	Concentrated	1	0	1
C2	Yi2Ding4	Must	0	1	1
C3 ₁	Sha1Fa1	Sofa	0	0	1
C3 ₂	Dong1Xi1	East west	0	0	1
C3 ₃	Gang1Cai2	Just	1	1	1
C4	Gang1Cai2	Steel material	1	1	1

Note that 0 means focus on the specific position is unnatural.

3. EXPERIMENT 1

3.1 Materials and analysis procedure

Five pairs of disyllabic words from different categories were selected as target words (Table 2). Each pair of words has the same tone combination and the same or similar phonemic structures. For control of the narrow focus position, we used the carrier sentence, “Zhe4 shi4 _ _ er2 bu2 shi4 _ _ . (This is _ _ but not _ _)”, into which target and contrastive words were inserted. For example, given the word ‘steel material’ (gang1cai2), we have the following derivatives.

- “This is steel material but not wood.”
- “This is steel material but not a steel product.”
- “This is steel material but not a wood product.”
- “This is steel material but not anything else.”

Three sentences with different focus conditions and a neutral sentence, listed above are prepared for the word ‘steel material’ . Three native Chinese speakers born in Beijing, two males and one female, were asked to read one sentence two times. All of the utterances were recorded on the DAT in a soundproof room. Recorded utterances were then digitized at a sampling rate of 11.025 kHz. The syllable and vowel segmentation were carried out manually with waveform and sound spectrogram. Fundamental frequency (F0) contours were automatically analyzed and valued with a software on a PC. The F0 minimum and F0 maximum within the vowel duration were recorded and the F0 range was defined as the difference between the two values.

3.2 Results

The relative variations of vowel duration and F0 range between the focused and neutral utterances were analyzed in two-factors ANOVA, with Word A and Word B as a between-items factor, and focus positions as a within-items factor. In addition, one-factor ANOVA was used for each focus position.

The one-factor ANOVA analysis result of speaker MZJ in Table 3 shows that there is no effect in words from different categories on the variation of either vowel duration ($F < 1$) or F0 range ($F < 1$) of speaker MZJ. In addition, the analysis results by two-factors ANOVA are that: (1) There is a significant effect of focus positions both in the variation of vowel duration ($p < 0.05$) and in the F0 range ($p < 0.05$); (2) There is no interaction effect ($F < 1$).

Table 3 Means of the variations of vowel duration and the F0 range of the Speaker MZJ.

F_P	Syllable 1				Syllable 2			
	W_A	W_B	F	p	W_A	W_B	F	p
Ratio of vowel duration of focus to neutral								
Fp12	1.66	1.67	0.003	0.958	1.25	1.29	0.417	0.536
Fp1	1.79	1.76	0.023	0.882	0.94	0.95	0.032	0.862
Fp2	1.11	1.19	0.146	0.711	1.32	1.41	0.295	0.601
Difference of F0 Range of focus to neutral (semitone)								
Fp12	2.27	1.97	0.148	0.709	4.24	3.91	0.025	0.877
Fp1	2.21	2.53	0.049	0.831	1.58	1.81	0.036	0.854
Fp2	1.03	1.05	0.001	0.974	5.42	4.79	0.094	0.766

Table 2: Target words for experiment 1.

Word pairs (A, B)	Sense	Cat.
Gang1guo3 Gan1guo3	Congo Dry fruit	C3 ₁ C4
Ten2 tong4 Tou2 tong4	Pain Head-ache	C3 ₃ C4
Yi2 ding4 Yi2 ding4	Must A tablet	C2 C4
Gang1 cai2 Gang1 cai2	Just Steel material	C3 ₃ C4
Zhuan1yi1 Zhong1 yi1	Concentrated Traditional Doctor	C1 C4

3.3 Discussion

The results show that the meaning of the character in Chinese words has a direct relationship to the focus positions, which affect the prosodic correlates significantly, but it is also an independent factor to the variations of prosodic correlates (Figure 1), i.e. there is no reason to assume that prosodic correlates of words with different semantic structure differ systematically under the same tonal context and focus conditions. This serves as an answer to the first major question mentioned above, and a foundation for the further investigation in experiment 2.

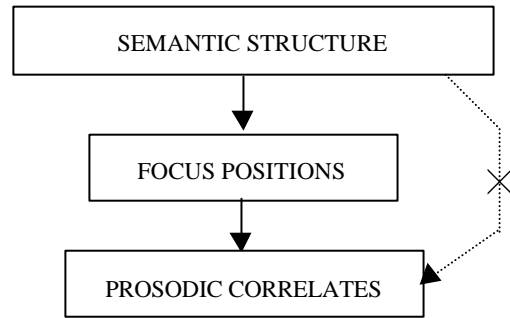


Figure 1: Relationship between semantic structure of words and prosodic correlates.

4. EXPERIMENT 2

4.1 Materials and analysis procedure

Most of the process is the same as experiment 1. The difference is that disyllabic words have all of the 15 tonal (Chinese 4 tones: H, R, L, F) combinations, except for tones L-L¹. Two words were selected for each tone combination (See Appendix for a full list of all 30 target words). Thus 720 (=2×15×4×3×2) utterances were recorded and the prosodic variation was examined syllable-by-syllable through analysis of vowel duration, F0 maximum, F0 minimum, and F0 range.

4.2 Results

Vowel Duration

Figure 2 illustrates the distribution of vowel duration (r_1, r_2), where r_1 and r_2 are the ratios of narrow focus to neutral utterance vowel duration, in syllable 1 and syllable 2, respectively. The mean values of the ratios are listed in Table 4. As can be seen in Figure 2, there is a significant asymmetry of vowel duration, because firstly, most points of Fp1 lie in the left region of the ordinate, which indicates that the vowel duration of syllable 2 is shortened in the case of the narrow focus on the syllable 1; secondly, most points of Fp2 lie above the abscissa, which indicates that the vowel duration of syllable 1 is lengthened in the case of the narrow focus on syllable 2. Table 4 shows that the vowel duration of the second syllable is reduced significantly to an average of 84% in Fp1 and increased remarkably to an average of 154% in Fp2, whereas the duration of the first syllable is not reduced but increased to an average of 111% and 143% in Fp2 and Fp1, respectively. In addition, the difference between the mean post-focus duration and the mean pre-focus duration is significant, $t(169)=12.02, p<0.05$, indicating that narrow focus on the syllable of disyllabic words not only strengthens the length of the on-focus syllable but also affects the length of the neighboring syllable systematically. The other substantial difference is that the variation of vowel duration in a narrow focus condition is greater than in a broad focus condition.

F0 Range

In Table 5, we find that F0 range is increased more in the narrow focus condition than on the broad focus condition. The differences are significant, $t(178)=2.7, p<0.05$ in syllable 1, and $t(171)=3.67, p<0.05$ in syllable 2. The difference between the two narrow focus Fp1 and Fp2 is also significant, $t(164)=2.08, p<0.05$.

Figure 3 illustrates the distribution of F0 range (r_1, r_2), where r_1 and r_2 are the differences of narrow focus to neutral utterance F0 range, in syllable 1 and syllable 2, respectively.

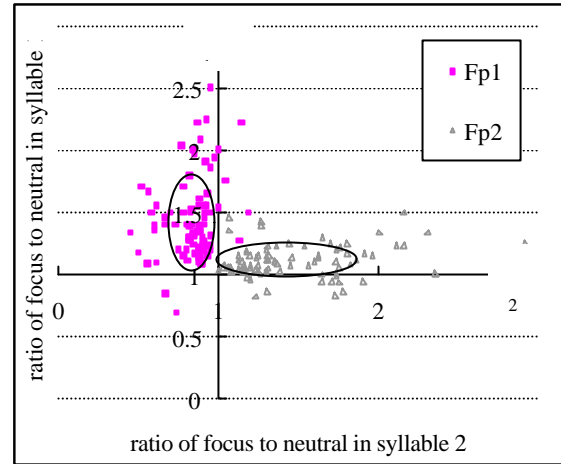


Figure 2: Distribution of vowel duration (r_1, r_2), where r_1 and r_2 are the ratios of narrow focus to neutral utterance vowel duration, in syllable 1 and syllable 2, respectively.

Table 4: Mean variation of vowel duration in syllable 1 and 2.

F_P	Syllable 1	Syllable 2
Fp12	1.36	1.38
Fp1	1.43	0.84
Fp2	1.11	1.54

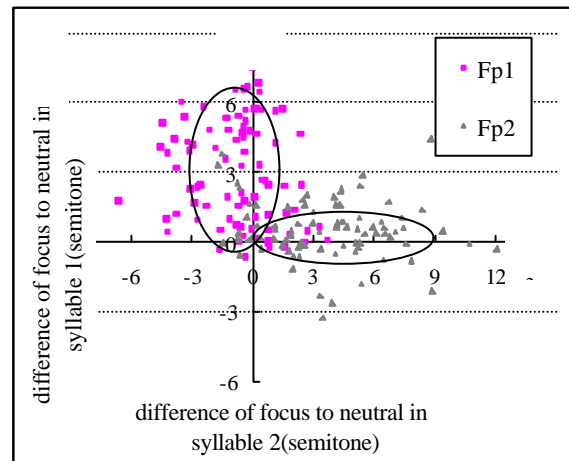


Figure 3: Distribution of F0 range (r_1, r_2), where r_1 and r_2 are the differences of narrow focus to neutral utterance F0 range, in syllable 1 and syllable 2, respectively.

¹ The L tone will be changed into a R tone when it precedes another L tone.

Table 5: Mean F0 range variation in syllable 1 and 2.

F_P	Syllable 1(semitone)	Syllable 2(semitone)
Fp12	2.04	2.23
Fp1	2.93	-0.66
Fp2	0.50	3.74

As can be seen, firstly, most points of Fp1 lie in the left region of the ordinate, which indicates that the F0 range of syllable 2 is compressed in the case of the narrow focus on syllable 1; secondly, only a few points of Fp2 lie below the abscissa, which indicates that the pre-focus range of syllable 1 is not likely to be compressed while syllable 2 is in a narrow focus. It seems that these overall results are consistent with the word-level asymmetry of F0 range, but there seems to be local fluctuation of carryover and anticipatory effects [6] caused by tonal combinations and syllable positions.

4.3 Discussion

The substantial local fluctuation phenomenon may indicate that the carryover and anticipatory effect is significant between the two syllables. For example, when the first syllable is focused on with a high target, such as tone H or tone R, the following tone is always shifted upward while the F0 maximum is increased much larger than the F0 minimum. As a result, the F0 range of a post-focus syllable is not compressed. It might be the reason why some square points of Fp1 lie in the right region of the ordinate.

5. SUMMARY

The results from our two experiments reveal that (a) there is no direct link between the semantic structure of two-syllable on-focus words and the prosodic correlates of the different words, but an indirect link through the mediation of focus positions, (b) the prosodic correlates of the second syllable are more variable than those of the first syllable. In more detail, the vowel duration and the F0 range of the second syllable are not only increased but also decreased more than those of the first syllable. These findings imply that different strategies are employed in the task of narrow focusing in the first and the second syllable, i.e., emphasizing the first syllable as well as weakening the second syllable for the former, and stressing the second syllable only for the latter.

6. REFERENCE

1. Ladd, D.R., "Intonational phonology", Cambridge: CUP, pp.160–204, 1996.
2. Zhang, J.L., "Acoustic Parameters and Phonological Rules of a Text-to-Speech System for Chinese", *IEEE ICASSP*, pp.2023–2026, 1986.

3. Garding, E., "Speech Act and Tonal Pattern in Standard Chinese Constancy and Variation", *Phonetica*, 44, pp.13–29, 1987.
4. Xu, Y., "Effects of tone and focus on the formation and alignment of f_0 contours", *Journal of Phonetics*, 27, pp.55–105, 1999.
5. Shih, C., and Sproat, R., "Issues in Text-to-Speech Conversion for Mandarin", *Computational linguistic and Chinese Language Processing*, 1(1), pp.37–86, 1996.
6. Xu, Y., "Contextual tonal variations in Mandarin", *Journal of Phonetics*, 25, pp.61–83. 1997.

APPENDIX

Target words for experiment 2

Target word	Sense	Tones
Zhong1Xing1	Center	H-H
Zhong1Xing1	Cordial	H-H
Zhang1Jie2	Chapters and sections	H-M
Zheng1Jie2	Chastity	H-M
Zhong1Gu3	Bell tower	H-L
Zhong1Gu3	Medieval times	H-L
Zhong1Ri4	China and Japan	H-F
Zhong1Ri4	All day	H-F
Nan2Fang1	South area	M-H
Nan2Fang1	Husband's side	M-H
Chang2Chang2	Often	M-M
Chang2Qing2	Reason	M-M
Tou2Nao3	Brains	M-L
Hou2Nao3	Monkey head	M-L
Hua2Gui4	Luxurious	M-F
Hua2Gui4	Slide cabinet	M-F
Shi3Zhong1	All along	L-H
Shi3Guan1	Official historian	L-H
Ta3Tai2	Collapse	L-M
Ta3Tai2	Control tower	L-M
Ma3Ke4	Mark	L-F
Ma3Ku4	Riding breeches	L-F
She4Shi1	Missing a shot	F-H
She4Shi1	Facilities	F-H
Dong4Liang2	Ridgepole and beam	F-M
Dong4Ling2	Icicle	F-M
Kan4Fa3	View	F-L
Kan4Hao3	Take a fancy to	F-L
Hou4Bei4	Posterity	F-F
Hou4Bei4	Reserve	F-F