Research on Vowel Production by EFL Children Learners from the Shanghai Dialect Area

Wanying Xiao

Department of Literature, Faculty of Humanities, The Open University of China, Beijing, China
Email: Xiaowy@ouchn.edu.cn (W.Y.X.)
Manuscript received August 30, 2024; revised October 11, 2024; accepted November 27, 2024; published February 28, 2025.

***Abstract*—In order to explore the acquisition of English vowels by children with Shanghai dialect, this paper investigates 20 children whose native language is Shanghainese, and also investigates the pronunciation of 20 native American native English speakers. Considering the results of the acoustic analysis, we examined the accuracy of pronunciation of English learners in Shanghai, and the order of learning of English vowel /i, ɪ, ɛ, æ, ɑ, ɔ, u, ʊ, ʌ, ɜ/ and the influence of Shanghainese. Based on the analysis, the possible acquisition order of Shanghai children could be /i/ , then comes /ɑ, ɔ, ɪ, æ, ɜ, ʌ/ and / ɛ, u, ʊ/.. The experimental results partially support Flege's Speech Learning Model theory (SLM). The similarity of vowels may have less influence on the English vowel acquisition, in fact, the process of language learning is actually the process of getting rid of the influence of mother tongue interference. This paper can be a reference for the English vowel teaching and learning in different dialect areas in China and also other countries in the world, especially in Shanghai. Besides, this paper also provides a research reference for Chinese students' second language acquisition of English with different dialects.**

***Keywords*—vowel acquisition, EFL children learners, Shanghai dialect area, experimental study**

I. Introduction

Many researchers explore English vowel acquisition. Different languages have different sound systems, mainly reflects on features of sound and rules of pronunciation. The pronunciation of English among Chinese learners may be influenced by their mother tongue. As a special dialect different from Mandarin characterized by its unique pronunciation, vocabulary, and grammar, Shanghai dialect, also known as Shanghainese, is a variety of Wu dialectal region spoken primarily in the city of Shanghai and its surrounding regions.

Therefore, through acoustic analysis and phonological comparison of English vowels between Shanghai children English learners and native English speakers, we hope to investigate the accuracy of pronunciation of Shanghai children English learners, so as to analyze the influence of Shanghai dialect on English vowel pronunciation and provide references for English teaching. The section should be organized as:

(1) What are the phonological system of Shanghai dialect and American English?
(2) What are the materials and method of this research?
(3) What are the results of accuracy of Shanghai children English learners?

II. Literature Review

There are lots of researches on English as a second language in the world, such as the vowel production and perception by native Korean adults and children [1]. In China, there are also plenty of researches about the acquisition of English vowel with people from different dialect areas. For people in Shanghai dialect area. Li (1957), Wu (2002) and Ji (2013) focused on the acquisition of Shanghai English learners [2–4].

There are several hypothesis of second language acquisition, the Speech Learning Model created by Flege (1995) might be one of the most influential theory [5]. He suggested that unfamiliar phones are easier to acquire than similar phones. While Markedness Differential Hypothesis (MDH) by Eckman (1977) suggested that unmarked phones were learned earlier than marked ones, the weakly marked phones were learned earlier than strongly marked phones [6]. Whereas, as the MDH showed, /i, u, ɑ/ are unmarked phones in many languages in the world, are these three phones all easy to learn? How much influence does the similarity of two languages have on the acquisition of vowels? Selinker [7] proposed the concept of "Interlanguage" in 1969 and published in 1972. His hypothesis regarded the language system constructed by learners independently as interim grammar which is a dynamic language system that gradually transitions from mother tongue to target language. Are there an "Interlanguage" system for Shanghai children? Does the system different from the adults? Besides, many scholars supported the hypothesis of the critical period of language learning which was put forward by Lenneberg (1967) [8]. Recent studies of the human brain have also shown that Broca's area, the region responsible for language learning, is sensitive between the ages of 4 and 12. However, are there any specific characteristics of children's vowel acquisition at this stage in Shanghai dialect? To what extent does Shanghai dialect affect children's English vowel pronunciation? There are still many unanswered questions.

Although the existing researches have produced fruitful results, most of them were based on contrastive analysis and error analysis, and few involve the new transfer theory and the MDH theory. In addition, the research object is relatively simple, and the influence of dialect background, age, gender, learning time and other factors on vowel acquisition is not fully considered. Most of the research methods are subjective experience introduction, judgment accuracy or simple acoustic analysis, lacking the support of statistical data and accurate measurement of acoustic parameters.

III. Materials and Methods

A. Participants

This experiment mainly investigated the English vowel pronunciation of 20 Shanghai children and 20 Americans. Participants of Shanghainese were recruited locally, participants of Americans were recruited both locally and

online. Shanghainese are natives born and raised in downtown Shanghai. All Americans were born and raised in the United States excluding native speakers of New England and Southern dialects (Group A). Children are from age 11 to 14 randomly selected from primary and secondary schools in Shanghai and are consists of 10 boys and 10 girls (Group B). (see Table 1)

Table 1. Introduction of participants

Group	Participants	Age	Number of people
A	Native American speakers	20–45	20
B	Children	11–14	20

B. Reading Materials

The pronunciation material of English is 30 words in three groups of 10 syllables containing English vowels, beginning with [h] and ending with [d]. The subjects were asked to read the sentence "Say__twice" three times. The first group of words is preferred in the data analysis, if there is a pronunciation error or the data cannot be analyzed, then the other two groups of words are used. (See Table 2)

The pronunciation materials for Shanghai children learners include both English word lists and Shanghai dialect word lists, consisting of 27 words in three groups of nine syllables containing monophthong in Shanghai dialect. The same tone 34, preceded the vowels with stop or fricative are used in order to reduce the influence of consonants and tones on vowels. Shanghai dialect word lists are presented in Chinese characters, the subjects were asked to read the sentence "I say to you three times." (See Table 3)

Table 2. English word lists

Vowel	/i/	/ɪ/	/ɛ/	/æ/	/ɑ/	/ɔ/	/u/	/ʊ/	/ʌ/	/ɝ/
Word	heed	hid	head	had	hard	hawed	who'd	hood	hud	heard
	bee	bid	bed	bad	bar	door	too	food	but	bur
	see	lid	dead	dad	car	four	do	could	cut	fur

Table 3. Shanghai dialect word lists

Vowel	/i/	/ɛ/	/ʌ/	/u/	/ɔ/	/ø/	/ɤ/	/ɿ/	/y/
Word	pi^{34}	$pɛ^{34}$	$pʌ^{34}$	pu^{34}	$pɔ^{34}$	$pø^{34}$	$pɤ^{34}$	$sɿ^{34}$	y^{34}
	ti^{34}	$tɛ^{34}$	$tʌ^{34}$	tu^{34}	$tɔ^{34}$	$tø^{34}$	$tɤ^{34}$	$ts^{h}ɿ^{34}$	$ɕy^{34}$
	pi^{34}	$kɛ^{34}$	$kʌ^{34}$	ku^{34}	$kɔ^{34}$	$kø^{34}$	$kɤ^{34}$	$tsɿ^{34}$	$tɕ^{34}$

C. Procedure

The experimental equipment was a Lenovo laptop installed with Cooledit Pro, which used an external sound card and capacitor microphone. The sampling frequency of mono digital voice signals was 44100HZ and the sampling accuracy was 32 bits floating point. Cooledit software was used for recording, and the obtained data was analyzed by Praat. The vowel data is first annotated by praat, and then the modified script Analyze tier created by Daniel Hirst is adopted. According to the practice of Morrision & Escudero (2007) [9], F1 and F2 values of the stationary segment between 25% and 75% of the vowel duration are taken. In order to exclude the influence of the rollover effect on the value of the formant, the values of F1 and F2 in the stable segment between 25% and 50% of the vowel duration are obtained in retroflex vowel, and the mean value of the three pronunciations of each sound is obtained. After that, Excel, SPSS and other software are used for experimental calculation and statistical diagram of the data.

D. Data Analysis

Firstly, according to Flynn and Adank, the most effective ways of normalization are vowel-extrinsic normalization, formant-intrinsic normalization and speaker-intrinsic normalization. Among them, the z-score method of Lobanov (1971) [10] has the best effect on eliminating biological differences such as gender and age of individuals. Therefore, z-score method were used to normalize the original Hertz value of a single speaker, so as to eliminate the influence of gender and age of men and women on the data of the speaker, while retaining the necessary linguistic information. The formula is (F-μ) / δ, where μ and δ represent the mean and standard deviation of all vowels, respectively.

Secondly, in order to make the normalized data more truly reflect the actual frequency values of F1 and F2, we convert the normalized values into frequency values according to the formula proposed by Thomas and Kendall (2007) [11].

F'1=250 + 500 (FN1-FN1 MIN) / (FN1MAX-FN1 MIN)

F'2=850 +1 400 (FN2-FN2 MIN) / (FN2MAX-FN2 MIN)

FNi is the normalized value of Fi, FNiMIN and FNiMAX are the minimum and maximum values of Fi for all speakers in each group, respectively.

Finally, in order to get close to the actual perception of speech by human ears, we then adopted the method of converting frequency value (Herz) to Bark value, and normalized the formant data by NORM. This method was first proposed by Traunmuller (1990) [12] and further optimized by Bladon, Henton, Pickering (1984) [13] and Thomas and Kendall (2007):

$$F_iN = 26.81/(1+1960/F_i)) - 0.53$$

"Fi" represents the measured formant (i= 1, 2, 3), and N represents the normalized value.

IV. Result and Discussion

The Chinese pronunciation level of foreigners can be evaluated by the accuracy of the pronunciation. Therefore, we then compare the English vowels of Shanghai English learners and Americans, and carry out an accuracy analysis

We first measured the English vowel pronunciation data of 20 Shanghai children, including 10 boys and 10 girls. In the process, stem-leaf graph method was used to eliminate an outlier value to ensure the overall objectivity of the data. On this basis, the average values of F1 and F2 of each vowel are calculated respectively. As a control group, we also measured data of 20 American speakers, including 10 male and 10 female speakers. After that, we compared the English vowel pronunciation data of both groups with statistical method and analyzed the acoustic vowel graphs, in order to find out the vowel production by EFL children learners from the Shanghai Dialect area. One-way ANOVA was used between Group A and Group B1 in order to investigate the accuracy of vowel pronunciation. (See Table 4).

Table 4. One-way ANOVA results of group A and group B

Vowel	One-way ANOVA results		
	df	F	P
i	1.38	0.702	0.407
	1.38	1.612	0.212
u	1.38	29.387	0.000**
	1.38	8.163	0.007**
ɑ	1.38	10.244	0.003**
	1.38	2.366	0.132
ɛ	1.38	28.391	0.000**
	1.38	9.985	0.003**
ɔ	1.38	15.539	0.000**
	1.38	0.049	0.827
ɪ	1.38	10.445	0.003**
	1.38	0.080	0.779
æ	1.38	15.228	0.000**
	1.38	2.180	0.148
ɜ	1.38	32.042	0.000**
	1.38	3.057	0.088
ʌ	1.38	11.984	0.001**
	1.38	2.192	0.147
ʊ	1.38	45.802	0.000**
	1.38	55.640	0.000**

Note: * means significant difference at 0.05 level, ** means significant difference at 0.01 level)

Then we plotted acoustic vowel graphs of English vowels of Americans (Group A) and children from Shanghai dialect (Group B), as shown in Fig. 1.

As shown in the charts and figures, the vowel /i/ is similar to those of Americans in both the height of the tongue position and the anterior and posterior positions of the tongue, with no statistically significant differences. It can also be seen from the acoustic vowel graph that the acoustic space almost coincides, and there is no significant difference.

There are no significant differences in F2 (> 0.05) of vowels /ɑ, ɔ, ɪ, æ, ɜ, ʌ/, but significant differences in F1 (< 0.01), which means Group B are similar to those of Americans in the anterior and posterior positions of the tongue, but are different in the height of the tongue position. Other vowels, including /ɛ/ /u/ /ʊ/ and F1 of /ɑ, ɔ, ɪ, æ, ɜ, ʌ/, all have significant differences with Americans.

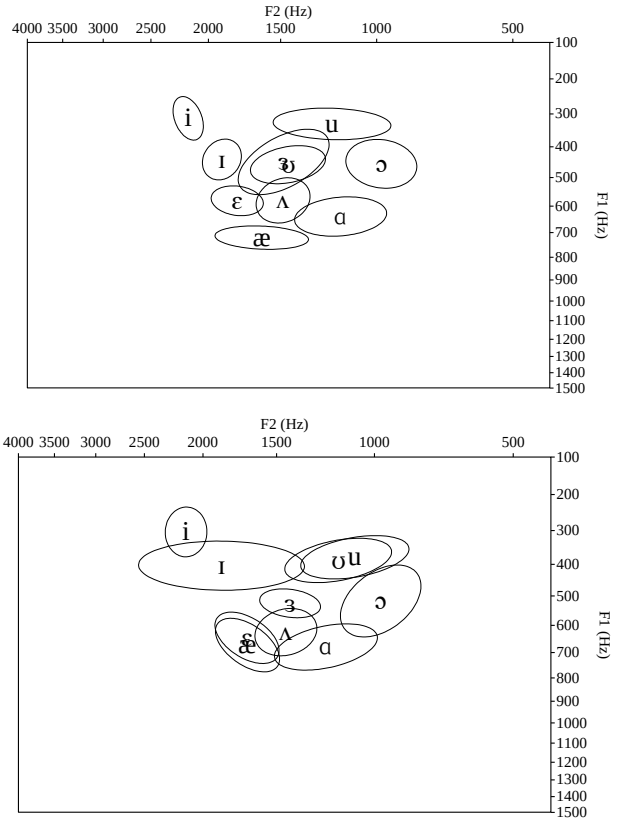

Fig. 1. Acoustic vowel graph of Group A (left) and Group B(right).

V. Conclusion

From the combined results of accuracy, the pronunciation of /i/ and F2 of /ɑ, ɔ, ɪ, æ, ɜ, ʌ/ show good acquisition outcomes, while the acqusition of /ɛ, u, ʊ/ seems to be not good enough. Based on the above analysis, the possible acquisition order of Shanghai children could be /i/ > /ɑ, ɔ, ɪ, æ, ɜ, ʌ/ > /ɛ, u, ʊ/.

Through the above research, we also find that the English pronunciation of children is at a high level which might be a support for the critical period of language learning. Pre-adolescent learners have better language acquisition ability, start learning English at an earlier age and use it frequently which are pretty helpful for English study. Besides, some of their vowels are already very close to native English speakers such as /i/ /ɜ/ and /ɑ/. As for the hypothesis of SLM, /i, u, ɑ, ɪ, ɔ/ are similar phones, while /ɛ, æ, ʊ, ʌ, ɜ/ are new phones, but both /ɜ/ and /ɑ/ show good learning outcomes. The influence of mother tongue might plays an important role in second language pronunciation acquisition, and the influence of first language transfer precedes that of universal grammar. The process of second language acquisition is actually the process of getting rid of the interference of

mother tongue, which may not have much to do with similar phones and new phones. Flege, in his latest study, also pointed out that because the classification criteria for the same phone, similar phone and new phone are difficult to agree on, this division was abandoned actually, and stress the quantity and quality of language input instead in his SLM-r model with the research of language perception.

The experimental results provide empirical data support for more effective English pronunciation teaching, and can be targeted to learn the difficulties and errors in the process of second language acquisition of people from Shanghai dialect, as well as other dialect areas in China and the other countries in the world, so as to acquire the pronunciation of the target language efficiently and accurately.

The paper mainly focuses on children's vowel production, while vowel perception and the relationship between production and perception remain to be studied as well. Besides, middle and old aged people's characteristics of acquisition might differ from children which are also noteworthy.

Conflict of Interest

The author declares no conflict of interest.

Funding

This research was funded by China Center for Language Planning and Policy Studies, grant number WYZL2023BJ0021.

References

[1] K. Tsukada, D. Birdsong, E. Bialystok *et al.*, "A developmental study of English vowel production and perception by native Korean adults and children," *Journal of Phonetics*, vol. 33, no. 3, pp. 263–290, 2005.
[2] Z. Li, Y. Zhang, and W. Dong, "A preliminary study on pronunciation errors of students in Shanghai Dialect Area," *Fudan Journal (Humanities Edition)*, pp. 53–72, January, 1957.
[3] H. Wu, "An analysis of recognized vowel pronunciation characteristics of middle school students in Shanghai," Ph.D. dissertation, Dept. Elect. Eng., Dalian Maritime University, 2002.
[4] X. Y. Ji, "A study on the correlation between lack of English pronunciation teaching and students' listening comprehension Impairment in middle school," M.S. thesis, Shanghai Normal University, China, 2013.
[5] F. J. Emil, "Two procedures for training a novel second language phonetic contrast," *Applied Psycholinguistics*, vol. 16, no. 4, pp. 425 –442, 1995.
[6] F. R. Eckman, "Markedness and the contrastive analysis hypothesis," *Language Learning*, pp. 315–330, 1977.
[7] L. Selinker, "Interlanguage," *International Review of Applied Linguistics in Language Teaching*, vol. 10, pp. 209–231, 1972.
[8] L. Eric, *Biological Foundations of Language*, New York: Wiley, 1967.
[9] G. S. Morrision and P. Escudero, "A cross-dialect comparison of peninsular and Peruvian-Spanish vowels," in *Proc. the 16th International Congress of Phonetic Sciences*, Saarbrücken, 2007.
[10] B. M. Lobanov, "Classification of Russian vowels spoken by different speakers," *Journal of the Acoustical Society of America*, pp. 606–608, 1971.
[11] E. R. Thomas and T. Kendall, "The vowel normalization and plotting suite," *Norm*, 2007.
[12] H. Traunmüller, "A note on hidden factors in vowel perception experiments," *Journal of the Acoustical Society of America*, vol. 88, no. 4, pp. 2015–2019, 1990.
[13] R. A. W. Bladon, C. G. Henton, and J. B. Pickering, "Towards an auditory theory of speaker normalization," *Language & Communication*, pp. 59–69, April 1984.