Joint Effect of Dialect and Mandarin on English Vowel Production: A Case Study in Changsha EFL Learners

Xinyi Wen [1,2]*, Yuan Jia*[1]

[1] Institute of Linguistics, Chinese Academy of Social Sciences
[2] Graduate School of Chinese Academy of Social Sciences
wenxy1994@gmail.com, summeryuan_2003@126.com

Abstract

Phonetic acquisition of English as a Foreign Language (EFL) for learners in dialectal areas has been increasingly regarded as an important research area in second language acquisition. However, most existing research has been focused on finding out the transfer effect of dialect on English production from a second language acquisition point of view, but ignores the impact of Mandarin. The present research aims to investigate the joint effect of dialect and Mandarin on Changsha EFL learners' vowel production through acoustic analysis, from both spectral and temporal perspectives. We will further explain the results with the Speech Learning Model (SLM). Three corner vowels, i.e., /a/ /i/ /u/, are studied, and the results show that: English vowels /i/ and /a/ produced by Changsha learners are significantly different from those of American speakers; specifically, /i/ is more affected by Mandarin, and /a/ is more affected by Changsha dialect, which can be explained by SLM. While /u/ produced by Changsha learners is similar to that of American speakers. Besides, Changsha learners produce shorter vowels in duration, due to dialect and Mandarin's transfer effect, but can still make tense-lax contrasts in /i-ɪ/ and /u-ʊ/ pairs.

Index Terms: cross-language influence, vowel production, acoustic analysis, Speech Learning Model

1. Introduction

1.1. Previous Studies

Globalization has made English teaching and learning an inevitable trend, and speaking English idiomatically is among the top agendas. Therefore, researchers have long been interested in second language (L2) speech learning, especially on first language (L1)'s transfer effect on L2 vowel production and perception. For example, Ingram and Park [1] investigated vowel production and perception of Japanese and Korean learners of English, observing the effects of L1 phonological learning on vowel perception in Japanese. Rallo Fabra and Romero [2] found that vowels produced by Catalan English learners were less peripheral than the native versions. Wester [3] investigated how L1 vowel characteristics influence the production of /æ/ /i/ /u/ in L2, among English learners of Spanish and Spanish learners of English, and found that learners were able to realize differences between languages to different degrees, in spite of limited exposure to L2.

Considering the diversity of Chinese dialects as well as the great variance in their phonemic systems [4], domestic research attention on how L1 influences EFL learners' vowel production has shifted from Mandarin to the learner's own dialect [5]. For example, Jiang [6] studied 5 typical English monophthongs produced by EFL learners of Min and Wu dialectal areas, finding that both groups failed to acquire English vowels with equivalences in their dialects. Wang [7] analyzed the acquisition of four English monophthongs, /i/ /ɪ/ /ɛ/ /æ/, by Shandong EFL learners based on AESOP-CASS-SD, and found that dialect has positive transfer on /i/ and /æ/, while negative transfer on /ɪ/ and /ɛ/. Apart from mere segmental studies, several researches focused on the supra-segmental level, especially the intonation pattern of yes-no questions and tag questions [8][9][10][11], and found that Chinese EFL learners from dialectal areas differ from native speakers in several aspects of narrow focus realization.

Be it domestic or abroad, segmental or supra-segmental, previous studies mentioned above were majorly conducted within the framework of transfer theory, which regarded English as L2 and accounted for EFL learners' performance with L1 phonetic system.

1.2. A Shift in View

Due to the mandate use of Mandarin as the official language in China, Chinese EFL learners first exposed to their dialect went through a process of Mandarin acquisition in their childhood. Therefore, the unanimous assumption adopted by previous studies, which regards dialect as L1 and English as L2, indeed neglect Mandarin's role on English learning.

Recently, a few researchers noticed the breach, and began to study both dialect and Mandarin's influence on English learning. But they took different views toward the role Mandarin plays, among which [12] and [13] are representatives. The former regarded both dialect and Mandarin as L1, and English as L2, in the study of Beijing and Jining English learners; while the latter regarded Mandarin as L2, hence English as third language (L3), in the study of Changsha English learners. The divergence might arise from the different language space between Mandarin and the dialects involved.

Third language acquisition is relatively under-explored compared with second language acquisition, which has long been a hot issue of cross-linguistic influence [14]. But there has been increased interest in the interaction of three languages during language learning process, and learners were found to produce interlanguage in either partially or completely L2 forms, as was shown in [14][15][16][17]. Though relevant studies were mainly on lexical and syntactic acquisition, there is good reason to assume an influence of both L1 and L2, on English vowel production of EFL learners from dialectal areas.

1.3. The Present Study

The review of previous studies showed that, researches on vowel production of EFL learners were limited to certain dialectal areas, and they all assumed English as L2 without considering Mandarin's effect.

Changsha is the capital city of Hunan province, where Changsha dialect, a representative of new Xiang dialect, is mainly used in daily communication. Up to now, few acoustic studies on Changsha EFL learners' vowel production has been conducted. Besides, Changsha dialect and Mandarin are phonemically distinct and acquired sequentially [13], which qualifies English as L3 to Changsha English learners. Since Norman once pointed out in [19] that, "The Chinese dialectal complex is in many ways analogous to the Romance family in Europe," and Hammarberg defined L3 as the language currently being acquired, and any other languages acquired after L1 as L2 [18], it is legitimate to adopt the "English-as-L3" view in this situation.

Therefore, the present research would study Changsha EFL learners' monophthong production, and tries to answer the following questions: a. What are the differences between English monophthongs produced by Changsha EFL learners and native speakers? b. What are L1 and L2 vowel system's impact on L3 vowel production?

And the influence of L1 and L2 is further accounted for by Speech Learning Model (SLM) [20] within the framework of transfer theory in the discussion section.

2. Method

2.1. Data collection

2.1.1. Participants

Two groups of participants with equal gender ratio were employed in this research: 6 Changsha EFL learners and 2 Americans. The 2 Americans are native speakers of general American English, without regional accent. Changsha EFL Learners are undergraduates from Hunan University, aged between 18 and 22. They were all born and raised in Changsha with enough exposure to their dialect, and have spoken Mandarin for over ten years and English for over five years. They have all passed College English Test (Band 6), so their English proficiencies can be regarded as at a similar level. None of them presented any deficiencies on speaking or hearing.

2.1.2. Materials

The materials of this research consisted of English data, Mandarin data, and dialect data. English data were composed of 144 CVC combinations, covering 8 cardinal vowels in different consonantal context, and were collected from both native speakers and Changsha learners. Mandarin and dialect data were designed for the investigation of L1 and L2's transfer effect, covering all monophthongs in both languages. Considering the syllable structure of mandarin and Changsha dialect, those data were in CV form instead of CVC.

2.1.3. Procedure

The recording was conducted in a quiet room, and the equipment was a laptop computer and a Sennheiser PC 166 microphone with built-in sound card. All sounds were recorded with CUHK-SIAT recording tool, a software designed for this project, at a 44 kHz sampling rate. The recording materials were presented on the computer screen in sequence, and each participant was instructed to pronounce the word they saw clearly and naturally.

2.2. Data Analysis

The speech data were automatically segmented and annotated with software, and manually adjusted by trained annotators based on acoustic cues and auditory impression. Then, acoustic parameters of vowels, namely F1, F2 values and duration, were extracted by Praat script and collected into Microsoft Office Excel. To eliminate the physiological differences caused by age and gender, original F1 and F2 values are normalized based on Lobanov's formula [21]:

$$F_{n(V)}^N = (F_{n(V)} - MEAN_n)/S_n$$

where $F_{n(V)}^N$ is the normalized formant value for formant n of vowel V, $F_{n(V)}$ is the raw formant value, $MEAN_n$ is the mean value of formant n for the speaker, and S_n is the standard deviation of formant n for the speaker.

For the similar reason, the durations were normalized with the following formula:

$$D^N = D/MEAN_D$$

where D^N is the normalized duration, D is the original duration, while $MEAN_D$ is the mean value of duration for the speaker.

3. Results

Acoustic analysis of the vowels produced by Changsha EFL learners and native speakers were conducted to account for L1 and L2's joint effect on L3. This section revealed the differences between English vowels produced by the 2 groups, and further discussed L1 and L2 vowel systems' impact on L3 vowel production, from both the spectral and temporal perspective. Due to the space limit, the discussion was mainly focused on 3 corner vowels, i.e., /a//i//u/, which can reflect the general pattern of the vowel systems [3].

3.1. Spectral measurements

The mean F1 and F2 values of the 3 vowels were plotted with NORM [22] on Figure 1 after normalization. Though all the vowels are included in normalization to avoid skewed normalized values, the vowels not discussed in this paper are removed from the plot for the purpose of simplification. From the plots below, several differences between the 2 groups can be spotted. In terms of general pattern, the backward tendency of learners' vowel system is quite noticeable, especially for male learners. In terms of individual vowels, it is obvious that /a/ and /i/ produced by both male and female learners are far away from their native counterparts: /i/ is pronounced in a more backward manner, while /a/ is more forward and downward.

To account for Changsha learner's performance, the vowel systems of their L1 and L2 were considered. It is observed that, L1 vowels are obviously more backward and downward than English vowels, which may to a large extent affect L3 pronunciation of /i/; while L2 vowels are comparatively forward, which might cause the forwardness of L3 /a/.
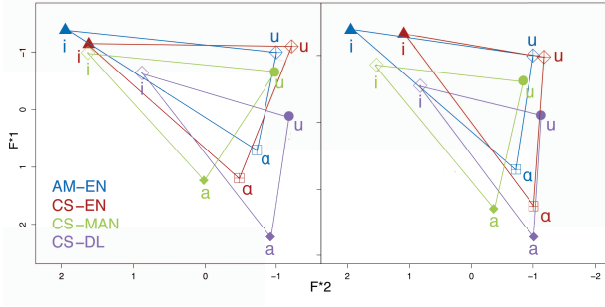
Figure 1: *Vowel plots of native speakers' English, Changsha learners' English, Mandarin, and dialect (the left for female, the right for male).*

To provide statistical evidence for the above-mentioned observation, several independent sample T-tests were conducted, examining the differences between vowels produced by native speakers and Changsha learners. T-test results indicate significant differences in F1 values of /a/ (p=.003), F2 values of /a/ (p=.033), F1 values of /i/ (p=.000), and F2 values of /i/ (p=.000) between the 2 groups, while there were no significant differences in F1 (p=.123) and F2 (p=.551) values of /u/, which confirms the observation above.

Then, 3 separate one-way ANOVAs were conducted to further prove L1 and L2's impact on English, by examining whether F1 and F2 of L3 are significantly different from those of their L1 and L2. Tukey's HSD was used to determine the nature of the differences, as were listed in Table 1. The statistical results show that, the F1 values of English /i/ and /a/ produced by Changsha learners have no significant differences from their dialect, and the F1 values of English /i/ and /a/, F2 values of English /i/ and /u/ produced by them have no significant differences with their Mandarin counterparts.

Table 1: *Results of Tukey's HSD.*

	F1	F2
i	**L1 (p=.998)**	L1 (p=.012)
	L2 (p=.118)	**L2 (p=.169)**
u	L1 (p=.008)	L1 (p=.058)
	L2 (p=.003)	**L2 (p=.083)**
a	**L1 (p=.657)**	L1 (p=.000)
	L2 (p=.067)	L2 (p=.000)

Though the results of one-way ANOVAs could present both Changsha learners' F1 and F2 performance, they merely indicated no significant differences, without telling to which extend L3 vowel production resembles L1 and L2 vowels. Therefore, Euclidean Distance was employed to compare the degree of similarity between them, as is adopted in [23][24]. The formula is as follows:

$$ED = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

where x_1 and x_2 refer to the normalized F1 values of the 2 vowels involved, and y_1 and y_2 refer to the normalized F2 values. Table 2 lists the Euclidean Distance of /i/, /u/ and /a/ between L1 and L3, L2 and L3 respectively, and reveals the following patterns: English vowels /i/ and /u/ produced by Changsha learners are more similar to their counterparts in Mandarin, while /a/ is closer to its counterpart in Changsha dialect.

Table 2: *Euclidean Distance between L1-L3, L2-L3 produced by Changsha learners.*

	ED: L1-L3	ED: L2-L3
i	0.485	**0.292**
u	0.547	**0.220**
a	**0.258**	0.448

To sum up, the L3 vowel /i/ is produced by Changsha EFL learners in a significantly backward and downward manner, /a/ in a significantly forward and downward manner, while /u/ is similar to that produced by native speakers without significant differences. In terms of transfer effect, L3 /i/ is more influenced by L2, Mandarin, and shares similar F1 and F2 with it; L3 /a/ is more influenced by L1, Changsha dialect, and shares similar F1 with it. And the correctly pronounced /u/ is closer to L2 than L1.

3.2. Temporal measurements for the tense-lax pairs

Though duration is not the only cue for the discrimination of English tense-lax vowels, it is an essential condition of such contrast, hence regarded as a criterion for English learners' phonetic acquisition.

The mean durations of vowels /i/-/ɪ/ and /u/-/ʊ/ produced by native speakers and Changsha EFL learners were represented on Figure 2 after normalization. It is obvious from the bar chart that Changsha EFL learners produce much shorter vowels than native speakers, which is further proved by a series of independent-sample T-tests. There were significant differences between mean duration of /i/ (p=.002), /i:/ (p=.005), /u/ (p=.000) and /u:/ (p=.000) produced by the two groups.
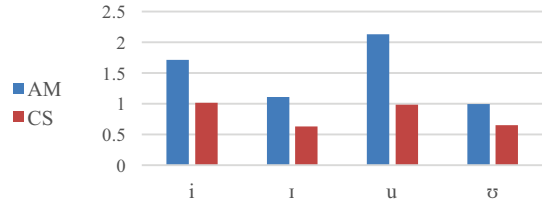

Figure 2: *Duration of 2 tense-lax vowel pairs produced by native speakers and Changsha learners.*

To figure out the transfer effect of L1 and L2 on duration, English /i/ and /u/'s counterparts in Mandarin and Changsha dialect were considered.
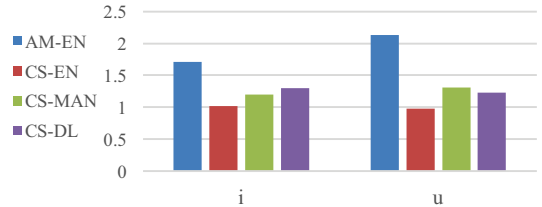

Figure 3: *Duration of /i/ and /u/ in native speakers' English, Changsha learners' English, Mandarin, and dialect.*

As is shown on Figure 3, /i/ and /u/ produced in both Changsha dialect and Mandarin are much shorter than those produced by native speakers, though slightly longer than their

L3 English. Therefore, it can be inferred that the short duration of /i/ and /u/ in L1 and L2 affect the duration of L3.

The data above show that Changsha learners tend to produce much shorter vowels than native speakers do, but are not sufficient in assessing their performance in vowel contrast, when the duration of tense and lax vowels are shortened simultaneously. Therefore, the duration ratios of tense-lax pair were calculated, which can indicate the degree of tense-lax contrast from the temporal perspective [2]. The /i-ɪ/ duration ratio of native speakers is 1.54, and that of Changsha learners is 1.61; the /u-ʊ/ duration ratio of native speakers is 2.13, and that of Changsha learners is 1.51. It is noticed that Changsha learners can contrast /i-ɪ/ in duration, with a similar duration ratio to native speakers. Though their duration ratio of /u-ʊ/ is smaller than that of native speakers, it is similar to native speakers' /i-ɪ/ duration ratio, so we still regard them to be able to contrast /u-ʊ/.

4. Discussion

The results in section 3 have revealed the differences between English vowels produced by Changsha learners and native speakers. In this section, L1 and L2 vowel systems' joint effect on L3 vowel production are further explained within theoretical framework.

In terms of spectral features, this research adopts Speech Learning Model (SLM) [20], which insisted that if the target phones are similar to those of previous acquired languages, learners are not able to distinguish these categories. Both L1 and L2's vowel systems are compared to L3, with the 3 vowels involved in this research plotted in Figure 4. It can be noticed from the plot that, English /i/ is similar to L2's /i/, English /u/ is similar to L1's /u/, while English /a/ is similar to L1's /a/, which is further proved by the Euclidean Distance listed on table 3.
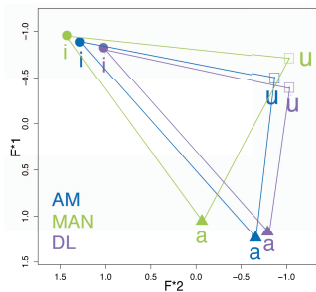

Figure 4: *Vowel plot of English, Mandarin, and dialect.*

Table 3: *Euclidean Distance between different vowel systems (English-dialect, and English-Mandarin).*

	ED: AM-DIAL	ED: AM-MAN
i	0.267	**0.157**
u	**0.191**	0.265
a	**0.147**	0.623

According to SLM, all the 3 vowels should be pronounced incorrectly and categorized into L1 or L2 vowel systems, since all of them have similar counterparts in previous acquired languages. And /i/ should be pronounced closer to L2, /u/ and /a/ closer to L1. The results in section 3 show that, Changsha EFL learners' production of /i/ and /a/ conforms to SLM's prediction, since both of them are mispronounced: L3 /i/ is categorized into L2 /i/, and L3 /a/ is categorized into L1 /a/. However, /u/ is produced without significant differences between Changsha learners and native speakers, and in a manner more similar to L2 rather than L1, which is unexplainable with SLM, and deserves further investigation.

In terms of duration, Changsha learners tend to produce shorter vowels than native speakers do, and they can make a general contrast in English tense-lax pairs, though their duration ratio of /u-ʊ/ is smaller than that of native speakers. Their performance in duration can be partially accounted for by L1 and L2 transfer, as is indicated in section 3. But there is also possibility that the difference is caused by native speakers' idiosyncratic variations (especially the long duration of /u/), since the native speaker corpus is small compared with that of Changsha EFL learners.

The analyses of English vowels produced by Changsha EFL learners so far come to an end, but the joint effect of dialect and Mandarin spotted in this research could offer some inspirations to relevant studies. International research on English accent would always investigate the participants' language backgrounds carefully ahead of further analysis, and make out strategies on pronunciation correction accordingly. By now, most domestic research on vowel production of EFL learners from different dialectal areas have focused on dialect's transfer effect. However, when dialects are generally losing their dominance in daily communication, the studies are not comprehensive without taking Mandarin into consideration. After all, the ultimate goal of such research is to provide suggestions for pronunciation correctness to EFL learners, the majority of whom are speakers of both dialects and Mandarin. The language backgrounds of Chinese EFL learners from different dialectal areas are more complex than previously assumed, so they should be investigated more carefully, together with the role English plays: L2 as adopted in previous studies, or L3 as claimed in the present research.

5. Conclusions

This research investigated the differences between English vowels produced by Changsha learners and native speakers, and further discussed dialect and Mandarin's joint effect on English vowel production, from both spectral and temporal perspectives. An "English-as-L3" view is adopted in this research, with Changsha dialect as L1 and Mandarin as L2. The results are as follows: Spectrally, Changsha English learners would produce English vowel /i/ in a significantly backward and downward manner, whose F1 and F2 resemble its Mandarin counterpart; /a/ is produced in a significantly more forward and downward manner, which is influenced by Changsha dialect with a similar F1; while /u/ is similar to the American version without significant differences. Among the 3 vowels, Changsha learners' performance of /i/ and /a/ can be explained with Speech Learning Model, while L3 /u/ production shows unconformity with the model, which deserves further investigation. Temporally, Changsha learners tend to produce shorter vowels due to L1 and L2's transfer effect, but still, they can make a contrast in English tense-lax pairs, even though there is no such contrast in their L1 and L2.

6. Acknowledgements

7. References

[1] J. C. Ingram and S.-G. Park, "Cross-language vowel perception and production by Japanese and Korean learners of English," *J. Phon.*, vol. 25, no. 3, pp. 343–370, 1997.

[2] L. Rallo Fabra and J. Romero, "Native Catalan learners' perception and production of English vowels," *J. Phon.*, vol. 40, no. 3, pp. 491–508, 2012.

[3] M. Wester, L. M. Luisa García, and M. Cooke, "/u/-Fronting in English Speakers' L1 but not in their L2," in *proceedings of ICPhS*, 2015.

[4] H. Zhai and J. Zhao, "An introduction on the influences of Chinese dialects on English phonetic acquisition," *Foreign Lang. World*, no. 1, pp. 88–95, 2015.

[5] H. Chen, "Theory, methodology, and direction of second language speech learning," *Foreign Lang. Their Teach.*, no. 6, pp. 12–15, 2011.

[6] Y. Jiang, "Acousitic analysis of English learners from Min and Wu dialectal areas," *Foreign Lang. Res.*, no. 4, pp. 36–40, 2010.

[7] F. Wang, H. Zhai, and Y. Jia, "An experimental phonetic study on acquisition of English vowel segments by learners of Shandong dialect: a case study of Weifang, Mengyin, Qingdao and Rizhao," in *proceedings of O-COCOSDA*, 2015, pp. 325–329.

[8] D. Qian, Y. Jia, A. Li, and L. Xu, "A Comparative Experiment on the Prosodic Features between Ningbo EFL Learners and American Native Speakers—in the Case of Yes-No Questions," in *proceedings of ISCSLP*, 2014.

[9] H. Wang, Y. Jia, A. Li, and Y. Tang, "Intonation Patterns of Yes-No Questions with Narrow Focusing of Chinese Northeastern EFL Learners ---A Case Study of Changchun and Dalian," in *proceedings of O-COCOSDA*, 2014, pp. 212–217.

[10] P. Shao, Y. Jia, and A. Li, "Intonation Patterns of Tag Questions for Chinese EFL Learners from Shandong Dialect," in *proceedings of NCMMSC*, 2013, pp. 47–52.

[11] X. Lian, Y. Jia, W. Duan, and A. D. Roeck, "Phonetic Realization of Narrow Focus by Beijing EFL Learners in English Yes-No Sentence," in *proceedings of O-COCOSDA*, 2013.

[12] N. Zhi, A. Li, and Y. Jia, "The role of L1 production compactness on the L2 production accuracy," in ICPhS, 2015.

[13] M. Wu and J. van de Weijer, "The perception of English nasal coda contrast by bilingual Changsha Chinese learners of English," in *proceedings of O-COCOSDA*, 2015, pp. 284–288.

[14] S. Murphy, "Second language transfer during third language acquisition," *Work. Pap. TESOL Appl. Linguist.*, vol. 3, no. 1, pp. 1–21, 2005.

[15] J.-M. Dewaele, "Lexical inventions: French interlanguage as L2 versus L3," *Appl. Linguist.*, vol. 19, no. 4, pp. 471–490, 1998.

[16] J. C. Amaro and J. Rothman, "On L3 acquisition and phonological permeability: A new test case for debates on the mental representation of non-native phonological systems," *IRAL - Int. Rev. Appl. Linguist. Lang. Teach.*, vol. 48, no. 2–3, pp. 275–296, 2010.

[17] C. Bardel and Y. Falk, "The role of the second language in third language acquisition: the case of Germanic syntax," *Second Lang. Res.*, vol. 23, no. 4, pp. 459–484, 2007.

[18] B. Hammarberg, "Roles of L1 and L2 in L3 production and acquisition," in *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*, 2001.

[19] J. Norman, Chinese. Cambridge University Press, 1988.

[20] J. E. Flege, "Second language speech learning: Theory, findings, and problems," *Speech Percept. Linguist. Exp. Issues cross-language Res.*, pp. 233–277, 1995.

[21] B. M. Lobanov, "Classification of Russian vowels spoken by different speakers," *J. Acoust. Soc. Am.*, vol. 49, pp. 606–608, 1971.

[22] T. Erik R and T. Kendall, "NORM: The vowel normalization and plotting suite," 2007. [Online]. Available: http://ncslaap.lib.ncsu.edu/tools/norm/.

[23] H. Feng and L. Zhao, "An Empirical Study of Phonetic Transfer in English Monophthong Learning by Tibetan (Lhasa) Speakers," in *proceedings of O-COCOSDA*, 2015, pp. 181–185.

[24] K. Yazawa, Y. Ozaki, G. Short, K. Mariko, and Y. Sagisaka, "A study of the production of unstressed vowels by Japanese speakers of English using the J-AESOP corpus," in *proceedings of O-COCOSDA*, 2015, pp. 96–100.