PHONETIC REALIZATION OF ACCENT FROM CHINESE ENGLISH LEARNERS IN VARIOUS DIALECTAL REGIONS

[1]Yuan Jia [1]Aijun Li

[1] Phonetics Lab, Chinese Academy of Social Sciences, Beijing, China

summeryuan_2003@126.com, liaj@cass.org.cn

ABSTRACT

The present study investigates the accent realization of English learners from Chinese dialectal regions, i.e., Beijing, Jinan, Zhenjiang and Hangzhou(hereinafter BJ, JN, ZhJ and HZ, respectively) within the frame of 'Focus-to-Accent' FTA theory. The words within the focus scope are designed to have different numbers of syllables and various word stress distributions. Results have demonstrated that English learners exhibit problems on accent realization, specifically, they cannot realize F_0 raising on the word stress and F_0 compression after the accented syllable. The production error becomes more obvious when the number of the syllables is increased. Among the learners of the four regions, they perform different magnitude of accent error, the BJ speakers show the highest level of accent realization and the JN speakers perform the worst prosodic features of accent realization.

Index Terms: English learners, dialectal region, accent realization, FTA theory.

1. INTRODUCTION

Chomsky [1] and Jackendoff [2] defined focus in terms of discourse notion of presupposition, and *focus* is the non-supposed part in the sentence. Focus is usually phonetically marked by sentential accentuation. However, there remains a number of disagreements about how focus is conveyed by accent (Ladd [3]). From the phonetic and phonological perspective, the relation between focus and accent has long been addressed. Gussenhoven [4] proposed 'Focus-to Accent (FTA)' approach to describe this phenomenon. Within the FTA theory, it contains two major branches, i.e., structure-based approach and highlighting-based approach. The former approach maintains that the words and constituents in utterances can be focused for various reasons, however, the focused parts are not always marked by pitch accents in the sentential form (Ladd [3]). The latter approach emphasizes the bi-directional relation between focus and accent, therefore, within this approach, the focused item is the accent bearing unit (Bolinger [5]).

In Chinese literatures, the relation between focus and accent in Standard Chinese (Hereinafter, SC) has been systematically discussed. The studies proposed that the relation between focus and accent in SC shows *symmetry* and *non-symmetry* features: under single and dual focus conditions, the accent distribution corresponds with the focus placement, however, under multiple focus conditions, the accent distribution does not associate with focus in number and distributing positions (Jia [6][7]).

In regard with the study of learners from Chinese, it mainly concerns the error patterns of supra-segmental features. The studies mainly discusses the following aspects: i) various types of English sentences, i.e., the imperative sentence, exclamatory sentence, *wh*-questions, and *yes-no* questions (Wang [8], Sun [9] and Ji [10]), etc.; ii) the distribution of accents, e.g., Chen [11]; iii) segmental and prosodic features of the English learners from Chinese Dialectal regions, e.g., Jia [12], Hu [13] and Wang [14]. The results mentioned in i) and ii) show that English learners in China exhibit production problems on accents realization, i.e., phonological types and distributions. Moreover, the results of learners from dialectal regions demonstrated that productions of segmental and supra-segmental features of English are influenced by their mother dialects to a great extent. Specifically, speakers from different dialectal regions, e.g., Jinan and Zhenjiang, show different error patterns on the phonological features of identical English sentences.

From the overview, it can be obtained that the examination of the Chinese English learners has not discussed in detail the error patterns from Focus-to-Accent (FTA) features produced by speakers from different dialectal regions. The present study, in this regard, intends to investigate the realization of English accents produced by speakers from dialectal regions, and the cross-dialectal comparisons of the accent features are further conducted based on FTA theory to explore the *negative transfer* from different dialects. The study further endeavors to answer the following questions: i) What constitutes the accent differences among speakers from four dialectal regions on the production of the same focused words; ii) When the focused words are designed to observe different word stress, and how the accents are performed by the speakers from four areas?

2. GEOGRAPHIC DISTRIBUTION OF DIALECTS

The study selected Beijing, Jinan, Hangzhou and Zhenjiang as the target dialectal areas, the former two dialects belong to the Guan language while the latter two are branches of Wu dialect. *BJ*(北京), the capital city of China, is located in the northernmost part of the North China Plains. *Beijing*

dialect, or Pekingese, is the dialect of Mandarin spoken in the urban area of Beijing, China. It is the phonological basis of Standard Chinese. In fundamental structure, the phonology of the Beijing dialect and Standard Chinese are almost identical. *JN* (济南) is the provincial capital of Shandong province. It is located in the North-West of the province. *Jinan dialect* is a dialect of the Mandarin language family, which in turn constitutes one of the Sinitic language families. *ZhJ* (镇江) lies in the southwest of Jiangsu province. *Being* used in the city of Zhenjiang, *Zhenjiang dialect* belongs to Wu family and mixes with the features of Shanghai and Nanjing. *HZ*(杭 州) is geographically located at the south wing of the Yangtze River Delta The Hangzhou dialect, or Rhangzei Rhwa, is spoken in the city of Hangzhou and its immediate suburbs. It belongs to the Wu language family, which in turn constitutes one of the Sinitic language families. The distribution of the dialectal regions selected in the study is listed in Figure1, the cities are marked with red symbols.

Figure 1: *Dialectal distribution of Chinese languages*

3. METHODOLOGY

3.1 Materials

The aim of the study is to examine the accent distribution and phonetic features of English learners conveyed by identical focus conditions. In order to explore the accent variations, the focused words were designed to observe different word stress through which the FTA theory manifested in the English learners can be tested. The focus is realized by contrastive context that can be defined as contrastive focus (Chen [15]). The target sentence is always kept as 'I said __ ten times', the underline is the place to put into the focused words. All the words are designed to have two, three and four syllables, which are selected with varied word stresses. Table 1 lists some sample words in this study, and the placement of word stress and numbers of syllables are presented.

Table 1. *Target words in the study*

Word	IPA	Placement of word stress	Number of syllables
money	['mʌ ni]	first	two
January	['dʒ ænjuəri]	first	three
video	['vidiəu]	first	three
apartment	[ə'pɑːtmənt]	second	three
experience	[ik'spiəriəns]	second	three
afternoon	[ˌɑːftə'nuːn]	third	three
overnight	[ˌəuvə'nait]	third	three
supermarket	['sjuːpəˌmɑːkit]	first	four
misunderstand	[ˌmisʌ ndə'stænd]	third	four

3.2 Recording procedure

The recording was conducted in a quiet indoor environment. The recording equipments are the laptop and the head-wear with microphone and its type is Sennheiser PC166, with the built-in type sound card. The recording software is developed by Chinese University of Hong Kong. In accordance with the purpose of the recording experiment, the recording software of the project AESOP (Asian English Speech Corpus Project) ─ CASS (Chinese Academy of Social Sciences), made a preliminary improvement by phonetic lab of CASS. Recording was conducted in CASS, Shandong University of Science and Technology, Jiangsu University of Science and Technology and Zhejiang University for observations of speakers from Beijing, Jinan, Zhenjiang and Hangzhou. For each dialectal region, the study invited four speakers as subjects, i.e., two men and two women from English department. They were all born and brought up in their cities, moreover, they can speak their mother dialects fluently. As for the Standard English speakers, the study invited four American speakers, also two men and two women, to participate in the experiment. All of them come from the areas near New York.

3.3 Annotation

The 'wav' files were labeled with syllabic and phonemic boundaries by automatic segmentation software. Then, the supra-segmental features are annotated by the combined system of IViE and ToBI, the content and tiers are: i) *Orthographic Tier*: transcriptions of the spoken words; ii) *Prominence Tier*: location of prominent syllables (stressed and accented); iii) *Break Index Tier*: transcriptions of intermediate and intonational phrase boundaries; iv) *Target Tier:* Phonetic transcriptions, syllable-based, allowing transcribers to draw up a first set of hypotheses about accent alignment in phrase boundary; v) *Phonological Tier*: formal linguistic representations of speakers' intonational choices; vi) *Comment Tier*:

alternative transcriptions and notes. The pitch tier information and the supra-segmental features were extracted by praat script and the results were further tested by SPSS 18. The normalization of duration was approached through extraction of ten points of each vowel.

4. RESULTS AND DISCUSSION

In this part, the study examines the accent distribution within focus scope by learners from BJ, JN, ZhJ and HZ. The error patterns of the learners are explored through the comparison of pitch contour between American speakers and learners from the dialectal regions. Based on these results, the FTA theory can be further tested from the perspective of second language acquisition.

4.1 Production of accents in disyllabic focused words

In this sub-part, the study examines the accent realization of *di*-syllabic words produced by learners from four cities, i.e., Beijing, Jinan, Zhenjiang and Hangzhou. Figure 2 is the F_0 of the target sentence 'I said money ten times', within the figure, the bottom part shows the content of each contour, specifically, AM stands for the American speaker, BJ is adopted to mark Beijing speaker, JN is for Jinan learners, ZhJ for Zhenjiang, and HZ for Hangzhou, respectively. The top part lists the content of the sentence. The rectangle in the figure illustrates the F_0 contour of the focused word 'money'. The Y-axis shows the LZ-score values (Zhu [16]):

$$Z_i^t = \frac{y_i - m_y}{s_y} \quad (1)$$

Within the formula, $y=\log_m X_i$, and $m_y$and s_y are the means and Standard deviation y_i (i=1, 2, ⋯, n), respectively. Range of the LZ-score in the study is selected from '-3～3' based on the pitch variations from all the speakers.

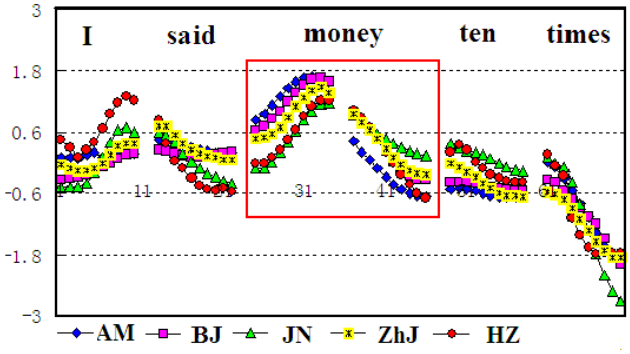

Figure 2: *F_0 contours of 'I said money ten times'*

In Figure 2, concerning with the F_0 pattern of the speakers from BJ, JN, ZhJ and HZ, they show similar tendency as the American speakers. However, the differences between English learners and AM speakers lie in the *register difference*. The AM speakers show the greatest pitch drop between the word stress bearing vowel and the following vowel, specifically, the F_0 of word stress bearing syllable is raised obviously. As for the English learners, BJ, JN, ZhJ, and HZ, the JN speakers perform the least magnitude of pitch drop and the BJ speakers perform the greatest magnitude of pitch drop in comparison with learners from other areas. The ZhJ speakers occupy the second position while the HZ speakers take the third place. As for the constituents after the focused words, it can be seen that the AM speakers show obvious F_0 lowering. It has been proposed that accent can compress the F_0 after focus (Xu [17]). However, the JN speakers show no F_0 lowering, which indicates that they do not put accent on the focused items. And, previous results show that they cannot realize word stress within the focus scope. The HZ speakers also show similar results with JN speakers and the ZhJ speakers perform medium results.

4.2 Production of accents in tri-syllabic focused words

The following part investigates the inter-differences of the English learners in the production of the tri-syllabic focused constituents. The focused word is 'afternoon', the other content in the figure keeps identical with Figure2.

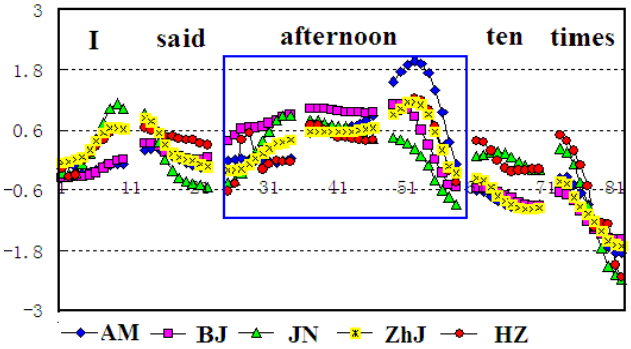

Figure 3: *F_0 contours of 'I said afternoon ten times'*

In the target sentence, the focused word is *afternoon* [ˌɑːftəˈnuːn], which observes two word stress, i.e., the primary stress and the secondary stress. The pitch contour of the AM speakers shows that the accent is realized obviously on the third syllable in the way that the F_0 is raised obviously, and the pitch contour after the focused words is compressed. As for learners from the four dialects, the obvious error is on the syllable deserving the primary stress. Learners from BJ and ZhJ perform the least magnitude of error, while the JN speakers show no obvious accent distribution within the whole word. Moreover, learners from JN and HZ show no obvious pitch compression after the focused word. Therefore, while the number of the focused item is increased the learners show more obvious errors on accent realization in the place of word stress distributing position.

4.3 Production of accents in multi-syllabic focused words

In this sub-part, the focused word has four syllables, through which the study explores differences of the accent within multi-syllabic constituents. The focused words are designed to have the word stress on the first syllable.

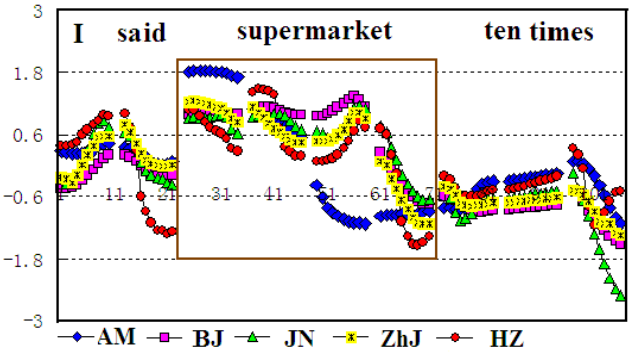

Figure 4: *F_0 contours of 'I said supermarket ten times'*

In the figure, the focused word is a four syllable constituents with the word stress on the first syllable. It can be seen that the F_0 of AM speakers exhibits great differences on the focused constituents. Specifically, the difference starts from the word stress bearing syllable which shows higher pitch register. The F_0 of the following items is compressed by the accented syllable. However, all the learners from the four areas do not perform this kind of tendency. Results demonstrate that the learners cannot put the accent on the word stress bearing syllable. The learners show problem on the F_0 raising on accented syllable and the compression of the F_0 on the following constituents. Therefore, they exhibit problem from focus to accent.

The following figure is adopted to examine the accent distribution on the word which has four syllables and the word stress on the third syllable.

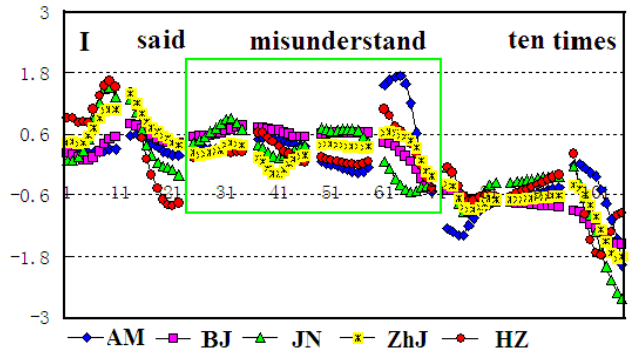

Figure 5: *F_0 contours of 'I said misunderstand ten times'*

In the figure, it can be seen that the focused word deserves the primary stress on the third syllable and the secondary stress on the first syllable, and IPA is [ˌmisʌ ndəˈstænd]. Previous study on tri-syllabic constituents demonstrates that learners from Chinese dialectal regions exhibit production problems on the accent realization place within the focus scope, and the problem is even more obvious on word stress. Regarding with the four-syllable focused constituent, the AM speakers also put the accent on the primary stress-bearing unit. However, all the learners from the four areas exhibit problems on the production of the accent within the focus scope. The JN learners also realize no accent within the focus scope. As for the constituents after the focus, the F_0 compression of learners is not as obvious as the AM learners. This result further indicates that learners from Chinese dialectal regions show accent problem on the word stress position.

5. CONCLUSIONS

The present study systematically investigates the accents realization of the English learners from Chinese dialectal regions. Learners from Beijing, Jinan, Zhenjiang and Hangzhou are selected as the samples. The production features of the accents are investigated within the FTA theory, specifically, the focus condition is designed identically, and the focused words have different numbers of syllables and varied word stress distribution. Major observations of the study reside in the following aspects: i) the English learners exhibit problem on the F_0 raising on word stress within focus scope; ii) the learners show less magnitude of F_0 compression after the accented syllable; iii) the production error of the learners is more obvious when the number of syllable is increased; iv) among the learners from the four dialectal regions, the BJ speakers show the highest level of accent realization, the Zhenjiang learners occupy the secondary position, while the Hangzhou speakers the third place, and the JN speakers show the greatest problem on accent realization. After all, the English learners show production problems within the FTA theory.

6. 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).

7. REFERENCES

[1] Chomsky, Noam "Deep structure, surface structure, and semantic interpretation. Semantics—a Interdisciplinary Reader in Philosophy, Linguistics and Psychology", Cambridge: Cambridge University Press, 183-216, 1971.

[2] Jackendoff, Ray, "Semantics Interpretation in Generative Grammar", Cambridge, Massachusetts: Massachusetts Institute of Technology Press, 1972.

[3] Ladd, D. Robert, "Intonational Phonology", Cambridge: Cambridge University Press,1996.

[4] Gussenhoven, "Focus, Mode and Nucleus", Journal of Linguistics 19: 377-417, 1983.

[5] Bolinger, Dwight, "Accent is predictable (if you're a mind-reader)", Language, 48: 633-644, 1972.
[6] Jia Yuan, Li Aijun, Xiong Ziyu "A Phonetic and Phonological Analysis of Dual and Multiple Focuses in Standard Chinese", Speech Prosody, the USA, 75-78, 2010.
[7] Jia Yuan, Li Aijun Relation between Focus and Accent in Standard Chinese, ISCSLP,IEEE Press, Taiwan, 348-352, 2010.
[8] Wang X., Sun J. and Li A. J., "A Comparative Study on Prosody of English Negative Imperative for Chinese EFL Learners", Report of Phonetic Research, Beijing, 2009.
[9] Sun J., Lu J. L., Li A. J. and Jia Y., "The Pitch Analysis of Imperative Sentences in Standard Chinese", The 6th International Symposium on Chinese Spoken Language Processing, Kunming: IEEE Press, 2008.
[10] Ji X. L., Wang X. and Li A. J., "Intonation Patterns of Yes-No Questions for Chinese EFL Learners", Report of Phonetic Research, Beijing, 2009.
[11] Chen Hua, "Yingyu Xuexizhe Langdu Kouyuzhong de Diaohe Weizhi", Journal of PLA University of Foreign Languages, (Placement of nucleus in reading speech from Chinese learners) 29(6): 32-38, 2006.
[12] Jia Yuan, Wang Xia, Li Aijun "The Influence of Shandong Dialects on the Acquisition of English Plosives". Oriental COCOSDA, IEEE Press, 79-84, 2011.
[13] Hu Na, "Phonetic and Phonological Realization of Narrow Focus in English Declarative Sentences by Zhenjiang EFL Learners", Speech Prosody, 2012.
[14] Wang Meng, "Prosodic Features of the Chinese EFL Learners as They Express New Information", Speech Prosody, 2012.
[15] Chen Y. Y., "Durational adjustment under corrective focus in Standard Chinese", Journal of Phonetics, 34: 176-201, 2006.
[16] Zhu X. N., "F_0 Normalization: How to Deal with Between-speaker Tonal Variations", Linguistic Sciences, 2004.
[17] Xu Yi, "Effects of tone and focus on the formation and alignment of F_0 contours", Journal of Phonetics, 27: 55-105, 2005.

[This paper was published in ISCSLP 2012]