Phonetic Realization of Narrow Focus by Beijing EFL Learners

Hu Na\textsuperscript{1,2}, Jia Yuan\textsuperscript{2}, Liu Bin\textsuperscript{1}

\textsuperscript{1}Foreign Languages Department, Graduate University of Chinese Academy of Sciences, Beijing
\textsuperscript{2}Phonetics Lab, Institute of Linguistics, Chinese Academy of Social Sciences, Beijing

huna-2010@hotmail.com, summeryuan_2003@126.com, liu.bin@gucas.ac.cn

Abstract

Focus is of great importance in English. The present study investigates how focus is phonetically realized in English declarative sentences by Beijing EFL learners. The goal is to explore the effects of focus on their \( F_0 \) contours, the shapes of \( F_0 \) contours and the correlations between \( F_0 \) variations and accent perception. Sixteen Beijing EFL learners and two native speakers of American English recorded short declarative sentences in different focus locations. Detailed \( F_0 \) inspection reveals that Beijing EFL learners realize focus in the similar way with native speakers do. However, Beijing EFL learners’ \( F_0 \) contours under focus show several distinctions, i.e., fewer delayed \( F_0 \) peaks and fewer post-focus \( F_0 \) peaks. Furthermore, the results of a perceptive grading show that \( F_0 \) rising of minor magnitudes on focus would not be perceived as accented and over-sized \( F_0 \) rising would result in unnaturalness.

Index Terms: Focus, Beijing EFL learners, phonetic realization, negative transfer

1. Introduction

In functional grammar, matters of focus are treated as assigning pragmatic function to linguistic expression. In languages like English, focus is always realized by highlighting a particular piece of information against information already shared by the conversation participants (Bolinger \cite{1}). Focused items are always considered as bearing new information in discourse, which is neither inferable from the context nor shared by the interlocutors (Schwarzchild \cite{2}).

From a phonological point of view, focus in English is realized primarily by giving prominence to the syllables that are lexically stressed and the language items that are emphasized bear pitch accents (Gussenhoven \cite{3}). On phonetic level, according to Cooper et al \cite{4}, focused words are featured as of higher pitch and longer duration. They also conclude that the \( F_0 \) contour in the post-focus region is reduced if compared with its unfocused counterpart. Similar findings are also reported in Xu \cite{5} and he further concludes that a narrow focus is realized by expanding the pitch range of the on-focus stressed syllables, suppressing the pitch range of post-focus syllables, and leaving the pitch range of pre-focus syllables largely intact. As for the characteristics of stress in Mandarin Chinese, the stressed place is featured as emphasized bear pitch accents (Gussenhoven \cite{3}). On languages like English, focus is always realized by assigning pragmatic function to linguistic expression. In

2. Methods

The present study intends to address this problem through visually comparing the \( F_0 \) contours produced by Chinese EFL learners in different focus conditions. Former studies also explore whether there are \( F_0 \) variations within the focused words. Xu \cite{5} examines short English declarative sentences and the data reveals that the alignments of \( F_0 \) peaks and valleys vary with word length, stress pattern, phonological length of stressed syllable and focus. Different \( F_0 \) alignments of \( F_0 \) peaks and valleys are also identified in Mandarin, as that an \( F_0 \) peak also occurs earlier as syllable duration increases (Xu \cite{7}). Thus another question arises: does English focus produced by Chinese EFL learners possess the similar pattern within focus? Answers to this question will be approached by means of comparing Chinese EFL learners’ and native speakers of American English’ \( F_0 \) contours under different focus conditions.

Chinese EFL learners’ performance on English intonation has long been studied. Adam \cite{8} proposes that English language learners from syllable-timed language always fail to reduce conventionally unstressed syllables. Wang \cite{9} compares Chinese speakers’ and English speakers’ duration and pitch in reading speech and finds that the duration of unstressed syllables in Chinese English speakers’ speech is often longer than the native English speakers’ and Chinese English speakers generally have narrower pitch range, less \( F_0 \) fluctuations and more stressed syllables than native speakers. Similar findings are also reported in Juffs \cite{10}. He reports that Chinese English learners’ speech are always perceived as ‘flat’ in intonation, probably due to too many stresses are assigned and he also reports that in many learners’ performance the stress is realized by extending the length of the syllable rather than adjusting the pitch range . Chen \cite{11} further investigates the problem from a phonological perspective. She compares 45 Chinese English learners’ and 8 native British English speakers’ prosody within the framework of Halliday’s intonation structure and the results show that (1) EFL learners’ speech includes more feet than that of their British counterparts; (2) strong and weak contrasts are less salient in the EFL learners’ speech than the native speakers, etc. Besides, studies also reveal that Chinese EFL learners always fail to place nucleus on the units that bear new information (Juffs \cite{10}, Chen \cite{11}).

However, the existing studies are limited in two aspects. First, studies involving Chinese EFL learners’ English focus describe the distribution of focus in their speech, but don’t examine the overall tendency of their \( F_0 \) contour changes. It is necessary to examine the changes of \( F_0 \) contours because there might be subtle distinctions between their \( F_0 \) contour shapes, though Chinese EFL learners may have the focus assigned on the right language unit as native English speakers do. Otherwise, there wouldn’t be the “foreignness” in Chinese EFL learners’ speech. Therefore, this study attempts to explore if there are any differences between Beijing learners’ and native speakers’ \( F_0 \) contours in and around focus. In addition, the inadequacy of the existing research also lies in subjects recruiting. In the research methodology of former
studies, subjects are recruited without considering their dialects. This criterion might be too general since Chinese has been classified by linguists into seven groups of related languages and within the broad classification, several subtypes have been further identified. Some of the varieties vary greatly and even are mutually unintelligible. Hence different dialects may affect English intonation in different ways. Considering this, the present study will only recruit Chinese EFL learners speaking Beijing dialect and further research is expected to investigate if the conclusions drawn in this study could be generalized across other dialect areas. Beijing dialect is focused on since it is from where Mandarin is developed and people with the Beijing dialect are generally considered as having no noticeable accent.

In this study, Beijing EFL learners’ $F_0$ contours in short English declarative sentences with a narrow focus, which varies in syllable numbers and word stress positions, will be examined to address the following three questions: (i) Do Beijing EFL learners’ $F_0$ contours move in the same tendency as the native American speakers’ in and around the narrow focus? (ii) Are alignments of $F_0$ peaks and valleys in and around focus produced by Beijing EFL learners the same with native speakers? (iii) Would listeners perceive the focus if $F_0$ contours are insignificantly adjusted from a focus condition to an unfocused one? Answers to this question will be provided by the results of grading of one experienced oral instructor.

2. Methodology

2.1. Material

The stimuli are 21 short declarative sentences with or without a narrow focus. Focus is controlled by having subjects say the target sentences as answers to prompt questions. The prompt questions are wh- questions that ask about specific pieces of information available in the target sentences. This method has been employed in Cooper et al. [4] and repeated successfully in Xu [5]. The compositions of all stimuli are shown in Table 1. The locations of narrow focus vary from sentence-initial, sentence-medial and sentence-final. Words under focus are different in word length and stress patterns. Word length varies from one syllable to three syllables. The lexical stress is located on either word-initial or word-final. Broad focus (i.e., neutral condition) is also included to examine how Beijing EFL learners’ $F_0$ contours vary from the focus condition to neutral focus condition. Some of broad focused stimuli are designed in response to the question “what happened?”, which, according to Cruttenden [12], is very often employed to prompt broad focused responses. All sentences are from “AESOP (Asian English Speech Corpus Project)/CASS_Beijing”.

<table>
<thead>
<tr>
<th>Focus locations</th>
<th>Focus conditions</th>
<th>Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Broad focus</td>
<td>Jane/Anna/Saran/Catherine comes with Tom.</td>
</tr>
<tr>
<td>Sentence medial</td>
<td>Narrow focus</td>
<td>– From where the ship departed on Sunday? – The ship departed from FRANCE/JAPAN/GERMANY on</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sunday (in the morning).</td>
</tr>
</tbody>
</table>

| Sentence final | Narrow focus | – What about Jane? Who can she come with? – JANE comes with TOM/MANNY/BURNELL/CHRISTOPHER. |
| Broad focus    | Jane comes with Tom/Manny/Burnell/Christopher. |

2.2. Subjects and recording

Sixteen Beijing EFL learners were recruited as subjects, eight females and eight males. They were all born and raised in Beijing and have learned English for more than ten years. Two native speakers of American English, one female and one male, were recorded as the referential group. The female currently works for Peking University as a professional voice recorder and the male is an experienced oral English instructor of Graduate University of Chinese Academy of Sciences. They all spoke general American English without noticeable abnormal accents. All speakers reported having no speech disorders. Recording was conducted in the sound-treated booth at the Phonetics Laboratory at Chinese Academy of Social Sciences. Subjects are instructed to produce each sentence at a normal rate and not pause in the middle of a sentence. Each sentence may be produced more than once but only one will be taken for further analysis. All speakers were paid after finishing recording. The recording software was CUHK-SIAT Recording Tool, designed by the cooperation of Chinese University of Hong Kong and CASS. The equipment of the recording was the laptop and the head-wear with microphone. Its type was Sennheiser PC166, with the built-in sound card. The sampling rate was 16 kHz.

2.3. Data annotation and extraction

All sound tracts were first automatically processed by a segmentation program to generate both word level and phone level transcriptions, which can be read by Praat. Inaccurate boundaries were mended manually. $F_0$ contour in each target sentence was modified manually to eliminate irregular points. All repetitions are examined to see if the focus is on the designed language items and those fail to focus the right words are excluded from further analysis. Since the locations of focus have been carefully controlled, all repetitions are perceived as having correct focus. $F_0$ of each vowel was extracted at ten equal proportional intervals to normalize duration and then averaged with the help of the SPSS to obtain the mean $F_0$ curves. The mean $F_0$ curves were used in exploring Beijing subjects’ focus realization in different focus conditions. In this way we could understand how focus is generally realized by Beijing subjects. The between-speaker tonal variations were ignored since all subjects speak Beijing dialect and the amount of subjects was considerably large.

2.4. Grading

One native speaker of American English was invited as the grader to evaluate each subject’s English pronunciations. The grader has been working in Graduate University of Chinese Academy of Sciences as an English instructor specializing in oral English for more than five years. The grades were given from four aspects, namely, intonation (rhythm), accentuation
(stressed words), focus (the emphasized parts) and fluency. “Intonation”, “accentuation” and “focus” were given three grades each, 1, 2 or 3 (1=poor, 2=fair and 3=good) and the fourth factor “focus” was graded as either 1 or 0 (1=perceivable, 0=unperceivable). Grades were given to each speaker individually according to the grader’s impression after listening to all stimuli produced by them. Grading was conducted by checking the corresponding grades in each box on the grading sheets and giving short comments.

3. Analysis and results

3.1. Focus effects on Beijing EFL learners’ $F_0$ contours

Figure 1 displays mean $F_0$ contours of all research sentences with and without focus produced by 16 Beijing subjects. In each graph, the ordinate is the mean $F_0$ in semitones averaged over repetitions by 16 subjects, and the abscissa is normalized time. The darker curves have narrow focus on the designed words as indicated by capital letters in the sentences. The lighter curves have broad focus (i.e., neutral focus). The breaks on the curves represent syllable boundaries. The research sentences are shown on top of each graph, with the focused constituents capitalized. Due to the limited length, only two of the stimuli are presented.

![Figure 1. Mean $F_0$ contours produced by 16 Beijing subjects](image)

From visual inspection of the mean $F_0$ contours produced by Beijing EFL learners under both focused and unfocused conditions, the effects of focus on their $F_0$ contours are quite evident. The following patterns are observed: (i) Under a narrow focus, the $F_0$ contours peak at the lexically stressed syllables for all three word stress locations; (ii) For all three narrow focus locations, the $F_0$ contour peak of a word is generally higher if put under a narrow focus than not focused; (iii) $F_0$ contours at the post-focus places are generally lower than those under the unfocused condition; (iv) $F_0$ contours at the pre-focus places remain rather intact with the unfocused conditions; (v) $F_0$ contour peaks occur not only on a narrow focus, but also in post-focus regions.

3.2. Alignments and shapes of $F_0$ contours

In this section, $F_0$ contours produced by Beijing EFL learners and native speakers of American English will be compared to see if there are distinctions in alignments of $F_0$ contours between the two groups. Comparison will be made regarding three focus locations. Graphs in this section display mean $F_0$ contours of the three focus locations produced by Beijing EFL learners (darker curves) and native American speakers (lighter curves). The layout of each graph is identical with that of Figure 1.

3.2.1. Focus at sentence-initial

Through visual inspection, differences between Beijing EFL learners’ and native speakers’ alignments of $F_0$ peaks are quite evident. Beijing EFL learners’ $F_0$ peaks are generally located at the on-focus stressed syllables under all the examined conditions. However, when the focused word is stressed at the first syllable, native speakers’ $F_0$ peaks occur on the unstressed syllables following the stressed syllables as revealed in Figure 2b. Furthermore, native speakers’ $F_0$ peaks are generally much higher than those of Beijing EFL learners’ as best illustrated in all graphs in Figure 2. Additionally, both Beijing EFL learners’ and native speakers’ $F_0$ drop from the offset of the stressed syllables with native speakers’ $F_0$ rise and fall generally of a rather steeper slop than that of Beijing subjects.

![Figure 2. Mean $F_0$ contours of stimuli with focus on sentence-initial](image)

When the on-focus words are stressed at the word-final position (“Jane” and “Saran”), it is observed in both Beijing subjects’ and native speakers’ $F_0$ contours that $F_0$ peaks occur near the onset of the stressed syllables. However, when the focus is stressed at the word-initial position (“Anna” and “Catherine”), both Beijing EFL learners’ and native speakers’ $F_0$ peaks are close to the offset of the stressed syllables as revealed in Figure 2b and Figure 2d. It is also observed in native speakers’ $F_0$ contours that there are $F_0$ valleys near the onset of a stressed syllable whether or not the syllable is focused. This is also identified in Xu [5]. However, $F_0$ valleys near the syllable onsets are not always visible in Beijing subjects’ $F_0$ contours. No $F_0$ valleys are identified at the same place in their $F_0$ contours if the syllable is unfocused. This distinction is apparent at the unfocused word “Tom” in Figure 2a-b.

In the post-focus region, both Beijing EFL Learners’ and Native American speakers’ $F_0$ contours are observed with $F_0$ prominence at the sentence final word “Tom”, but the $F_0$ peaks are not as high as those at the focused words. However, there are subtle differences between Beijing EFL learners’ and native speakers’ $F_0$ adjustments. Native speakers’ $F_0$ rises at the point in discussion are much higher than Beijing EFL learners’.

3.2.2. Focus at sentence-medial

![Figure 3. Mean $F_0$ contours of stimuli with focus on sentence-medial](image)
When focus is at sentence-medial, both Beijing EFL learners’ and native speakers’ F0 peaks occur on the focused stressed syllables and their alignments of F0 peaks are identical across all word stress locations. Specifically, when the focus is stressed on the first syllable (“Germany”), F0 peaks occur near the offset of the lexically stressed syllables as shown in Figure 3b. When the stress of the on-focus word is on the second syllable (“Japan”), F0 peaks are located near the onset of syllables as observed in Figure 3a.

Although sharing the aforementioned features, Beijing EFL learners’ and native speakers’ F0 contours are different in several patterns. When the on-focus word is stressed on word-final (“Japan”), for Native American speakers, the F0 peak is reached through a gradual climbing up from the offset of preceding syllables, leaving a F0 valley close to the onset of the stressed syllables. However, no such F0 valleys are observed for Beijing EFL learners and their F0 peaks are aligned with the onset of the stressed syllable. This feature of Beijing EFL learners’ F0 peaks is generally observed over all word stress places. Furthermore, it deserves attention that for both Beijing EFL learners and native speakers, F0 contours remain rather constant before the onset of the stressed syllable, and then decreases sharply afterwards. However, as best exemplified in Figure 3, native speakers’ F0 contour decreases are of a much steeper slope.

As to the post-focus region, when the focus is at the sentence-initial position, there are some visible F0 peaks corresponding to the key words (“morning”) in F0 contours produced by Native American speakers. However, Beijing EFL learners’ F0 contour has no F0 prominence at the same word. Reasons for this phenomenon will be explained later.

3.2.3. Focus at sentence-final

Through visually inspecting the F0 contours produced when focus is at the sentence-final position, it is found that F0 peaks are still located on the stressed syllables and before the offset of the stressed syllables, as observed for both Beijing EFL learners and native speakers. However, minor distinctions within the focus are identified when the word stress is at word-final. For Native American speakers, F0 peak occurs earlier when the vowel of the stressed syllable is phonologically long (“Tom” in Figure 4a) than when the vowel is short (“Burnell” in Figure 4c). However, this variation is not observed in Beijing EFL speakers’ F0 peaks. Their F0 peaks don’t vary with the phonological length of the stressed syllables and are generally located near the vowel onset no matter the stressed syllable is long or short.

![Figure 4. Mean F0 contours of stimuli with focus on sentence-final](image)

Furthermore, considering the F0 contours of the on-focus stressed syllables, native speakers’ F0 contours generally go through a process of rising up to the peak and falling down afterwards, thus a valley close to the onset of stressed syllables is observed. Contrastively, when stress is word-final (“Tom” and “Burnell”), Beijing EFL learners’ F0 peaks occur much earlier, almost immediately at the onset of stressed syllables, without the gradual climbing up as observed in native speakers’ F0 contour.

3.3. Results of grading

All sixteen native speakers’ proficiency was evaluated by an experienced oral English instructor according to the following four criteria: intonation, accentuation, fluency and focus. Significantly, eight out of sixteen speakers were perceived as weak focus for either speaking with “a flat sameness” or focusing wrong syllables. In addition, some correlations between the four factors were found: (a) speakers who were graded as “unfocused” generally received poor grades in “accentuation”, while speakers graded as “focused” received “fair” even “good” in “accentuation”; (b) speakers graded “focused” were generally graded as “good” in “intonation”, while those graded as “unfocused” mostly received “poor” in intonation. It is noticeable that six of them received comments as “utterly flat, unvarying pronunciation and unaccented intonation” and the rest two are perceived as “lacking any appropriate stresses”. It also deserves attention that the unvarying intonation and lack of proper focus directly contribute to a difficulty in understanding. This is concluded from the fact that subjects perceived as missing the focus generally received comments as “making her difficult to understand, very tough to understand and impossible to understand”. However, such comments were not made for speakers who had difficulty in fluency and intonation.

To better illustrate the differences between good pronunciation and deficient one, F0 contours of one stimulus produced by a good speaker and a low achiever are presented in Figure 5.

![Figure 5. Mean F0 contours produced by a good speaker (M8, a) and a low achiever (F13, b)](image)

Generally speaking, good speakers are all perceived as having the focus on the designed word and their F0 contours are observed with regular ups and downs. According to the results of stress perception, they generally received comments like “very fine intonation and accents” and “easy to understand”. In addition, good speakers are consistent in the trend of F0 changing from a neutral focus condition to a focused one. Their F0 peaks are on the on-focus stressed syllable and the F0 peaks are significantly higher under focus than those of the same word when unfocused. Besides, the pitch range of the post-focus region largely remains unchanged, which can be observed in Figure 5a.

However, for low achievers, F0 contours remain somewhat flat if compared with those of the high achievers’ and their F0 contours don’t vary much with focus. Besides, F0 peaks are either insignificantly higher or even lower under a narrow
focus than in the neutral-focus sentences. It is worth noting that speaker F13’s peak is expanded at the focused syllable and the post-focus region lowered, but they are perceived as without proper focus. This is probably due to the reason that the \( F_0 \) changes are too small to be perceived. Moreover, extensively changed \( F_0 \) would cause unnaturalness to hearing as evidenced by comments the subject M9 received. The \( F_0 \) peak in the speaker M9 is drastically higher under focus than unfocused and the speaker is described to have “exaggerated stresses” and “sound artificial at times”. Owing to the limited space, his \( F_0 \) contour is not shown.

4. Discussion and conclusions

4.1. Focus realization by Beijing EFL learners

Through visual inspection of detailed \( F_0 \) contours in short English declarative sentences with different focus conditions and word stress locations produced by Beijing EFL learners, it is found that the English focus realization in Beijing EFL learners’ \( F_0 \) contours is similar to that of native American speakers. Specifically, the pitch range of the focused item is expanded, the pitch range of the post-focus items is compressed and lowered, and the pitch range of the pre-focus items remains neutral. Such systematic pitch range adjustments generate \( F_0 \) contours that have been described by the British nuclear tone tradition and the American AM theory in terms of nuclear tone or nuclear pitch accent combined with low tail or phrase accent. This study also agrees with the notion that \( F_0 \) movements are rather independent of focus, proved by the exists of post-focus \( F_0 \) peaks.

4.2. \( F_0 \) contour shapes of Beijing EFL learners

The observations of Beijing EFL learners’ \( F_0 \) contour shapes in Section 3.2 reveal the following patterns: (i) no or unapparent \( F_0 \) valley near the onset of a stressed syllable occurs if the focus is (a) both word-final and sentence-medial or (b) both word-final and sentence-final. (ii) an \( F_0 \) peak occurs near the offset of a stressed syllable when the focus is not word-final. (iii) \( F_0 \) peaks are located on the focused stressed syllables in all examined circumstances. (iv) the shape of \( F_0 \) peak doesn’t vary much with different phonological length of syllables. (v) Beijing EFL learners’ \( F_0 \) peaks are generally much lower than those of native speakers and the former’s falling slope is much smaller. (vi) some minor \( F_0 \) peaks occur in the postfocus region, but are not always apparent. The \( F_0 \) contour shapes that are unique to Beijing EFL learners can be summarized as: fewer \( F_0 \) variations within the focused items, no delayed \( F_0 \) peaks and fewer post-focus \( F_0 \) peaks. Possible reasons for these differences are discussed in the following sections.

Generally speaking, the shapes of \( F_0 \) contour of Beijing EFL learners are of fewer \( F_0 \) variations within the focused word if compared with native speakers’ and more importantly, the former’s \( F_0 \) variations are found inconsistent with the latter at times. These inconsistencies can be probably explained in terms of the contribution of negative transfer of Mandarin Chinese. In a tone language like Mandarin, metrical structure is built on alternating strong and weak articulatory strength rather than high and low pitch targets (Kochanski & Shih [13]). Pitch in Mandarin is to construct different types of tones. For instance, a rising tone is constructed by a succession of a low target and a high target (LH), a third tone by HLH and a falling tone by HL. Furthermore, Mandarin tones are not mere pitch target combinations. Durations also play a contributive role. As specified in Feng [14], the more pitch variations there are, the longer the tone will be. For example, the third tone (HLH) in Beijing dialect is longer than the falling tone (HL) because the former is constructed by more pitch targets than the latter. In English, however, the intonation is mainly constructed by alternations of local pitch targets: high versus nonhigh, as reported in Xu [7], with syllable strength playing a secondary role (Liberman & Prince [15]). In addition, there is no such notion of “tone” assigned to individual syllables in English and pitch varies in a much larger sense called “intonation domain”, which plays a significant role in English intonation, but is rather foreign to language learners whose native language is a tone language. In this way, without adequate knowledge of intonational variations in English, Beijing EFL learners produce English with rather unvaried intonation under the effects of negative transfer from Mandarin Chinese.

Some minor \( F_0 \) peaks are observed in the post-focus region in native speakers’ \( F_0 \) contours. However, they are either invisible or of much smaller size in Beijing EFL learners’ \( F_0 \) contours. The existence of postfocus local \( F_0 \) peaks is identified in English, though of much smaller magnitudes (Xu [5]). Post-focus local \( F_0 \) peaks are also found not only in English, but also in other languages, such as Mandarin (Jia [16]). A few factors may contribute to these off-focus \( F_0 \) peaks. Firstly, they may be caused by realizing a lexical stress, which are of significance in intonation languages like English as it distinguishes word meanings. Besides, the off-focus \( F_0 \) peaks may help construct intonation as English intonation is structured by alternating high and low pitch targets. Yet, a third function may have to do with information load of the word (Hirschberg [17]). However, further research is expected to explore these possibilities.

Beijing EFL learners’ \( F_0 \) rise at the sentence-final word is not as high as that of native speakers’. It could be either that these stresses are assigned on purpose, but are made less prominent by the surrounding syllables, or that they are not intended assigned at all but driven by the structural rules. For the first possibility, the word “Tom” may be intentionally stressed by the learners, but the \( F_0 \) peak on it has been made less prominent from the surrounding because the preceding syllables are equally stressed due to an effect of negative transfer from Mandarin. It is reported in previous studies that learners with a syllable-timed language as their native language have problems in differentiating strong and weak syllables in stress-timed languages like English. They tend to assign stresses on almost every syllable, including those normally unstressed ones. In light of this, the \( F_0 \) peak on “morning” is made less prominent and Chinese learners’ English speech is always perceived as overly stressed.

In addition to the aforementioned circumstance, it also could be that the pitch prominence on the word “Tom” is not intentionally assigned, but driven by the Nuclear Stress Rules (NSR). The Nuclear Stress Rules were first reported by Chomsky and Halle (Schmerling [18]). It is stated that under an unmarked condition (or “neutral condition), the right-most primary-stressed vowel in a major constituent will receive the primary stress. This right-prominent rule is also concluded in a number of studies (Cruttenden [12], Crystal [19]). Ladefoged [20] extends the notion into both pronunciation of individual words and longer intonational phrases and divides the accent into an obligatory primary accent and an optional secondary accent. This rule also applies to Mandarin Chinese. According to Feng [14], Mandarin also has weak-strong distinction among a series of continuous syllables with the right-most syllable being the strongest, the left-most secondary stressed and the medial least stressed. As for the examined stimulus “JANE cones with Tom”, the narrow focus is at the sentence-initial word with the rest of the sentence largely “unmarked”. 
Thus driven by the Nuclear Stress Rules, Beijing EFL learners assign stresses on the sentence-final word “Tom” besides the on-focus word “Jane”.

However, this explanation fails to apply to other circumstances. For the stimulus “the ship departed from JAPAN in the morning” and “the ship departed from GERMANY in the morning”, no local F0 peaks are observed in the post-focus region in Beijing EFL learners’ F0 contours. It is observed from Figure 3a and Figure 3b that there are F0 peaks corresponding to the word “morning” in the post-focus region of native speakers’ F0 contours, but not at the same place in Beijing EFL learners’ F0 contours. However, after listening to all 32 repetitions (2 words × 16 subjects), the underlying reason for this variation is made available. It is found in nearly all repetitions produced by Beijing subjects that the article “the” preceding the word “morning” is almost of equal prominence with “morning”. However, the article “the” is hardly audible in Native Americans’ repetitions. The word “the”, being one of the English articles, which is one category of “function words” (other categories are auxiliary verbs, conjunctions, prepositions, pronouns, relative pronouns and articles), is more likely to be unaccented, unless meaning requires a pitch accent on it (Ladefoged (20)). Unaccented function words are of “weak forms” with vowels reduced. The vowels that are reduced are called “schwa”, while the vowels in stressed syllables are called “full vowels” (Avery & Ehrlich (21)). Being a tone language, Mandarin depends on tone variations to distinguish word meanings. Stressed syllables in Mandarin, however, are hard to identify in contrast to unstressed ones in continuous speech. Mandarin function words also have “weak forms”, which maintain their full forms and are only signaled by the loss of their original lexical tones (Feng (14), Wang (9)). Furthermore, the syllables in weak forms are approximately half as long as those in content words and function words have smaller relative amplitudes than content words (Baker et al (22)). However, it seems that a great number of Beijing EFL learners are unaware of this difference and features of Mandarin function words transferred into Beijing EFL Learners’ English productions, which result in “few weak forms” in their speech and “all syllables are produced clear and strong” (Chen (11)). In this study, no focus is designed on the article “the”. However, a dominant number of Beijing EFL learners utter “the” in its cited form.

4.3. Are these F0 variations in Beijing EFL learners’ F0 contours perceivable?

As to whether or not those minor F0 variations on some Beijing EFL learners’ F0 contours from a focus condition to an unfocused one are perceivable, the discussion in Section 3.3 has provided a negative answer. Although F0 peaks are higher under a narrow focus than in the unfocused sentence, these repetitions are still perceived as unfocused if the F0 peaks are of minor magnitudes. Additionally, an abrupt F0 rise would be perceived as unnatural and less fluent. Therefore, there may be a scale, within which F0 varies to trigger listeners’ perception and meanwhile, maintain naturality. Further studies are expected to give answers to this question.

Additionally, there seems a cause and effect relationship between focus and comprehension. Speakers who stress too many syllables are difficult to understand. In an intonation language like English, words that are prominent in continuous speech always bear new information (Schwarzschid (2)), which requires extra attention from listeners. The large amount of stresses produced by Chinese EFL learners would mean an equal amount of new information from a native speaker’s point of view, which is too much to process at one time and causes an understanding burden for listeners. What is more, drastic pitch variations would be perceived as unnatural. Therefore, students are expected to make the focused parts prominent and meanwhile maintain fluent globally when speak English.

5. Acknowledgements

This research is financially supported by Innovation Program of Chinese Academy of Social Sciences (AESOP—Corpus and Phonetic Study of English Learners from Chinese Dialect Regions), Chinese Social Sciences Foundation for Youth ‘Phonetic Characteristics and Phonological Expression of Chinese Discourse’ (No. 10CYY036), CASS Youth Foundation ‘Level of Accents in Standard Chinese’. Sincere gratitudes go to the anonymous reviewers for their insightful comments which help improve this paper.

6. References