PHONETIC REALIZATION OF NARROW FOCUS BY CHINESE EFL LEARNERS

Xiaoyu Zhou, Binbin Sun, and Hui Feng

School of Foreign Languages, Tianjin University
zxy8899000@126.com, sunbinbin_tju@163.com, fenghui@tju.edu.cn

ABSTRACT

Focus bears the function of information transmission and is regarded as a notion closely related to the discourse function of prosody. This study takes 20 Chinese EFL learners and 8 native English speakers as the subjects to investigate whether the difference between Mandarin and English focus realization patterns affect the narrow focus realization of Chinese EFL learners. Findings are as follows. (1) Learners of lower English proficiency exhibit similar phonetic realization to Mandarin Chinese, with a post-focus intensity decrease in both initial and medial focus sentences and a significant post-focus duration reduction than English speakers. (2) Learners show new features not found in English narrow focus phonetic realization, that is, the loss of post-focus pitch range compression (PFC) in lower English proficiency learners, and the significantly higher pitch rising of final-focus words than the other focus positions in higher English proficiency learners.

Keywords: Focus, Phonetic realization, Chinese EFL learners, L2 Intonation

1. INTRODUCTION

As a notion closely related to the discourse function of prosody [1], the phonetic realization of focus embodies the combination of semantics and prosody.

Previous studies have observed the several cues related to focus realization within the phonetic level and discovered the existence of a close relationship between on-focus constituents and higher F0, longer duration and stronger intensity [2-3]. Based on detailed F0 analyses, Xu and Xu [4] revealed that the narrow focus is realized by expanding the pitch range of the on-focus stressed syllables and suppressing the pitch range of post-focus syllables, namely "PFC" phenomenon. The existence of PFC can be found in many languages and dialects, for instance, English [5], Korean [6], Beijing dialect [7], Uyghur [8], and Lhasa Tibetan [9].

Though there have been numerous constructive explorations on the way native speakers realize the focus, studies concerning language learners are fewer. The differences in focus realization between learners and native speakers are mainly reflected in the different roles of acoustic cues, the less obvious changing range, and the loss of PFC (eg., [10-12]). Besides, L2 focus realization have been discussed from the perspective of L1 influence (eg., [13-15]).

Based on the research background, the present study attempts to detect the phonetic realization of narrow focus by Chinese EFL learners, and explore cross-language distinction on the narrow focus realization.

2. RESEARCH DESIGN

2.1. Research questions

In an effort to detect the phonetic realization of narrow focus by Chinese EFL learners, this study aims to anwser the following questions:

1) Does the difference between Mandarin and English focus realization patterns affect the narrow focus realization of Chinese EFL learners?

2) Do different proficiency levels affect the phonetic realization strategies of narrow focus by Chinese EFL learners?

2.2. Stimuli

A total of 28 subjects participated in the experiment, including 20 Chinese and 8 Native American English speakers. The 20 Chinese learners are divided into two groups with equal numbers of higher and lower proficiency levels based on the CEFR grade derived from Dialang vocabulary test results [16].

To ensure that speakers naturally produce target declarative sentences with narrow focus, this experiment adopts the combination of guiding questions and corresponding target declarative sentences as the main form of pronunciation stimuli.

The English stimuli design refers to [4], which contains short declarative sentences with keywords based on sonorant onsets and limited coda consonants (reproduced in Table 1 in the Appendix). The pairs of English sentences are divided into four groups according to various focus conditions.

The Chinese recording materials are selected from the stimuli of [17]. Each Chinese sentence contains three keywords with the same number of Chinese characters and consistent internal tones (reproduced in Table 2 in the Appendix). They are produced under

the four focus conditions described above. Thus, a total of 1980 tokens are taken into consideration in the following experiment process, with 28 subjects * 3 keywords * 21 sentences for English stimuli and 8 subjects * 3 keywords * 9 sentences for Chinese stimuli.

2.3. Data collection and processing

During recording process, the recording materials were presented to the subjects in turn through the software Speech Recorder [18] and the subjects were required to operate the recording independently. After recording, all recordings were manually marked under ProsodyPro script files [19]. On this basis, ProsodyPro automatically modified pitch contour and extracts F0, duration and intensity.

Due to the variety of pronunciation styles and physical properties of speech signals, normalization is needed to eliminate interpersonal random differences and extract constant parameters to obtain information of linguistic significance [19-20]. Therefore, Formula (1) is used for F0 values to convert Hz to semitones (st) with 1Hz as reference value.

(1) $F_{st} = 12 \times \log_2(F0)$.

As for the parameters of duration and intensity, in order to explore the impact of discrete variables on duration and intensity and eliminate individual differences, the analysis of duration and intensity in this study mainly focuses on comparison rather than conducting quantitative analysis of absolute values. Hence in terms of the statistical analysis, this paper adopts the processing method of time length and intensity data normalization by Zhao et al. [21], and takes the duration ratio of syllables and relative intensity ratio as the statistical analysis parameter. The formulas are presented below:

(2) $D_r = D/\overline{D_i}$
(3) $INT_r = INT/\overline{INT_i}$

where D_r and INT_r refer to the duration ratio and the relative intensity ratio of the stressed syllable within a keyword respectively, and they are calculated based on the ratio between *D/INT* (the duration/intensity of the syllable) and $\overline{D_i}$ /INT_i (the average syllable duration/intensity of the sentence in which the syllable is located).

3. RESULTS AND DISCUSSION

3.1. Comparison of focus realization between English and Chinese

3.1.1. Pitch analysis

As is illustrated in Fig. 1, American English speakers (AE) show in-focus MaxF0 values increase (t=2.058, *p*=0.049), and in-focus pitch range expansion (t=3.298, *p*=0.013) when realizing narrow focus. Moreover, there are post-focus pitch drop and pitch range compression for initial (t=-4.977, *p*=0.002; t=-3.237, *p*=0.013) and medial focus sentences (t=-2.325, *p*=0.053; t=-2.74, *p*=0.029). Pre-focus pitch range compression also exists, but within medial focus sentences only (t=-3.723, *p*=0.008). Besides, one-way ANOVAs with focus location as influencing factors confirms that in-focus MaxF0 rising values are significantly larger than that of post-focus and pre-focus words (F(2,23)=5.352, *p*=0.013).

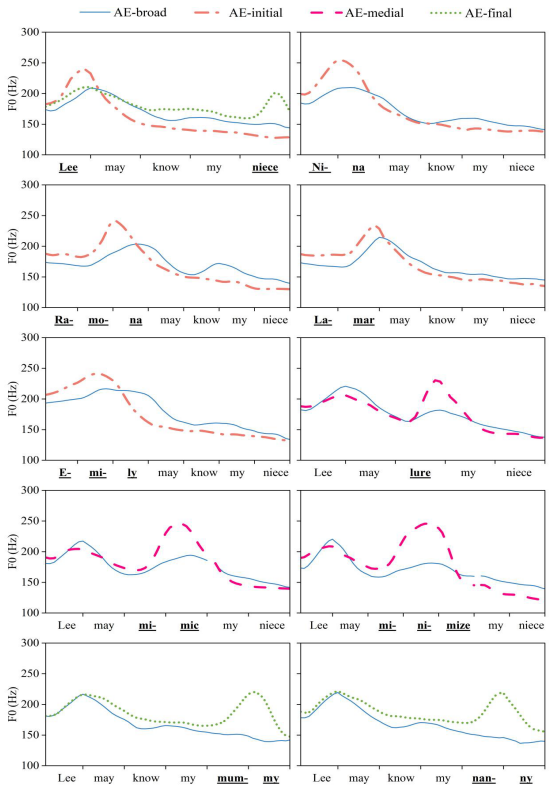

Figure 1: Mean F0 contours of native American English speakers.

Similar to AE, as is reflected in Fig. 2, in-focus pitch rise (by 1.37 semitones, t=2.158, *p*=0.049) and pitch range expansion (by 0.77 semitones, t=2.607, *p*=0.035) are both confirmed among Mandarin Chinese speakers (MC). Besides, the post-focus pitch drop (t=5.281, *p*=0.001; t=-3.278, *p*=0.014) and pitch range compression (t=-6.436, *p*=0.000; t=-2.909,

p=0.023) are inspected significantly from data analysis. However, paired-samples t-tests' results suggest that the pitch changing values of MC are larger in post-focus words than in-focus (by 1.95 semitones, *p*=0.000) and pre-focus words (by 2.78 semitones, *p*=0.000), which is distinct from AE.

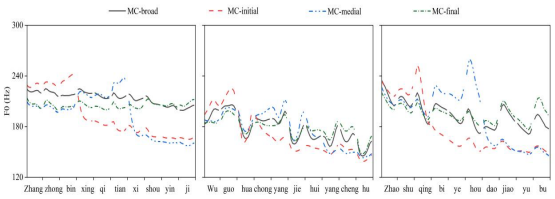

Figure 2: Mean F0 contours of Mandarin Chinese speakers.

3.1.2. Duration analysis

As is shown in Fig. 3, the duration extension was exhibited by both native AE and MC under all focus conditions. Besides, as is proved by the data analysis, the in-focus rise of MC is significantly greater than that of AE (t=-2.769, *p*=0.015), and post-focus decline of AE is lower than MC (t=4.207, *p*=0.001). It is also observed that pre-focus declining amplitude in AE is larger than MC (t=-2.393, *p*=0.031).

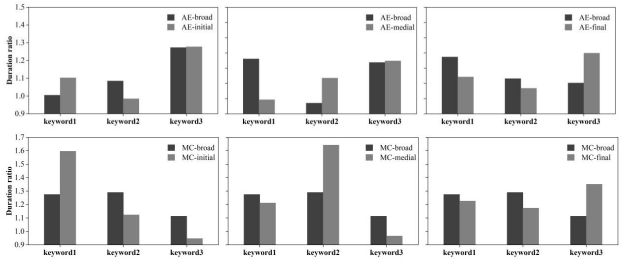

Figure 3: Mean duration ratio of AE and MC.

3.1.3. Intensity analysis

As illustrated in Fig. 4, intensity ratio increase is found in native AE speakers. Moreover, in-focus intensity increase is also found in MC. The pre-focus intensity ratio changes in AE are also significantly lower than that of the MC (t=-4.041, *p*=0.001).

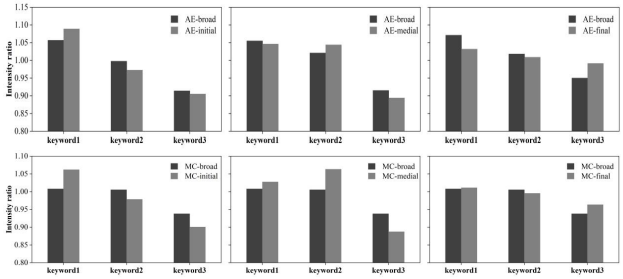

Figure 4: Mean intensity ratio of AE and MC.

3.2. Analysis of focus realization in Chinese EFL learners

3.2.1. Pitch analysis

As is shown in Fig. 5, similar to AE, in-focus MaxF0 rise, pitch range expansion and post-focus pitch drop were visually observed and statistically confirmed to be significant in both levels of Chinese learners.

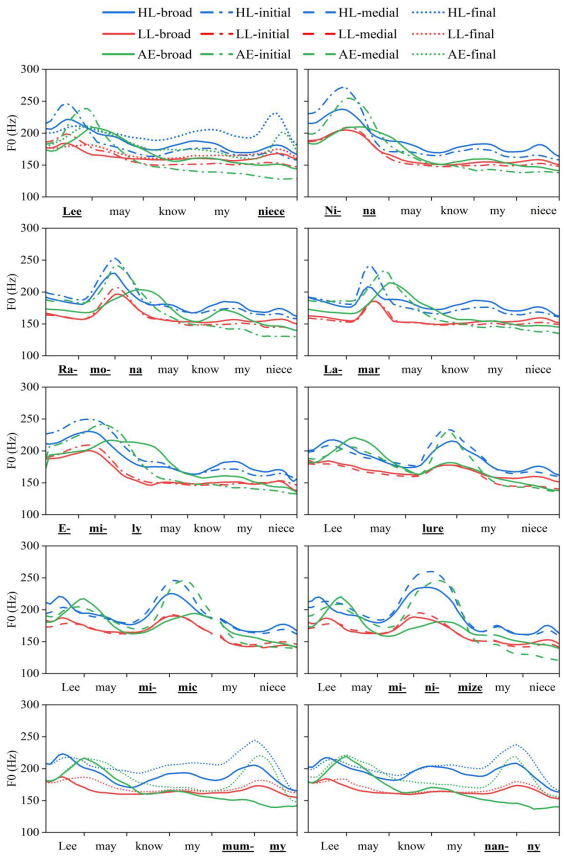

Figure 5: Time-normalized F0 contours (Hz) by HL, LL and AE group.

However, the magnitude of pitch rise does not appear to be consistent across different focus positions when higher proficiency learners (HL) use pitch as the clue to realize narrow focus. As is observed in Fig. 5, the in-focus rising range under final focus condition shown by the distance between dotted blue lines and solid blue lines appeare to be greater than that under the other two focus conditions. To statistically verify the influence of focus positions on in-focus pitch rising values, one-way ANOVAs were conducted and results showed that the main effect of focus positions is significant on the in-focus increase magnitude of MaxF0 values (F(2,29)=5.157, *p*=0.013), manifesting that in-focus rising values in final focus sentences was significantly higher than that in initial focus by 2.05 semitones (*p*=0.007) and

than in medial focus sentences by 1.86 semitones (p=0.014).

As for F0 contours of lower proficiency learners (LL) (red lines in Fig. 5), the post-focus pitch range compression was not found either in initial or in final focus sentences. Paired-sampled t-tests confirmed the absence of post-focus pitch range compression in initial focus (t=-1.782, p=0.108) and medial focus sentences (t=0.493, p=0.634). Hence it is speculated that lower proficiency learners lost PFC in the process of learning English as a foreign language, which proved that PFC was not easy to be transferred from one language to another [20]. Although pitch range was found to have insignificant changes in post-focus words produced by LL, it exhibited significant difference in pre-focus words.

3.2.2. Duration analysis

As is illustrated in Fig. 6, and is proved by repeated ANOVAs and post-hoc independent t-tests, the in-focus duration change ratio of AE was significantly higher than that of HL and LL (a near significant difference between HL and AE (t=-2.001, p=0.063) and between LL and AE (t=-1.192, p=0.064)). For post-focus duration changing, significant difference was found between LL and AE (t=-2.490, p=0.024).

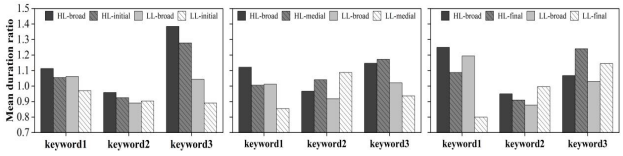

Figure 6: Mean duration ratio of higher proficiency (HL) and lower proficiency learners (LL) .

3.2.3. Intensity analysis

As is illustrated in Fig. 7, and is verified by a set of repeated measures ANOVA, both HL and LL showed in-focus intensity increase, but the in-focus intensity ratio changes (t=-2.709, p=0.015 for HL; t=-2.638, p=0.018 for LL) were significantly lower than AE.

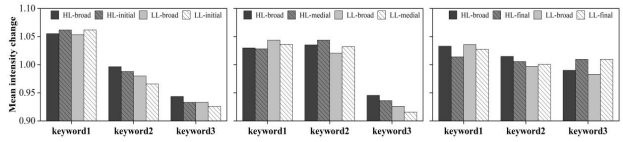

Figure 7: Mean intensity ratio of higher proficiency and lower proficiency learners.

Besides, HL showed significant post-focus intensity decrease in initial focus sentences (t=-2.858, p=0.019) and pre-focus intensity decrease in finial focus sentences (t=-3.020.p=0.014). LL showed significant post-focus intensity decrease in both initial (t=-4.001, p=0.003) and medial (t=-2.421, p=0.039) focus sentences. However, the pre-focus intensity change was found not significant both in medial (t=-1.658, p=0.132) and final focus conditions (t=-0.640, p=0.538) among LL.

4. CONCLUSION

This study aims to explore the the narrow focus realization of Chinese EFL learners under the influence of the differences between Mandarin and English focus realization patterns and English proficiency.

In terms of the differences between English and Mandarin focus realization, it is found that English mainly uses in-focus cues rather than those before and after it, and have a larger pitch rise, smaller duration extension for in-focus words, and less pitch drop, more duration reduction, smaller intensity declines for post-focus words than Mandarin Chinese.

As to focus realization by Chinese EFL learners, it is revealed that higher proficiency learners tend to use the consistent parameter adjusting methods as native American English speakers, while lower proficiency learners tend to use the categorical strategies and gradient features similar to Mandarin Chinese.

It is also found that both speakers of higher and lower proficiency show new features different from their L1 and L2 focus phonetic realization. Learners of lower proficiency lose PFC, the important categorical strategy in Mandarin Chinese, when learning English as a foreign language, which proved that PFC is easy to lose but difficult to learn [12]. Learners of higher proficiency show significantly higher in-focus pitch rising values in final focus sentences than initial and medial focus sentences. In terms of the in-focus duration extension, learners' value is significantly less than American English speakers, while the duration extension of American English speakers is less than Mandarin Chinese speakers, which indicates the overuse of L2 strategies.

Future works are needed to prove the effect of perception performance on the phonetic realization of narrow focus and to conduct long-term research to further confirm the positive development as their English proficiency increased.

5. ACKNOWLEDGEMENTS

This research is partly supported by National Major Social Sciences Foundation of China under Grant 15ZDB103, the Innovation Program of Chinese Academy of Social Sciences, and the Social Science Foundation of Tianjin, China (No. TJWW19-009), and the National Key R&D Program of China under Grant 2018YFB1305200.

6. REFERENCES

[1] Chun, D. M. 2002. *Discourse intonation in L2: From theory and research to practice,* Vol. 1. John Benjamins Publishing.

[2] Cooper, W. E., Eady, S. J., Mueller, P. R. 1985. Acoustical aspects of contrastive stress in question–answer contexts. *The Journal of the Acoustical Society of America*, *77* 6, 2142-2156.

[3] Wells, W. H. 1986. An experimental approach to the interpretation of focus in spoken English. In *Intonation in discourse,* 53-76. Routledge.

[4] Xu, Y., Xu, C. X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, *33* 2, 159-197.

[5] Xu, Y. 2005. Speech melody as articulatorily implemented communicative functions. *Speech communication*, *46* 3-4, 220-251.

[6] Lee, Y. C., Xu, Y. 2010. Phonetic realization of contrastive focus in Korean. In *Speech Prosody 2010-Fifth International Conference*.

[7] Liu, F. 2009. *Intonation systems of Mandarin and English: A functional approach*. The University of Chicago.

[8] Çetinkaya, E. 2022. The linguistic realization of focus in Uyghur: can the two focusing strategies be used interchangeably? *Poznan Studies in Contemporary Linguistics*.

[9] Zhang, X. X., Wang, B. 2018. The prosodic coding of focus and question in Tibetan Lhasa dialect. *Journal of Tsinghua University (Science and Technology)*, *58* 4, 368-373.

[10] Ueyama, M., Jun, S. A. 1996. Focus realization of Japanese English and Korean English intonation. *UCLA Working Papers in Phonetics*, 110-125.

[11] Yeou, M. 2004. The realization of accentual focus in Moroccan learners of English. In *Speech Prosody 2004, International Conference*.

[12] Chen, Y., Xu, Y., Guion-Anderson, S. 2015. Prosodic realization of focus in bilingual production of Southern Min and Mandarin. *Phonetica*, *71* 4, 249-270.

[13] Nava, E., Zubizarreta, M. L. 2010. Deconstructing the nuclear stress algorithm: Evidence from second language speech. *The sound patterns of syntax*, 291-316.

[14] Zubizarreta, M. L., Nava, E. 2011. Encoding discourse-based meaning: Prosody vs. syntax. Implications for second language acquisition. *Lingua*, *121* 4, 652-669.

[15] Li, A., Wan, X., Zhao, C., Zhu, L. 2020. Phonological and phonetic realization of narrow focus in declarative sentences by Jinan EFL learners. *In proceedings of the 10th International Conference on Speech Prosody*, 876-880.

[16] Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

[17] Chen, Y., Shi, F. 2011. Pitch performance of intonation of focus sentences in Putonghua. *Nankai Linguistics*, 1, 12.

[18] Draxler, C., Jänsch, K. 2020. *SpeechRecorder - a Universal Platform Independent Multi-Channel Audio Recording Software* (Version 6.2.6). [Computer software].

[19] Xu, Y. 2013. ProsodyPro—A tool for large-scale systematic prosody analysis. *Laboratoire Parole et Langage*, France.

[20] Zhu, X. 2004. Fundamental frequency normalization - how to deal with random differences in pitch?. *language Science*, *3* 2, 3-19.

[21] Zhao, J., Yang, X., Yang Y. 2012. The correspondence between focus and stress in Chinese -- a preliminary study based on Corpus. *Studies in Language and Linguistics*, 32 4, 5.

APPENDIX

Focus	Type	Sentences
Broad Focus	Question	What did you say?
	Answer	*See listed sentences below.*
Initial Focus	Question	Who may know your niece?
	Answer	***Lee / Nina / Lamar / Emily / Ramona*** may *know* my *niece*.
Medial Focus	Question	What may Lee do to your niece?
	Answer	*Lee* may ***lure / mimic / minimize*** my *niece*.
Final focus	Question	Who may Lee know?
	Answer	*Lee* may *know* my ***niece / nanny / mummy***.

Table 1: Prompt questions and answers for eliciting focus in the English stimuli (reproduced from [4])

Focus	Tones	*Sentences*
Broad Focus	high level	*Zhangzhongbin Xingqitian Xiu Shouyinji. (Zhangzhongbin fixes the radio on Sunday.)*
	high rasing	*Wuguohua Chongyangjie Hui Yangchenghu. (Wuguohua returns to Yangcheng Lake on the Double Ninth Festival.)*
	high falling	*Zhaoshuqing Biyehou Dao Jiaoyubu. (Zhaoshuqing went to the Ministry of Education after graduation.)*
Initial Focus	high level	***Zhangzhongbin*** *Xingqitian* Xiu *Shouyinji.*
	high rasing	***Wuguohua*** *Chongyangjie* Hui *Yangchenghu.*
	high falling	***Zhaoshuqing*** *Biyehou* Dao *Jiaoyubu.*
Medial Focus	high level	*Zhangzhongbin* ***Xingqitian*** Xiu *Shouyinji.*
	high rasing	*Wuguohua* ***Chongyangjie*** Hui *Yangchenghu.*
	high falling	*Zhaoshuqing* ***Biyehou*** Dao *Jiaoyubu.*
Final focus	high level	*Zhangzhongbin Xingqitian* Xiu ***Shouyinji***.
	high rasing	*Wuguohua Chongyangjie* Hui ***Yangchenghu***.
	high falling	*Zhaoshuqing Biyehou* Dao ***Jiaoyubu***.

Table 2: Target declarative sentences for eliciting focus in Chinese stimuli (reproduced from [17])

Note. The words in italics are keywords, and the non-italic words are non-keywords in this experiment. The in-focus words are located in bold.