Prosodic Comparisons of Two Types of Realization of Focus in Mandarin

Ying Chen
School of Foreign Studies, Nanjing University of Science and Technology, China
ychen@njust.edu.cn

Abstract

Two types of narrow focus that differ in discourse context in Mandarin are investigated in terms of prosodic change from broad focus, namely, “complete focus” as the answer to a question with no contrastive meaning and “replacing focus” as the correction to reject the presupposed information with contrastiveness [1]. The acoustic analysis indicates that there are in-focus expansion of duration, F0 and intensity and post-focus compression of F0 and intensity in both types of narrow focus. However, the detailed statistic comparisons reveal that complete focus is acoustically realized with more salient prosodic variations than replacing focus.

Index Terms: prosodic change, comparison, sentence focus, complete focus, replacing focus

1. Introduction

Focus is the pragmatically unpredictable element of an utterance and the element of information whereby the presupposition and the assertion differ from each other. It is the complement of the presupposition in an utterance and makes an utterance into an assertion [2]. According to Bolinger [3], words in English utterances can be ‘focused’ or ‘highlighted’ to new, contrastive, or other special information and focused words are marked by pitch accents. Guessenhoven’s [4] Focus-to-Accent theory points out that focus can span portions of utterances and pitch accents can mark both focused words and focused constituents in English. Selkirk [5] defines the Prosody-Focus Relation that the focus structure of a sentence in English is directly constrained by the distribution of intonational pitch accents in the sentence. Ladd [6] also elaborates pitch accents on words that convey new information in English as well as utterance-level accentuation cross-linguistically.

In spite of lexical tones, focus in Mandarin Chinese [7, 8, 9, 10], similar to English [11, 12, 13], has been found to be acoustically encoded with prosodic expansion on focused constituents and prosodic compression on post-focus constituents. In other words, the focused word is prosodically realized with increased values of the acoustic parameters—duration, F0 and intensity—compared to the no-focus version, and the post-focus constituent with decreased values of F0 and intensity.

To examine the prosodic realization of focus, most of the previous studies elicited focus in Question-Answer conditions [8, 9, 10, 11] and a few directed participants to read aloud sentences with font-highlighted focal words [13, 14]. These studies generally investigated only one type of focus in terms of discourse context. Research comparing the prosody of two types of focus remains scarce. The current study attempts to examine the acoustics of prosodic focus in different contexts in Mandarin Chinese. The focus types involve “sentence focus” [2] as the baseline and “complete focus” and “replacing focus” [1] as dependent variables.

According to Lambrecht [2], sentence focus refers to those clausal structures without any pragmatic presupposition. It is usually the answer to ‘what happened’ or ‘what’s happening’. Therefore, the assertion is the whole sentence that constitutes the answer and the focus domain covers all items in the sentence. Words in a sentence-focus construction are regarded as broad and non-contrastive foci. Complete focus, as Dik et al. [1] defines, is the answer to a “Q-word question”. “The Focus information is meant to fill in a gap in the pragmatic information of the addressee”. Therefore, complete focus is considered as narrow and non-contrastive focus. There are two subtypes of complete focus – “Focus on term” and “Focus on predicate”. The current study only takes account of “Focus on term”. Replacing focus occurs “in the case in which a specific item in the pragmatic information of the addressee is removed and replaced by another correct item” [1]. A replacing-focus construction is usually composed of two pragmatic constituents—“rejection” and “correction”. The former rejects the information presumably presupposed and thus removes the incorrect information; the latter inserts presumably correct information or substitutes the correct information for the incorrect information. Therefore, replacing focus conveys both narrow and contrastive information.

Based on these defined focus types, the current study explores two research questions: (1) Will complete focus and replacing focus demonstrate the typical patterns of prosodic variations (i.e. in-focus expansion and post-focus compression) compared to sentence focus? (2) Which type of focus will demonstrate more salient prosodic change, complete focus or replacing focus?

2. Methods

2.1. Participants

Twelve native speakers of Beijing Mandarin who do not speak other Chinese languages were recruited. They were born in northern cities of China and raised in Beijing and were undergraduate students at the University of Oregon at the time of testing. There were six males and six females.

2.2. Stimuli

The stimuli were designed based on focus type. The target sentence is restricted with the same characters and only the context varies in accordance with the focus type. Tones of the target sentence frame were arranged to avoid conflicting tonal contexts [15, 16]. That is to say—the offset of the tone and the onset of the one that follows it should be either high-high or low-low in order to achieve natural tonal coarticulation. Tone 3 was placed at sentence-final position for the possibility that a complete pitch contour (dipping) is produced as it is normally
produced with the first half pitch contour (low falling) at sentence-medial position. The target sentence frame is tā (Tone 1, high-level) zài (Tone 4, falling) chū (Tone 2, rising) cáo (Tone 3, dipping). Table 1 shows its focused versions and the contexts.

Table 1: Target sentences and the contexts.

<table>
<thead>
<tr>
<th>Character</th>
<th>Pinyin</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>S F c</td>
<td>Ni shuo shenme? Ta zài chu cao.</td>
<td>‘What do you say? He is mowing the lawn.’</td>
</tr>
<tr>
<td>C O M P L E T E V E R E F O C U S</td>
<td>Shuí zài chu cao? TA zài chu cao.</td>
<td>‘Who is mowing the lawn? He IS mowing the lawn.’</td>
</tr>
<tr>
<td>F O C U S</td>
<td>Ta zài chu shenme? Ta zài chu CAO.</td>
<td>‘What is he mowing? He is mowing the lawn.’</td>
</tr>
<tr>
<td>S F r</td>
<td>Wo shuo de shì: ta zài chu cao.</td>
<td>‘What I say is: He is mowing the lawn.’</td>
</tr>
<tr>
<td>R E P L A C I N G</td>
<td>Wo bu zài chu cao, TA zài chu cao.</td>
<td>‘I’m not mowing the lawn; HE is mowing the lawn.’</td>
</tr>
<tr>
<td>A C I N G</td>
<td>Shuí shuo ta bu chu cao, TA zài chu cao.</td>
<td>‘Who says he is not mowing the lawn? He IS mowing the lawn.’</td>
</tr>
<tr>
<td>F O C U S</td>
<td>Ta bu zài zhong cao, TA zàichu CAO.</td>
<td>‘He is not planting the grass; he IS MOWing the lawn.’</td>
</tr>
<tr>
<td>U S E R</td>
<td>Ta bu zài chu chóng cao, TA zài chu CAO.</td>
<td>‘He is not killing the insects; he IS mowing the LAWN.’</td>
</tr>
</tbody>
</table>

Note: SFc = Sentence focus for Compleitive focus to compare; SFr = Sentence focus for Replacing focus to compare.

2.3. Recording

Recording took place in the sound-attenuated booth of the phonetics lab at the University of Oregon. The stimulus sentences written in Chinese characters with no highlighting were displayed in PowerPoint with two blocks—compleitive focus and replacing focus. Sentences in each block were displayed in three pseudo-random orders (sentence focus always displayed first in both blocks) to record three repetitions. Compleitive-focus sentences including its sentence-focus version were elicited after answering the pre-recorded prompt questions. Replacing-focus sentences were recorded by reading aloud the entire context.

2.4. Analysis

Acoustic data were collected by using ProsodyPro (Version 6.1.3 Beta), a Praat script for prosody analysis [17]. Duration (ms), intensity (dB), mean F0 (Hz), F0 range (semitone), and F0 value (Hz) at 10 even points of the vowel in each syllable interval in all the target sentences with the three repetitions were taken into account in the current study. Differentials in mean F0, F0 range, intensity and duration between in-focus, pre-focus and post-focus words in the compleitive-focus sentences and the replacing-focus sentences and their sentence-focus counterparts were calculated. The in-focus differential of duration, intensity and F0 range was calculated as the mean of measured values of the focused syllables in the compleitive- or replacing-focus sentences minus that of their sentence-focus counterparts. The pre-focus change and post-focus change were calculated as the mean differences between the pre-focus or post-focus syllables and their sentence-focus counterparts. Mean F0 differences were converted from Hertz (Hz) to semitones (ST): F0 change = 12 * log2(F0/fto), where F0 refers to the F0 value in Hz with compleitive or replacing focus and fto refers to the F0 value in Hz with sentence focus. The differentials of mean F0, F0 range, intensity and duration change were examined in repeated measures ANOVAs, which assume subject as a random factor, with two fixed factors—focus type (two levels: compleitive focus and replacing focus) and focus condition (three levels: pre-focus, in-focus and post-focus).

3. Results

3.1. F0 contours

Time-normalized F0 contours (plotting the 10 points in each syllable) of the stimulus sentences were first examined for an observation of the general differences across focus types. Figures 1 and 2 illustrate the overall mean time-normalized F0 contours associated with different focus conditions for compleitive focus and replacing focus versus their corresponding sentence focus.

![Figure 1: Time-normalized F0 contours with sentence focus and compleitive focus.](image)

The F0 contour of sentence focus in Figure 1 demonstrates a similar pattern to that in Figure 2. However, other F0 contours with the focus on the same word do not indicate the same shape in Figures 1 and 2. The in-focus expansion and post-focus compression of F0 also show noticeable differences in the two figures. These differences call for in-depth statistical analysis.
3.2. Mean F0 comparisons

The ANOVA results of mean F0 change show no interaction between focus type and focus condition and no main effect of type. However, there is a significant main effect of condition \[ F(2, 22) = 26.789, p < 0.001 \]. Six post hoc paired-samples t-tests were then conducted to compare the mean F0 change in terms of focus condition. The significance reference level of \( p \) value was adjusted to 0.05/6 = 0.0083. The results of completive focus show significant differences between in-focus change and pre-focus change \[ t(11) = 4.150, p = 0.002 \] and between in-focus change and post-focus change \[ t(11) = 3.263, p = 0.008 \]. The results of replacing focus show a significant difference between in-focus change and pre-focus change \[ t(11) = 4.225, p = 0.001 \]. For the three pairs of comparison in type, the significance reference level of \( p \) value was adjusted to 0.05/3 = 0.0167. The results show significant differences between completive focus and replacing focus in pre-focus condition \[ t(11) = 3.054, p = 0.011 \] and in-focus condition \[ t(11) = 2.881, p = 0.015 \]. Figure 3 illustrates these statistical results.

3.3. F0 range comparisons

Due to tonal coarticulation, the in-focus F0 raise mostly results in an anticipatory raise in the preceding syllable and a sequential raise in the following syllable as Figures 1 and 2 demonstrate. The F0 range of pre-focus change and post-focus change will be thus expanded. Therefore, pre-focus F0 range and post-focus F0 range are calculated excluding the adjoining syllables of the in-focus words.

The ANOVA results indicate no interaction between focus type and focus condition but significant main effects of type \[ F(1, 11) = 9.401, p = 0.011 \] and condition \[ F(2, 22) = 8.299, p = 0.002 \]. Post hoc paired-samples t-tests were then conducted to compare the F0 range change in terms of focus type and focus condition. For the six pairs of comparison in condition, the significance reference level of \( p \) value was adjusted to 0.05/6 = 0.0083. The results of completive focus show significant differences between in-focus change and pre-focus change \[ t(11) = 4.150, p = 0.002 \] and between in-focus change and post-focus change \[ t(11) = 3.263, p = 0.008 \]. The results of replacing focus show a significant difference between in-focus change and pre-focus change \[ t(11) = 4.225, p = 0.001 \]. For the three pairs of comparison in type, the significance reference level of \( p \) value was adjusted to 0.05/3 = 0.0167. The results show significant differences between completive focus and replacing focus in pre-focus condition \[ t(11) = 3.054, p = 0.011 \] and in-focus condition \[ t(11) = 2.881, p = 0.015 \]. Figure 4 illustrates these statistical results.

3.4. Intensity comparisons

Similar to mean F0 change, the ANOVA results of intensity change show no interaction between focus type and focus condition and no main effect of type. However, there is a significant main effect of condition \[ F(2, 22) = 22.521, p < 0.001 \]. Six post hoc paired-samples t-tests were then conducted to compare the intensity change in terms of focus condition. The significance reference level of \( p \) value was adjusted to 0.05/6 = 0.0083. The results of completive focus show significant differences between in-focus change and pre-focus change \[ t(11) = 4.837, p = 0.001 \] and between in-focus change and post-focus change \[ t(11) = 6.292, p < 0.001 \]. The results of replacing focus show a significant difference between in-focus change and post-focus change \[ t(11) = 6.616, p < 0.001 \]. Figure 5 illustrates these statistical results.
3.5. Duration comparisons

The ANOVA results for duration change indicate that there is no interaction between focus type and focus condition. However, the main effects of type \( F(1, 11) = 22.278, p = 0.001 \) and condition \( F(2, 22) = 40.004, p < 0.001 \) are both significant. Post hoc paired-samples t-tests were then conducted to compare the duration change in terms of focus type and focus condition. For the six paired comparisons in condition, the significance reference level of \( p \) value was adjusted to 0.05/6 = 0.0083. The results of complete focus show significant differences between in-focus change and pre-focus change \( t(11) = 10.362, p < 0.001 \) and between in-focus change and post-focus change \( t(11) = 7.167, p < 0.001 \). The results of replacing focus also show significant differences between in-focus change and pre-focus change \( t(11) = 10.334, p < 0.001 \) and between in-focus change and post-focus change \( t(11) = 5.444, p < 0.001 \). For the three paired comparisons in type, the significance reference level of \( p \) value was adjusted to 0.05/3 = 0.0167. The results show significant differences between complete focus and replacing focus in pre-focus condition \( t(11) = 2.979, p = 0.013 \) and in-focus condition \( t(11) = 5.165, p < 0.001 \). Figure 6 illustrates these statistical results.

4. Discussion

Time-normalized F0 contours of complete focus vs. sentence focus (Figure 1) and replacing focus vs. sentence focus (Figure 2) indicate in-focus F0 raise on level tone tā, in-focus expansion of F0 range on contour tones zhā, chū and cāo, and F0 lowering on the post-focus constituents. The statistical analysis of mean F0 change and intensity change proves the in-focus expansion and post-focus compression in both complete focus and replacing focus as Figures 3 and 5 indicate. No main effect of focus type on the change of mean F0 and intensity implies that the pattern of change from sentence focus to complete focus is similar to that from sentence focus to replacing focus. However, detailed pairwise comparisons in t-tests by focus condition reveal that the changes of mean F0 and intensity in complete focus are more robust than those in replacing focus.

F0 contours in Figures 1 and 2 also indicate some differences in F0 range of each syllable in complete-focus sentences vs. its counterparts in replacing-focus sentences. The statistical analysis of F0 range changes from sentence focus to complete focus and from sentence focus to replacing focus shows clear in-focus expansion in both focus types and plausible post-focus change. For complete focus, there is somewhat post-focus expansion of F0 range, while for replacing focus, there is somewhat post-focus compression of F0 range. However, as the error bars in Figure 4 indicate, the post-focus change in both focus types deviates in a large domain and thus seems unstable. This result may be due to that the last syllable in the target sentence is underlyingly a dipping tone and realized either dipping or falling, even with a creaky voice. Unlike mean F0, the comparison of F0 range was found with the main effect of focus type. The in-focus expansion in complete focus is greater than that in replacing focus. Focus constituents in complete focus show anticipatory F0 range expansion even though the data of the syllables that precede the focused words were excluded. All of these results indicate again that the change of F0 range in complete focus is more robust than that in replacing focus.

The duration change also shows main effects of focus condition and focus type in the statistical analysis. Robust in-focus shortening can be seen in both complete focus and replacing focus. However, the in-focus increase of duration is more salient in complete focus than in replacing focus. There is almost no pre-focus and post-focus change of duration in complete focus while somewhat decrease of duration on the pre-focus and post-focus constituents can be seen in replacing focus.

The above analysis answers the two research questions in the current study. For Question 1: Yes, both complete focus and replacing focus demonstrate in-focus expansion of duration, F0 and intensity and post-focus compression of F0 and intensity. However, for Question 2, complete focus results in more robust change of these prosodic parameters. This kind of difference may be attributed to the discourse context. For complete focus, the target sentence was produced by directly answering a question [1]. The participants only needed to produce such a short target sentence. The prosodic prominence of focused words is more straightforward. As for replacing focus, the participants had to produce both the presupposed “rejection” and the informative “correction” [1]. The target sentence was placed after the contextual clause. Despite strong contrastiveness, the prosodic prominence of focused words in the second half of a long sentence may not be as robust as that in a short target sentence. Instead, it could be realized with an intonational downtrend [18]. Therefore, even though complete focus is regarded as narrow and non-contrastive focus whereas replacing focus is regarded as narrow and contrastive focus, complete focus demonstrates more salient prosodic variations than replacing focus.

5. Conclusions

The current study investigates prosodic change from sentence focus to complete focus and from sentence focus to replacing focus in Mandarin Chinese. The acoustic data indicate in-focus expansion of duration, F0 and intensity and post-focus compression of F0 and intensity in both types of narrow focus. However, the statistical comparisons between them reveal that complete focus is realized with more robust prosodic variations than replacing focus.

6. Acknowledgements

This work was supported by the National Science Foundation of China, approval number 61573187. The author thanks the subjects at the University of Oregon and the research assistants at Nanjing University of Science and Technology.
7. References