INTONATIONAL PROMINENCE OF “SHI...(DE)” CONSTRUCTION IN STANDARD CHINESE

Yuan Jia¹, Aijun Li², Ziyu Xiong³

¹ English Department, Nankai University

^{2,3} Phonetics Lab, Institute of Linguistics, Chinese Academy of Social Sciences

ABSTRACT

The present study mainly deals with the phonetic realization of the intonational prominence in the *shi...(de)* (是...(的)) construction in Standard Chinese. Results of acoustic and perceptual experiments demonstrate that the prominence placement bears corresponding relationship with the focused constituents marked by *shi...(de)* structure, specifically, the appearance of intonational prominence is symbolized by the focus marker *shi*. The phonetic realization of the intonational prominence lies in the expansion of the pitch range of the focus-bearing constituent and the compression of the pitch registers of the successive syllables.

Index Terms: *shi...(de)* construction, intonational prominence, focus marker, focus-bearing constituent

1. INTRODUCTION

In the Chinese literature, the *shi...(de)* structure is usually adopted as a device to reverse the normal end-focus pattern in a comment to maintain the basic syntactic structure (Wu[1]). “*shi*” is the closest equivalent of English copula “*be*”, and “*de*” is a particle with various functions: a modification marker (Ross [2]), a nominalizer (Chao [3], Li & Thompson [4]), and a past-tense marker (Song [5]), etc.

Previous studies of *shi...(de)* construction in Chinese concentrate on grammatical aspects, e.g., the syntactic functions and the focus marking function of this structure. Syntactically, Yuan [6] proposes that the shift of *de* preceding the object can reduce the grammatical focus scope. For the focus marking function, Fang [7] highlights that the *shi...(de)* construction in Chinese is applied to mark contrastive focus. The focused constituents are immediately following the focus marker *shi* and bearing the phonetic prominence.

In addition to the grammatical approaches to *shi...(de)* construction, in phonetic aspect, most of the researchers confine their attentions to monosyllabic, disyllabic, or quadri-syllabic focus constituents when dealing with the phonetic realization of Chinese focused parts. An example can be drawn from Xu [8] who adopts monosyllables and disyllables as the basis to observe the pitch range variance. Results demonstrate that the pitch range gets obvious expansion in the under-focus and compression in the following focus position. Furthermore, Jia [9] provides a detailed exploration of the underlying causes for the pitch range variances in three focus positions, e.g., under-focus, pre-focus, and post-focus, through the investigation of the

focus constituents with exhaustively various tonal combinations. It points out that the variation of pitch range is restricted by the “H” tone of the focused parts. Chen [10] observes the durational distribution of quadri-syllabic focus constituents and argues that the durational distribution is influenced by both prosodic structure and edge effect in Standard Chinese.

It can be obtained from the above analysis that, grammatical analysis of *shi...(de)* construction of Standard Chinese did not provide an explicit description of the phonetic realization of intonational prominence in this structure; whereas, the phonetic analysis on the focused constituents did not take the syntactic marker *shi* into consideration. In this regard, the study of the phonetic realization of intonational prominence in the *shi...(de)* construction of Standard Chinese is important, firstly, it can describe the specific “manner” and “position” of the intonational prominence realization in this structure; secondly, the relationship between the “focus scope” and “prominence placement” can be further discussed based on the experimental results; thirdly, the application of the results obtained from the experiments can help to improve the naturalness of speech synthesis and prosodic modeling.

2. METHODOLOGY

The acoustic experiment, perceptual experiment and statistical analysis are adopted in this study through which we explore the specific phonetic realization of the “manner” and “position” of the intonational prominence in the structure of *shi...(de)*, specifically, how it is phonetically realized and where it is located.

2.1. Acoustic experiment

2.1.1. Materials

Based on the objective of the present study, factors to be considered in the experiment are: the syntactic component of *shi*, the number of focus marker, and the variance of the grammatical focus position. Therefore, this study selects the following four sentences as the target samples for both acoustic and perceptual investigations:

(1) Xiao3 Wang2 Shi4 Qū4 Nian2 Chang2 Chang2 Qū4 Tai2 Wan1. (小王是去年常常去台湾。)

Xiao Wang is last year usually go Taiwan.
(It was last year that Xiao Wang usually went to Taiwan).

(2) Wo3 Shi4 Zuo2 Tian1 Cai2 Ting1 Shuo1 Shi4 Lao3 Wang2 Tui4 Xiul Le0. (我是昨天才听说是老王退休了。)

I is yesterday just hear is
Lao Wang retired already

(It was yesterday that I just heard the retirement of Lao Wang).

(3) Shi4 Tian1 Tian1 Jiao4 Tai1 Qü4 Bei3 Jing1 De0.

is tian tian ask he go Beijing

(It was tiantian who asked him to go to Beijing). (是天天叫他去北京的。)

(4) Shi4 Xiao3 Wang2 Zhi1 Dao4 Shi4 Xiao3 Liu2 Da3 Sui4 De0 Na4 Ge0 Bei1 Zi0.

is Xiao Wang know shi Xiao Liu break down that cup

(It was Xiao Wang who knows that it was Xiao Liu who broken down that cup). (是小王知道是小刘打碎的那个杯子。)

2.1.2. Recording procedure

Eleven Standard Chinese speakers, six female and five males, aged within 20-45, were invited as the subjects. The informants all work in the Linguistic Institute, Chinese Academy of Social Sciences. The recording was conducted in the sound-treated booth in the Institute of Linguistics, CASS. During the recording procedure, each sentence appeared on the screen in totally random order. The subject was asked to read aloud the displayed sentence in normal speed without any irregular pause. In case of any mistake, the subject was instructed to repeat the sentence. All these sentences were recorded and saved directly into computer through sound recording software “CASSRecorder” as “wav” files.

2.1.3. Data labeling and extraction

The data for the acoustic study is obtained via the following steps: 1) the speech data was first labeled by an automatic segmentation software for the boundary of each syllable in the target sentence, and then the syllable boundaries were modified by hand; 2) three scales of prosodic boundaries were then labeled, “2” stands for prosodic word, “3” means intermediate phrase and “4” indicates intonational phrase; 3) the “pitchtier” file for each target sentence was modified automatically by praat script; 4) the extraction of F_0 data was based on the pitchtier files with each syllable in the target sentence being selected ten points; 5) SPSS was adopted to get the F_0 means from eleven speakers; 6) One-Way ANOVA was conducted to test the differences of the maximum pitch values of each prosodic word in the target sentence.

2.2. Perceptual experiment

The goal of the perceptual experiment is to further test the judgment of the “number” and “position” of the intonational prominence from the subjects. Sentence recorded in the acoustic experiment were all selected as the perceptual experiment samples. In total, three females and four males were invited to participate in the perceptual experiment. They are all standard Chinese speakers and show sensitive perception. Each subject was asked to finish the experiment individually, without any interpersonal consultation. During the course of the experiment, the subject was expected to choose the position(s) of the intonational prominence, the numbers and the units (i.e. syllable, prosodic word, or prosodic phrase) were not restricted. The perceptual results were finally obtained from all the data being summed up.

3. PHONETIC REALIZATION OF INTONATIONAL PROMINENCE IN SHI...(DE) STRUCTURE

3.1. Intonational prominence marking by shi

In this part, we intend to investigate the overall pitch performance of the target sentence with *shi* as the focus marker. The following Figure 1 is the F_0 means of the target sentence “xiao3 wang2 shi4 qü4 nian2 chang2 chang2 qü4 tai2 wan1”. The abscissa axis describes the content of each syllable in the sentence while the ordinate axis illustrates the F_0 range with the “Unit” as “Hz”(The coordinates in Figure 3, 5, 7 displays same content).

Figure 1: F_0 means of “xiao3 wang2 shi4 qü4 nian2 chang2 chang2 qü4 tai2 wan1”.

Based on the analysis of Fang [7], the constituent immediately following the marker *shi* is the focus-bearing unit, which can function as subject, adverbial, and predicate, etc. Therefore, the focused element in this sentence dwells on the word “qü4 nian2” and the result of prosodic boundary labeling also tells that there exists “3” boundary after the word “qü4 nian2” which implies the existence of the intermediate phrase boundary after the marked focus. Examination of the above figure shows that the pitch range of “qü4 nian2” observes great enlargement and the registers of the following syllables are compressed successively. The compression phenomenon can spread to the very end of the target clause. In contrast with the expanding and lowering of the pitch range variation, the focus marker *shi* in this sentence achieves its full pitch contour, however, it will not be taken as the prominence bearing unit in this study. This result is obtained according to Xu [8] that the focus only exerts effect on the pitch in the under and post focus position, while leaving the pre-focus pitch much the same. Thereafter, we treat the focus-maker *shi* as the pre-focus constituent following Xu [8] and it deserves no obvious pitch range changes. Based on the acoustic analysis, the study further conducts One-Way ANOVA to test the differences of the maximum pitch values among each word¹ in the sentence. Results of the Bonfironi post hoc test demonstrate that the maximum pitch value of “qu4nian2” is significantly different from the words “xiao3wang2”(P=0.014), “chang2chang2”(P=0.002), “qu4”(P=0.002) and “tai2wan1”(P=0.002), whereas, it shows no obvious difference with “shi4”(P=0.069).

The Figure 2 beneath is the illustration of the perceptual results of the distribution of the intonational prominence of the sentence “xiao3 wang2 shi4 qü4 nian2 chang2 chang2 qü4 tai2 wan1”. The abscissa axis also shows the content of the words or syllables which were chosen as prominent elements, whereas the ordinate axis demonstrates the number of the perceptual results (The coordinates in Figure 4, 6, 8 display same content).

¹ The classification of the word is based on the boundary of prosodic word which is labeled as “1”.

Figure 2: Perceptual results of “xiao3 wang2 shi4 qü4 nian2 chang2 chang2 qü4 tai2 wan1”.

It can be seen from the results that there are totally five choices for the prominent constituents, namely, “chang2 chang2”, “qü4 nian2”, “shi4”, “tai2 wan1”, and “xiao3 wang2”, among these results, the word “qü4 nian2” deserves the greatest numbers of choices. This result is in accordance with the F_0 realization in the way that the focus-bearing unit observes the intonation prominence. The focus position marked by *shi* is the distributing place of the prominence.

The following Figure 3 is the F_0 means of “wo3 shi4 zuo2 tian1 cai2 ting1 shuo1 shi4 lao3 wang2 tui4 xiu1 le0”. The sentence is applied to investigate the corresponding relationship between the number of the focus marker *shi* and the number of intonational prominence, furthermore, the different levels of intonational prominence can also be examined within this sentence.

Figure 3: F_0 means of “wo3 shi4 zuo2 tian1 cai2 ting1 shuo1 shi4 lao3 wang2 tui4 xiu1 le0”.

From the concept of focus-mark *shi* proposed by Fang [7], in this sentence, it contains two focused constituents, concretely, “zuo2 tian1” and “lao3 wang2”. Analysis of prosodic boundaries shows that the “3” boundaries locate after “zuo2tian1” and “lao3wang2”, and the “4” boundary dwells after “ting1shuo1”. This result demonstrates that the sentence contains intermediate phrase boundaries after the marked focus and the intonational phrase boundary before another focus marker *shi*. Closer examination of the above Figure 3 illustrates that the pitch ranges of the former mentioned two words are expanded obviously. The declining phenomenon of the successive words is blocked by the appearance of another *shi*. And the compression of the pitch register after the second focused element extends to the end of the sentence. Therefore, each intonational phrase observes a prominence constituent. Then, we test the maximum pitch value differences among each word, the result the Bonfroni post hoc test exhibits that the word “zuo2tian1” is significantly different from the word “wo3”(P=0.012), “shi4”(P=0.02), “cai2”(P=0.013), “ting1shuo1”(P=0.017), and “tui4xiu1”(P=0.01), moreover, “lao3wang2” also significantly differs from the word “wo3”(P=0.032), “shi4”(P=0.032), “cai2”(P=0.023), “ting1shuo1”(P=0.025), and “tui4xiu1”(P=0.02), however, “zuo2tian1” shows no obvious difference from the word “lao3 wang2”(P=0.053).

The following graph is the perceptual results of the sentence “wo3 shi4 zuo2 tian1 cai2 ting1 shuo1 shi4 lao3 wang2 tui4 xiu1 le0” through which its scales can be clearly examined.

Figure 4: Perceptual results of “wo3 shi4 zuo2 tian1 cai2 ting1 shuo1 shi4 lao3 wang2 tui4 xiu1 le0”.

It demonstrates in Figure 4 that except for the syllable “le0” in the sentence, all the items are selected as intonational prominence-bearing unit in terms of prosodic words, i.e. “cai2”, “lao3wang2”, “shi4”, “ting1shuo1”, “tui4xiu1”, and “zuo2tian1”. However, the perceptual results do not get equal distribution, specifically, the prosodic word “zuo2tian1” observes the greatest number while “lao3wang2” occupies the secondary position. Although the other elements are judged as intonational prominence the number is too small to be taken as intonational prominence. In this regard, it exists scales differences between focus-bearing unit in different intonational phrase. The one distributes near the beginning of the sentence gets the primary prominence and the one locates far from the beginning of the sentence deserves the secondary. This result is partly due to the declination of the pitch tour.

3.2. Intonational prominence marking by *shi...de*

The above analysis concentrates on the phonetic realization marked by *shi* structure, in this part, the prominence pattern will be mainly discussed in the *shi...de* construction.

The following is Figure 5 displaying the F_0 means of the sentence “shi4 tian1 tian1 jiao4 ta1 qü4 bei3 jing1 de0”.

Figure 5: F_0 means of “shi4 tian1 tian1 jiao4 ta1 qü4 bei3 jing1 de0”.

Fang [7] points out that a sentence is divided into two parts by *shi...de*, the contrastive focus only locates after the marker *shi*. Prosodic boundary labeling of this sentence tells that there is an intermediate phrase boundary after the marked focus “tian1tian1”. Moreover, it is displayed in Figure 5 that the pitch ranges of “shi4” and “tian1tian1” are expanded and raised obviously, and the pitch register of the following syllables are lowered successively. In this sentence only the word “tian1tian1” is taken as the prominence-bearing unit with *shi* being left out. The causes can be found in the examination of Figure 1. However, results of the Bonfroni post hoc test show that the differences of maximum pitch values among each word show that “tian1tian1” exists obvious difference from “jiao4”(P=0.002), “ta1”(P=0.014), “qü4”(P=0.018), and “bei3jing2”(P=0.041), whereas, “tian1tian1” shows no obvious differences from the word “shi4”(P=0.68).

Figure 6 is the perceptual result of the sentence “shi4 tian1 tian1 jiao4 ta1 qü4 bei3 jing1 de0”.

Figure 6: Perceptual results of “*shi4 tian1 tian1 jiao4 tai qü4 bei3 jing1 de0*”.

Similar case is found in the Figure 6 that the judgment of the intonational prominence accord with the F_0 means that the syllable “*tian1 tian1*” is both the focus and intonational prominence bearing element.

The graph below is the F_0 of the target sentence “*shi4 xiao3 wang2 zhi1 dao4 shi4 xiao3 liu2 da3 sui4 de0 na4 ge0 bei1 zi0*” which includes two focus marker *shi* in the *shi...de* construction, in which the correlation between the marker and the intonational prominence can be seen.

Figure 7: F_0 means of “*shi4 xiao3 wang2 zhi1 dao4 shi4 xiao3 liu2 da3 sui4 de0 na4 ge0 bei1 zi0*”

Similar with Figure 3, the results of prosodic boundaries demonstrate that the intermediate phrase boundaries are after the marked focus and the intonational phrase locates before the second marker *shi*. It can be obtained from figure 7 that, the constituents in the focus-marking position sees pitch range expansion, specifically, the prosodic words “*xiao3 wang2*” and “*xiao3 liu2*” get pitch range enlargement. The pitch registers of the syllables between the two *shi* and after the second *shi* are also compressed obviously Results of the Bonferroni post hoc test tell that “*xiao3 wang2*” is quite different from “*shi4*”(P=0.02), “*zhi1 dao4*”(P=0.03), “*da3 sui4*”(P=0.013), “*na4 ge0*”(P=0.025), “*bei1 zi0*”(P=0.027), and “*xiao3 liu2*” also differs significantly from “*shi4*”(P=0.015), “*zhi1 dao4*”(P=0.025), “*da3 sui4*”(P=0.017), “*na4 ge0*”(P=0.027), “*bei1 zi0*”(P=0.029), and “*xiao3 wang2*” shows no obvious differences with “*xiao3 liu2*”.

Figure 8 is the results of the further judgments of the prominence positions of the above sentences.

Figure 8: Perceptual results of “*shi4 xiao3 wang2 zhi1 dao4 shi4 xiao3 liu2 da3 sui4 de0 na4 ge0 bei1 zi0*”

It obtains similar result from the Figure 7 that the words “*xiao3 liu2*” and “*xiao3 wang2*” are judged as intonational prominence. To put it further, the scales of the differences between the two words can be seen from the Figure 8 that “*xiao3 wang2*” observes the primary prominence and “*xiao3 liu2*” the secondary. This result is also due to the declination of the pitch contour of the target sentence.

4. CONCLUSIONS AND DISCUSSIONS

In general, the present study adopts the prominence realization of *shi...de* construction as the research perspective and systematically investigates the relationship between the focus-bearing unit and the intonational prominence-bearing unit. Major findings of this study are: 1) In the *shi...de* structure, the intonational prominence has highly corresponding relationship with the marked focus, when there is an focus marker in *shi...de*, there appears an intonational prominence, the focus bearing-unit accords with the prominence-bearing unit; 2) The phonetic realization of the intonational prominence is the enlargement of the pitch range of the under-focus position, and the compression of the pitch registers of the post-focus syllables; 3) When there exists two focus-markers, it contains the scale differences between two intonational prominences which causes by the declination of the pitch contour; 4) There contains intermediate phrase boundaries after the marked focus constituent and the intonational phrase boundaries before the second focus-marker *shi*.

5. ACKNOWLEDGEMENTS

This research is financially supported by “National Hig-tech Research and Development Project” (863 Program), “New Methodology for Speech Generation in Natural Human-Machine Interaction” 2006AA01Z138 and “Fundamental Speech Corpus and Pitch Pattern of Standard Chinese” YZDN50-05050.

6. REFERENCES

- [1] Wu Guo. *Information Structure in Standard Chinese*, Peking University Press. Beijing. 1999.
- [2] Ross, Claudia. “On the functions of Mandarin DE”, *Journal of Chinese Linguistics*, 11.2: 214-216, 1983.
- [3] Chao, Yuanren. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press, 1968.
- [4] Li, Charles N. & Sandra A. Thompson. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley and Los Angeles: University of California Press, 1981.
- [5] Song Yuzhu. “On tense markers *de* and *laizhe*”, *Zhongguoyuwen*, 4: 271-276, 1981.
- [6] Yuan Yulin. “From the perspective of focus theory to observe the syntactic and semantic function of *de*”, *Zhongguoyuwen* 1:3-16, 2003.
- [7] Fang Mei. “The syntactic approach of Chinese contrastive focus”, *Zhongguoyuwen* 4: 279-287, 1995.
- [8] Xu, Yi. “Effects of tone and focus on the formation and alignment of F_0 contours”, *Journal of Phonetics*, 27: 55-105, 1999.
- [9] Jia, Yuan, Xiong Ziyu, Li Aijun. “Phonetic and Phonological Analysis of Focal Accents of Disyllabic Words in Standard Chinese”, The 5th International Symposium on Chinese Spoken Language Processing, Springer press, 55-66, 2006.
- [10] Chen, Yiya. “Durational adjustment under corrective focus in Standard Chinese”, *Journal of Phonetics*, 34:176-201, 2006.