BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

English Vowel Pronunciation Errors Research among Chinese K-9 Learners

Yanlu Xie [1], Yilin Tao [1*], Xiaoli Feng [1], Xinyue Gao [1], Xiaoli He [1], Ziqing Wang [1], Liyuan Wang [2]

[1] School of Information Science, Beijing Language and Culture University, 15 Xueyuan Road, Beijing 100083, China

[2] Tencent Technology Co., Ltd., Shenzhen 100193, China

ABSTRACT

Among bilingual Mandarin-English children, spoken-English learners of different ages have different pronunciation characteristics. This paper studies English vowel pronunciation errors of Chinese learners from kindergarten to 9th Grade (K-9). We divided K-9 learners into three age groups: Kindergarten Children (KC), Primary-school Students (PS), and Middle-school students (MS). The results show that: (1) The overall English vowel acquisition error ratio of the KC group is higher than that of the PS group, and the PS group is higher than that of the KC group. (2) It is easier for learners to master the phonemes that have the same pronunciation in English and Chinese, and difficult to accurately distinguish the phonemes that have similar pronunciation in English and Chinese, and difficult to master the phonemes that do not exist in Chinese. (3) English vowel pronunciation errors of the KC group tend to be concentrated and specific, and English vowel errors of the PS and MS groups tend to be scattered and diverse. Understanding the English vowel pronunciation errors in learners of different ages can be more targeted to correct learners' pronunciation errors. According to the English vowel mispronunciation and the Chinese mother tongue influence of different aged learners, we analyze the characteristics and causes of learners' errors and try to provide some references for oral English teaching.

Keywords: L2 Acquisition; Mispronunciation Patterns; Vowel Contrast; Learner Age Differences

*CORRESPONDING AUTHOR:

Yilin Tao, School of Information Science, Beijing Language and Culture University, 15 Xueyuan Road, Beijing 100083, China;
Email: tyilin034725@163.com

ARTICLE INFO

Received: 3 November 2024 | Revised: 15 November 2024 | Accepted: 21 November 2024 | Published Online: 7 January 2025
DOI: https://doi.org/10.30564/fls.v7i1.7670

CITATION

Xie, Y., Tao, Y., Feng, X., et al., 2025. English Vowel Pronunciation Errors Research among Chinese K-9 Learners. Forum for Linguistic Studies. 7(1): 566–577. DOI: https://doi.org/10.30564/fls.v7i1.7670

1. Introduction

1.1. Research Background

With the acceleration of globalization and the growth of international communication, foreign language learners attach more importance to clear, natural, and accurate pronunciation to better communicate with native speakers and the international community[1–5]. A large body of research has shown that L2 speakers who begin to acquire an L2 at a young age are better able to produce and differentiate L2 sounds than late L2 speakers who begin to acquire an L2 in adulthood[6–11]. For Chinese students learning English, young bilingual Mandarin-English children have attracted wide attention, as more learners hope to weaken their foreign language accent and make their pronunciation more similar to that of native speakers. Pronunciation problems of English learners are reflected in various aspects, such as segmental or suprasegmental accuracy[12, 13], fluency[14], etc. Among them, vowel pronunciation errors cannot be ignored. In addition, the specific performance of pronunciation errors of learners of different ages has differences[15]. Therefore, exploring the characteristics of vowel pronunciation errors of learners of different age groups is an essential link in the study of second language pronunciation acquisition.

On the English vowel pronunciation error of Chinese native speakers, previous studies mainly focused on the error ratio of different phonemes and tried to analyze the errored vowels from various aspects like the influence of native phonology and second language learning experience[16–18]. On the acquisition of English vowels by Chinese learners at different ages, existing studies have shown that learners not only have differences in the acquisition difficulty of different phonemes but also have different pronunciation error patterns[19–21]. Therefore, the characteristics of vowel pronunciation errors of learners at different ages need to be further explored. Understanding the extent and pattern of vowel errors of learners in different age groups can help learners correct vowel errors in a more targeted way.

This paper studies the vowel acquisition of Chinese K-9 (Kindergarten to 9th Grade) English learners in three groups: Kindergarten Children (KC), Primary-school Students (PS), and Middle-school students (MS). We use the intermediate corpus to analyze and summarize the characteristics of English vowel pronunciation errors and combine the native Chinese pronunciation characteristics of different aged learners to discuss the causes of mispronunciation.

1.2. Research Status

Many studies have been conducted on the acquisition of English vowels by Chinese learners. For children's English learners, Zou Linlin[22] found that substitution errors were the most likely to occur, such as the substitution of monophthongs /i:/ with monophthongs /i/ and diphthongs /ei/, and the confusion of monophthongs /e/ and /æ/. For adolescent English learners, Hong Liu[23] believes that Chinese middle school students are lacking in cognition and practice in oral English learning and suggests that teachers carry out long-term practice in pronunciation, intonation, rhythm, and other aspects. Cheng Chunmei and He Anping[12] found that the error frequency of Chinese senior English learners for loose vowels /uh/, /e/, and /i/ is much higher than that of other vowels. Xin Yan[24] summarized the common problems of Chinese students in English vowel pronunciation, and he found native Chinese speakers tend to have pronunciation problems such as poor vowel pronunciation, stress, and intonation in the process of English learning.

According to the common pronunciation errors of English vowels among Chinese K-9 learners, Peng Po[25] and Li Li[26] believe that the negative transfer of the Chinese mother tongue has a great impact on the acquisition of English pronunciation, and learners need to overcome the influence of their mother tongue. Guo Minghua[27] and Wang Jun[28] believe that the mother tongue has both positive and negative influences on the process of foreign language acquisition. Zheng Guifang[29] suggested from the aspect of pronunciation teaching that English teachers should initially understand the similarities and differences between the two language systems, and they try their best to promote the positive teaching guidance of positive transfer while avoiding the influence of negative transfer on pronunciation.

It is not difficult to find out that researchers who study Chinese students' English vowel pronunciation almost focus on the analysis of learners at one stage. There is a lack of comparative studies on adjacent age groups, and the research on early childhood is relatively scarce. Therefore, we conduct a systematic analysis and research on English vowel errors of K-9 learners in three adjacent age groups. In this

paper, we summarize the tendency and characteristics of English vowel pronunciation errors in different aged learners. Based on the characteristics of learners' English vowel pronunciation errors and their native Chinese pronunciation, we analyze the causes of learners' English vowel pronunciation errors at different ages and try to provide references for oral English teaching.

2. Materials and Methods

2.1. Data Sources

All the data used in this study were collected from the online English learning APP English Jun, and the learners were native Chinese speakers in Beijing. According to the age, grade, and other information filled in by the learners during online registration, about 500 learners were divided into three groups: Kindergarten Children (KC), Primary-school Students (PS), and Middle-school Students (MS). Learners in the KC group are 3–6 years old, learners in the PS group are 6–12 years old, and learners in the MS group are 12–15 years old. The number of learners in the three groups was relatively balanced. Since the APP can only record the registration time of learners, and the information of the starting learning age and learning duration of learners is not recorded, this study only analyzes the pronunciation error characteristics of the three groups of learners from the perspective of the age.

The "English Jun" APP mainly collects data from learners reading aloud or reading followed with them online. The content of reading aloud or reading along is words, texts, and exercises in the textbooks used by the corresponding learners' grades. Learners mainly use smartphones to record their pronunciations. The data sampling frequency is 16000Hz, and the bit rate is 16bit. The data used in the experiment were recorded by learners in a low-noise environment, with full coverage of phonemes taken into account, gender and grade balance of learners as far as possible, to ensure that the data reflected the real pronunciation level of learners of all ages as far as possible. After manual screening and deletion of invalid phonemes, 1074 phonemes in the KC group, 1066 phonemes in the PS group, and 936 phonemes in the MS group were finally selected.

2.2. Data Annotation

In terms of phonetic transcription, The CMU Dictionary[30] was selected as the reference dictionary for data annotation. In terms of segments, the dictionary is divided into two parts: English vowel phonemes and consonant phonemes, among which vowel phonemes include AA, AE, AH, AO, AW, AY, EH, ER, EY, IH, IY, OW, OY, UH, UW. Chinese learners are easily influenced by their mother tongue in second language learning, and vowel pronunciation is no exception. Chinese vowels include monophthongs and diphthongs, there are ten monophthongs in the Chinese language family, including a, o, e, ê, i, u, ü, -i (the vowel i after the Chinese initial consonants z, c, and s), -i (the vowel i after the Chinese initial consonants zh, ch, sh and r), and er; there are thirteen diphthongs in the Chinese language family, including ai, ei, ao, ou, ia, ie, ua, uo, üe, iao, iou, uai, uei. There are great differences in vowels between English and Chinese languages[31]: the number of vowels in English is different from that in Chinese, and some vowels exist in English but do not exist in Chinese, some vowels exist in Chinese but do not exist in English, and there are differences in phoneme, syllable, tone, intonation, rhythm, and transition between the two languages[32].

In the experiment, we used Praat software to mark pronunciation errors in the speech data[33]. The data annotation was completed by five graduate students majoring in English at Beijing Language and Culture University, and an English native speaker from the United States conducted a random check on the data of all the annotators according to 20% of the proportion. The accuracy of the annotation was more than 90%, which was considered qualified. All the data taggers and quality inspectors were proficient in English phonological structure and phonetic symbols. Before tagging, the project leader conducted CMU dictionary and Praat software training and tests for all the researchers involved in the project tagging. The annotation interface is shown in **Figure 1**.

We use Praat software to annotate pronunciation errors, and the annotation of each audio has five layers. The first layer is the word regularization layer, the second layer uses the Montreal-force-Aligner tool[34] to force the phoneme level alignment, the third layer is the phoneme error annotation layer, the fourth layer is the pronunciation accuracy layer, and the fifth layer is the comment layer. The anno-

tation file is automatically generated by the script of Praat software, and the generated file format is TextGrid. The core part of annotation is the third layer, that is, the phoneme error layer. In this paper, the pronunciation error is divided into three types: substitution, insertion, and deletion [35, 36]. For substitution errors, the tagger would mark the actual pronunciation in the position corresponding to the third layer. For example, in **Figure 1**, the speaker pronounced the vowel AY(/ai/) instead of AE(/æ/), thus marking "AE" under "AY". For insertion errors, the third layer marks the "+" and the inserted phoneme where the insertion error occurred. For example, the "+H" in the third layer marks in **Figure 1** indicates that the pronouncer inserted the sound HH(/h/) at the corresponding position of the audio. For deletion errors, the third layer marks the "-" and the removed phoneme where the deletion error occurred. For example, the phoneme T(/t/) in the audio in **Figure 1** is missing, so marking "-T" at the corresponding position in the third layer.

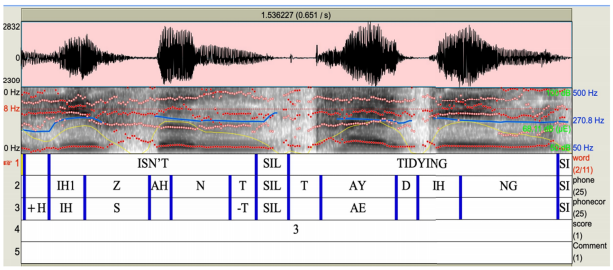

Figure 1. Example of the marking interface.

3. Results

3.1. Vowel Errors in the KC Group

The KC group in this paper mainly refers to the 3–6-year-old children of English pronunciation learners, which is an important stage of speech development in human growth [37]. Studies have shown that young children's Chinese phonological awareness develops rapidly at the age of 3–4 [36], and their English phonological awareness will gradually increase with the growth of age [38]. Therefore, it is of great significance to analyze the vowel pronunciation errors in early childhood.

Based on the English vowel pronunciation errors of the KC group, this study uses the phoneme error ratio (the ratio of the error number of the target phoneme to the total number of occurrences) to quantitatively evaluate the pronunciation difficulty of each phoneme. The higher error ratio of a phoneme indicates that it is more difficult to master the phoneme for learners. The statistical results show the vowel error ratios of the KC group in **Figure 2**.

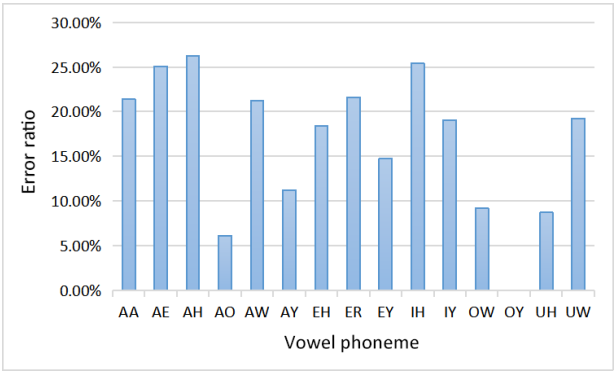

Figure 2. Vowel phoneme error ratios in KC.

According to the statistics, the average error ratio of vowel phonemes in the KC group is about 16.5% (variance is 0.0062). As shown in **Figure 2**, the top three phonemes in order of error ratios from highest to lowest are AH, IH, and AE, with error ratios of about 26.3%, 25.4%, and 25.1%. From **Figure 2**, we find that the error ratio of vowel phoneme OY is 0, which does not mean that children in the KC group have a good grasp of phoneme OY, but that there are fewer data samples with OY phoneme in the processed data. In the KC group, the total number of occurrences of phoneme OY is only 3, and the number of errors of phoneme OY is 0, so the error rate of OY is 0. We will expand the data scale to focus on the OY phoneme errors in the KC group in subsequent studies.

From the perspective of the error pattern, the phoneme AH with the highest error ratio may be mispronounced as the other 14 different phonemes, and AH is easily mispronounced as AA, AE, and AO, with error proportions of 24.6%, 15.9%, and 11.6%, respectively. For example, the phoneme AH in "sun", "run", and "bug", may sometimes be incorrectly articulated as AA, AE, and AO, respectively. The phonemes IH are prone to be mispronounced by the KC group as EY, IY, and EH. The phonemes AE are prone to be mispronounced as AA, AH, and IY. In addition, some phonetic errors are mainly manifested as confusion between pairs, such as AE and AH, AY and EY, and other phonemes. For example, the KC group may mispronounce the phoneme AE as AH when reading the word "backpack", and they may also mispronounce the phoneme AH in the word "sun" as AE.

From the perspective of the types of vowel pronunciation errors, the distribution of the types of errors in the KC group is in **Table 1**. For the KC group, substitution is the main pronunciation error of vowel phonemes, and insertion and deletion are less. The pronunciation errors in the phonemes AA, OW, UH, and AO are all substitution errors, and the KC group tends to mispronounce AA(/ɑ/) for AO(/ɔ/), OW(/oʊ/) for AO(/ɔ/), UH(/ʊ/) for AO(/ɔ/), and AO(/ɔ/) for AW(/ɑʊ/). In the pronunciation of the phoneme ER, 94% of the errors are substitution type, which is most easily replaced with AH. For example, the phoneme ER(/ər/) in "tiger" may be mispronounced as AH(/ʌ/). In the pronunciation of the phoneme AW, 93% of the errors are substitution type, and the easiest replacement is AO. For example, the phoneme AW(/ɑʊ/) in "flower" may be mispronounced as AO(/ɔ/). In the pronunciation of the phoneme EH, 89% of the errors are the substitution type, which is most easily replaced by AE. For example, the phoneme EH(/e/) in "ten" may be mispronounced as AE(/æ/). In the pronunciation of the phoneme AE, 88% of the errors are the substitution type, which is most easily replaced by AA. For example, the phoneme AE(/æ/) in "apple" may be mispronounced as AA(/ɑ/).

Table 1. Proportion of vowel error types in the KC group.

Error Type	Proportion
Substitution	86.7%
Deletion	9.7%
Insertion	3.6%

3.2. Vowel Errors in the PS Group

The PS group in this paper mainly refers to the Chinese learners aged 6–12 years old, who have begun to use daily speech to actively express themselves. Meanwhile, the PS group is good at imitation and sensitive to English pronunciation, intonation, and pronunciation [39]. We analyze the characteristics of vowel errors in the PS group according to the vowel phoneme errors.

3.2.1. Vowel Mispronunciation in the PS Group

The error ratios of different vowel phonemes in the PS group are presented in **Figure 3**. The mean ratio of errors in the PS group is about 14.9% (variance 0.0109). According to **Figure 3**, we find the error ratio of phoneme OY is as high as 50%, which is much higher than that of other vowels. Compared with the KC group, the overall vowel error ratio of the PS group shows a downward trend.

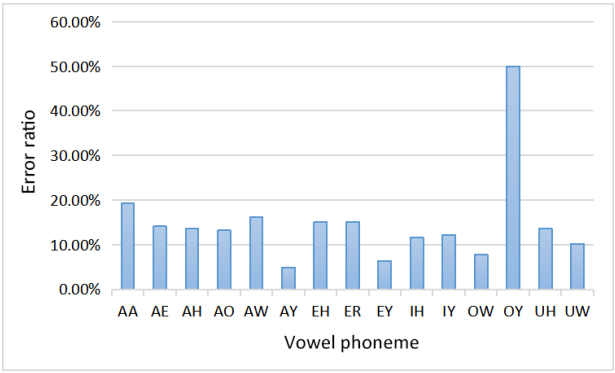

Figure 3. Vowel phoneme error ratios in PS.

From the perspective of the error pattern, the phoneme OY is prone to error in the four vowel phonemes of EY, AO, ER, and AY. A comparison of the Chinese and English vowel phonology shows that OY corresponds to /ɔi/ in the international phonetic alphabet, while there is no exact pronunciation of the word as "ɔ" in Chinese pinyin, so the phoneme OY may be easily mispronounced. The PS group has the lowest pronunciation error ratio of phoneme AY, which is about 4.9%. As the English pronunciation of the phoneme AY(/aɪ/) sounds similar to that of the Chinese vowel "ai", most primary-school students can master the pronunciation of the AY phoneme well.

From the perspective of error types (**Table 2**), similar to the KC group, the PS group has the highest probability of substitution errors, and deletion and insertion errors account for a relatively small proportion. The PS group's pronunciation errors in the phonemes UH and OY are all substitution errors, and they tend to mispronounce UH(/ʊ/) for AO(/ɔ/) or OW(/oʊ/), and they are likely to mispronounce OY(/ɔi/) for AO(/ɔ/) or EY(/eɪ/). In the PS group, 90% of the AE pronunciation errors are substitution types, and AE(/æ/) is most easily pronounced as AA(/ɑ/). In the pronunciation of the phoneme AA, 89% of the errors are substitution type, and AA(/ɑ/) is most easily pronounced as AO(/ɔ/). Among the substitution errors, the phoneme AH appears most frequently, and the PS group tends to mispronounce AH(/ʌ/) as ER(/ər/), AA(/ɑ/), or AE(/æ/). In the deletion errors, the phoneme AH appears at the highest frequency, and the PS group easily ignores the pronunciation of /ʌ/. Among the insertion errors, the phoneme IY appears most frequently, which shows that /i/ is more easily mixed by the PS group in English pronunciation.

Table 2. Proportion of vowel error types in the PS group.

Error Type	Proportion
Substitution	81.7%
Deletion	9.8%
Insertion	8.5%

3.2.2. Characteristics of Vowel Errors in the PS Group

When Chinese native primary school students learn English vowel pronunciation, they tend to replace English pronunciation by referring to the pronunciation of their mother tongue through transfer learning[25]. If they encounter a pronunciation that does not exist in their mother tongue, it is difficult for the PS group to master the pronunciation skills. It is easy to apply the pronunciation patterns of Chinese to English in an emergency[13], which often leads to pronunciation errors. By analyzing the data, we can find some rules, which can be roughly divided into the following situations:

1. Both English and Chinese languages exist, the symbols are the same, and the pronunciation is similar but different, and the PS group is prone to errors;
2. Both English and Chinese languages have the same symbol, but there are obvious differences in pronunciation, which may be difficult for the PS group to identify, easily confuse or misread;
3. It exists in English but not in Chinese, and it is difficult for the PS group to master the pronunciation.

In the first case, the phonemes exist in both English and Chinese, the symbols are the same, and the pronunciation is similar but different. The common error situations are as follows: the symbol is "a", and the English pronunciations /ʌ/ and /ɑ:/ are easily confused with the pronunciation of "a" in Chinese Pinyin; the symbol is "u", and the English pronunciations /u:/ and /u/ are easily confused with the pronunciation of "u" in Chinese pinyin; the symbol is "o", and the English pronunciations /ɔ:/ and /ɔ/ are easily confused with the pronunciation of "o" in Pinyin[25]. English vowels are usually divided into long vowels and short vowels, but there are no existing long and short vowels in Chinese, so the PS group is prone to errors. The second case is that the phonemes exist in both English and Chinese, and they have the same symbol, but there are obvious differences in pronunciation. For example, the symbol "e" corresponds to /e/ in the English phonetic alphabet and corresponds to the phoneme EH, but the pronunciation of "e" in the Chinese pinyin corresponds to the English phonetic alphabet of as /ʌ/, the corresponding phoneme is AH, which may be difficult for the PS group to identify, easily confuse or misread. In the third case, the phonemes exist in English but not in Chinese. For example, if the phonemes OY, ER, and IH do not have the same or similar pronunciation in Chinese pinyin, then the PS group will try to transfer learning with the help of the vowel pronunciation that indicates similar symbols in their mother tongue, it is easy to appear inaccurate pronunciation and confusion, so it is difficult for the PS group to master such pronunciation.

3.3. Vowel Errors in the MS Group

The MS group in this paper mainly refers to English learners aged 12–15. Although most middle-school students in China have been learning English for many years, they are generally unable to use English flexibly in oral English, and there are still problems with oral pronunciation.

3.3.1. Vowel Mispronunciation in the MS Group

The error ratios of vowel phonemes in the MS group are shown in **Figure 4**. The average error ratio for the MS vowel phonemes is 9.9% (variance 0.0012). According to **Figure 4**, the phoneme AA has the highest error ratio of 16.6%, followed by UH, OY, and ER, which have higher error rates of 16.0%, 13.3%, and 12.5% respectively.

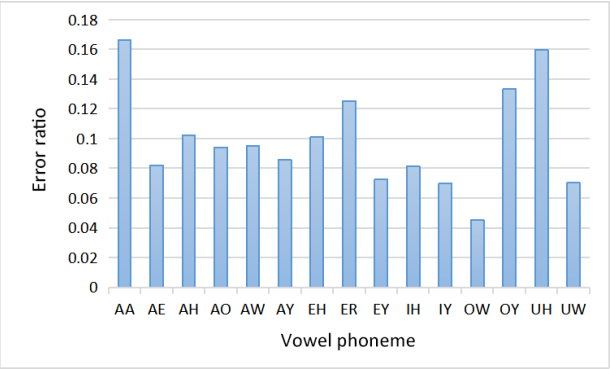

Figure 4. Vowel phoneme error ratios in MS.

From the perspective of the error pattern, for the phonemes AA with the highest error ratio, the MS group tends to error AA(/ɑ/) as AO(/ɔ/), such as reading

was(/wɒz/) as /wɔz/ and reading dot(/dɒt/) as /dɔt/. For the phonemes UH, OY, and ER with high error ratios, the MS group tends to mispronounce UH as OW or UW, for example, to pronounce look(/lʊk/) as /loʊk/ or /luk/. The MS group is easy to pronounce OY(/ɔɪ/) as AO(/ɔ/), for example, boil(/bɔɪl/) is pronounced as /bɔl/, misread voice(/vɔɪs/) as /vɔs/. And the MS group is prone to mispronounce ER(/ər/) as AH(/ʌ/), for example, misread water(/ˈwɔːtər/) as /ˈwɔːtʌ/.

From the perspective of error types (**Table 3**), similar to the KC group and the PS group, the MS group has the highest probability of substitution errors, followed by deletion and insertion errors. Among the substitution errors, the top three phonemes with the highest error ratio were AH(/ʌ/), IH(/ɪ/), and AA(/ɑ/). In insertion errors, the top three phonemes with the highest error ratio were AH(/ʌ/), UW(/u/), and ER(/ər/). In the deletion error, the top three phonemes with the error ratio are AH(/ʌ/), IH(/ɪ/), and ER(/ər/). It can be found that the MS group has a poor mastery of the phoneme AH(/ʌ/), and all three types of errors are likely to occur.

Table 3. Proportion of vowel error types in the MS group.

Error Type	Proportion
Substitution	79.6%
Deletion	10.8%
Insertion	9.6%

3.3.2. Characteristics of Vowel Errors in the MS Group

In the MS group, the overall error ratio of vowel phonemes was lower than in the PS group, but the same vowel phoneme could be mispronounced as more other phonemes. The error ratio of most vowel phonemes in the MS group is less than or equal to 0.1, but the error ratios of AA, UH, OY, and ER are more than 0.1, and the error ratio is 0.166, 0.160, 0.133, and 0.125. The total frequency of AA, UH, OY, and ER phonemes in the corpus is 494, 144, 15, 607. The frequency of these four phonemes being mispronounced into other phonemes is in **Tables 4–7**.

Common mispronunciations include the phoneme AA being substituted with AO, OW, or AH; the phoneme UH being replaced with UW or OW; the phoneme OY being pronounced as AO; and the phoneme ER being articulated as AH, AO, or R. Among them, since the total occurrence number of the phoneme OY in the corpus is small and the frequency of errors is also low, the statistical results may be accidental. We find the MS group generally has a good mastery of most phonemes. For individual phonemes with a high error ratio, the MS group needs to pay more attention to them and correct them with the help of standardized pronunciation recordings.

Table 4. Vowel phoneme AA error distribution.

Mispronounced Phonemes	Error Frequency
AO	47
OW	7
AH	6
AE	3
EY	2
IH	2
L	2
NG	2
AW	1
EH	1
ER	1
N	1
R	1
UH	1

Table 5. Vowel phoneme UH error distribution.

Mispronounced Phonemes	Error Frequency
OW	11
UH	8
AH	1
AW	1

Table 6. Vowel phoneme OY error distribution.

Mispronounced Phonemes	Error Frequency
AO	2

Table 7. Vowel phoneme ER error distribution.

Mispronounced Phonemes	Error Frequency
AH	24
AO	9
R	8
EH	3
UH	3
AA	2
AY	2
B	2
IH	2
L	2
N	2
EY	1
S	1
W	1

4. Discussion

4.1. Comparison of Vowel Errors

The vowel phoneme error ratios of the KC group, the PS group, and the MS group are shown in **Figure 5**. On the whole, except for some individual phonemes, the phoneme error ratio of the KC group is generally higher than that of the PS group, and the phoneme error ratio of the PS group is generally higher than that of the MS group. With the growth of age, the accumulation of learning time can improve learners' ability to correctly pronounce vowel phonemes.

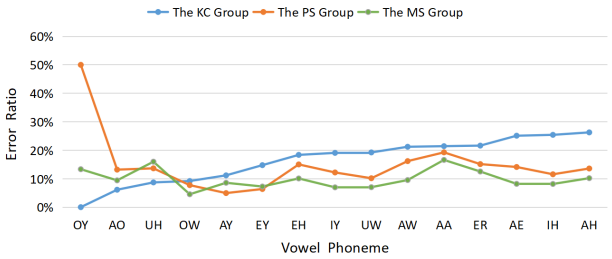

Figure 5. Vowel phoneme error ratios in different age groups.

The error ratios of phonemes AY, EY, and UH in the MS group are higher than that of the PS group, indicating that learners need to pay attention to these phonemes with age. The phoneme OY has an unusually high error rate in the PS group. We speculate that the phoneme OY does not have the same or similar pronunciation in Chinese pinyin, and the PS group will try to apply it harshly with the help of Chinese pronunciation, which is prone to negative transfer of pronunciation errors [32]. The error ratio of phoneme AA in all stages is not low, which shows that English learners at the three stages find it difficult to master the pronunciation of AA, and learners need to strengthen targeted practice. The error ratio of phoneme OW in the three stages is low, indicating that learners of all ages have a good mastery of the pronunciation of phoneme OW. In addition to the pronunciation error ratio of phonemes, we analyze the error types of vowel phonemes in the three stages, as shown in **Table 8**.

Table 8. Proportion of vowel error types in different ages.

Error Type	KC Group	PS Group	MS Group
Substitution	86.7%	81.7%	79.6%
Deletion	9.7%	9.8%	10.8%
Insertion	3.6%	8.5%	9.6%

By comparing the error types of vowel phonemes in the three periods, the substitution type is the core error type of learners in each stage, and the deletion type occupies the least proportion in each stage. With the increase of age, the proportion of substitution errors gradually decreases, and the proportion of deletion and insertion errors both show an increasing trend. We speculate that through continuous learning, learners will gradually improve the accuracy of pronunciations, but there may be an increasing trend of swallowing or adding pronunciations when reading sentences.

Learners at different stages have different pronunciation errors on the same phoneme. **Figure 6** shows the number of wrong phonemes that learners at different stages mispronounce the same vowel phoneme into other wrong phonemes. If learners mispronounce the same consonant phoneme into more wrong phonemes as they grow, it indicates that learners' mispronunciation of this consonant tends to be dispersed and diversified. For example, according to our statistics, for the phoneme AH, the KC group mispronounces the phoneme AH into fifteen wrong phonemes, the PS group mispronounces the phoneme AH into eighteen wrong phonemes, the MS group mispronounces the phoneme AH into twenty-three wrong phonemes. Therefore, learners tend to have diversified pronunciation errors in the phoneme AH with age. As a whole, the KC group may have fewer objects of errors for the same vowel phoneme, and the errors tend to be concentrated and specific. The PS and the MS may have more error objects for the same vowel phoneme, and the errors tend to be diversified and dispersed. This phenomenon may also have a certain correlation with the development of learners' first language phonetic accuracy, which is only explained from the perspective of speculation.

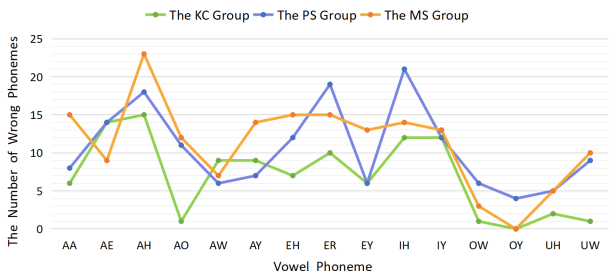

Figure 6. The number of wrong phonemes in learners.

4.2. Characteristics of Vowel Errors

The Chinese native language system of the PS group is not mature and stereotyped, it is easy to have vowel errors

when learning a second foreign language, and the reasons are complex and various. When learning English vowels, the PS group will consciously imitate the pronunciation that is slightly similar to Chinese vowels for transfer learning, so the average correct ratio is improved compared with that of the KC group, but there are also errors caused by rote [32]. Compared with the KC group and the PS group, the average error ratio of vowel phonemes in the MS group is smaller. We can see that the MS group generally has a good mastery of vowel pronunciation and needs to pay attention to the pronunciation of individual phonemes.

When Chinese K-9 learners learn English vowel sounds, they learn faster when they encounter English sounds that are the same as Chinese vowel sounds [36]. When encountering English pronunciations that are similar to Chinese vowels, refer to the "Native Language Magnet Model (NLMM)" proposed by Kuhl. The magnetic effect of the phonetic prototypes in Chinese will interfere with learners' second language learning [40], so it is difficult to distinguish similar pronunciations in Chinese and English. When English vowel pronunciations are missing in the Chinese language family, Chinese people learning English lack relevant pronunciation experience [41], which makes it difficult to learn [42] and slow to learn.

Based on the above analysis, the characteristics of pronunciation error in three age groups are summarized:

1. The overall English vowel acquisition error ratio of the KC group is higher than that of the PS group, and the PS group is higher than that of the MS group;
2. Due to the influence of the native Chinese language, it is easier for learners to grasp the phonemes with the same pronunciation in English and Chinese, but difficult to accurately distinguish the phonemes with similar pronunciation in English and Chinese. For phonemes that are absent in Chinese pronunciation but exist in English pronunciation, it is difficult for learners to master pronunciation skills;
3. The English vowel errors of the KC group tend to be more focused and specific, while the English vowel errors of the PS group and the MS group tend to be more scattered and diverse.

5. Conclusions

With the help of the speech data information of the intermediate corpus, this paper analyzes the rules of English vowel errors in the three adjacent stages of the KC group, the PS group, and the MS group. With the increase in age, learners' vowel error ratios showed an overall decreasing trend. The cases of vowel errors in the KC group tend to be more focused and specific, and the situations of vowel errors in the PS group and the MS group tend to be more scattered and diverse. Among the different types of mistakes, learners of various ages are most likely to make substitution-type English vowel pronunciation errors. Combined with the influence of the mother tongue of Chinese, we discuss the partial causes of English vowel errors and summarize the characteristics of English vowel errors in Chinese K-9 learners. During learners' growth, the construction of bilingual space and the characteristics of native pronunciation will influence English vowel pronunciation. In the future, we plan to use the research data to train the English vowel automatic detection model to provide a more intelligent and accurate platform for Chinese K-9 learners to learn English vowels. We will develop various phonetic models for learners of different ages and add targeted language acquisition strategies and teaching methods on the intelligent learning platform according to the characteristics of pronunciation errors of age groups to help learners correct English vowel pronunciation errors more quickly and accurately. We hope that our future research can provide reference information for the theory and practice of multilingual education and contribute to the field of applied linguistics and language education.

Author Contributions

All authors have made a substantial, direct, and intellectual contribution to the work. Conceptualization, Y.X. and X.F.; methodology, Y.X. and X.F.; software, X.F., Y.T., X.G., X.H., and Z.Q.; validation, Y.T., X.G., X.H., and Z.Q.; formal analysis, Y.T., X.G., X.H., and Z.Q.; investigation, Y.T., X.G., X.H., and Z.Q.; resources, Y.X. and L.Y.; data curation, X.F., Y.T., X.G., X.H., and Z.Q.; writing—original

draft preparation, Y.T., X.G., X.H., and Z.Q.; writing—review and editing, Y.X. and Y.T.; visualization, Y.T.; supervision, Y.X. and Y.T.; project administration, Y.X.; funding acquisition, Y.X. All authors have read and agreed to the published version of the manuscript.

Funding

This research project is supported by the Key Projects of the National Language Commission (ZDI145-101), Humanities and Social Sciences Research Planning Fund of the Ministry of Education (23YJA740012), Key Research Project on International Chinese Language Education (22YH49B).

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

All participants provided informed consent before participating in the study. The anonymity and confidentiality of the participants were guaranteed, and participation was completely voluntary.

Data Availability Statement

Data can be given upon a reasonable request.

Acknowledgments

We gratefully acknowledge the schools, children and companies that made this research possible. We also thank the researchers at Beijing Language and Culture University who provided statistics and annotations during the analysis.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Yang, J., Resendiz, M., 2023. Assessing Speech Production in Mandarin–English Bilingual Children: Comparison of Mandarin and English Sound Systems and Special Considerations. Perspectives of the ASHA Special Interest Groups. 9(3), 715–737.

[2] Guo, Q., Qi, R.Y., 2023. A Study of Bilingual Acquisition of Wh-Questions of a Mandarin-English Bilingual Preschool Child from China to Australia. Education Sciences. 14(9), 978. DOI: https://doi.org/10.3390/educsci14090978

[3] Zhang, H., Huang, F.W., Wang, C., 2024. Language choice and code-switching in bilingual children's interaction under multilingual contexts: evidence from Mandarin-English bilingual preschoolers. International Journal of Multilingualism. 2024(4), 1–26. DOI: https://doi.org/10.1080/14790718.2024.2340749

[4] Brien, B.A., Arshad, N.A., Ng, S.C., 2024. Bilingual families and the home literacy environment: An examination of English and ethnic language activities and outcomes. Journal of Research in Reading. 47(4), 580–600. DOI: https://doi.org/10.1111/1467-9817.12465

[5] Zheng, Z.J., Degotardi, S., Djonov, E., 2024. Translanguaging practices in infant rooms: case studies of Chinese-Australian bilingual infant-educator interactions. International Journal of Bilingual Education and Bilingualism. 27(8), 11021–1035. DOI: https://doi.org/10.1080/13670050.2024.2320843

[6] Miralpeix, I., 2010. Age effects on second language acquisition: Critical issues under debate. Language Teaching. 44(1), 122–132. DOI: https://doi.org/10.1017/S0261444810000364

[7] Aoyama, K., Guion, S.G., Flege, J.M., et al., 2008. The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English. International Review of Applied Linguistics in Language Teaching. 46(1), 61–90. DOI: https://doi.org/10.1515/IRAL.2008.003

[8] Smemoe, W.B., Trofimovich, P., Flege, J.E., et al., 2008. Child-adult differences in second-language phonological learning: the role of cross-language similarity. Language and speech. 51(Pt 4), 317–42. DOI: https://doi.org/10.1177/0023830908099068

[9] MacLeod, A.A.N., Stoel-Gammon, C., Wassink, A.B., 2009. Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. Journal of Phonetics. 37(4), 374–387. DOI: https://doi.org/10.1016/j.wocn.2009.07.001

[10] Meador, D., Flege, J.E., MacKay, I.R.A., 2021. Factors affecting the recognition of words in a second language. Bilingualism: Language and Cognition. 3(1), 55–67.

[11] Marinova-Todd, S.H., Zhao, J., Bernhardt, M.B., 2010. Phonological awareness skills in the two languages of Mandarin–English bilingual children. Clinical Linguistics and Phonetics. 24(4–5), 387–400. DOI: https://doi.org/10.3109/02699200903532508

[12] Cheng, C.M., He, A.P., 2008. Analysis of speech segment errors in Advanced English learners: a corpus-

based study. Journal of Foreign Languages Institute of PLA. 31(1), 38–42.

[13] Hong, M., 2004. The negative transfer effect of English and Chinese phonological Differences on English phonological acquisition. Journal of Zhejiang Normal University. 29(2), 108–111.

[14] Fulcher, G., 1996. Does thick description lead to smart tests? A data-based approach to rating scale construction. Language Testing. 13(2), 208–238. DOI: https://doi.org/10.1177/026553229601300205

[15] Llanes, À., 2018. Reading in English as a foreign language: Examining differences in reading speed, comprehension, efficacy and L1 cross-linguistic influence across grades. Investigaciones Sobre Lectura. 2018(9), 1–49. DOI: https://doi.org/10.24310/revistaisl.vi9.11099

[16] Best, C.T., Tyler, M.D., 2007. Nonnative and second-language speech perception: Commonalities and complementarities. Language experience in second language speech learning: In honor of James Emil Flege. John Benjamins: Amsterdam, The Netherlands. pp. 13–44.

[17] Flege, J.E., 1995. Second-language speech learning: Theory, findings, and problems. Speech perception and linguistic experience: Issue in cross-language research. York Press: London, UK. pp. 229–273.

[18] Jia, G., Strange, W., Collado, J., et al., 2006. Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. Journal of the Acoustical Society of America. 119(2), 1118–1130. DOI: https://doi.org/10.1121/1.2151806

[19] Yang, J., 2021. Vowel development in young Mandarin-English bilingual children. Phonetica. 78(3), 241–272.

[20] Yulia, L., 2020. Acquisition of English vowels by Chinese EFL students. In Proceedings of the XLIX International Philological Conference; Saint Petersburg, Russia, 16–24 November 2020. pp. 215–224.

[21] Wang, X.C., 1997. The acquisition of English vowels by Mandarin ESL learners: a study of production and perception [Master's Thesis]. Simon Fraser University: Burnaby, BC, Canada. pp. 1–92.

[22] Zou, L.L., 2021. A Study on Common Errors in Phonological Acquisition of Lower grade Children English Learners. Yan Bian College of Education. 35(5), 160–163.

[23] Hong, L., 2005. How to Improve Middle school students' Oral English Ability. Friends of Teachers. 2005(1), 30–31.

[24] Xin, Y., 2010. On Common pronunciation Problems in English Pronunciation Teaching and their Countermeasures. Business Culture (Science and Technology Edition). 2010(2), 174–175.

[25] Peng, P., 2012. Analysis of Chinese Students' English Vowel Pronunciation Errors. Journal of Chu Xiong Normal University. 27(10), 81–85.

[26] Li, L., 2006. The influence of Chinese on English Pronunciation Teaching and its Countermeasures. Journal of Foreign Languages College of Shandong Normal University (Basic English Education). 8(2), 3–7.

[27] Guo, M.H., 2002. On the role of mother tongue in foreign language class. Foreign Languages and Foreign Language Teaching. 2002(4), 24–27.

[28] Wang, J., 2011. A study on English Phonetics Teaching in Chinese Environment: Based on the influence of Chinese phonemes on English phonetics. Modern commercial trade industry. 23(18), 215–216.

[29] Zheng, G.F., 2011. The effect of mother tongue transfer on English Phonetics Acquisition and its countermeasures. Journal of Jilin Radio and Television University. 2011(2), 122–123.

[30] The CMU pronouncing dictionary. Available from: http://www.speech.cs.cmu.edu/cgi-bin/cmudict/ (cited 1 November 2024).

[31] Song, H.L., 2002. A Probe into the relationship between English Phonetic Learning and Mandarin. Journal of Dandong Teachers College. 24(S1), 130–131.

[32] Liu, N.H., 1988. Comparison of Main Characteristics of Chinese and English Speech Systems. Journal of Nanjing Normal University (Social Sciences Edition). 1988(3), 79–84.

[33] Boersma, P., Weenink, D., 2019. Praat: Doing phonetics by computer (version 6.0.49). Available from: http://www.praat.org/ (cited 1 November 2024).

[34] McAuliffe, M., Socolof, M., Mihuc, S., et al., 2017. Montreal Forced Aligner: trainable text-speech alignment using Kaldi; In Proceedings of the Annual Conference of the International Speech Communication Association; Stockholm, Sweden , 20–24 August 2017; pp. 498–502. DOI: https://doi.org/10.21437/Interspeech.2017-1386

[35] Zhao, G.L., Sonsaat, S., Silpachai, A., et al., 2018. L2-ARCTIC: A Non-Native English Speech Corpus. In Proceedings of the Annual Conference of the International Speech Communication Association; Hyderabad, India, 2–6 September 2018; pp. 2783–2787.

[36] Yang, J., Xu, F., Ren, P., et al., 2007. The development of Chinese phonological awareness and English phonological awareness in early childhood. Psychological Science. 30(6), 1418–1421.

[37] Li, K., Qian, X.J., Meng, H.M., 2017. Mispronunciation detection and diagnosis in L2 English speech using multi-distribution deep neural networks. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 25(1), 193–207. DOI: https://doi.org/10.1109/TASLP.2016.2621675

[38] Zhang, C.Y., Zhou, A.B., Zhang, M., et al., 2004. A study on the developmental characteristics of children's English phonological awareness and Implicit Learning. Journal of Jiangsu Normal University of Technology.

10(1), 71–75.

[39] Feng, L., 2023. The present situation and improvement strategies of English pronunciation teaching in primary schools under the new curriculum standard. Jilin Education. 2023(12), 88–90.

[40] Kuhl, P.K., Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. Perception and Psychophysics. 50(2), 93–107. DOI: https://doi.org/10.3758/BF03212211

[41] Li, L., Li, Y.X., Kang, Y., et al., 2020. The role of tonal language experience in Mandarin tone perception among Chinese second speakers. Journal of South China Normal University (Social Science Edition). 2020(01), 83–91.

[42] Luo, T.F., 2020. Common problems and coping strategies of Chinese students' English pronunciation. Overseas English. 2020(17), 22–23.