Typology of English Monophthongs by EFL Learners from Wu Dialectal Region- A Case Study of Ningbo and Shanghai

Yuan Jia¹, Yu Wang¹,²

¹ Institute of Linguistics, Chinese Academy of Social Science, Beijing, China
² Department of foreign languages and Literature, Zhengzhou Chenggong University of Finance and Economics, Zhengzhou, China

Abstract

The present paper investigates acoustic features of English vowels by EFL learners (English as a Foreign Language) from Ningbo (NB) and Shanghai (SH) dialectal regions, both of which belong to the Wu dialect. Eleven English monophones, i.e., /ɪ/, /æ/, /a/ etc. are selected as target samples and their corresponding F1&F2 formants are employed as parameters to approach the research aim. Through the acoustic results, we focus on exploring the degree of phonetic transfer of dialects (L1) onto English (L2). The Speech Learning Model (SLM) is further adopted to examine the differences caused by the dialectal accent. Results show that, with regard to the tongue position of vowels, EFL learners from these two dialectal regions and American (AM) native speakers do show a great divergence. Specifically, /ɪ/ is affected by NB and SH dialects, which can be explained by SLM. On the other hand, /æ/ and /a/ produced by SH learners is similar to that of American speakers. Besides, NB and SH learners produce shorter vowels in duration, due to dialects’ transfer effect, but can still make tense-lax contrasts in /i/-/ɪ/, /u/-/u/.

Index Terms: Wu dialect, Vowel formant, Convergence, Divergence, Language transfer

1. Introduction

English teaching has been implemented in various parts of China. Because of the importance of spoken English, researchers have been interested in second language (L2) speech learning. Furthermore, in the field of second language (L2) acquisition, learners’ knowledge of their first language (L1) has long been regarded as an important factor in phonological acquisition [1]. The feature hypothesis posits that if a phonetic feature in L2 is not meaningful in L1 phonology, then learners acquire the feature with lower accuracy [2]. The similarities and dissimilarities between L1 and L2 are considered to influence the process of L2 phonology acquisition, which is known as L1 transfer.

Previous literature on the English acquisition of Chinese speakers from the phonetic perspective mainly explored the super-segmental features [Duan [3]; Wang [4]; Qian [5]; Hu [6]]. More recent research has been focused on the acoustic characteristics of English vowels produced by EFL learners in dialectal regions (Jia [7]; Tang [8]; Li [9]; Ruan [10]; Chen [11]; Wang [12]; Jiang [13]; Shi [14]; Yang [15]).

The overview of the previous studies demonstrated that many researches concentrated on the second language acquisition in supra-segmental features or the influences from Mandarin or dialect. Moreover, the past study on dialects mainly focused on a particular dialect, and little attention had been paid to the convergences and divergences among dialects from the typological perspective.

In the late 1950s, Greenberg [16] proposed a new field of linguistic research-linguistic typology, which has evolved into an important subject of contemporary linguistics. Based on the cross-linguistic comparison, linguistic typology explores the universal laws of languages through observation, data collection, and comparison of a large number of languages, and discusses major constraints in forming these universal laws. In the process of language research and acquisition, if we ignore the similar and dissimilar features across languages, it would not only be difficult to find the universal laws, preventing a greater contribution to the language theory, but also hamper the language acquisition effectiveness. Recently, the application of language typology theory and research method to English acquisition has become a trend gradually. The main approach is to explore the common factors and differences between English and Chinese, so as to uncover the universal laws between each other.

The present study explores the transfer of L1 onto the L2 acquisition within the formwork of (SLM) [17]. We aim to investigate the convergences and divergences in the acoustic features of English pure vowels by EFL learners from the Wu dialectal region and native speakers from the typological perspective. Specifically, we aim to find out whether the vowel systems of dialects would influence the English vowel output of EFL learners.

2. Methodology

2.1. The vowel inventory of Guanhua dialect

The selected dialect regions in this study are Ningbo (NB) and Shanghai (SH), both of which belong to the Wu dialect. According to its internal differences within Wu dialect, it could be divided into eight sub-dialects, specifically, Taihu, Taizhou, Oujiang, Wuzhou, Chuqu, Xuanzhou, Yanzhou, Huizhou [18]. Ningbo and Shanghai dialects are important branches of Taihu dialect. Compared with English, NB dialect vowels include ten pure vowels [i u e o ɛ a ɔ u ə ɪ], which can appear in (C)V structure. Additionally, there are nine vowels [ʌʊ ɛ uɜ ɤ y u] in SH dialect [19]. In contrast, English has twelve oral monophthongs. Phonologically speaking, Chinese dialects do not contrast in a long vs. short vowel distinction, while there are distinctions in English. Such observations, along with further statistical analysis of data
collected from the experiments, can be expected to be more self-evident and reliable in guiding the L2 learners to acquire a more native like production of English pure vowels.

2.2. Data collection

2.2.1. Stimuli

The stimuli for this study consisted of English data and dialect data. Covering 11 English pure vowels, English data were composed of 110 CVC combinations, preferably everyone with an initial of plosive, fricative and affricate and the same were true of the syllable codas. All of them collected from native speakers, Ningbo and Shanghai learners. Dialect data were designed for the investigation of L1’s transfer effect. Due to space limit, the similar vowels were mainly focused on 3 corner vowels, i.e. /i/, /u/, /æ/, which can reflect the general pattern of the vowel system [20]. Considering the syllable structure of dialects, those data were in CV form instead of CVC.

2.2.2. Participants

Twelve English learners from Ningbo and Shanghai regions participated in this experiment, including 3 females and 3 males in each dialect, are all born and raised in their own dialectal city where they have been learning English for more than 10 years. In addition, they are skillful in using their dialect without speech and hearing disorders. 6 American native speakers, 3 females and 3 males, spoke general American English without a regional accent.

2.2.3. Recording

Recording was conducted in a sound-treated booth, with headset microphone (Sennheiser PC 166) connected to a laptop. The sampling frequency was set at 16000Hz, and bit depth 16 digits. The speech data were automatically segmented and annotated with software, and manually corrected by trained annotators based on acoustic cues and auditory impression.

2.3. Data analysis

Acoustic parameters of vowels, namely F1, F2 values and duration were extracted using a PRAAT script and collected into EXCEL. To eliminate the physiological differences caused by age and gender, the data of vowels were plotted in NORM [21], and the Bark Difference Metric method was used to plot the vowels trajectories. The formula is as following:

\[ Z_i = 26.81/(1 + 1960/F_2) - 0.53 \]  

Where, \( F_i \) (i=1, 2, 3) denotes the value of vowel formant. \( Z_i \) refers to the normalization value of formant frequency extracted from the sampling points of the vowel. It then computes the differences \( Z_3 - Z_4 \) and \( Z_3 - Z_2 \); \( Z_3 - Z_2 \) is used to plot the normalized front-back dimension and \( Z_3 - Z_1 \) is used to plot the normalized height dimension.

For the similar reason, all duration are calculated through the following formula:

\[ X_{i} = (X_{i} - X_{\text{min}})/(X_{\text{max}} - X_{\text{min}}) \]

Within the formula, X is the duration, and \( X_i \) (i=1, 2, ..., n) is individual’s duration. While, \( X_{\text{max}} \) and \( X_{\text{min}} \) are the maximum and minimum of the duration in each utterance respectively.

3. Results and Findings

In this part, the acquisition of English pure vowels produced by leaners in Wu dialectal region and native speakers was explored from a typological perspective. Further, acoustic analysis of the vowels produced by EFL learners in the two dialects and native speakers were conducted to account for the L1 effect on L2. This section revealed the divergences between English vowels produced by 2 groups, and further discussed L1 vowel systems’ impact on L2 vowel production, from both spectral and temporal perspective. Based on this observation, as for the production of English vowels, the common and different characteristics between the two dialects can be found out.

3.1. Spectral properties

Illustrated in figure 1 is the formant chart of the same set of English pure vowels plotted to show the systematic differences between the outputs of EFL learners in NB and SH dialectal regions and native speakers, in which the Y-axis represents the height of the vowels, and the X-axis is the vowels’ front-back.

![Figure 1: Vowel Plots of native speakers’ and NB and SH learners’ English](image)

Firstly, the result of the t-test reveals that, in terms of tongue position, Ningbo EFL learners do differ noticeably from the native English speakers in front vowels (/i, ɪ, ɛ, æ/) (pF1<.05, pF2<.05); extremely for /i/, their front vowels take a lower and more backward tongue position, especially the high front vowel /i/ (pF1<.05, pF2<.05) and the low front vowel of /æ/ (pF1<.001, pF2<.05).

Secondly, great disparities were found with central vowels. EFL learners take obviously lower and more backward position than native speakers for the production of the central vowel /a/ (pF1<.001, pF2<.01) and /u/ (pF1<.01, pF2>0.05).

Thirdly, a close look at the high back vowels reveals noticeable inter-group disparity in the frequency value of F2 among subjects chosen for the experimental study (pF2<.05). Specifically speaking, the high back tense vowel /u/ (pF1<.001, pF2<.05) articulated by EFL learners is lower and further front in tongue position. By contrast, the central-back vowel /æ/ (pF1<.05, pF2<.01), with its output values lower in F1 and higher in F2 than the native speakers respectively, they are found to hold a much higher and backward tongue position.

Similarly, great disparities are to be found between learners from SH dialect and native speakers. And it is interesting the characteristics of vowels produced by English
learners in SH dialectal region are the same as those of learners in Ningbo to some extent. Especially, the vowels pronounced by EFL learners take lower position than native speakers. In addition, for the front vowels, the more backward tendency can be observed, while concerning the central except for /ɔ/ and back vowels articulated by EFL learners, it shows more frontward in tongue position.

On the other hand, the distribution of vowels is more scattered than that of NB EFL learners and the most striking aspect of figure 1 is the narrowed separation of English tense and lax vowels produced by EFL learners. That is to say, NB EFL learners cannot distinguish between /i/-/ɪ/, /u/-/ɛ/ and /a/-/ɑ/ in the same way as native speakers do. Rather, they distinguish /i/-/ɪ/ (pF1<.01, pF2>.05), /u/-/ɛ/ (pF1<.01, pF2>.05) by the vowel height, while rely on the backwardness of the tongue position to mark the difference between /a/-/ɑ/ (pF1>.05, pF2>.05). Besides, statistics also indicate that EFL learner can hardly distinguish /a/ from /ɑ/ (pF1>.05, pF2>.05).

However, there is showing the sign of tense-lax vowels merging in SH dialect. Concretely, it is not difficult for SH learners to differentiate /i/-/ɪ/ (pF1<.001, pF2<.05), /u/-/ɛ/ (pF1<.01, pF2>.05). Moreover, SH EFL learners cannot differentiate /i/-/ɪ/ (pF1>.05, pF2>.05), /u/-/ɛ/ (pF1>.05, pF2>.05), /a/-/ɑ/ (pF1>.05, pF2>.05).

In contrast, native English speakers make more notable differentiations between tense and lax vowels. They were witnessed to integrate both factors of height and backwirdness to distinguish /i/-/ɪ/ (pF1<.01, pF2<.001) and /u/-/ɛ/ (pF1<.01, pF2>.05). In addition, when distinguishing the vowel of /i/-/ɪ/ (pF1<.05, pF2>.05) and /a/-/ɑ/ (pF1<.01, pF2>.05), American native speakers mainly focus on the vowel height.

3.2. Transfer from dialects

Phonologically speaking, the vowels /i/, /u/ and /ɑ/ in English, the vowels /ɪ/, /ɛ/ and /ɑ/ (/a/ in Shanghai dialect equals to /ɑ/) in dialects are regarded as similar vowels [20]. This part is mainly to investigate the influence of the similar sounds of dialect happened in English vowel inventory on the acquisition of English. Within the figure, the X-axis represents the vowels’ front-back, and the Y-axis is the height of the vowels.

Figure 2: Vowel Plots of native speakers’ and NB and SH learners’ English, dialect

As shown in figure 2, the distribution of /i/, /u/ and /ɑ/ in English produced by EFL learners is close to the development of the system of dialects, that is to say, /i/, /u/ and /ɑ/ can find their counterparts in dialects respectively.

Firstly, a series of one-way ANOVA was conducted to test the impacts on English vowel production resulted from Ningbo dialect, by examining whether the F1 or F2 values of English vowels produced by Ningbo learners are significantly different from their counterparts. If the tests show no significant differences, it is to be assumed that the English vowels under observation have been assimilated to their counterpart either in Ningbo dialect.

The results show that there is no great divergence between NB dialect and English bearing local accent in terms of their F1 & F2 values on the production of /i/-/ɪ/ and /u/-/ɛ/ (pF1(/i/-/ɪ/)>.05, pF2(/i/-/ɪ/)>.05, pF2(/u/-/ɛ/)>.05). Additionally, both F1 & F2 of English /ɑ/ produced by Ningbo learners are significantly different from those of the similar vowels /ɑ/ found in their dialect (pF1<.01, pF2<.01). Therefore, it could be reasoned out that, among the English monophthongs targeted by the experiment for acoustic analysis, /i/ and /u/ tend to be drawn to their counterparts in Ningbo dialect.

Secondly, in view of the intra-lingual transfer induced by SH dialect, its dialects’ /i/ show great sign of phonetic transfer since the results indicate that there is no significant difference of the F1 and F2 value on the production of /i/ (p<.05) between English and dialects spoken by EFL learners. While about the examination of English /u/ and /ɑ/ articulated by SH learners, the analysis result manifests the significant difference of F1 and F2 values on the production of /u/ (p<.05) and /ɑ/ (p<.05) with the dialect vowels /ɛ/ and /ɑ/, which makes it safe to say that the English /ɑ/ cannot be explained in terms of phonetic transfer. Therefore, it could be reasoned out that, among the English monophthongs targeted by the experiment for acoustic analysis, /i/ tends to be drawn to their counterparts in Shanghai dialect.

To better prove the phonetic transfer from the four dialects respectively, a measure of Euclidean distance (ED) was derived on the basis of rescaled normalized formant values.

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

Where \(x_1\) and \(x_2\) refer to the normalized F1 of the two vowels, \(y_1\) and \(y_2\) correspond to the normalized F2 of the two vowels. If ED of an English vowel by EFL learners to dialect is bigger than that of native speakers, it indicates that EFL learners can pronounce the sound well. On the contrary, it may be influenced by dialect.

| Table 1. Euclidean Distance between L1-L2 produced by NB and SH EFL learners. |
|-------------------------------|---|---|
|                              | NB | SH |
| i                             | 0.45 | 0.50 |
| u                             | 0.67 | 1.94 |
| a                             | 1.73 | 1.99 |

| Table 2. Euclidean Distance of English produced by NB and SH EFL learners and native speakers. |
|-------------------------------|---|---|
|                              | NB | SH |
| i                             | 1.28 | 1.05 |
| u                             | 1.37 | 1.26 |
| a                             | 0.61 | 0.30 |

Table 1, 2 provide statistical evidence for the above-mentioned observation. We can see clearly that, English
vowels /i/ and /u/ produced by NB learners are more similar to their counterparts in dialect except /ɑ/. For SH learners, only /ɪ/ can be found similar counterpart in dialect, that is to say, /u/ and /ɑ/ cannot be explained in terms of phonetic transfer, deserve further investigation. To sum up, it is reasonable to argue that the Wu dialectal students’ L2 learning, L1 could affect their L2 vowel system.

3.3. Temporal structure

The normalized durations of English tense-lax vowels and the three similar vowels from NB and SH dialects were summarized in the bar chart in figure 3.

![Figure 3: Temporal structures of English tense-lax vowels and similar vowels in English learners and native speakers.](image)

It is obvious from the bar chart that Ningbo and Shanghai EFL learners produce much shorter vowels than native speakers, which is further proved by a series of independent-sample T-tests. There are significant differences between mean duration of /ɪ/ (p<0.05), /u/ (p<0.01), /ɑ/ (p<0.05), /ɑ̃/ (p<0.05), /a/ (p<0.05), /u̯/ (p<0.05). What is more, it can be seen also clearly that the distinction of duration between tense and lax vowels by American native speakers is not as remarkable as that of EFL learners (.01<p<.05). This can be further demonstrated by a close observation of the mean duration of each individual vowel segment, as listed in the following table.

Additionally, as is shown on figure 3, the similar vowels /ɪ/, /u̯/ and /ɑ/ /a/ (or /ɑ̃/) in Shanghai dialect equals to /a/ produced in both NB and SH dialect are much shorter than those produced by native speakers. Therefore, it can be inferred that the short duration of tense-lax in L1 affect the duration of L2 for NB and SH learners.

When the duration of tense-lax vowels are shortened or lengthened simultaneously, the data above are still not sufficient in assessing their performance in vowel contrast. Therefore, the duration ratios of tense-lax pair were calculated, which can indicate the degree of tense-lax contrast from the temporal perspective.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>AM</th>
<th>NB</th>
<th>SH</th>
</tr>
</thead>
<tbody>
<tr>
<td>i-I</td>
<td>1</td>
<td>1.60</td>
<td>2.22</td>
</tr>
<tr>
<td>u-o</td>
<td>1.53</td>
<td>1.68</td>
<td>1.43</td>
</tr>
<tr>
<td>ɔ-ɒ</td>
<td>1.05</td>
<td>1.48</td>
<td>1.06</td>
</tr>
</tbody>
</table>

To sum up, after the analysis of the formant and duration, it indicates that, native speakers make the distinction between tense-lax vowels mainly by the tongue position; while for Wu dialectal EFL learners, they exhibit significant difference on the duration of vowels.

4. Discussions

The present paper looks into research Wu dialectal students’ foreign accent of English from the perspective of language transfer. Based on the calculation of formant frequency and duration between English and dialects produced by Ningbo and Shanghai learners, the acoustic similar and dissimilar features can be obtained.

Specifically, the results have presented a downward tendency in producing English pure vowels when compared with native speakers. Meanwhile, it also revealed the difficulties in distinguishing English tense-lax vowels produced by EFL learners. What causes this phenomenon? Feige [22] claims that if the target phones are similar to those of previous acquired languages, learners cannot distinguish these categories. As in the case of /i/, /a/ and /ɑ/ of English can find their counterparts in NB and SH dialect, except for /ɑ/ and /ɑ̃/ articulated by SH learners. According to the above analysis, the realization of similar vowels in English from NB EFL learners’ /ɪ/, /u̯/ and /ɑ/ always fall into the categories already established in their L1. For SH learners, only /ɪ/ can be found similar counterpart in dialect, that is to say, /u̯/ and /ɑ/ cannot be explained in terms of phonetic transfer. This deserves further investigation.

In terms of duration, parametric analysis of the six target English vowels articulated by EFL learners indicates that Ningbo and Shanghai EFL learners are not doing well in discriminating vowels defined by tenseness, instead they can make a general contrast from the temporal perspective. Additionally, in terms of the duration of similar vowels in L1 and L2 from NB and SH learners is shorter than that of vowels in L2 produced by native speakers, which inferred the short duration of tense-lax in L1 can affect the duration of L2.

This study still has limitations in many aspects. In further study, a larger corpus and more regions in Wu dialectal area would be built up and other similar vowels between Wu dialect and English need to be conducted. Furthermore, the specific method to help students in Wu dialect to improve their pronunciation would be also established.

5. Conclusions

With a comparative study on the formant features and duration of English monophthongs, this research illustrate the convergences and divergences in vowel production of EFL learners from Wu dialectal region, in a case study of Ningbo and Shanghai. Meanwhile, the influences from Dialect are researched. It not only provides quantifiable hints that could guide the L2 learners to improve their pronunciation, but also offers a foundation for further investigations into the effects resulted from phonetic transference.

6. Acknowledgements

This research is financially supported by the National Major Social Sciences Foundation of China under Grant 201515ZDB103.
7. References


