International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 6514-6536

https://doi.org/10.70517/ijhsa464553

Phonetic Errors in English Vowel Articulation by Undergraduates from Shanxi Province: Analysis and Countermeasures

Yingping Liang[1,*], Yuehong Yao[2], Liya Ma[3] and Shujie Xu[4]
[1] Qiushi College (Zongfu College), Mingxiang Campus, Taiyuan University of Technology, Taiyuan, Shanxi, 030000, China
[2] School of Foreign Languages, Beijing Institute of Technology, Beijing, 100000, China
[3] College of Humanities Law and Foreign Languages, Mingxiang Campus, Taiyuan University of Technology, Taiyuan, Shanxi, 030000, China
[4] School of International Education, Shanxi University of Finance and Economics, Taiyuan, Shanxi, 030000, China
Corresponding authors: (e-mail: liangyingping@tyut.edu.cn).

Abstract Vowels are an essential part of the speech system, and their accurate pronunciation plays a fundamental role in English speech acquisition. In recent years, Chinese college students have generally struggled to master English vowels due to negative transfer from their native language, perceptual imitation learning methods, and insufficient attention to speech instruction in colleges and universities. Existing studies have primarily focused on error analysis at the national level or specific speech points, and there is still a lack of empirical research on undergraduates in Shanxi. To this end, based on the experimental phonetics method of "perception + acoustics", this paper focuses on 10 first-year students (5 males and five females) from Taiyuan University of Technology to conduct identification and cause analysis of English vowel pronunciation errors. In the experiment, the subjects read 20 English words containing vowels. The results were compared with standard British pronunciation (RP). Through experimental analysis, it was found that the subjects experienced difficulties, including problems with high and low tongue positions, confusion between front and back sounds, unclear distinctions between long and short vowels in the pronunciation of single vowels, and inadequate understanding of sliding and pronunciation transitions in the pronunciation of diphthongs. The study suggests that colleges and universities should strengthen explicit pronunciation teaching and develop a regional dialect-adaptive training mechanism to enhance learners' English pronunciation abilities.

Index Terms English vowel articulation; phonetic errors; phonological acquisition; countermeasures

I. Introduction

As research on second language acquisition continues to deepen, speech ability has been widely regarded as a crucial pillar in developing language communication skills. Speech is the physical carrier of language, carrying multiple functions of emotion, structure, and meaning. In the speech system, vowels are the core components of syllables, and their accurate mastery is of fundamental significance for learners' semantic analysis and clarity of expression [1]. However, in reality, Chinese college students face significant difficulties in vowel acquisition, which are manifested in problems such as fuzzy pronunciation, deviation of sound value, and insufficient system stability, severely restricting the development of their oral communication ability [2].

From the perspective of existing teaching mechanisms and learning models, most scholars believe that current college speech teaching is still at the primary stage, focusing on phonetic explanation and perceptual imitation, and lacks systematic training, speech recognition feedback, and acoustic tool assistance. From the perspective of cognition and input paths, most learners currently rely on the "perception-imitation" path to learn English speech. This method, which relies on subjective feelings and auditory impressions, fails to provide systematic and accurate pronunciation feedback, thereby amplifying the adverse transfer effect of the native phonetic system [3]. Shanxi dialect has significant differences from standard English in vowel system structure. Its tendency to close the mouth, the aggregation of sound values, and the lack of sound length distinction make learners more likely to make systematic errors in tongue position control, vowel delay, and sliding transitions. At the same time, current phonetics teaching practices have not yet effectively responded to regional dialect backgrounds [4].

At the research level, existing literature primarily focuses on analyzing phonetic errors among college students nationwide, as well as in the South and East. In contrast, empirical research on learners in Shanxi Province is exceptionally scarce. Existing research on the relationship between the Shanxi dialect and English phonetics primarily remains at the theoretical level, lacks empirical acoustic support, and has not yet formed a systematic

summary of error types. Therefore, there are still two gaps in the "identification of specific regional phonetic migration mechanisms" and "construction of localized phonetics teaching paths" in the current research map [5].

Given the disconnection between theoretical research and teaching practice, this paper proposes a research conception for identifying and analyzing the causes of English vowel errors under dialect transfer, based on the experimental phonetics path. In selecting research subjects, considering the representativeness of Shanxi dialects and the urgent need for phonetics teaching intervention, first-year students of Taiyuan University of Technology are chosen as the research subjects [6]. The dual paths of perceptual evaluation and acoustic analysis are integrated to systematically identify learners' standard errors in English vowel pronunciation and analyze their roots in generation from the perspective of pronunciation characteristics and language transfer mechanisms. In terms of experimental design, this paper selects 10 subjects to perform a 20-word reading task [7]. Audition collects the recording data, and Praat extracts key parameters, such as F1 and F2 frequencies and durations. Combined with RP (Received Pronunciation) as a reference standard, quantitative comparison and visual analysis are carried out to achieve the location, classification, and explanation of errors [8]. The research contributions of this paper are mainly reflected in the following three aspects:

(1) First, the paper introduces the "acoustic empirical method" into studying English speech in the context of dialects. Modeling specific tongue position parameters and duration indicators helps to promote the expansion of language transfer theory to the observable level and supplement the existing research framework based on theoretical assumptions.

(2) This paper presents an integrated analysis path of "dialect identification-error location-speech modeling", which not only facilitates error visualization but also enhances the operability of cross-dialect speech comparison, thereby enriching the empirical technology system of second language speech research.

(3) At the practical level, the error types and their generation paths revealed in this paper can provide a basis for college English teachers to optimize teaching content, adjust evaluation indicators, and offer localized references to regional English textbook compilation and teaching policy formulation.

The rest of this paper is organized in the following order: Chapter 2 reviews the research results and deficiencies of relevant literature on English vowel acquisition and dialect transfer; Chapter 3 introduces the experimental design, sample composition and analysis methods; Chapter 4 presents the empirical results and error type classification, and explores its generation logic in combination with acoustic parameters; Chapter 5 proposes teaching suggestions for teachers, students and institutions based on the experimental results; finally, Chapter 6 summarizes the entire paper and discusses the research limitations and prospects.

II. Literature Review

As an important topic at the intersection of linguistics and education, the acquisition of second language pronunciation has gradually developed into an independent research direction since the mid-20th century [9]. In the early stages, researchers primarily used error phenomena as a starting point, focusing on the impact of language input and mother tongue transfer on the development of the pronunciation system. They then gradually developed a research tradition that emphasized both theoretical construction and empirical analysis [10]. The core issues in this field include the type of induction of pronunciation errors, the cognitive mechanism of error generation, and the identification of pronunciation interference based on mother tongue transfer. Among them, vowels are the core units of language segment structure, and their pronunciation errors not only affect semantic identification and communication effects but also become a key variable for evaluating second language oral ability. Therefore, the acquisition path of English vowels, the causes of errors, and their correction mechanisms constitute an important research direction in this field [11].

In the gradual improvement of second language acquisition research, pronunciation errors, as an essential manifestation of the asymmetry between language input and output, have become a necessary focus of foreign language pronunciation teaching and acquisition research [12]. Corder (1967) clearly distinguished between "mistakes" and "errors" from the perspective of language acquisition, proposing that errors have systematic characteristics and reflect incomplete or misunderstood language knowledge. This distinction provides a theoretical framework for language error research [13]. French (1949) systematically summarized the phonetic errors of multilingual learners in the early stage, and Lee (1957) further analyzed more than 2,000 types of errors that occurred in the process of Czechoslovak language acquisition, pointing out that error classification helps teachers optimize teaching content and time allocation [14], [15]. Richard (1971) proposed a three-part classification of error causes on this basis, emphasizing the role of negative transfer from the mother tongue as one of the dominant factors [16], while the "Speech Learning Model" (SLM) constructed by Flege (1995) takes phonological similarity as the starting point and proposes that the distance between the mother tongue and the target phonemes determines the difficulty and path of second language acquisition [17].

As phonetic research has evolved from theoretical discussion to empirical modeling, error recognition technology has gradually expanded from auditory perception analysis to a multi-method system, including acoustic measurement and spectral analysis [18]. In the 1980s and 1990s, second language phonetics gradually emerged in European countries, developing in conjunction with traditional phonetics, psycholinguistics, sociolinguistics, and other disciplines, to form a complex research system centered on segmental contrast. Deterding (2006) found in his study of Chinese EFL learners that they had common problems in segmental pronunciation, such as replacing /θ/ with [s] and /n/ with /l/, and incorrectly inserting [x] instead of /h/ at the end of some speech sounds [19]. However, the empirical literature on Chinese learners remains scarce in international speech error research. Although the number of related corpus studies has gradually increased, most are based on European language construction, with limited coverage and adaptability [20].

In Chinese speech acquisition research, a research paradigm with theoretical construction and contrastive analysis as the core has gradually formed since the 1950s. Early research mainly focused on the concept of phonetic teaching and the comparison of Chinese and English phonology [21]. For example, Wu Qianzhi (1963) systematically compared the differences in intonation between Chinese and English, providing a theoretical basis for comparing pronunciation mechanisms between language systems [22]. Subsequently, Wu Qianguang (1979) proposed using phonetic errors as a crucial means to identify learning bottlenecks, marking a shift in Chinese phonetic error analysis from theoretical abstraction to problem-oriented analysis [23]. Since then, some scholars have tried to combine contrastive analysis with phonetic error research. For example, Sun Fali (1979) noted through a study of learners in Sichuan dialect areas that their phonetic errors primarily manifest as sound value deviations and high variation rates [24]. Ma Chuandong (1997) further systematically compared the differences in tongue position, sound length, and pitch control between the Sichuan dialect and English, finding that the influence of dialects has a significant interference effect on multiple levels of the phonetic system [25]. Although the research at this stage was mainly theoretical, it provided a fundamental corpus and theoretical assumptions for developing subsequent empirical methods [26], [27].

Since the beginning of the 21st century, the research path of Chinese speech errors has gradually shifted from theoretical discussion to constructing acoustic empirical models based on experimental phonetics [28]. Scholars have introduced pronunciation tasks, spectral analysis, and frequency extraction techniques in their research to quantitatively reveal the influencing mechanism of mother tongue transfer in segmental acquisition. For example, in his experimental study on the pronunciation of [w] and [v] by Chinese learners, Sun Chengkun (2019) combined acoustic data with perceptual evaluation to clarify the transfer path of specific speech pairs and proposed corresponding teaching strategies [29]. At the same time, the research objects began to expand to a broader area, covering multiple dialect belts in North China, Southwest China, Northeast China, and East China, and the proportion of research on learner variables and external environment as influencing factors gradually increased. According to statistics from Song Huiping and Zhou Weijing (2015), a total of 32 error analysis research papers were published between 1980 and 1999, of which 15 included dialect transfer as the primary variable in the research path [30], indicating that the importance of language input background is constantly increasing [31].

However, significant gaps remain in the current research landscape. On the one hand, although existing studies have covered most dialect areas in my country, research on English phonetic acquisition under the influence of the Shanxi dialect is exceptionally scarce [32]. Among the 189 relevant kinds of literature retrieved, only 3 involve the Shanxi region, and there is a lack of systematic modeling of the influence mechanism of the Central Plains dialect. On the other hand, existing studies primarily focus on enumerating and classifying segmental error phenomena, lacking micro-identification and visualization analysis based on acoustic parameters [33]. In addition, teaching suggestions are mostly empirical summaries, and a strategic intervention path for specific dialect backgrounds has not yet been formed [34]. Based on this, this paper intends to use Shanxi undergraduate first-year students from Taiyuan University of Technology as the subject, and adopt a dual method of perceptual judgment and acoustic experiments to identify and model their English vowel pronunciation errors, thereby revealing the influence mechanism of Central Plains dialect transfer in English phonetic acquisition, and providing empirical support for regionalized teaching practice and phonetic transfer theory [35]-[37].

III. Method

III. A. Participants

The participants in this study were ten first-year undergraduate students from Shanxi Province in TYUT. Their ages ranged from 20 to 23, and they came from diverse regions in Shanxi, including both rural and urban areas, as shown in Table 1. These participants had an average of 12 years of English study experience, and none had undergone systematic phonetic training. The choice of the samples was based on two primary considerations [38]: First, pronunciation proficiency is crucial for college students in their academic pursuits and future professional endeavors.

Secondly, they represent a diverse group of undergraduate students from Shanxi Province, offering a balance in terms of gender and geography. Therefore, exploring the factors affecting their English pronunciation learning holds excellent practical significance [39].

Table 1: Background of the Participants

	Native Place	Residence	Gender	Age
Participant 1	Taiyuan City	Urban	Female	22
Participant 2	Qingxu County	Rural	Female	21
Participant 3	Datong City	Urban	Female	22
Participant 4	Hunyuan County	Rural	Female	20
Participant 5	Shuozhou City	Rural	Female	23
Participant 6	Lvliang City	Urban	Male	22
Participant 7	Dingxiang County	Rural	Male	21
Participant 8	Yuanping City	Rural	Male	22
Participant 9	Xinzhou City	Urban	Male	21
Participant 10	Shanyin County	Rural	Male	23

III. B. Materials

For this study, a wordlist containing 20 English vowels was utilized. A commonly used English word accompanied each vowel [40]. The chosen words were familiar to the participants and intended to facilitate natural pronunciation. For instance, the word "is" was used to reflect the monophthong /i/and "boy" for the diphthong /ɔɪ/. The participants read these words aloud, and their pronunciations were recorded.

III. C. Equipment

The experimental equipment consisted of a Lenovo Air14 laptop, Adobe Audition 2023 audio processing software, an Edifier K750W headset equipped with a microphone, and Praat acoustic analysis software. Adobe Audition offers advanced audio editing and processing, making it ideal for recording and initially processing the participants' articulations. Praat, a free and open-source speech analysis tool developed by Professor Paul Boersma and Assistant Professor David Weenink from the University of Amsterdam, Netherlands, was utilized to analyze the recorded audio data.

III. D. Procedure

Initially, the author selected ten undergraduate students as willing participants. These individuals agreed to read the given material, and they did so in a quiet classroom in the College of Foreign Languages, TYUT. Before reading, they were given two minutes to acquaint themselves with the material and raise any questions, which the author would promptly address [41].

The participants read the target words aloud, articulating them clearly in their natural way. There were pauses of 1 to 2 seconds between each word. If any participant mispronounced a word, they were instructed to reread it. The entire reading session was recorded in the format of a WAVE file, captured using Audition software, and securely stored on the author's laptop. The recording settings were carefully calibrated to ensure high-quality audio, with a sampling accuracy of 44.1 kHz and a sampling rate of 16 bits, and all recordings were made in mono format. The author created individual voice files for each participant to facilitate the subsequent processing and analysis.

When analyzing the data, the author employed integrated perception with spectrum analysis using the Praat software. This comprehensive method was chosen to thoroughly investigate the causes of phonetic errors and develop practical solutions to address them. The author imported the processed and validated audio segments into the Praat software to annotate the target phonemes. Furthermore, formant 1 (F1) and formant 2 (F2) were extracted for the target vowels. It is worth noting that the formants are a dynamic acoustic characteristic; thus, the author adhered to the principle of extracting the most stable part of the vowel's midsection, as this provides the most reliable experimental outcomes [42].

To enhance the accuracy of the Praat spectrum analysis, the author established gender-specific maximum formant frequencies, setting the threshold at 5,000 HZ for male participants and 5,500 HZ for female participants. Additionally, Received Pronunciation (RP) pronunciations were extracted for both genders to serve as a reference. All the extracted values were recorded in an Excel worksheet and visualized in graphical form. This structured approach can ensure the accuracy and comprehensiveness of the data analysis, laying a solid foundation for insights into the phonetic errors among the participants [43].

IV. Results and Discussion

IV. A. Acoustic Features of Vowels

Vowels are composed of the monophthongs and the diphthongs. Monophthongs can be further divided into three categories according to the front or back position of the tongue, namely, the front vowels, the central vowels and the back vowels, while diphthongs are divided into two large groups by following the manner of articulation, including the closing diphthongs and the centering diphthongs. For monophthongs, their acoustic characteristics include formants and durations. Formant is the peak point shown by the vocal cord vibration on a spectrogram [44]-[46]. In phonetics, formants refer to the effect produced by the sound cavity, or more specifically, a set of frequency-doubling peaks in the sound cavity that correspond to the air frequency. Formants are essential for distinguishing human-produced vowels. The lowest frequency formant is designated as F1, followed by F2, then F3, and so on. The first and second formants are the most valuable expressions of vowel features in a language, and the values of these two formants determine the quality of a vowel. F1 determines the height of the tongue. The smaller the F1 value is, the higher the tongue is. F2 determines the front or back position of the tongue. The higher the value of F2, the more forward the tongue position is (Cruttenden, 2001) [47]. Whether the pronunciation is standard or not is closely related to the front-back position and the height of the tongue, which is reflected by the frequency of the formant, so F1 and F2 are also the key points of this experiment. The relationship between F3 and vowel tongue position is not very close, so it is not discussed in the thesis [48].

In general, since the vocal tracts of females are smaller than those of males, the vocal cord vibration frequency is higher, and correspondingly, the formant value is higher than that of males. The formants extracted in this thesis also conform to this rule, which also shows the validity of the data. Praat was used to extract F1 and F2 values of the stable segment in vowel articulation. The results were then imported into Excel to obtain the average values, and Table 2 was subsequently generated.

Table 2: Statistical Results of Monophthong Formants

		Female		Male	
		Participants	RP	Participants	RP
i	F1	385.75	391	359.67	576
	F2	2337.25	1936	2156.5	2398
i:	F1	382.25	329	312.25	372
	F2	2253.75	2476	2021.75	2664
e	F1	739	539	478	775
	F2	1727.25	1894	1738.5	2127
æ	F1	546.75	733	527.5	1006
	F2	1233.75	1540	1776	1728
ʌ	F1	819.75	962	667.5	924
	F2	1791	1399	1221.5	1134
ɜ:	F1	429.25	769	399.25	550
	F2	1439.25	1625	932	1419
ə	F1	538.75	920	657.5	662
	F2	1980.5	1503	1326.25	1577
ɔ:	F1	343.75	534	541	542
	F2	942.5	743	1062.67	748
ɒ	F1	420	523	460.5	519
	F2	1306	980	745	979
ʊ	F1	426.25	379	447.3	558
	F2	1514	880	1226	450
u:	F1	389.25	558	387.3	381
	F2	1742.25	969	882	884
a:	F1	887	790	755	746
	F2	1415	1136	1230	1068

The table shows that F1 from output of English vowels by the female college students can be ordered from highest to lowest, namely the lowest position to the highest of the tongue, as: /ɑ:/>/ʌ/>/e/>/æ/>/ə/>/ɜ:/>/ʊ/>/ɒ/>/u:/>/ɪ/>/i:/>/ɔ:/. F1 from the output of English vowels by male college students can similarly be ordered as: /ɑ:/>/ʌ/>/ə/>/ɔ:/>/æ/>/e/>/ɒ/>/ʊ/>/ɜ:/> /u:/>/ɪ/>/i:/. According to Table 2, F2 from output of female students' English vowel

pronunciation can be ordered from highest to lowest, namely from the most front position to the most back position of the tongue, as: /ɪ/>/i:/>/e/>/ʌ/>/u:/>/e/>/ʊ/>/ɜ:/>/ɑ:/>/ɒ/>/æ/>/ɔ:/. Similarly, F2 of male students can be ordered as: /ɪ/>/i:/>/e/>/ə/>/ɑ:/>/ʊ/>/ʌ/>/ɔ:/>/ɜ:/>/u:/>/ɒ/. From the obtained data, it can be seen that there are certain differences between male and female speakers, whether in terms of height or the front-back position of the tongue. This is also why male and female students should be discussed separately. However, it can also be seen that the difference is not very large, so it does not affect the unity of the conclusion.

IV. B. Analysis of Monophthongs

According to the position of the tongue, monophthongs can be categorized into three groups. The front vowels, characterized by their forward position of the tongue in articulation, include /i:/, /ɪ/, /e/ and /æ/. Central vowels, distinguished by their central positioning, consist of /ʌ/, /ə/ and /ɜ:/. Lastly, back vowels are designated by the posterior position of the tongue, encompassing /ɒ/, /ɔ:/, /ʊ/, /u:/ and /ɑ:/. These groups are analyzed individually as follows.

IV. B. 1) Formant Values of Front Vowels

(1) Female Participants

Here are the results of the formant values for the participants' front vowels.

As can be seen from Figure 1 and Figure 2, in the pronunciation of /i/, F1 value of female participants is the same as that of native speakers, indicating that the height of tongue position is identical with that of native speakers, but F2 value is higher than that of native speakers, suggesting that there are differences in the front and back of the tongue position, that is, female participants put their tongues too forward, resulting in errors in their overall pronunciation. In the pronunciation of /i/, the F1 values of female participants closely approximate those of RP speakers, indicating a comparable tongue height during articulation. However, their F2 values are significantly higher, suggesting a more forward tongue placement—this forward shift in articulation results in noticeable deviations from the standard pronunciation. In contrast, for the long vowel /i:/, participants demonstrated relatively minor deviations. Though F1 values were slightly elevated and F2 values slightly reduced—indicating a lower and more backward tongue position—the overall acoustic divergence was much smaller than that observed for other front vowels [49].

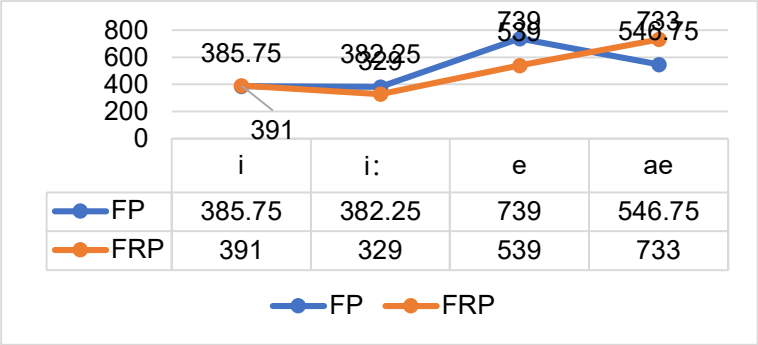

Figure 1: F1 of front vowels by FPs and FRPs. Note: FP=Female Participant FRP=Female RP Speaker

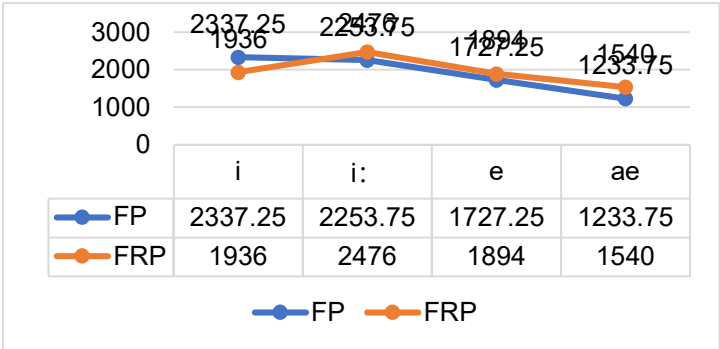

Figure 2: F2 of front vowels by FPs and FRPs

In the case of /e/, participants exhibited marked pronunciation errors. The F1 values were substantially higher than those of RP speakers, while F2 values were slightly lower, revealing that the tongue was placed too low and

slightly back. Notably, the F1 value of /e/ even exceeded that of /æ/, which is theoretically implausible given their articulatory distinctions: /e/ is a mid-front vowel and should naturally yield a lower F1 than /æ/, a low-front vowel. The RP data confirmed this expectation, yet the participants reversed this trend, reflecting a fundamental misunderstanding of vowel height differentiation. Furthermore, the similarity in F2 values between /e/ and /æ/ among participants reinforces the conclusion that they failed to perceive or produce the contrastive features of these vowels accurately.

Additionally, male participants exhibited specific difficulties in articulating/æ/, frequently failing to open their mouths adequately and tending to advance the tongue excessively. This combination of insufficient jaw opening and excessive tongue fronting led to further deviation from native pronunciation norms. Taken together, these findings highlight a pronounced challenge in the participants' phonetic discrimination of closely related front vowels, suggesting both perceptual insensitivity and articulatory inaccuracy as key factors underlying their production errors.

(2) Male Participants

Here are the results of the formant values for the participants' front vowels, as shown in Figure 3 and Figure 4.

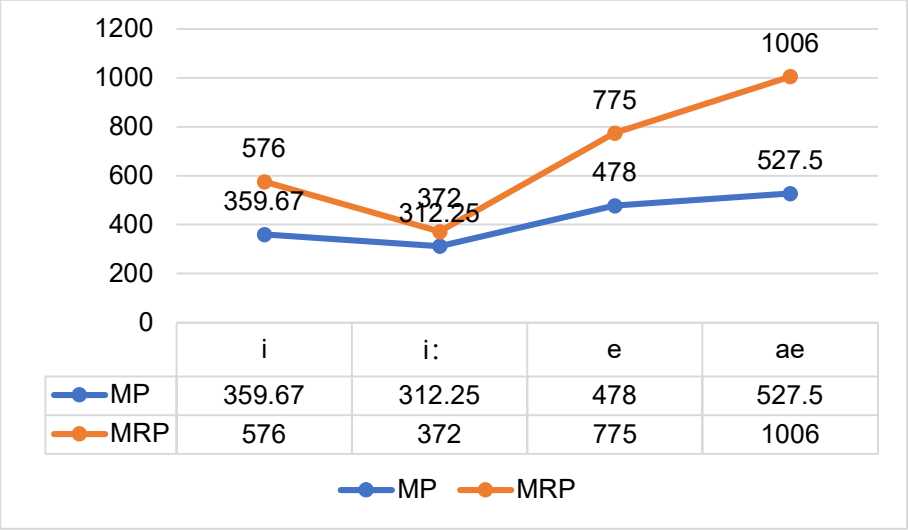

Figure 3: F1 of front vowels by MPs and MRPs. Note: MP=Male Participant MRP=Male RP speaker

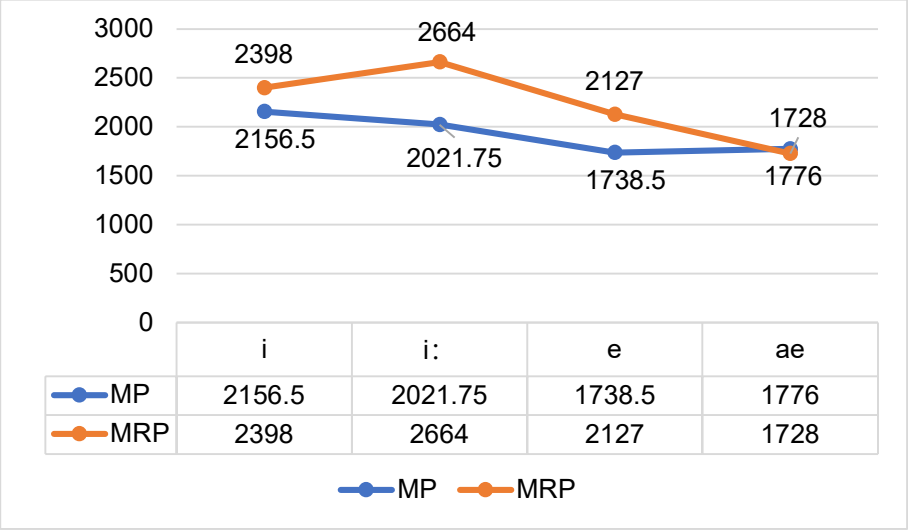

Figure 4: F2 of front vowels by MPs and MRPs

The male participants had some problems with all four vowel sounds. On the whole, except for the value of F1 for /i:/, which is close to that of native speakers, F1 values for the other three vowels are lower than those of native speakers, indicating a common problem of higher tongue position in pronunciation. In addition to /æ/, the F2 values for the other three words are also lower than those of the native speakers, indicating that the male participants tend

to place the vowel in the lower position. The high and backward position of the tongue leads to serious errors in the pronunciation of the front vowel. However, the male participants do better than the females in distinguishing between /e/ and /æ/. The F1 value of the former is lower than that of the latter, and F2 is the same. It shows that they have noticed the different heights of tongue position to distinguish the two vowels. However, in terms of performance, there is still a gap with native speakers, and it is necessary to continue strengthening practice for improved performance.

(3) Analysis of the Causes of the Errors for Front Vowels

On the whole, in the pronunciation of /i:/ and /i/, the common problem of college students from Shanxi is that there is no distinction between short and long sounds. This is greatly affected by Chinese pronunciation, which does not distinguish between long and short vowels. Some students do not realize the difference between the two sounds in the pronunciation of the tongue position. Instead, they tend to use the Chinese "yi" sound to replace /i:/ and /i/ with friction in the tongue surface, resulting in significant errors. This error is standard among many Chinese learners of English. Shang Chunyu and Bo Youhong pointed out that when native English speakers read “fill” and “feel”, the acoustic features have noticeable differences: the former is stronger than the latter and also bears a difference in tongue position. [50] They tested three other groups of words - read/rid, heel/hill, and deed/did - and the experimental results show that many Chinese students do not distinguish between these two sounds. Many of them confused these two sounds with the Chinese “yi” sound. The pronunciation of /i:/ directly starts with a vowel rather than a semi-vowel /j/.

The vowel /æ/ is similar to /e/. However, there is a slight difference in the position of the jaw and, often, in the lip spreading as well. In Mandarin Chinese, there is no counterpart to /æ/ or /e/, and students are not sensitive to the difference between these two sounds, which can be explained by Contrastive Analysis Theory, holding that learners would find the second language easier if there are some relations to their first language, vice versa. Therefore, students have a relatively poor command of these two vowels. Additionally, when some students pronounce these two vowels, a phenomenon of nasalization occurs, which is due to the presence of a nasal vowel in many local dialects of Shanxi, particularly the Taiyuan dialect, similar to the pronunciation of /æ/ in English, students from these areas would replace the English /æ/ with their mother dialect. However, it is worth noting that students from Datong have a good grasp of the/æ/ pronunciation, which is likely due to the presence of sounds similar to /æ/ in the Datong dialect. This is because the Mandarin word "ban" is pronounced like /bæ/ in the Datong dialect. Hence, students in this region benefit from some positive transfer of their mother tongue.

IV. B. 2) Formant Values of Central Vowels

(1) Female Participants

Here are the results of the formant values for the participants' central vowels, as shown in Figure 5 and Figure 6.

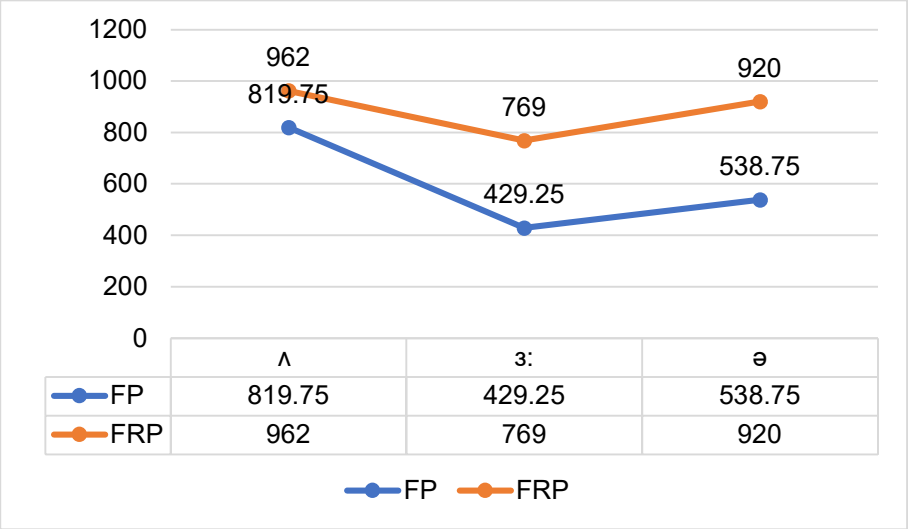

Figure 5: F1 of central vowels by FPs and FRPs

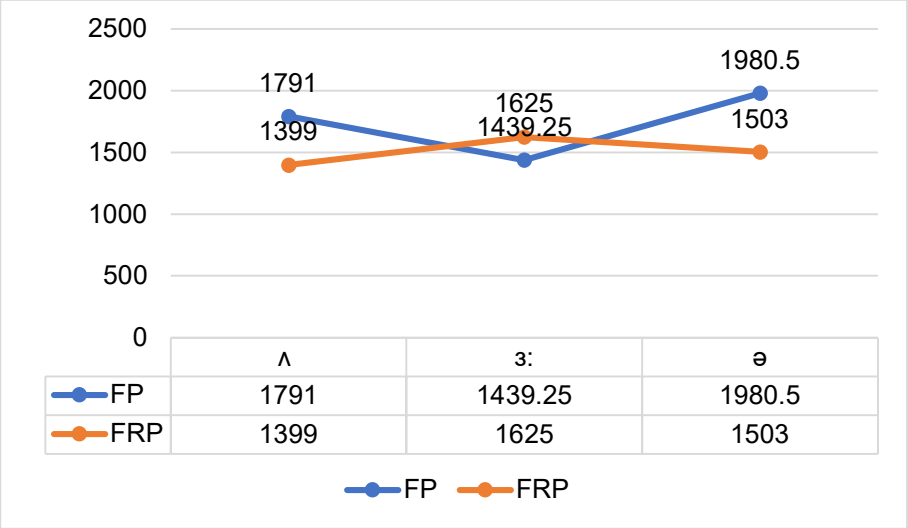

Figure 6: F2 of central vowels by FPs and FRPs

The F1 value of female participants was lower than that of native speakers when central vowel sounds were voiced, indicating that they tended to place their tongues higher. When pronouncing /ʌ/, the F1 value of female participants was lower than that of native speakers, close to that of the/ɑ:/ sound, indicating that the female participants did not notice the difference between the two sounds at the height of the tongue and pronounced them at the same height of the tongue. F2 was higher than that of the native speaker, surprisingly, even higher than that of /ɑ:/, indicating that the tongue was too far back when they articulated /ʌ/. This could be because the sound of /a/ in Chinese is similar to the sound of/ɑ:/ in English. The participants tend to pronounce similar sounds in their native language, so they didn't notice the significant difference in tongue position, which led to a grave error.

For /ə/ and /ɜ:/, F1 was lower than that of native speakers, indicating that participants had higher tongue positions when pronouncing both sounds. F2 value for /ɜ:/ was lower. In contrast, the F2 value is higher, indicating that participants have a poor grasp of the tongue position for the two sounds, and even the opposite position of the tongue occurred, which is quite different from that of native speakers. The results further reveal that female participants had limited sensitivity to the subtle differences between the central vowels /ə/ and /ɜ:/. Specifically, the values of F1 and F2 for these two vowels remain relatively similar. Consequently, they tend to mix up the height and back or front positioning of the tongue required for the accurate pronunciation of these vowels, often leading to confusion between the two sounds.

(2) Male Participants

Here are the results of the formant values for the participants' front vowels, as shown in Figure 7 and Figure 8.

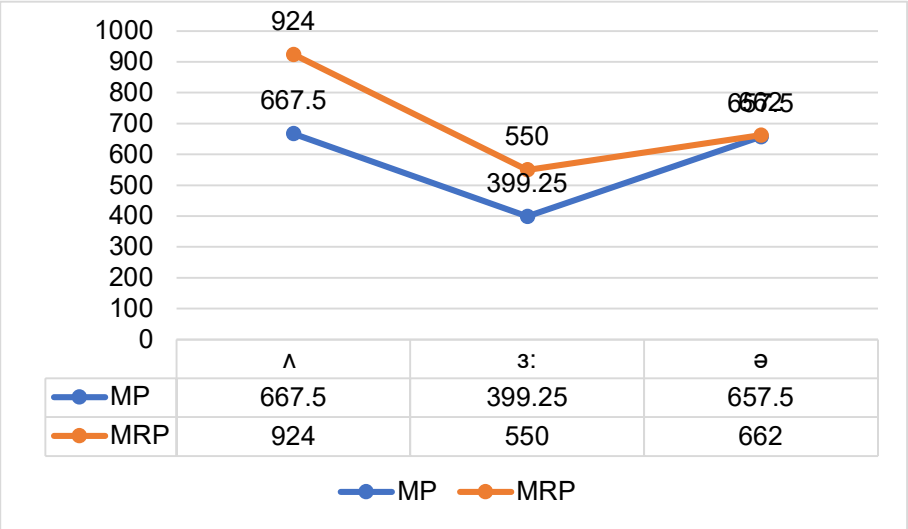

Figure 7: F1 of central vowels by MPs and MRPs

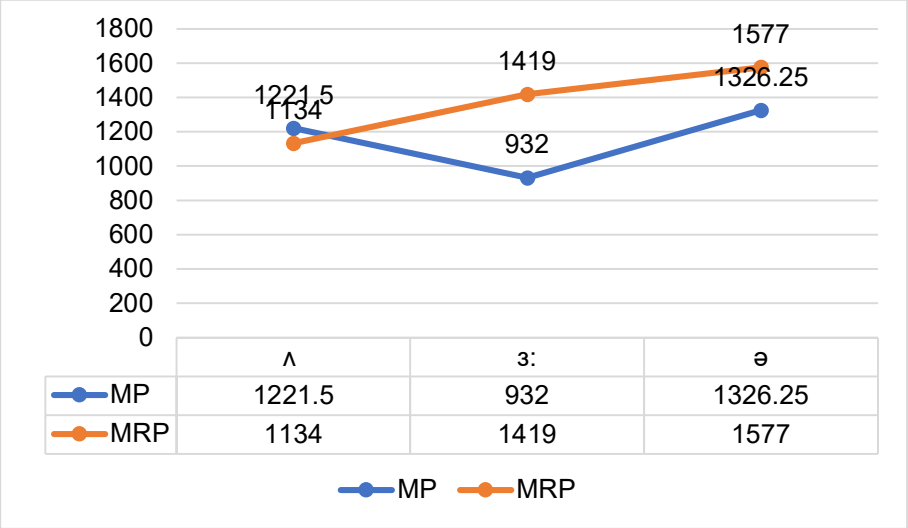

Figure 8: F2 of central vowels by MPs and MRPs

In pronouncing/ʌ/, male participants, like their female counterparts, showed lower F1 values than native speakers, indicating a higher tongue position. However, their F2 values were much closer to the RP standard, suggesting relatively accurate control over the front-back dimension of tongue placement. Compared to female participants, males demonstrated better articulatory distinction between /ʌ/ and /ɑ:/, as evidenced by a more precise separation in F1 values and closer F2 values to the native norm. This suggests that male participants were more aware of the phonemic contrast and consciously tried to differentiate them acoustically.

Male participants exhibited mixed performance for the central vowels /ɜ:/ and /ə/. In /ɜ:/, F1 and F2 were significantly lower than native speakers, reflecting a tendency to raise and retract the tongue, likely influenced by the articulation of the Chinese /e/, indicating a strong negative transfer from the mother tongue. In contrast, their pronunciation of /ə/ was more accurate: F1 aligned closely with the native level, while F2 was slightly lower, reflecting some residual influence of Chinese /e/. Male participants struggled more than females in terms of vowel duration, often failing to establish a clear distinction between short and long vowels. Notably, in some cases, the long vowel /ɜ:/ was pronounced even shorter than the short vowel /ə/, indicating a lack of awareness regarding the temporal features of vowel production, which could affect intelligibility in real communication contexts.

(3) Analysis of the Causes for the Errors of Central Vowels

Overall, college students from Shanxi were not good at distinguishing /a:/ and /ʌ/. In English, /a:/ and /ʌ/ are distinct phonemes. The difference is enough to change the meaning. While Many students pronounce /ʌ/ with their mouths too open and their tongues too low, some use the Chinese "ah" sound to replace it. This is because in Chinese, there is no distinction between /a:/ and /ʌ/, and only one "ah" sound similar to them. Therefore, students often fail to understand the distinction between these two vowels, and their native language has a significant influence on their pronunciation, resulting in specific errors. This is also a common problem among Chinese learners.

In the pronunciation of /ɜ:/ and /ə/, participants are unaware of the difference between these two sounds. Some might stress/ə the/ə /ə/, and some add the/r/ sound. As for the confounding of the distinction, it is similar to the pronunciation of /i:/ and /ɪ/ mentioned above. This problem is primarily caused by the lack of differentiation between short and long sounds in Chinese, and Chinese students often struggle to understand the concept of long and short sounds. However, there is a particular error among students from Shanxi Province, that is, the inappropriate stress of /ə/, which is because in some areas of Shanxi Province, when pronouncing "le" sound at the end of a sentence in Chinese, it is often pronounced as /la:/, which has little influence on the recognition of meaning in Chinese. However, vowels play an essential role in distinguishing meanings in English. For example, pronouncing today /tə dei/ like /tadei/ will lead to misunderstanding. Another unique error by Shanxi students is that they often add the /r/ sound when the /ə/ sound is at the end of the word. For example, in the pronunciation of "data," some participants pronounced the last syllable as /ter/ because there are many rhoticities in the Shanxi dialect, and they transfer this feature to English pronunciation.

IV. B. 3) Formant Values of Back Vowels

The five back vowels /ɒ/, /ɔ:/, /u:/, /ʊ/ and /ɑ:/ are discussed in this section.

(1) Female Participants

Here are the results of the formant values for the participants' back vowels, as shown in Figure 9 and Figure 10.

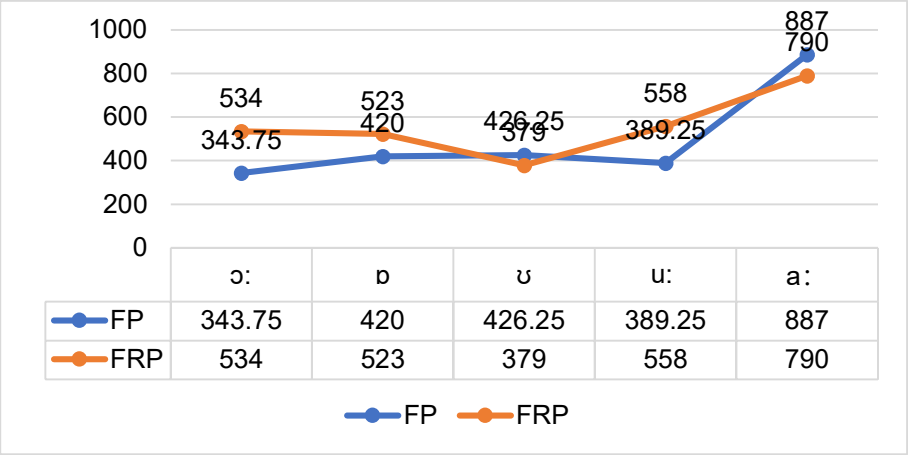

Figure 9: F1 of back vowels by FPs and FRPs

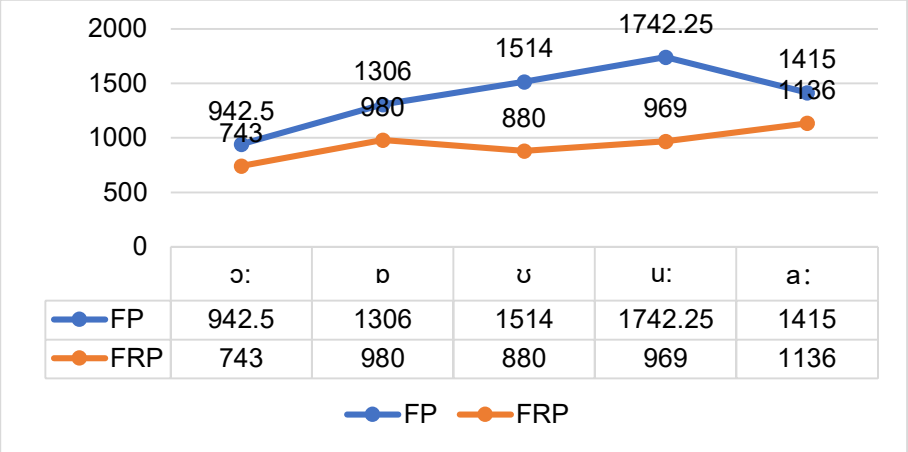

Figure 10: F2 of back vowels by FPs and FRPs

The F2 values of the five back vowels of female participants were higher than those of native speakers. The difference was significant, indicating that the tongue position was generally forward when the back vowel was pronounced. The pronunciation characteristics of the back vowel were overlooked, resulting in almost no difference between the front and back positions of the tongue when pronouncing the central vowels, which led to serious errors. The F1 value for /ʊ/ in the pronunciation of the female participants was relatively close to that of the native speakers. There was no error in the tongue height, but when pronouncing the rest of the back vowels, there were gaps between participants' tongue height and native speakers', except /ɑ:/, and the remaining ones were all on the high side.

In the pronunciation of /ɔ:/ and /ɒ/, the articulatory distinction primarily lies in tongue backness and vowel duration, as the tongue height remains relatively similar in native speech. However, for /ɒ/, female participants exhibited elevated F1 values compared to RP speakers, suggesting a lower tongue position than appropriate. This deviation led to a vague and delayed vowel quality, acoustically falling between a long and short vowel. The lack of distinction in tongue height and insufficient control over duration indicate incomplete mastery of this vowel pair.

For the short back vowel /ʊ/, female participants demonstrated relatively accurate control over tongue height, as reflected by F1 values close to those of native speakers. However, F2 values were significantly higher, indicating a forward tongue placement, contradicting the back vowel's articulatory requirement. Similarly, in the pronunciation of /u:/, F1 values were low—implying a raised tongue—but F2 remained high, further confirming the forward placement of the tongue. This suggests that participants failed to distinguish the articulatory depth between /ʊ/ and /u:/, leading to the convergence of the two sounds and phonetic confusion.

Regarding the vowel /ɑ:/, F1 and F2 values among female participants exceeded those of RP speakers. This indicates that their tongue position was both lower and more forward than required, deviating from the acoustic characteristics of a low back vowel. Native speakers typically produce /ɑ:/ with a lower F1 than /ʌ/, reflecting a higher

tongue position, but participants failed to perceive this contrast, showing confusion in vertical tongue movement. The increased F2 also underscores their limited control over back vowel articulation, with the produced sound acoustically falling between a central and back vowel.

(2) Male Participants

Here are the results of the formant values for the participants' back vowels, as shown in Figure 11 and Figure 12.

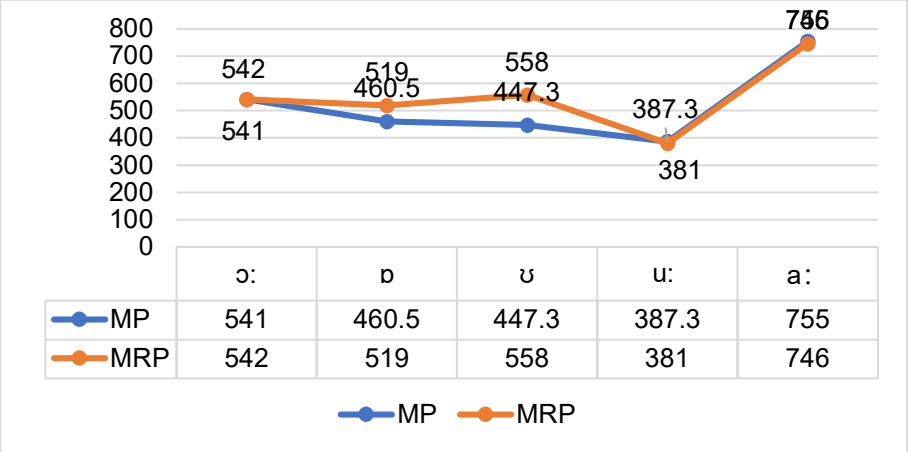

Figure 11: F1 of back vowels by MPs and MRPs

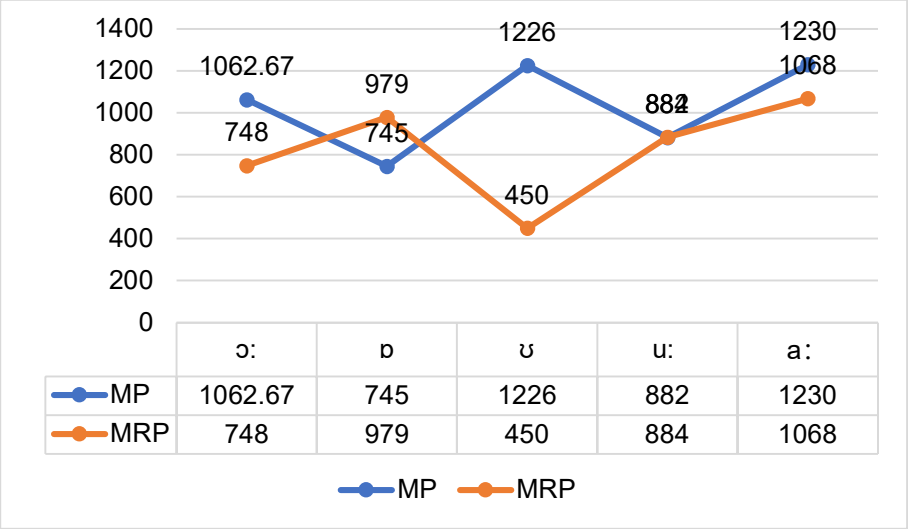

Figure 12: F2 of back vowels by MPs and MRPs

In general, male participants mastered the tongue height of the back vowel better than female participants, and the gap between the F1 of males and native speakers was smaller than that between female participants. Male participants performed better than female participants when pronouncing/u:/ and /ɑ:/. There was no problem with the pronunciation of /u:/ because F1 and F2 were identical with native speakers. In the case of /ɑ:/, the F1 value was also essentially similar to that of native speakers. In contrast, the F2 value was slightly higher than that of native speakers, indicating that male participants had a better understanding of the height and front-back position of the tongue for these two sounds. Also, they tried to reflect this in pronunciation, although when pronouncing /ɑ:/, the tongue position was still forward, which was a common problem in pronouncing back vowels.

The male participants' command of the height of the tongue remains relatively better compared with females in the pronunciation of /ɔ:/ and /ɒ/. The F1 value of /ɔ:/ and /ɒ/, compared with native speakers, is very close to that of native speakers. However, the male participants have a big problem in mastering the front-back tongue position. The F2 value of the native speakers when /ɒ/ is higher than that of /ɔ:/; but the opposite is true for the participants, indicating that the male participants have a poor grasp of the tongue position of the two sounds, leading to confusion.

Moreover, the male participants were inaccurate in distinguishing the sound length of /ɔ:/ and /ɒ/, as the difference in pronunciation between the two is slight, indicating that the participants overlooked the difference in sound length.

The male participants experienced difficulties when pronouncing /ʊ/; F1 is lower than that of native speakers, and F2 is significantly higher, indicating that the participants' tongues were positioned higher and forward as they articulated the pronunciation. In particular, there is a big gap in F2. As for the native speakers, F2 is lower than the pronunciation of /u:/, while the participants, on the contrary, indicate that they have a poor grasp of the front-back position in pronouncing/ʊ/. The tongue is too far forward and pronounced as a front vowel, resulting in a significant error.

(3) Analysis of the Causes of the Errors of Back Vowels

Overall, the participants had difficulties with pronouncing/ɔ:/ and /ɒ/, which is a common issue among Chinese students. Chinese pronunciation lacks a direct equivalent to the vowel sound /ɔ:/, making it challenging for native Chinese speakers to pronounce it accurately. Some attempt to approximate it with the Chinese sound /a/, but this fails to capture the higher tongue position and rounder mouth shape characteristic of /ɔ:/. Similarly, participants faced difficulties when attempting to produce the back vowels /ɔ:/ and /ɒ/, reflecting the complexities of mastering these non-native sounds. Two participants pronounced /ɔ:/ like Chinese /o/because they could not control their jaws and tongues properly. When pronouncing /ɔ:/, some participants pronounced it as short /ɒ/, and one participant pronounced it like a diphthong /ou/, because the participants could not keep their lips rounded enough.

When pronouncing /u:/ and /u/, participants experience difficulties with duration and tongue position. /u:/ has a long interval and a low pitch, while /u/ has a short interval and a high pitch. This is also a common problem among Chinese learners of English. Many students experience friction between their lips when pronouncing the two phonemes, such as the Chinese sound “wu”. This error also occurs because there is no difference between short and long sounds in Chinese, and only one "wu" sound is similar to it, which makes it difficult for students to distinguish between these two sounds in English pronunciation, despite using a similar sound in Chinese instead.

In the case of /a:/, as mentioned above, since there is no corresponding sound in Chinese to distinguish, and only one "ah" sound is similar to it, participants were not good at distinguishing /ʌ/ and /a:/, and they tended to replace it with "ah" in Chinese, which is also a common problem among Chinese university students.

IV. C. Analysis of Diphthongs

IV. C. 1) Acoustic Features of Diphthongs

The diphthong in English consists of two monophthongs. Different from the combination of two monophthongs, the main pronunciation characteristic of diphthongs is sliding from the first to the second, and the pronunciation should be complete when sliding. As for the classification of diphthongs, eight diphthongs in English fall into two categories. The first group is closing diphthongs, which means that the second component of the diphthong is /ɪ/ or /ʊ/, including /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, and /əʊ/. The pronunciation of these sounds changes from opening to closing. The other group is central diphthongs, whose second component is /ə/, which mainly includes /ɪə/, /eə/, and /ʊə/, and all of these sounds slide into the /ə/ sound. These two types of diphthongs are forward diphthongs; that is, the pronunciation characteristics of these diphthongs are mainly reflected by the pronunciation of the first vowel with heavy attributes of the former part and light characteristics of the latter, and long duration of the former part and shorter one of the latter. The pronunciation of the second vowel is fuzzy, and the two vowels should be combined into one after the change of mouth shape and tongue position. Therefore, the formant phenomenon of the two vowels should be observed separately for the analysis of diphthongs.

IV. C. 2) Formant Values of Closing Diphthongs

It needs to be explained that in Figure 13, 1-4 in the horizontal axis are related to /ei/, 5-8 are related to /aɪ/, 9-12 are related to /ɔɪ/, 13-16 are related to /əʊ/, and 17-10 are related to /aʊ/.

(1) Female Participants

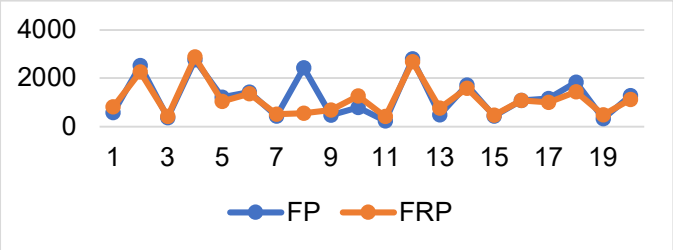

Figure 13: F1&F2 of closing diphthongs in FPs and FRPs

As shown in Figure 13, the analysis of diphthongs reveals that female participants had a relatively accurate pronunciation of /ei/ and /əʊ/. In /ei/, the articulation of the second vowel /i/was closely aligned with that of RP speakers, while the first vowel /e/ showed slight deviations—lower F1 and higher F2—indicating a higher and more forward tongue placement. This is inconsistent with their earlier monophthong /e/performance, suggesting a lack of awareness that diphthongs involve gliding between two vowels rather than producing an isolated new sound. Similarly, in /əʊ/, participants pronounced the initial /ə/ with high accuracy, while /ʊ/ was slightly misarticulated with a raised tongue, though F2 was close to native values. This shows positive transfer from segmental to diphthongal contexts, especially in tongue backness.

Conversely, the diphthong /aɪ/ posed the most significant challenge. Although the articulation of /a/ was relatively acceptable, the second part /ɪ/ showed a significantly higher F2, indicating excessive tongue fronting. Notably, since participants performed well on isolated /ɪ/ and also in /ɔɪ/, the error here likely stems not from segmental inaccuracy but from negative L1 transfer. The tendency to treat /aɪ/ as a standalone sound, influenced by the Chinese “ai,” led to a breakdown in vowel transition and segmental integrity. For /ɔɪ/, the error was concentrated in the first part /ɔ/, which showed reduced F1 and F2 values, implying a tongue position that was too high and retracted. The wide articulatory gap between /ɔ/ and /ɪ/ may have hindered smooth vowel gliding.

Regarding /aʊ/, participants consistently displayed an issue in articulating/a/, with both F1 and F2 frequencies exceeding native levels, reflecting a tongue position that was too low and forward, echoing similar problems in /a:/. In the second vowel /ʊ/, F1 was lower, and F2 was higher, suggesting a forward and slightly raised tongue. The inaccurate backness mirrors previous errors in /ʊ/, while the reduced tongue height may be affected by co-articulatory influence from /a/. These patterns indicate that while some diphthongs benefited from positive segmental transfer, others—especially those involving more complex glides or L1 phonetic interference—posed greater challenges for learners.

(2) Male Participants

Male participants exhibited more pronounced difficulties in diphthong pronunciation than female participants, with more significant acoustic deviation from native speakers. For /ei/, they performed relatively well in articulating the initial vowel /e/, showing F1 and F2 values consistent with RP norms—likely due to their better differentiation of /e/ and /æ/ noted earlier. However, the second part /i/ showed reduced F2 values, indicating a more retracted tongue position and confirming their limited mastery of front vowel backness. This aligns with previously observed issues in their production of /i/ as a near-central rather than front vowel, see Figure 14.

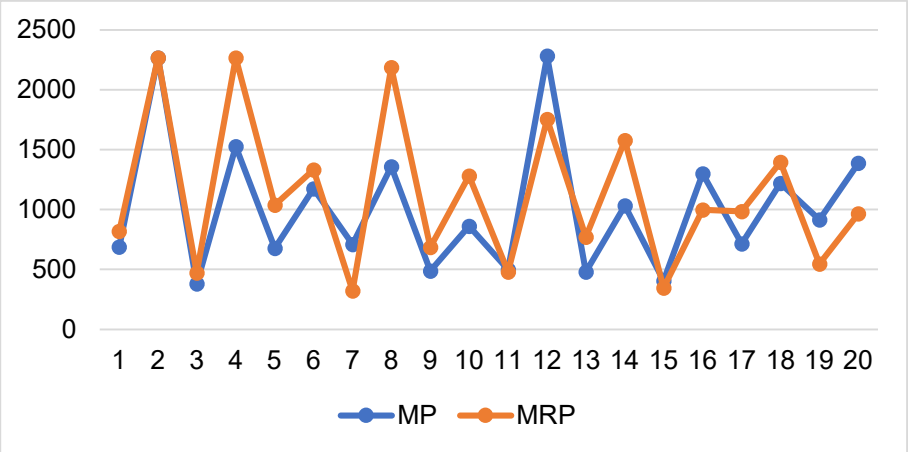

Figure 14: F1&F2 of closing diphthongs in MPs and MRPs

Pronunciation of /aɪ/ revealed significant errors. Native speakers typically show distinct formant shifts between /a/ and /ɪ/, reflecting a glide from a low-back to a high-front tongue position. Yet, male participants showed minimal F1 and F2 changes across the two segments, indicating that they failed to realize the transitional nature of diphthongs. Instead, they treated /aɪ/ as a static unit, likely under the influence of the Chinese "ai", mirroring errors also seen in female participants.

In /ɔɪ/, the first vowel /ɔ/ was produced with higher F1 and F2 values than native norms, suggesting a tongue position that was too low and forward. This reflects consistent issues in monophthong and diphthong contexts, where participants struggled with the articulatory control of back vowels. While the second part /ɪ/ showed minor F1 deviations, F2 was again overly high, suggesting continued confusion in front-back tongue positioning and failure to achieve the expected glide.

For /əʊ/ and /aʊ/, male participants also failed to demonstrate smooth formant transitions. In /əʊ/, both parts exhibited higher F1 and F2 values, reflecting a forward and lower tongue position throughout, likely influenced by the Chinese "o" sound. Similarly, in /aʊ/, F1 and F2 patterns indicated reversed tongue movement: higher and retracted in the first part, lower and fronted in the second, eliminating the glide. This suggests an articulatory compromise, where learners substituted continuous transitions with static, hybridized forms, revealing deep-seated perceptual and articulatory misunderstandings of diphthong structure.

(3) Analysis of the Causes of the Errors of Closing Diphthongs

The vowel /eɪ/ has a similar sound in Mandarin Chinese, namely "ei" with a shorter duration. As such, most speakers do not encounter significant difficulties in pronouncing /e ɪ/ with a reduced gliding element, often approximating it to the Chinese sound "ei".

However, for /aɪ/, the situation is more nuanced. While the initial sound /a/ has a counterpart in Mandarin Chinese, it is typically pronounced longer than the Chinese "a". On the other hand, the vowel /ɪ/ lacks a direct equivalent in Mandarin Chinese, leading students to commonly substitute it with the shorter Chinese "i" or attempt to approximate it with a sound resembling the Pinyin "ye". The challenge here is that /ɪ/ is a monophthong without the gliding /y/ sound at its end, requiring a different articulation where the muscles relax somewhat and the jaw drops slightly. Consequently, some speakers may confuse it with the Chinese "ai" sound.

Due to the confusion among specific learners between the English vowels /i:/ and /ɪ/ and the Chinese vowel /i/, pronouncing the diphthong /ɔɪ/ presents a challenge. When pronouncing the sound /ɔ/ in /ɔɪ/, many Chinese learners mistakenly equate it with the Chinese sound "ao". Additionally, some learners pronounce the ending vowel /i/ as the Chinese "yi", resulting in an amalgamation of the two sounds that resembles "aoyi". This issue is a common problem for Chinese EFL learners.

Furthermore, there were instances where learners pronounced the diphthong /əʊ/ as the Chinese "ou". It's important to note that /əʊ/ begins in a mid-position and then glides upward towards the vowel /u/, with the latter having a longer duration than /e/. Although Mandarin Chinese has a similar sound to /əʊ/, namely "ou", this substitution is problematic as the Chinese "ou" does not involve a rounded lip position.

Moreover, while some learners are aware that /əʊ/ is composed of the monophthongs /ə/ and /ʊ/, they lack proficiency in pronouncing these individual vowels correctly. This often leads to a mispronunciation resembling the Chinese sounds "e" and "wu" when attempting to produce /əʊ/.

Overall, pronouncing the diphthongs /ɔɪ/ and /əʊ/ accurately can be challenging for Chinese learners, primarily due to the differences in vowel systems and articulatory positions between English and Chinese. Some participants may replace it with the Chinese pinyin "ao". Students in Shanxi are also susceptible to the effect of the local dialect when pronouncing /au/. One participant from Qingxu pronounced the word "about" as /əbɔt/, because in the dialect of Qingxu County, Shanxi Province, people tend to pronounce the sound of the Chinese character "dao" /dau/ as /dɔ/, an error attributed to the size of the mouth.

IV. C. 3) Formant Values of Central Diphthongs

It is necessary to note that in Figure 15, points 1-4 on the horizontal axis are related to /ɪə/, points 5-8 are related to /eə/, and points 9-12 are related to /ʊə/.

(1) Female Participants

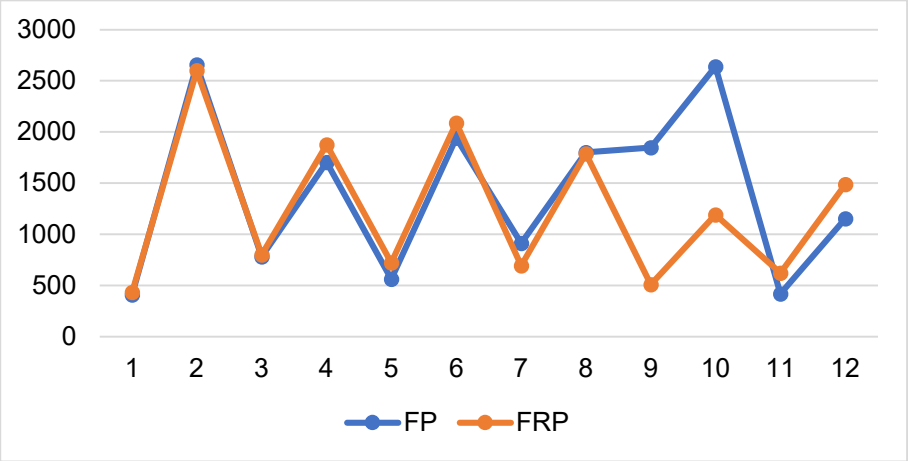

Figure 15: F1&F2 of central diphthongs in FPs and FRPs

The female participants have fewer problems in the sound of /ɪə/ and /eə/, with F1 and F2 having little difference with native speakers, indicating that the female participants can better identify the height and position of the tongue for each sound due to the comparison of the former sound and the latter one, further demonstrating that they have a particular understanding of the tongue position of different sounds. However, due to the lack of references, it is easy to produce errors in a single pronunciation. Because of the contrast between the first and second parts of the diphthong pronunciation, the tongue's position is closer to that of natives.

The female speakers experience several pronunciation issues, primarily with the/ʊə/sound, which is mainly affected in the first half of /ʊ/. F1 and F2 are significantly higher than those of native speakers, indicating that the tongue is relatively low and forward. This is the same as the abovementioned problem when pronouncing/ʊ/. It is not recognized that it is a back vowel, the tongue position is too far forward, and it is more evident in the diphthong, and even affects the height of the tongue position.

(2) Male Participants

The male participants demonstrate a poorer mastery of the three vowels than the female participants, as shown in Figure 16. The main problem with the pronunciation of /ɪə/lies in the second part. F1 is slightly lower than that of native speakers. At the same time, F2 is significantly higher than that of native speakers, indicating that the tongue position is higher and more forward than that of native speakers. This is mainly because the male participants tend to add the sound of /r/ at the end of /ɪə/, resulting in more significant errors.

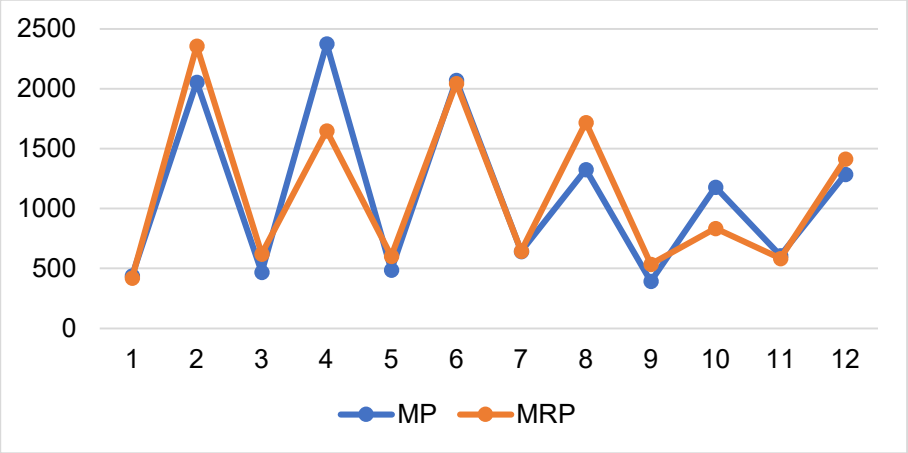

Figure 16: F1&F2 of central diphthongs in MPs and MRPs

Male participants have a relatively good grasp of the pronunciation of /eə/. In the first half, F1 and F2 differ slightly from native speakers and are relatively standard. In the second half, there is only a particular gap in F2, which is relatively lower, and it is the same as the problems in the pronunciation of /ə/ vowel alone, indicating that male participants recognize the nature of diphthong pronunciation in the pronunciation of /eə/ vowel. It is only due to the unclear knowledge of the front-back tongue of /ə/ pronunciation that the slight phonetic error is caused.

As for the pronunciation of /ʊə/, the problem primarily occurs in the first half, which is similar to the situation with the female participants; however, the specific errors differ. For the male participants, F1 is relatively lower and F2 is relatively higher, indicating that the tongue is relatively high and forward, similar to the problem in the pronunciation of /ʊ/ described above. It also shows that the male participants generally understand the nature of pronunciation, and a misperception of monophthong pronunciation is the cause of the errors. It is worth noting that males performed better in pronouncing /ʊ/ in /ʊə/ compared to monophthong /ʊ/, indicating that the contrast in diphthongs may help them better identify the correct tongue position.

(3) Analysis on the Causes for the Errors of Central Diphthongs

Generally speaking, there are fewer problems in participants' pronunciation of /eə/, except for a lack of gliding and misunderstanding of the essence of diphthongs in pronunciation. Pronouncing the vowel /ʊə/ poses specific challenges for Chinese speakers. The vowel /ʊə/ shares a degree of similarity with the Chinese sound "u", yet its accurate pronunciation differs subtly from the latter. It is crucial to avoid conflating /ʊə/ with /u/, as the two sounds possess distinct characteristics. The vowel /ʊə/ is characterized by a longer duration and a lower tongue position compared to /u/. This distinction often arises from the absence of a smooth transition or gliding motion that is typically present in the pronunciation of other vowels. This lack of transition and gliding is a common issue among Chinese learners attempting to master the pronunciation of /ʊə/. To accurately pronounce /ʊə/, learners need to focus on lengthening the vowel sound and slightly lowering their tongue position. Additionally, practicing the smooth

transition into and out of the vowel can help enhance the overall pronunciation quality. By carefully noticing these details, Chinese learners can improve their ability to pronounce /ʊə/ more naturally and accurately.

A unique problem for college students from Shanxi Province is that when the sound of /ɪə/ is at the end of a word, they tend to add the sound /r/ after. This problem is similar to that in the sound of "data" described above, and it is also due to the presence of more pediatrics in the Shanxi dialect. This indicates that the negative transfer of dialect rhotic sounds is common among English learners from Shanxi, which may be related to the teacher's pronunciation errors.

IV. C. 4) Sliding within Diphthongs

In the primary analysis, the author extracted several crucial parameters about diphthongs. Firstly, the focus was put on the overall duration of the diphthong, which serves as a fundamental metric for understanding its acoustic characteristics. Additionally, the duration and proportion of the sliding portion of the diphthong were measured, which represents the transitional aspect of the vowel sound.

Furthermore, the author delved into the duration and proportion of the stable segments of the first and second phonemes within the diphthong. These stable segments capture the steady portions of the vowel sounds before and after the sliding transition. By analyzing the values of these parameters, the study aims to investigate the existing issues with the diphthong transition and duration.

(1) Female Participants

Female participants have a poor grasp of diphthong length (Figure 17). The main performance is that the first vowel has a longer, more stable pronunciation time. The transition stage is shorter, and the second vowel has a longer stable pronunciation time. The pronunciation of /eɪ/, /ɔɪ/, /eə/, and /ʊə/ in female participants' pronunciation is longer than that of native speakers, while /aɪ/, /əʊ/, /aʊ/, and /ɪə/are shorter. These results show that females perform poorly in the length of pronouncing diphthongs (Table 3). The second problem is an insufficient diphthong sliding process. When pronouncing diphthongs, it is imperative to attend not only to the variation in the level of the tongue position before and after a single vowel but also to the smooth transition between the two sounds. This sliding process reflects the natural and fluid movement of the tongue. However, if the sliding characteristics are missed, it becomes easier to mispronounce the diphthong as a monophthong. From the perspective of the duration of the sliding portion, it is evident that the female participants exhibit significantly shorter durations and lower proportions of the sliding part compared to RP speakers. Among the diphthongs, the most significant differences were observed in the pronunciation of /aɪ/, /aʊ/, and /ʊə/, indicating an evident lack of sufficient sliding in the diphthong transition.

Table 3: Duration of Every Phase in Females' Diphthong Pronunciation

	FP			FRP		
	First Stable Phase	Gliding	Second Stable Phase	First Stable Phase	Gliding	Second Stable Phase
/ei/	0.12	0.084	0.18	0.085	0.152	0.075
/aɪ/	0.226	0.116	0.085	0.087	0.175	0.061
/ɔɪ/	0.26	0.1	0.109	0.091	0.17	0.067
/əʊ/	0.108	0.083	0.041	0.08	0.151	0.074
/aʊ/	0.095	0.137	0.048	0.085	0.155	0.077
/ɪə/	0.105	0.108	0.1	0.089	0.188	0.056
/eə/	0.22	0.09	0.08	0.077	0.189	0.062
/ʊə/	0.131	0.103	0.122	0.089	0.189	0.066

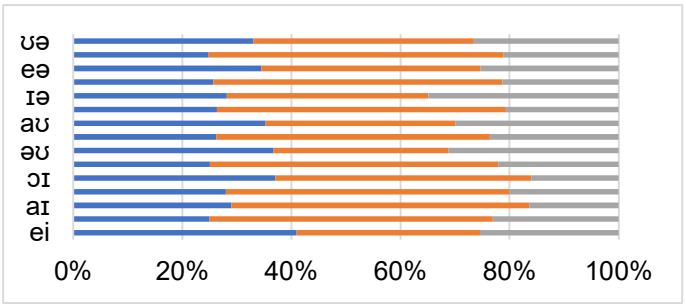

Figure 17: Contrast of the duration between FPs and FRPs

The third problem is that the significance of the first and second phonemes is misplaced. Another characteristic of diphthong pronunciation is that the former is heavy and the latter is light, with the length of the first part being extended and the length of the second part being short. Judging from the length of the first phoneme stable segment and the second phoneme stable segment in the RP native speakers which are the proportion of the total length of the phoneme stable segment, the length of the first phoneme stable segment and the proportion of the total length of the RP native speakers are larger than the second phoneme. The ratio of the stable period and the total period of the first phoneme is lower than that of the second phoneme. Females encounter significant challenges when attempting to pronounce diphthongs such as /eɪ/ and /ʊə/. This is primarily due to their tendency to adopt a pronounced short duration for the first segment and an excessively long duration for the second segment. They failed to accurately capture the inherent characteristics of diphthongs, which typically involve a shorter vowel in the first session, a smooth transition, and a longer vowel in the second session. In contrast, when pronouncing other diphthong sounds, the duration of the first phoneme is generally longer than the second, thereby adhering to the expected diphthong pronunciation characteristics.

(2) Male Participants

Male participants perform slightly better than women in the diphthong transition (Figure 18). Concerning the male participants' pronunciation of diphthongs, the total duration of all diphthongs, excluding /ɔɪ/ and /ɪə/, is notably shorter than that of native speakers. This significant difference indicates that the participants exhibit an insufficiency in the length of their diphthong pronunciations, showing the tendency to shorten the duration of long vowels (Table 4). This trend can be attributed to the absence of clear distinctions between long and short vowels in Chinese. Consequently, this lack of distinction can easily lead to pronunciation confusion for the participants.

Table 4: Duration of Every Phase in Males' Diphthong Pronunciation

	MP			MRP		
	First Stable Phase	Gliding	Second Stable Phase	First Stable Phase	Gliding	Second Stable Phase
/ei/	0.095	0.078	0.059	0.078	0.162	0.072
/aɪ/	0.076	0.143	0.043	0.094	0.175	0.067
/ɔɪ/	0.139	0.176	0.06	0.084	0.177	0.074
/əʊ	0.086	0.075	0.073	0.08	0.153	0.072
/aʊ/	0.092	0.091	0.078	0.106	0.213	0.083
/ɪə/	0.122	0.16	0.151	0.1	0.206	0.083
/eə/	0.109	0.127	0.08	0.1	0.218	0.085
/ʊə/	0.112	0.137	0.09	0.083	0.202	0.072

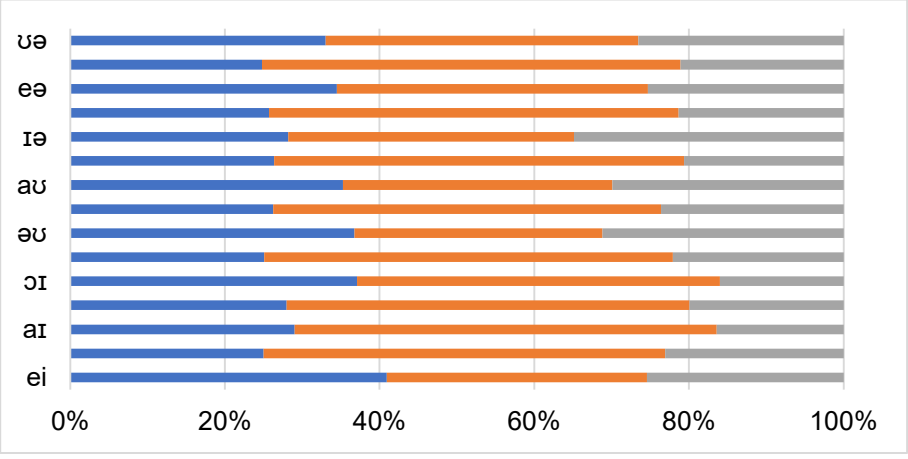

Figure 18: Contrast of the duration between MPs and MRPs

Male participants perform slightly better than female participants in terms of the lack of diphthong sliding process, but there are still deficiencies. Regarding the length of the sliding part and the proportion of the total length, except for/eə/, the value of male participants is also significantly lower than that of RP native speakers, indicating that almost every pronunciation has a problem with insufficient sliding length. Generally speaking, the male participants

have a good grasp of the pronunciation characteristics of the long first and short second parts. Except for /ɔɪ/, other diphthongs show this characteristic.

(3) Analysis of the Causes for the Gliding Errors

Lack of gliding is a common problem among Chinese learners, primarily due to the influence of the Chinese language. The sliding process of Chinese diphthongs is faster, whereas English requires a more gradual sliding process. The diphthong in English combines two vowels, sliding from the first vowel to the second vowel without any pause. It is characterized as follows: the first vowel is long, loud, and clear. The second vowel is short, weak, and ambiguous. Therefore, the English diphthong is a compound sound with a larger pronunciation stroke. In comparison, the Chinese compound sound has a much more minor pronunciation stroke, and many Chinese dialects even shorten it to a single sound. Chinese English learners tend to ignore this and thus produce diphthongs that are not long enough or glide enough to sound like Chinese diphthongs. Their mother's dialect influences some students, and they even use the monophthongs of their mother's dialect to replace the more complex compound sounds of English.

V. Countermeasures

V. A. Causes for Phonetic Errors

For vowel analysis, the closer the average parameter value of the participants aligns with that of native RP speakers, the superior their acquisition level and the more precise their pronunciation. Consequently, through analysis of the data, it can be concluded that the main errors of English vowel pronunciation of college students from Shanxi are as follows:

(1) The grasp of the tongue position when pronouncing vowels is poor. Regardless of whether the pronunciation is monophthong or diphthong, both male and female participants experience significant difficulties in recognizing and performing the tongue position, and tend to use the Chinese sound instead of the English sound, which is a good example of negative transfer.

(2) Since there is no distinction between short and long sounds in Chinese, Chinese learners generally have the problem of poor distinction between short and long sounds, basically pronouncing them into similar sounds, and even replacing them with a particular sound in Chinese, such as "ah" to replace /ɑ:/ and /ʌ/, "eh" to replace/ɜ:/ and /ə/, etc.

(3) The duration error is obvious. There exists a gender consistency, where both male and female non-native speakers tend to pronounce long vowels and diphthongs considerably shorter than their native counterparts when compared with Received Pronunciation.

(4) The diphthong gliding is insufficient, and the characteristics of a long first and short second parts are not realized. Both male and female participants exhibit a deficiency in diphthong gliding, which is evident in the proportion of gliding time to the total pronunciation time, differing significantly from that of native speakers. At the same time, the characteristics of diphthongs, with a long first part and a short second part, are less pronounced in female participants, and the short first and long second parts appear more frequently. In addition, both male and female participants had a problem with an unclear understanding of the principle of diphthong pronunciation. They tended to use a particular sound in Chinese instead of diphthong pronunciation.

(5) Additionally, some unique pronunciation problems exist among Shanxi learners. Influenced by local dialects such as Taiyuan, students from these areas add nasal sounds after the/æ/sound. Influenced by the dialects of northern Shanxi, students usually stress the /ə/ sound. Influenced by local dialects such as Shuozhou and Datong, students will likely confuse/eɪ/ and /aɪ/. Influenced by the local dialect, such as the Qing Xu dialects, students can confound /aʊ/ and /ɒ/. Influenced by the rhoticity of the Shanxi dialect, students tend to add the sound /r/ after the sound /ə/ and /ɪə/ if it occurs at the end of a word.

The causes for these errors can be summarized as follows:

(1) Traditional, exam-oriented education often leads to the neglect of pronunciation teaching in English instruction. Most English teachers only pay attention to students' academic performance during the teaching process, ignoring their phonetic errors; some even fail to correct them, resulting in students' lack of basic pronunciation knowledge and the development of bad pronunciation habits. Some students do not understand the essentials of English pronunciation and tend to use similar Chinese pronunciations as replacements.

(2) Local dialect accents seriously affect the pronunciation of English. Although some students can overcome this problem, most are unable to do so. Specifically, in the pursuit of mastering English as a second language, Chinese pronunciation habits can potentially exert a notable interfering influence on learners. This phenomenon is referred to as "negative transfer" within the framework of language transfer theory. Furthermore, this interference is not limited to Mandarin but also extends to dialects, the regional variations of a language, especially at the phonetic level. Learners from dialectal regions are conditioned to employ the pronunciation system characteristic of their

local dialects. Consequently, they often use the pronunciation of local dialects to replace English pronunciation when learning the Language, which inevitably leads to pronunciation errors.

(3) Students lack good sound discrimination ability, so they are not aware of their pronunciation, which significantly hinders their English pronunciation learning process. In the English class, it is evident that some students struggle to distinguish between self-pronunciation and standard pronunciation, and they are unaware of their mistakes, which makes it difficult for them to correct them.

V. B. Suggestions

Based on the above analysis, the common problems of English learners from Shanxi in English vowel articulation and their causes have been identified. To address these issues and enhance the phonetic acquisition of English learners, targeted suggestions are presented in this section for educational institutions, teachers, and English learners.

V. B. 1) Suggestions for Educational Institutions

One of the root causes of current pronunciation errors is that students lack a clear understanding and adequate training in standard pronunciation. Educational institutions should incorporate the study of the International Phonetic Alphabet (IPA) into their basic pronunciation teaching systems and focus on a teaching method that combines imitation with lip shape analysis, guiding learners to master the rules of tongue position changes and the skills of distinguishing long and short vowels. At the same time, teaching diphthongs should emphasize the natural sliding and coherence in the pronunciation process. Students should pay more attention to pronunciation accuracy at the educational policy level. Since pronunciation has little impact on grades, many students lack the motivation to correct it. Education administrators should guide students in recognizing the fundamental importance of pronunciation ability in language literacy through curriculum standards. For example, the latest version of the "General High School English Curriculum Standards" (2017) emphasizes the importance of phonetic intuition and phonetic awareness in language application ability. In addition, the compilation of textbooks should strengthen the systematic nature of phonetic content and learn from the integration of segmental and sound change knowledge in the 2019 textbooks published by the People's Education Press, aiming to build a clear phonetic system for students.

In addition, educational institutions should pay attention to the "critical period" effect of speech acquisition, especially during the primary stage, and allocate more resources to speech training. They should also construct teacher speech evaluation standards to ensure that students are initially exposed to standard speech. As the key node of speech reconstruction, the university stage should incorporate speech error correction content into course design and adjust the teaching syllabus according to the actual situation of students. This can be achieved, for example, by conducting a comparative course on the differences between Chinese and English pronunciation to enhance students' awareness of mother tongue transfer. In addition, dialect teaching should also retain a place in the classroom, as it reflects cultural heritage and helps students recognize the potential impact of dialects on English pronunciation, thereby facilitating positive transfer from the dialect to English.

V. B. 2) Suggestions for Teachers

Teachers play a key role in phonetics teaching. Their phonetics level directly affects the quality of students' imitation, so they need solid phonetics knowledge and teaching ability. Teachers should master visual phonetics teaching tools (such as Praat and Pronunciation Power), simplify abstract concepts with the aid of visual aids, and enhance classroom efficiency. Studies have shown that visualization technology helps improve students' phonics recognition and imitation abilities. In addition, teachers should have a deep understanding of the pronunciation mechanism, accurately identify students' common errors, and conduct targeted analysis and design of correction strategies. In kindergarten and primary school, teachers should utilize professional equipment to calibrate students' pronunciation in real-time and guide them in forming correct phonetic habits. University teachers need to correct and reconstruct students' phonetic systems, especially in the comparative teaching of Chinese and English phonetics, to guide students in grasping the core differences and avoiding the negative transfer of their mother tongue. In addition, teachers should also understand dialect knowledge and utilize dialect resources in their teaching to achieve positive phonetic transfer and enhance teaching effectiveness.

V. B. 3) Suggestions for Learners

Learners themselves are the core of improving pronunciation ability. Learning English should lay a sure foundation in pronunciation, pronunciation discrimination, and correction of pronunciation errors. First, we should fully recognize the fundamental role of accurate pronunciation in language communication and skill development and actively strengthen pronunciation practice and self-correction awareness. When learning vowel pronunciation, we need to master the coordination mechanism of the pronunciation organs (such as tongue position, mouth shape, lip

shape, and sound length) and improve accuracy through imitation, recording, and comparison with standard sounds. At the same time, students should be more sensitive to the differences in pronunciation between Mandarin, dialects, and English to avoid interference from their mother tongue. Secondly, students should focus on learning to identify and distinguish similar vowels, such as /e/ and /æ/, /aɪ/ and /eɪ/, and use phonemes repeatedly in authentic contexts. For example, students can enhance voice input and improve listening ability by watching English movies and listening to English songs. The auxiliary role of online resources cannot be ignored, as they help students perceive voice changes and natural speech flow. Finally, students should approach corrections from others with an open and positive attitude, correct voice errors promptly, actively imitate a standard voice, and continually optimize their pronunciation system. Through continuous practice, diversified input, and effective feedback, English learners can steadily improve their voice ability and overall language proficiency.

Overall, EFL learners in China should make an effort to grasp effective learning techniques, practice regularly, and actively seek to correct their errors. By exploring diversified resources and adopting the right approach, they can surely enhance their pronunciation skills and achieve excellent linguistic proficiency.

VI. Conclusion

Guided by experimental phonetics, this study examines the vowel pronunciation errors and underlying challenges faced by English learners from Shanxi Province, with a specific focus on ten undergraduate students at Taiyuan University of Technology (TYUT). To address persistent phonetic difficulties among EFL learners in northern China, the study employs an empirical approach, utilizing controlled recording tasks and spectrographic analysis to examine participants' articulation of monophthongs and diphthongs. The findings reveal systematic pronunciation deviations, including misplacement of tongue height and front-back positions, confusion between long and short vowels, and insufficient gliding in diphthongs. These issues stem from multiple factors: insufficient emphasis on pronunciation training in formal education, the negative transfer of regional dialects, and limited phonemic awareness. On this basis, the study provides targeted pedagogical suggestions for educational institutions, language instructors, and learners, emphasizing the importance of early phonetic intervention, dialect-informed instruction, and the enhanced use of visual and auditory feedback technologies to improve learning outcomes.

However, the study acknowledges several limitations that constrain the generalizability of its findings. The small sample size, limited to ten learners, may not fully capture the broader diversity of pronunciation patterns across Shanxi or other regions. Moreover, the reference materials used—recordings of a single male and female RP speaker—might contain individual phonetic biases, potentially influencing the baseline for analysis. The scope of the research also remains confined to isolated vowel sounds, without extending to pronunciation in connected speech or contextualized discourse. Future research should expand the sample population, incorporate more objective and standardized pronunciation benchmarks, and explore suprasegmental features and sentence-level articulation. Such efforts would provide a more comprehensive understanding of Chinese learners' phonetic development and support the design of more effective pronunciation training programs.

Acknowledgements

None.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Appendix

Please read the following words consistently.

Words for Testing	
/i:/	leap
/ɪ/	is
/e/	guess
/æ/	dad
/ɜ:/	word
/ə/	data
/ʌ/	such
/ɑ:/	father
/ɔ:/	more
/ɒ/	tall
/u:/	who
/ʊ/	put
/eɪ/	age
/aɪ/	drive
/ɔɪ/	boy
/əʊ/	joke
/aʊ/	mouth
/ɪə/	idea
/eə/	chair
/ʊə/	sure

References

[1] Ashby M, and Maidment J. Introducing Phonetic Science [M]. Cambridge University Press, 2005.
[2] Bent T, Bradlow A R, and Smith B. Phonemic Errors in Different Word Positions and Their Effects on Intelligibility of Non-native Speech [M]. 2007.
[3] Brown A. Lessons from Good Language Learners: Pronunciation and good language learners [J]. 2008.
[4] Chan K Y, and Hall M D. The Importance of Vowel Formant Frequencies and Proximity in Vowel Space to the Perception of Foreign Accent [J]. Journal of Phonetics, 2019.
[5] Corder S P. Idiosyncratic Dialects and Error Analysis [J]. IRAL- International Review of Applied Linguistics in Language Teaching, 1971, 9(2): 147-160.
[6] Fant G. Acoustic Theory of Speech Production: with Calculations Based on X-ray Studies of Russian Articulations [M]. Walter de Gruyter, 1971.
[7] Frisch S A, and Wright R. The Phonetics of Phonological Speech Errors: An Acoustic Analysis of Slips of the Tongue [J]. Journal of Phonetics, 2002, 30(2): 139-162.
[8] Gimson, A. C. An Introduction to the Pronunciation of English [M]. 4th ed. London: Arnold, 1989.
[9] James, C. Errors in Language Learning and Use: Exploring Error Analysis [M]. Beijing: Foreign Language Teaching and Research Press, 2001.
[10] Ladefoged P, and Disner S F. Vowels and Consonants [M]. John Wiley & Sons, 2012.
[11] Lado R. Linguistics Across Cultures; Applied Linguistics for Language Teachers [J]. 1957.
[12] Munro M J, and Derwing T M. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners [J]. Language learning, 1995, 45(1): 73-97.
[13] Corder S P. The Significance of Learner's Errors [J]. IRAL- International Review of Applied Linguistics in Language Teaching, 1967, 5(1-4): 161-170.
[14] French F G. Common Errors in English: Their Cause, Prevention and Cure [J]. 1949.
[15] Lee, W. R. The Linguistic Context of Language Teaching [J]. Elt Journal, 1957, XI (3): 77-85.
[16] Richards J C. A Non-Contrastive Approach to Error Analysis [J]. ELT Journal, 1970, XXV (3): 204-219.
[17] Flege J E. Second-language Speech Learning: Theory, Findings and Problems [J]. Speech Perception & Lingual Experience Issues in Cross Language Research, 1995.
[18] Selinker L, and Gass S M. Second Language Acquisition [J]. Lawrence Erlhaum Ass, 2008.
[19] Deterding, and David. The Pronunciation of English by Speakers from China [J]. English World-Wide, 2006, 27(2): 175-198.
[20] Suenobu M. An Experimental Study of Intelligibility of Japanese English [J]. Iral, 1992.
[21] Z. esták. James, C. Errors in Language Learning and Use. Exploring Error Analysis [J]. Photosynthetica, 2000, 38(4): 538-538.
[22] Liu, Y. H. (2024). The application of mainstream intonation theories in college English phonetic teaching: Review and prospect. Modern Foreign Languages, 2024(3), 61-73.
[23] Wu, Q. G. (1979). On the application of error analysis in English teaching. Modern Foreign Languages, (4), 40-48.
[24] Sun, F. L. (1979). Analysis of common English pronunciation errors among Sichuan students. East Sichuan Journal (Educational Research Edition), 7(4), 1-11.
[25] Ma, C. D. (1997). Analysis of common English pronunciation errors among students in Sichuan dialect areas. East Sichuan Journal (Educational Research Edition), 7(4), 1-11.
[26] Wong J, and Celce-Murcia M. A Re-examination of (the) Same Using Data from Spoken English A Re-examination of (the) Same Using Data from Spoken English [J]. Ilha do Desterro, 2008, (41): 185.

[27] Chen, H. M. (2013). The commonality of chinese college students' errors in english phonetic acquisition and countermeasures. Economic and Social Development, 11(03), 220-223.
[28] Chen, Y. (2013). Theoretical models and pedagogical implications of second language speech perception. Journal of Foreign Languages, 36(3), 68-76.
[29] Sun, C. K. (2019). An experimental study on the pronunciation errors of [w] and [v] sounds in English by Chinese native speakers. English Teachers, 19(17), 39-42.
[30] Zhou, W. J., Shao, P. F., & Chen, H. (2010). An empirical study on English majors' perception of RP vowels. Journal of PLA University of Foreign Languages, 33(6), 45-49+128.
[31] Cheng, C. M., & He, A. P. (2008). An analysis of oral segmental errors of advanced English learners—A corpus-based study. Journal of PLA University of Foreign Languages, 31(1), 38-42.
[32] Du, Y. Y., Luo, Q., & Ren, X. Y. (2012). A survey and analysis of college students' English individual consonant pronunciation errors. Chenggong (Education), (18), 34.
[33] Gao, L. (2011). A review of Chinese learners' English phonetic acquisition research. Journal of Donghua University (Social Science), 11(1), 31-35.
[34] Jia, S. N. (2010). A report on the common English pronunciation mistakes made by the English majored Cantonese students of higher vocational college. Journal of Sichuan College of Education, 26(7), 103-109.
[35] Li, T. (2008). A survey and analysis of English and Mandarin Chinese pronunciation in the Cantonese dialect area. Academic Exchange, 168(3), 145-147.
[36] Li, Y. H. (2009). Phonetic errors often made by students from Shandong Province and their analysis. Journal of Liaocheng University (Social Science Edition), 2009(1), 118-121.
[37] Liu, X. J. (2007). Errors' analyzing of English pronunciation to mother dialects. Journal of Baoji University of Arts and Sciences (Social Sciences), 27(4), 108-110.
[38] Shang, C. Y. (2014). Analysis of English pronunciation errors of Gansu students based on experimental phonetics. Gansu Science and Technology, 30(13), 1-5.
[39] Shang, C. Y. (2020). English Phonetic Suprasegmental Phonemes Learning Errors on Experimental Analysis among College Students. Journal of North China University of Science and Technology (Social Science Edition), 20(4), 97-101.
[40] Xie, Q. Q., & Bi, R. (2024). Visualization Analysis of the Research Progress in Chinese English Phonetic Teaching under the Background of Artificial Intelligence. Journal of Foreign Languages, 34(1), 101-104.
[41] Song, J. (2016). A study on the factors affecting English pronunciation errors among college students in Shandong. Time Education, (13), 102.
[42] Tang, J. (2013). Error analysis of pronunciation of adult learning English speech in parts of South China in the dialect area. Journal of Guangdong University of Foreign Studies, (2), 166.
[43] Wang, C. M. (2001). Two major factors affecting foreign language learning and foreign language teaching. Foreign Language World, (6), 8-05.
[44] Xu, L., & Wang, Y. (2020). An experimental study of vowel acoustic features and speech errors by Chinese Mandarin-speaking EFL learners. Journal of Chinese Phonetics, (2), 114-122.
[45] Yang, E. H. (2020). An analysis of the phonetic errors of college students and the corresponding strategies. Curriculum Teaching, (2), 1-1.
[46] Wei, W. Y. (2024). A comparative analysis of English vowel weakening in "-man" and Chinese human-indicating suffixes. Overseas English, (2), 74-77.
[47] Cruttenden, and Alan. Gimson's Pronunciation of English [M]. Foreign Language Teaching, 2001.
[48] Zhang, S. R. (2022). A study on English pronunciation errors and correction strategies among English major students. Overseas English, (9), 105-107.
[49] ZHI, N., & LI, A. J. (2020). Phonetic Training Based on Visualized Articulatory Model. Journal of Foreign Languages, 43(1), 59-74.
[50] Shang, C. Y., & Bu, Y. H. (2009). Error analysis of English vowels based on experimental phonetics. Journal of Tianshui Normal University, 29(6), 1-6.