Phonological and Phonetic Realization of Narrow Focus in Declarative Sentences by Jinan EFL Learners

Aijun Li[1], Xinyuan Wan[2], Chenyang Zhao[3], Lin Zhu[4]

[1]Institute of Linguistics, Chinese Academy of Social Science
[2]Institute of Linguistics,
Division of Psychology and Language Sciences, University College London,
[3]University of Chinese Academy of Social Sciences,
[4]Institute of Linguistics,
Beijing International Studies University

liaj@cass.org.cn, zgczx20@sina.com, cyzhao16@jlu.edu.cn, zhulinnini@126.com

Abstract

The present paper investigates the different prosodic cues Jinan L2 learners and native speakers use in the focus production of English declarative sentences. Learners' L1, Jinan dialect, and their L2, English are phonetically similar to some extents in focus realization, by which we assume that this L1-L2 similarity would account for their good acquisition. To prove the assumption, learners and native speakers' phonetic and phonological features are contrasted from the aspects of pitch pattern, pitch range in different sentence positions, and duration. Results show that learners realize the sentence-initial and sentence-medial focus by applying patterns of L+H* and H*, similar to the native speakers', but for sentence-final focus, learners' pitch pattern L+H* is different from their native counterparts. Phonetically, ANOVA analysis is employed to compare duration, on-focus and post-focus F0 variation elicited from the two speaker groups. By these analyses we know that native speakers apply different prosodic cues to realize focus in different sentence positions, while learners are found not able to distinguish sentence-initial and sentence-medial foci from the broad focus, and on-focus pitch expansion is observed in their sentence-final focus. All of these indicate that learners are not capable of using the same prosodic cues as the native speakers to realize these foci due to the transfer effect of their mother tongue dialect.

Index Terms: declarative sentence, narrow focus, Jinan English learners

1. Introduction

Focus bears an important communicative function in discourse by enhancing mutual understanding between interlocutors [1]. It usually refers to a set of new information which is neither shared by both interlocutors nor inferable from the context [2]. As a pragmatical element in a discourse, focus is realized by emphasizing the particular blocks of information which is against the common ground shared by the interlocutors [3]. Broad focus and narrow focus are the two focus types categorized on the basis of the size of focus constituents [4] [5]. Specifically, broad focus refers to the condition where all the information in the sentence is new to the hearers, so the focus domain is the entire utterance [6], while for narrow focus, only a single constituent in the utterance is made prominent [5].

From the perspective of phonetics, focus in a connected speech is essentially represented by sentence stress [7]. Focus in different languages, is realized through varied prosodic cues. For languages like English or Mandarin Chinese, a sentence is divided by focus into three zones which are characterized by different prosodic features. Specifically, on-focus elements demonstrate a significant increase in pitch range, while a clear compression in this aspect can be discovered on elements in the post-focus position, but for those in pre-focus position, pitch range remains nearly unchanged compared with the elements in broad focus condition [8]. For languages like Japanese, no significant post-focus compression (PFC) is observed after the accented syllables [9], and PFC is even absent in languages like Taiwanese or Southern Min [10] [11]. By these it is inferred that the F0 variation contributes a significant effort in focus realization of the world languages. In addition to this prosodic feature, intensity and duration to some degree, according to Chen et al. [11] are also correlated to the focus condition of a sentence. In this regard, it is assumed the learners with an L1 background of a certain language, would be influenced by the transfer effect in the process of L2 acquisition due to the similar or different prosodic cues for focus realization.

In present research, we pay close attention to the focus production in English sentences by L2 learners from China. Learners' mother tongue, Chinese, though bears some resemblance in this aspect with English, still does not always impose a positive effect on learners' acquisition. Previous research found that Chinese L2 learners differ from the native speakers in the tri-zone pitch range adjustment in English declarative sentences, and specifically, they show a less significant on-focus expansion and post-focus compression compared with their native counterparts [12]. Hu et al. [13] discovered that learners with L1 background of Zhenjiang dialect do not always peak their F0 contour on the on-focus syllables as the native speakers do, in addition, the PFC they produced was only limited to syllables immediately after the focused one. Learners from Northern China tend to apply H*L in realizing the prominent constituents, which is different pitch patterns from the native speakers [14]. Apart from pitch representation, Chinese L2 learners also differ from the native speakers in the application of other prosodic cues. A production experiment shows that learners with an L1 background of Beijing Mandarin, Changsha dialect and Cantonese respectively have a higher pitch contour, shorter duration, and weaker intensity in the on-focus elements compared with native speakers [15].

China has 10 dialectal areas where hundreds of dialects are spoken, leaving us impossible to study dialectal effects on the focus production of learners with all these dialectal backgrounds. Therefore, we decide to take learners from Jinan as a case to mainly investigate in this research.

Jinan is the principal capital of Shandong Province, located in East China. Dialect spoken by learners from this city is one of the representative Shandong dialects. The variation of F0 contour is the direct correlate to focus realization [16]. Specifically, both post-focus compression and pre-focus stabilization were observed in this dialect, additionally, a significant on-focus expansion were also discovered as well [16], which was similar to the pitch representation of English. However, according to Gao, Xu & Mu [12], even learners with an L1 background, similar to the realization of their L2, still find it difficult to fully acquire the prosodic cues native speakers use to do focus production. Xu [17] proposed that PFC is a feature easy to lose but hard to acquire. Therefore, it is of significance to explore whether the above-mentioned situation is the case in Jinan L2 learners' acquisition of focus in English declarative sentences.

Transfer theory proposed by Lado [18] that similarity between learners' L1 and L2 would influence L2 acquisition positively and the differences negatively, which was objected by Flege [19] who held in SLM (speech learning model) that similarities probably led to deviated L2 acquisition. These two theories in this research are employed to explain the data.

2. METHODOLOGY

2.1. Materials

All the sound samples are from the corpus AESOP-CASS-Jinan [20]. Specifically, to explore the different prosodic cues that the two speaker groups use to do focus production, 40 experimental sentences are designed to elicit the sound samples. All of them are declarative sentences, preceded by a WH-question as an inducing sentence, with or without a narrow focus in different positions. In addition, words in the focus position vary from one another in their length from one to three syllables. Table 1 lists some experimental sentences. Words in full capitalized letters are the set semantic focus.

Table 1: *List of experimental sentences*

location	Focus type	sentence
Sentence-initial	Narrow focus	>What about Tom? Who can come with him? >JANE/ANNA/SARAN comes with Tom.
	Broad focus	Jane/Anna/Saran comes with Tom.
Sentence-medial	Narrow focus	>From where the ship departed on Sunday? >The ship departed from FRANCE/JAPAN/TEXAS on Sunday.
	Broad focus	The ship departed from France/Japan/Texas on Sunday.
Sentence-final	Narrow focus	>What about Jane? Who can she come with? >Jane comes with TOM/MANNY.
	Broad focus	Jane comes with Tom/Manny.

2.2. Subjects

Altogether 10 Jinan English learners (5 males, 5 females) and 6 American native speakers (3 males, 3 females) are recruited in this experiment. Learners, born and raised in Jinan, are fluent in Jinan dialect and have learned English for more than 10 years. The native speakers recruited are international teachers serving in Chinese Academy of Social Sciences (CASS), born and raised in the Midwest of the U.S., and are reported to show no great variability in their accents, and have no hearing or speaking impairment. The age groups of learners and the speakers are 18-22 and 19-23 respectively. Both of them have no speech disorder and eyesight impairment.

2.3. Data Extraction and Annotation

The recording was conducted in a sound-proof booth at CASS by a software Cool Edit at a sampling rate of 22050 Hz with a 16-bit resolution. Each speaker was asked to read all the experiment sentence for only once. A repetition was required if the produced sound was bad in voice quality. Altogether we have collected 640 sound samples (400 by learners and 240 by the native speakers). On this basis, a manual ToBI annotation and pitch tier trimming were made for further data extraction. Specifically, four tiers, i.e. word, phone, break index, and tonal tiers, were added to present the sentence realization in different aspects. Pitch generated by consonants and silence were removed to get a smooth curve. In addition, points around the sharp spikes and the nasal-vowel conjunctions are adjusted. The trimming helped to reduce the random variation of the pitch contour.

The parameters discussed in this research are duration, pitch range, break index, and pitch pattern, and among them, the first two are extracted by using a Praat script. To minimize the gender difference and personal characteristics embedded in production, all the data produced extracted (except break index and pitch pattern) were further normalized by the Z-Score equation as follows:

$$z_i' = \frac{y_i - m_y}{s_y}$$

$$= \frac{\log_{10} x_i - \frac{1}{n}\sum_{j=1}^n \log_{10} x_j}{\sqrt{\frac{1}{n-1}\sum_{j=1}^n (\log_{10} x_j - \frac{1}{n}\sum_{k=1}^n \log_{10} x_k)^2}}$$

Where y_i stands for the log value of every pitch point, m_y is the log version mean value and s_y represents the standard deviation based on the log value.

3. Results

The phonetic and phonological features of focus production between learners and native speakers are compared in this section. The two factors, i.e. focus position, and the focused word length are taken into consideration in the categorization.

3.1 Comparisons of phonological patterns used by the two speaker groups

Phonologically, it is known from the results that only the two pitch patterns, i.e. H* and L+H* are used by both learners and the native speakers. Specifically, for sentences with their focus positioned initially, both speaker groups are more likely to apply L+H* in focus realization regardless of its word length and position of the stressed syllable. When the focus is positioned word-medially, H* is employed to realize the focus

in these sentences, but this is only true for the monosyllabic words in focus position. Sentence-medial focus with multiple syllables and sentence-final focus in any word length are realized by the two speaker groups through different pitch patterns which are presented as follows in detail:

Table 2: *Pitch patterns of Focus in different sentence positions and with varied word length*

Locations	Focused word length	Speaker group	Pitch pattern	
			H*	L+H*
Sentence-medial	Monosyllabic	Native speaker	16	2
		learner	19	11
	Multisyllabic	Native speaker	9	13
		learner	28	12
Sentence-final	Monosyllabic	Native speaker	27	3
		learner	3	47
	Multisyllabic	Native speaker	19	5
		learner	9	31

According to Table 2, it is discovered that for sentence-medial focus with multiple syllables, learners tend to use H* in doing focus realization, different from the native speakers' L+H*. Focus in sentence-final position is realized as H* and L+H* by native speakers and learners respectively regardless of the focused word length.

Table 3: *Distribution of break indexes*

Speaker group		Numbers of boundaries		
		1	3	4
Native speaker	Total	1879	323	241
	Mean	313.7	53.83	40.17
Learner	Total	3408	565	551
	Mean	340.8	56.5	55.1

Apart from the different pitch patterns they use to realize the sentence focus, learners also differ from the native speakers in the aspect of pause which is reflected through the distribution of break index in Table 3. We use the numbers "1" "3" and "4" to indicate word boundary, intermediate phrase boundary, and intonation phrase respectively. It is observed that learners produce more boundaries than the native speakers in all the three levels. This is probably caused by learners' disfluency in production.

3.2 Comparisons of phonetic cues used by the two speaker groups

In this part, the phonetic features between the two speaker groups are compared to see if there is a significant difference between the phonetic cues that they use to realize the focus in different sentence positions. Word length in this part is also taken into consideration.

Fig.1 presents the pitch contours of sentence-initial focus and in multisyllabic forms respectively. "A" here represents the productions of the native speakers and C those of the learners. BF and NF are the abbreviations of broad focus and narrow focus respectively.
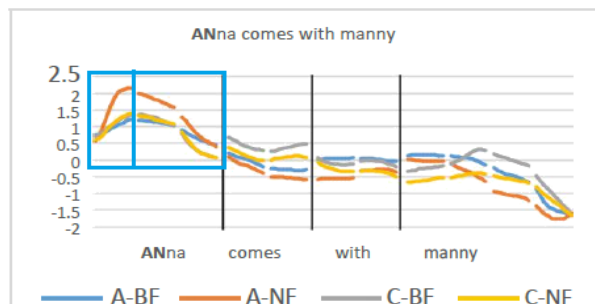

Fig. 1: *Pitch contours of Sentences with focus located initially*

By comparisons with the contours of sentence productions in broad focus, it is observed that for focused words in a multisyllabic form and with their stresses falling on the first syllable, native speakers show a significant on-focus pitch rising in their productions with focus positioned sentence-initially, which is then proved to be true by ANOVA test (P=0.022). But this is not the case for focused words in the other two possible syllable structures, i.e. focused words in monosyllabic form (P=0.180) and those in multisyllabic forms with the stress falling on the last syllable (P=0.752). Learners, in contrast, do not demonstrate a significant on-focus pitch variation in above-mentioned narrow focus condition, which goes true for stressed words in all the three syllable structures (monosyllabic: P=0.826, multisyllabic-1: P=0.814, multisyllabic-2: P=0.991, multisyllabic-1 and multisyllabic 2 here presenting the stress falling on the first and last syllable of the multisyllabic words respectively).

Apart from that, native speakers have a clear PFC which is absent in learners' contour. A further ANOVA test is made to find out if this phonetic cue is applied by the two group speakers differently. For focused word in monosyllabic forms, results show that the post-focus pitch range of learners in this narrow focus condition does not significantly different from that of their broad focus production (P=0.175), while native speakers demonstrate a significant PFC (P=0.005). This is in line with the case where the stressed syllable falls on the last syllable of a multisyllabic word, as the ANOVA test results are P=0.011 and P=0.359 for native speakers and learners respectively. But the PFC variations of the two groups are in an agreement when the stressed syllable is positioned initially in a multisyllabic word, with both of them showing insignificant change in this aspect (native speakers: P=0.115, learners: P=0.482).
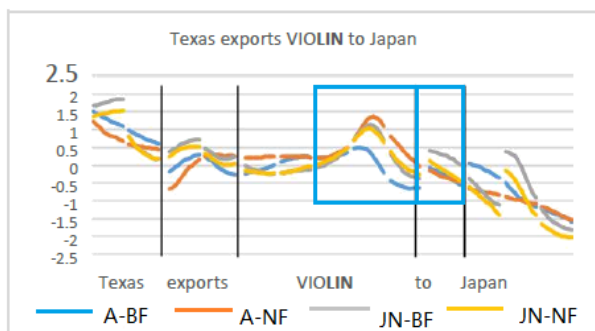

Fig. 2: *Pitch contours of Sentences with focus located medially*

Different from focus positioned sentence-initially, stressed words in sentence-medial position, are only divided into two syllable structure forms, i.e. monosyllabic forms and multisyllabic forms. According to pitch contours presented in Fig.2, native speakers demonstrate a clear on-focus pitch rising, as their F0 peak of the narrow focus is higher than that in broad focus. By this it is assumed that on-focus pitch variation is a phonetic cue used by native speaker to realize sentence-medial focus. This assumption is proved to be true by ANOVA test in both syllable structure possibilities of the stressed words (monosyllabic: P=0.005, multisyllabic: P=0.020). But learners on the contrast are discovered not able to distinguish sentences in this narrow focus condition and in broad focus condition through the same phonetic cue (monosyllabic: P=0.991, multisyllabic: P=0.881). With reference to PFC, it is known from the ANOVA test that both speaker group do not take it as contrastive feature to distinguish sentences of the above-mentioned two focus condition (monosyllabic: P=0.166 and P=0.278 for learners and native speakers respectively, multisyllabic: P=0.947 and P=0.278 for learners and native speakers respectively).
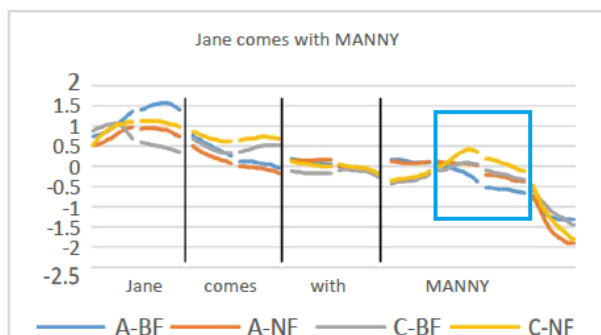

Fig. 3: *Pitch contours of Sentences with focus located finally*

Presented above in Fig. 3 is an example of sentence contour with the focus located finally, by which we could observe that different from the above two focus conditions, native speakers lack a significant on-focus pitch rising, learners however show a clearly higher F0 peak in this focus condition than that in broad focus. This visual judgment is proved by the ANOVA on the F0max (learners: P=0.035).

We then conducted ANOVA tests leon duration between the productions in broad focus and those in narrow focus by the two speaker groups. Results are presented below in Table 4.

Table 4: *Results of ANOVA analysis on duration between productions in broad and narrow focus*

Focus position	Syllable structure of focus words	Parameter	P values	
			Learner	NS
initial	Monosyllabic	Duration	0.535	0.262
	Multisyllabic-1	Duration	0.292	0.153
	Multisyllabic-2	Duration	0.363	0.931
medial	Monosyllabic	Duration	0.538	**0.047**
	Multisyllabic	Duration	0.456	0.714
final	Monosyllabic	Duration	0.341	**0.019**
	Multisyllabic	Duration	0.290	**0.012**

The above table shows that duration is not employed by learners as prosodic cues to distinguish foci in different sentence positions. As to native speakers, they apply longer duration to realize sentence-medial focus and sentence-final foci from those in other conditions.

To sum up, learners' productions of foci placed sentence-initially and sentence-medially do not significantly different from their broad focus in any prosodic aspects, with only a higher on-focus pitch is employed to represent the sentence-final focus, while native speakers are observed to use different prosodic cues to distinguish focus in different conditions. We infer that learners might not able to locate the correct semantic focus of a sentence when it is positioned word-initially and word-medially, result in the insignificant prosodic features on these words.

4. Discussions

This section offers further discussions of the prosodic representations of learners L2 focus productions in different sentence positions.

Specifically, PFC, as is observed, is not present in learners' productions with the narrow foci placed anywhere in the sentences. This is different from the prosodic features of their mother tongue which is discovered to be characterized by PFC in sentences with narrow foci [16]. The non-acquired PFC is in line with the proposal of Xu [17] that PFC is a feature easy to lose in L2 acquisition even it exists in learners' L1.

Similar to PFC, duration is used to do the distinction of focus conditions. This is probably due to the negative transfer of their mother tongue in which duration lengthening is not significant in sentences with narrow foci at any positions [16].

Compared with PFC, on-focus pitch representation is easier to be preserved in learners' L2 productions of narrow-focused sentences. But such prosodic cue is only used in sentence-final focus to distinguish it from the broad focus. According to Duan, Jia and Ran [21], on-focus pitch expansion actually is not a cue commonly used by learners' in Jinan dialect as it is only observed to be present in sentence-initial and sentence-medial focus, and does not make a more significant distinction with broad focus compared with PFC. For this reason, it is assumed that this feature is not recognized as a prosodic cue by learners able to do narrow-broad distinction and therefore is not transferred into the L2 production.

As to the on-focus pitch expansion on learners' production of sentence-final focus, it is known from Zhang and Pan [22] that compared with sentence-initial and sentence-medial foci in Mandarin Chinese, sentence-final focus is more perceptually prominent, and therefore, due to contact influence of the two varieties, learners are assumed to easily perceive the difference between syllables with this particular focus and those in broad focus condition. A prosodic cue then is applied to do the distinction. But as duration, the prosodic cue used by learners to highlight the prominence sentence-finally, and learners have no choice but to increase the on-focus pitch.

5. Acknowledgements

6. References

[1] P. Roach, *English Phonetics and Phonological: A Practical Course*, Beijing: Foreign Language Teaching and Research Press & Cambridge: Cambridge University Press, 2000.

[2] D. L. Bolinger, Accent is predictable (if you're a mind reader), *Language*, pp. 633-644, 1972.

[3] R. Schwarzschild, Givenness, avoid and other constraints on the placement of accent. *Natural Language Semantics*, pp. 141-177, 1999.

[4] C. Gussenhoven, Notions and subnotions in information structure. *Basic Notions inf. Struct.* pp. 185-203, 2007.

[5] M. Krifka, Basic notions of information structure. *Acta Linguistica Hungarica*, pp. 243-276. 2008.

[6] R. Ladd, *Intonational Phonology*. Cambridge, UK: Cambridge University Press, 2008.

[7] R. Ladd, *Intonational Phonology*. Cambridge, UK: Cambridge University Press, 1996.

[8] W. Cooper, S. Eady, and P. Mueller. Acoustical aspects of contrastive stress in question-answer contexts. *J. Acoust. Soc. Am.* pp. 2142-2156, 1986.

[9] A. Lee, and Y. Xu. Revisiting focus prosody in Japanese. In *Proceedings of Speech Prosody 2012*, pp. 274-277.

[10] H. Pan. Focus and Taiwanese unchecked tones. In Lee, C., Gordon, M., Buring, D. (eds.), *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*, pp. 195-213, 2007.

[11] S. Chen, B. Wang, and Y. Xu. Closely related languages, different ways of realizing focus. In *Proceedings of INTERSPEECH 2009*. pp. 1007-1010. 2009.

[12] W. Gao, Y. Xu, and F. Mu. An experimental study on teaching prosodic focus to Chinese EFL learners. *Foreign Language Teaching and Research (bimonthly)*, vol. 47, no. 6, pp. 861-960. 2015.

[13] N. Hu, Y. Jia, and B. Liu. Phonetic and phonological realization of narrow focus in English declarative sentences by Zhenjiang EFL learners. In *Proceedings of Speech Prosody, 2012*.

[14] X. Zhou, and S. Liu. A study on stress of declarative sentence productions by Northeastern EFL learners. *Journal of Changchun Normal university*. vol.38, no.9, pp. 95-100, 2019.

[15] X. Guo. English stress production and perception of EFL learners from Beijing, Changsha and Canton. Changsha: Hunan University, 2017.

[16] W. Duan, and Y. Jia. Contrastive study of focus phonetics realization between Jinan dialect and Taiyuan dialect. In *Proceedings of NCMMSC 2015*.

[17] Xu, Y. Post-focus compression: cross-linguistic distribution and historical origin. In W. Lee & E. Zee (eds). *Proceedings of the 17th ICPHs 2011*, pp. 152-155.

[18] R. Lado, *Linguistics Across Cultures*. AnnArbor: University of Michigan Press, 1957.

[19] J. E. Flege, and E. Wieke. Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, pp. 67-83, 1987.

[20] W. Duan, Y. Jia, and Q. Ran. Internal similarities and differences of phonetic realization of narrow focus in the Shandong dialect, 一 a case study of Jinan, Liaocheng and Zibo dialects. In *Proceedings of NCMMSC 2013*, pp. 229-232.

[21] Y. Jia, A. Li, and C. Tseng, The establishment of sound corpus of EFL learners in Chinese dialectal areas. *Chinese Journal of Phonetics*, pp. 38-45, 2013.

[22] H. Zhang, and H. Pan. 2019. Sentence-final information focus in Chinese and nuclear stress. *Contemporary Linguistics*. vol.21, no.1, pp. 23-42, 2019.