Journal of English Language Teaching and Applied Linguistics
ISSN: 2707-756X
DOI: 10.32996/jeltal
Journal Homepage: www.al-kindipublisher.com/index.php/jeltal

JELTAL
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

An Empirical Study on the Segmental Features of English Pronunciation of High School Students and Their Impact on Speech Intelligibility

HUANG Xiaohu
School of Foreign Studies, Nanjing University, China
Corresponding Author: HUANG Xiaohu, **E-mail**: Huangxu_study@smail.nju.edu.cn

| ABSTRACT

Oral examinations are now part of China's College Entrance Examination, yet many students lose points due to pronunciation errors that reduce intelligibility, largely stemming from the L1 transfer. High-school instruction seldom targets intelligibility-critical segmental features, partly because these features and their effects are under-researched. To address this gap, researchers analyzed the segmental pronunciation and intelligibility of 55 Grade-10 students in a Chongqing high school. Speech was collected via Voice Memo, annotated by trained raters for segmental errors, and rated for intelligibility; statistical analyses were conducted in IBM SPSS. Results show vowel errors outnumber consonant errors, with certain phonemes (e.g., /iː/, /ɪ/, /θ/) exceeding a 40% error rate. These errors frequently led raters to misidentify words, directly lowering intelligibility scores. Segmental error rates were significantly associated with intelligibility, indicating that targeted training on high-error vowels and consonants can yield measurable gains. The study offers practical guidance for pronunciation instruction, especially in dialect-influenced regions.

| KEYWORDS

English phonetics; pronunciation features; intelligibility

| ARTICLE INFORMATION

ACCEPTED: 20 October 2025 **PUBLISHED:** 30 October 2025 **DOI:** 10.32996/jeltal.2025.7.6.3

1. Introduction

Pronunciation is a core determinant of L2 speaking performance because it directly affects speech intelligibility. For many Chinese learners of English, substantial L1–L2 phonological distance fosters negative transfer, making segmental accuracy difficult to master and harder to self-correct than lexical or syntactic errors. When unfamiliar L2 sounds are encountered, learners often substitute the closest L1 categories, producing recognizable "foreign accent" (Ellis, 1999) that can reduce intelligibility, especially when both interlocutors are non-native speakers. Despite its centrality to communicative success, segmental accuracy has received less systematic attention in classroom practice than grammar or vocabulary.

Two major lines of research have addressed intelligibility: properties of the speech stimulus and listener factors (Munro, 2008). Using rating scales, transcription, and related techniques, studies have identified speech features linked to intelligibility (Kang, 2012; Munro & Derwing, 1995; Pérez-Ramón et al., 2022), while listener-oriented work highlights the roles of proficiency, experience, and other variables (e.g., Cao & Chen, 2023; Hayes-Harb et al., 2008; Li & Wang, 2015; Sun & Chen, 2022; Xie & Fowler, 2013; Yin, 2015; Zoghbor, 2018). Together, this scholarship underscores the complexity of intelligibility and has informed approaches to teaching and assessment. However, several gaps remain. First, many studies still privilege native-speaker judgments as a benchmark, which may not reflect typical L2–L2 communication. Second, within Chinese secondary education, research and pedagogy often emphasize global "accent" reduction rather than pinpointing which segmental errors most strongly depress intelligibility. Yet not all pronunciation deviations equally impede understanding; for instance, the frequent /θ/→/s/ substitution among Chinese learners does not always prevent successful comprehension in local L2 interactions.

Given these issues, there is a need for fine-grained, context-specific evidence on the segmental features that most affect intelligibility in Chinese high schools. The present study addresses this need by examining the English segmental pronunciation of Grade-10 students and quantifying how specific error patterns relate to intelligibility outcomes. By identifying high-impact vowels and consonants and their error rates, the study aims to provide actionable guidance for instruction and assessment that prioritizes intelligibility-critical segments over low-impact deviations.

2. Literature Review

Research on L2 pronunciation errors has expanded since the 1960s with richer methods and tools, yet several blind spots remain. Empirical work often targets teachers or undergraduates for convenience, leaving high-school learners underrepresented. Most studies foreground negative L1 transfer while giving comparatively less attention to intra-language variation such as regional dialect influence.

Within segmentals, vowels and consonants pattern differently. Kennedy (2003) notes vowel "cluster" behavior analogous to consonants, implying multiple loci for vowel errors within words. In a diagnostic study of 40 Saudi EFL learners, Alzainidi & Latif (2019) found lower-intermediate learners produced more consonant and cluster errors than intermediate learners, with greater variance word-initially.

In China, error analysis began later but emphasized systematic diagnosis to inform instruction (Wu, 1979). Dialect effects have been observed: Sun (1979) described common segmental errors among Sichuan learners; Yin (2001) analyzed causes across counties in southeast Gansu. These studies advanced descriptive accounts of dialect-conditioned segmental errors, but many were qualitative, localized, and not tightly linked to intelligibility outcomes.

Two research streams frame intelligibility: properties of the speech signal and listener factors (Munro, 2008). Using ratings and transcription, prior work related specific speech features to intelligibility (e.g., Kang, 2012; Munro & Derwing, 1995; Pérez-Ramón et al., 2022). The Lingua Franca Core (Jenkins, 2002) further argues for prioritizing segmental targets essential to mutual understanding in L2–L2 communication, challenging native-norm emphasis and aligning instruction with functional intelligibility. Listener-focused studies probe proficiency and experience effects and often report broadly comparable comprehension across diverse L1 pairings (e.g., Munro et al., 2006; Crowther et al., 2016), suggesting that not all segmental deviations equally reduce understanding.

First, dialect-specific segmental error profiles for Chinese high-school learners—especially in southwestern varieties such as Chongqing—are scarce. Second, many studies still privilege native-speaker judgments rather than typical L2–L2 interactions. Third, few projects quantify which segmental errors most strongly correlate with intelligibility in secondary-school contexts, partly due to access and time constraints with this population.

Based on the previous studies, this study aims at addressing these gaps by profiling Grade-10 learners with a Chongqing dialect background, mapping their segmental error patterns, and testing their associations with intelligibility, with the goal of generating actionable, dialect-aware priorities for pronunciation instruction and assessment.

3. Methodology

3.1 Research questions

Features of pronunciation errors and their correlation with intelligibility would be analyzed and summarized in the present study according to the three questions: (1) What exact errors did the participants make? (2) How did these segmental errors cause misunderstanding? (3) What are the correlations between speech intelligibility and segmental errors?

3.2 Participants and raters

Participants in the present study were randomly selected from a group of high school students in Chongqing, China. 60 students from Grade 10 participated in this project. All the participants have studied English for more than six years and some of them have already learned English for nine years. All the speakers were informed of the purpose of the present study and those who were fully aware of and consented to the purpose finished the recording.

To generalize the research results, both genders were invited with their age ranging from 14 to 16. The proportion of male and female participants in this study is different (35 males and 20 females) but in order to ensure that the participation and the sample size are appropriate, the ratio of gender was not adjusted.

Three professional raters (2 males and 1 female) were invited to listen to the recording of stimuli, identify the pronunciation errors, and rate the intelligibility of each sample. These raters have different Chinese dialect backgrounds (Chongqing dialect,

Jiangxi dialect, and Cantonese). All the raters have learned English for more than 13 years. They are in-service or pre-service English teachers who have received abundant linguistic and pedagogical training so as to guarantee the accuracy of identifying pronunciation errors. All raters had received specialized training as well. On the premise of ensuring consistency and reliability of ratings, raters were allowed to give ratings.

3.3 Research procedures

Participants read a standardized English passage aloud when they indicated readiness; speech was recorded on a smartphone in a quiet room. Readers could adopt their preferred posture and pace, and the researcher did not interrupt or cue pronunciation. All recordings were denoised to minimize ambient interference; clips with inadequate acoustic quality were excluded. Five recordings were removed, yielding 55 usable samples for analysis.

Before formal rating, a pilot pre-rating checked whether an online automatic speech recognition (ASR) system—trained on diverse accents—could reliably transcribe the material. Five student recordings were randomly selected for ASR transcription, alongside one control sample produced by a rater. The control was transcribed without error, but student recordings showed 5–13 transcription errors per passage, indicating that the ASR tool struggled with these productions and motivating reliance on human judgments for intelligibility.

For the main evaluation, trained raters first completed calibration on partial samples to establish consistent criteria. After satisfactory consistency was reached, raters listened to each full recording and assigned an intelligibility score based on overall ease of understanding. They then relistened and annotated segmental pronunciation errors only (vowels and consonants), marking substitutions, deletions, insertions, and misarticulations. No suprasegmental features were coded in this study.

Scores and annotations were compiled in Microsoft Excel 2016 and analyzed in IBM SPSS 26. Descriptive statistics summarized segmental error patterns (e.g., error counts and rates by phoneme category). To examine the relationship between pronunciation accuracy and understanding, Pearson's correlation coefficients were computed between segmental error measures and intelligibility scores. These analyses were used to characterize the intelligibility of Grade-10 learners' oral English and to quantify the extent to which specific segmental errors predict lower intelligibility.

4. Results and discussion

After screening, 55 valid recordings and 165 valid rating data were finally confirmed. In the field of statistics, the reliability of a sample is often judged by calculating Cronbach's alpha and if the coefficient is greater than 0.8, it could be considered that the sample is fairly reliable. The consistency test of the scores given by raters was also conducted to guarantee the reliability of the study based on the intraclass correlation coefficient (ICC). The ICC is often used to measure the consistency of the same judge when scoring multiple samples based on fixed criteria. A value greater than 0.75 indicates good internal consistency in the scoring data of the group. Table 1 revealed detailed information about the reliability and consistency of the samples.

Table 1 Verification of Reliability and Internal Consistency

Analysis Type	Numerical Value	
Cronbach's Alpha	0.91	
ICC	Rater 1	0.85
	Rater 2	0.88
	Rater 3	0.87

According to the data, the Cronbach's alpha obtained from analyzing the scoring data of the raters in this study is 0.91, which is greater than 0.8, indicating a high level of reliability of the sample. The ICC of each rater's ratings are all greater than 0.75, indicating high internal consistency.

In sum, it can be considered that the samples collected in this study are reliable, with high consistency in the scoring data, and can be used for further research. These statistical data reveal that listeners showed similar level of understanding to the same oral output, which means their dialect background did not obscure the process of understanding. Cao Yating and Chen

Hua (2023) suggested that different language backgrounds would not strongly affect the understanding to oral output. Listeners gave similar comments on the intelligibility of the same material when their L2 proficiency is relatively advanced.

It should be mentioned that all "error rates" in this study represent the percentage of participants who produced pronunciation errors in a specific type of sound compared to the total number of participants. For example, "the error rate of the vowel /i:/ is 50.9%," which means that 50.9% of the total number of participants made the same errors in this vowel during the material reading process.

Based on the Daniel Jones (DJ) phonetic symbols, there are three kinds of single vowels: front vowels, back vowels, and central vowels. The average error rates of these vowels are as follows. The rest vowels are called “diphthongs” which can be divided into two kinds: closing diphthongs and centering diphthongs.

4.1 Vowels

Table 2 shows the overall situation of pronunciation errors produced by participants when pronouncing vowels. It can be found that in the pronunciation of vowels, the average error rate of single vowels is 16.5%, and that of diphthongs is 5.2%, with the former being more than three times higher than the latter. It can be considered that the participants encountered greater difficulties when pronouncing single vowels than when pronouncing diphthongs.

Table 2 Average Error Rate of Single Vowels and Diphthongs

Vowel	Average Error Rate
Single vowels	16.5%
Diphthongs	5.2%

Table 3 Average Error Rate of each Type of the Vowels

Type of Vowel	Average Error Rate
Front Vowels	29.68%
Back Vowels	9.46%
Central Vowels	9.7%
Closing Diphthongs	2.56%
Centering diphthongs	9.7%

Table 3 shows the average error rates of each group of vowels. Participants had more prominent problems when pronouncing front vowels, with an average error rate of nearly 30%. The average error rates for back vowels, central vowels, and centering diphthongs are relatively close, fluctuating around 9.5%; And it performs best when pronouncing closing diphthongs, with an average error rate of only 2.56%.

These errors are basically consistent with those discovered by previous studies. Sun Fali (1979) found that learners with Sichuan dialect background would pronounce the word “sit” as “seat”, using /i:/ to replace /ɪ/. Luo Xuefei (2015) studied the impact of Sichuan dialect on the pronunciation errors made by students who have relative dialect backgrounds. It was mentioned in her study that some of the participants involved in the study had difficulties in distinguishing front vowels /i:/ and /ɪ/ for the misunderstanding of how to pronounce these vowels. They used the sound “i” in Chinese Pinyin to replace /i:/ and /ɪ/ in English and only distinguished these two different vowels with different lengths of the sound, ignoring the differences between the shapes of mouth when pronouncing.

It should be noticed that the participants performed well on most of the back vowels, with only a few participants encountered problems. But when being faced with pronouncing the back vowel /ɒ/, nearly a quarter of the participants had difficulties. Specific examples would be provided below.

The original text: *We got lost a few times…*

Correct phonetic symbols: /wi: gɒt l**ɒ**st ə fju: taɪmz/

Mispronounced phonetic symbols: /wi: gɒt l**ɔ:**st ə fjuː taɪmz/

In this example, participants had the problem of confusing /ɔ:/ and /ɒ/, which made them mispronounce the word "lost" as "law-st". Sun Fali (1979) stated that learners with Sichuan dialect backgrounds tended to confuse these two vowels, which is consistent with the findings in the present study. Besides, it is worth noticing that the differences between /ɔ:/ and /ɒ/ do not only involve the length. More than 70% of the participants held the view that the /ɒ/ sound is a shorter version of the /ɔ:/ sound, which might cause confusion between them.

The overall situation of pronunciation error of the closing diphthongs is relatively optimistic, with an average error rate of only 2.9%. The main problem with the pronunciation errors of diphthongs occurs in the section on centering diphthongs. No pronunciation error was found in /ɪə/ and /ʊə/, but the error rate of /eə/ is close to one-third.

Most of the participants who failed to pronounce the diphthong /eə/ correctly mispronounced the word "bear" as "beer", using /ɪə/ to replace /eə/. It is worth noticing that this phenomenon cannot be found in the previous studies and this might be a new finding or just for the deficiency of acquiring the correct pronunciation of this specific word. Sun Yun (2016) suggested that diphthongs need to be pronounced smoothly and fluently. One breath is crucial in the process of producing diphthongs. If there is an interruption between the transitions during the pronunciation process, it is easy to produce diphthongs as two single sounds. This situation mainly exists in northern dialects (including Sichuan and Chongqing dialects). By asking the raters, it could be confirmed that most of the participants who mispronounce diphthongs possess problems of pronouncing a diphthong as two separate vowels.

4.2 Consonants

Table 4 shows the average error rate of each group of consonants. It could be found that the overall performance of the participants in this section was better than their performance of pronouncing vowels.

In addition, the average error rate of the lateral sound is 40%, which is fairly prominent among all consonants. It may be due to the data density caused by its relatively small quantity.

Table 4 Average Error Rate of Each Type of the Consonant

Type of the Consonant	Average Error Rate
Plosives	2.72%
Fricatives	5.45%
Nasals	10.3%
Lateral	40%
Affricates	2.42%
Semi-vowels	1.82%

By comparing and confirming the materials, it was found that participants with this kind of pronunciation errors tended to add the schwa sound /ə/ at the end of the /t/ sound. In these examples, participants mispronounced the /t/ sound by adding a schwa sound after it, replacing "continent" and "want" with /ˈkɒntɪnəntə/ and /weɪtə/. As Sun Fali (1979) proposed in his research, it is rather common for beginners to add the schwa sound /ə/ after a plosive sound. The phenomenon of inappropriate epenthesis does harm to English pronunciation, not only causing inaccurate pronunciation of consonants but also seriously affecting the intonation. However, this occurrence does not exist in the speech produced by English learners from Sichuan or Chongqing only.

In the pronunciation of nasals, only the /n/ sound was found to be mispronounced in this study. Over 30% of the participants pronounce it incorrectly, mainly due to the tendency of some participants to mispronounce this sound, which appears in the middle or end of a word, as other nasal sounds. Examples are given below.

The original text: …*we managed to cross the Rocky Mountains*

Correct phonetic symbols: /wi: ˈmænɪʤd tə krɒs ðə ˈrɒki ˈmaʊntə**n**z/

Mispronounced phonetic symbols: /wi: ˈmænɪʤd tə krɒs ðə ˈrɒki ˈmaʊntə**m**z/

In the example above, participants mispronounced the word "mountains as /ˈmaʊntəms/, replacing the /n/ sound with /m/. It should be noticed that the /n/ sound which is placed at or close to the beginning of a word was not mispronounced in all the recordings.

Considering the fact that there is only one lateral sound in English, the table would be omitted and replaced with a description. Although there is only one consonant in this group, the lateral sound /l/ might appear at the beginning, middle, or end of a word, which causes a slight difference, i. e. the difference between /l/ and /ɬ/. Through several times of checking, it was confirmed that all the lateral sounds which appear at the beginning were correctly pronounced while the rest lateral sounds (/ɬ/) were mispronounced by 40% of the participants. Here are some examples.

The original text: *It's like another world*.

Correct phonetic symbols: /ɪts laɪk əˈnʌðə wɜːɬd/

Mispronounced phonetic symbols: /ɪts laɪk əˈnʌðə wɜː**əʊ**d/

or

The original text: *but all I saw was a small group of deer*.

Correct phonetic symbols: /bət ɔːɬ aɪ sɔː wəz ə smɔːɬ gruːp əv dɪə/

Mispronounced phonetic symbols: /bət ɔː aɪ sɔː wəz ə smɔː gruːp əv dɪə/

As underlined in the examples, participants who mispronounced this sound tended to use the closing diphthongs /əʊ/ to replace it, prono8uncing "world" as /wɜːəʊd/. In some cases, participants would even omit the lateral sound in a word, such as pronouncing "all" in a manner of "or". It was surprising to find that participants did not make errors in distinguishing the nasal sound /n/ and the lateral sound /l/ because it is relatively hard for speakers who have Chongqing or Sichuan dialect background to tell the difference between /n/ and /l/ even when they are speaking Chinese due to the deficiency of the /l/ sound in these dialects. By asking their teachers, it was revealed that most of the participants possess relatively good proficiency of Mandarin, which means they could tell the difference between /n/ and /l/ when speaking Mandarin although it is hard when speaking Chongqing dialect.

Based on the analysis above, it is revealed that single vowels have a higher error rate compared to diphthongs, with front vowels being particularly challenging. Specific errors include mispronunciation of /iː/ as /eɪ/, difficulty distinguishing between /ɪ/ and /iː/, and frequent misarticulation of the schwa sound. Among consonants, lateral sound and nasals like /n/ were often mispronounced, indicating a struggle with sounds absent in the speakers' native language.

4.3 Correlations Between Speech Intelligibility and Segmental Features

The present study converts the scores of the intelligibility of the recording materials rated into average values and conducts Pearson correlation analysis with various data of pronunciation errors, using IBM SPSS to obtain the data required for the present study.

Table 5 shows the data calculated by IBM SPSS. It can be found that there is a negative correlation between intelligibility and the other two variables respectively.

Table 5 Correlations Between Speech Intelligibility and Segmental Errors

Errors	Pearson Co.	Sig. (2-tailed)
VE	-.485**	.000
CE	-.395**	.003

**. Correlation is significant at the 0.01 level (2-tailed). N=55.

(VE = Vowel Errors, CE = Consonant Errors)

Table 6 Correlations Between Speech Intelligibility and Errors of Different Vowels

Errors	Pearson Co.	Sig. (2-tailed)
EOSV	-.369**	.006
EOD	-.341*	.011

**. Correlation is significant at the 0.01 level (2-tailed). N=55.

*. Correlation is significant at the 0.05 level (2-tailed).

(EOSV = Errors of Single Vowels, EOD = Errors of Diphthongs)

Table 6 shows the result of the correlation between intelligibility and errors of different vowels. The correlation coefficients are -0.369 ($p<0.01$) and -0.341 ($p<0.05$). It is revealed that there is a negative correlation between the errors of both kinds of vowels and intelligibility, which means that both single vowels and diphthong errors could have a certain negative impact on intelligibility, and the degree of their impact is relatively similar. The analysis would be conducted based on the classification mentioned earlier.

Table 7 shows the result of the correlation between intelligibility and errors of different single vowels (SVs). The correlation coefficient between intelligibility and error of front vowels is -0.334 (p=0.013). It can be considered that there is a significant negative correlation between intelligibility and errors of front vowels. By asking the raters, it was found that the pronunciation errors obstructed intelligibility most when they occurred in words or phrases which act as subjects or predicates in a clause. Examples will be shown after the table.

It is revealed that the correlation coefficient between intelligibility and errors of central vowels is -0.220 (p=0.106), which represents that there is a significant correlation between these two variables.

It can also be found that the correlation coefficient between intelligibility and errors of back vowels is -0.090 (p=0.512), which represents that there is no significant correlation between these two variables.

Table 7 Correlations Between Speech Intelligibility and E. of Different Types of SVs.

Errors	Pearson Co.	Sig. (2-tailed)
EOFV	-.334*	.013
EOCV	-.220	.106
EOBV	-.090	.512

*. Correlation is significant at the 0.05 level (2-tailed). N=55.

(EOFV = Errors of Front Vowels, EOCV = Errors of Central Vowels, EOBV = Errors of Back Vowels)

Table 7 shows the result of the correlation between intelligibility and errors of different single vowels (SVs). The correlation coefficient between intelligibility and error of front vowels is -0.334 (p=0.013). It can be considered that there is a significant negative correlation between intelligibility and errors of front vowels. By asking the raters, it was found that the pronunciation errors obstructed intelligibility most when they occurred in words or phrases which act as subjects or predicates in a clause. An example will be shown below.

The original text: *where we picked up our vehicle for the trip.*

What the raters heard: where we peeked up our vehicle for the trip

Table 8 Correlations Between Intelligibility and E. of Different Types of Diphthongs

Errors	Pearson Co.	Sig. (2-tailed)
EOCSD	-.331*	.013
EOCTD	-.210	.125

*. Correlation is significant at the 0.05 level (2-tailed). N=55.

(EOCSD = Errors of Closing Diphthongs, EOCTD = Errors of Centering Diphthongs)

Table 8 shows the result of the correlation between intelligibility and errors of different types of diphthongs. According to the table, it can be found that the correlation coefficient between intelligibility and errors of closing diphthongs is -0.331 (p=0.013), which represents that there is a significant correlation between these two variables.

It is revealed that the correlation coefficient between intelligibility and errors of centering diphthongs is -0.210 (p=0.125), which represents that there is no significant correlation between these two variables.

Table 9 Correlations Between Speech Intelligibility and E. of Different Types of Consonants

Errors	Pearson Co.	Sig. (2-tailed)
EOP	-.072	.599
EOF	-.338*	.012
EON	-.110	.425
EOL	-.155	.259
EOA	-.307*	.023
EOSMV	-.156	.255

*. Correlation is significant at the 0.05 level (2-tailed). N=55.

(EOP = Errors of Plosives, EOF = Errors of Fricatives, EON = Errors of Nasals, EOL = Errors of the Lateral, EOA = Errors of Affricates, EOSMV = Errors of Semi-vowels)

Table 9 shows the result of the correlation between intelligibility and errors of different types of consonants. It shows that the correlation coefficient between intelligibility and errors of plosives is -0.072 (p=0.599), which represents that there is no significant correlation between these two variables.

It is also revealed that the correlation coefficient between intelligibility and errors of fricatives is -0.338 (p=0.012), which represents that there is a significant correlation between these two variables. By asking the raters, it was found that the pronunciation errors of the /θ/ sound obstructed intelligibility the most.

The correlation coefficient between intelligibility and errors of nasals is -0.110 (p=0.425), which represents that there is no significant correlation between these two variables. Besides, the correlation coefficient between intelligibility and errors of the lateral is -0.155 (p=0.259), which represents that there is no significant correlation between these two variables. Although 40% of the participants produced pronunciation errors of this sound, it did not cause a factual barrier which interfered with understanding.

It can also be found that the correlation coefficient between intelligibility and errors of affricates is -0.307 (p=0.023), which represents that there is a significant correlation between these two variables. The correlation coefficient between intelligibility and errors of semi-vowels is -0.156 (p=0.255), which represents that there is no significant correlation between these two variables.

Based on the above data and analysis, it can be summarized that there are correlations between the intelligibility of participants' speeches and multiple variables, all of which are negatively correlated. The most relevant among them is the errors of vowels, which may be due to the fact that during the rating process, raters generally indicated that although the total number of consonants in the reading material is greater than vowels, vowels play a relatively greater role in expressing meaning in syllables, and the occurrence of errors of vowels further affects listener's understanding. In the errors of vowels, intelligibility is negatively related to front vowels and closing diphthongs. Consonants also have a significant negative correlation with intelligibility, which is mainly caused by the negative correlation between intelligibility and two variables, i.e. fricatives and affricates.

5. Conclusion

This study catalogued the segmental pronunciation patterns of Chongqing high-school learners and examined their relationship with intelligibility. Errors concentrated in vowels, with frequent merging and misarticulation of front vowels; confusions involving /iː/ (e.g., with /eɪ/ or /ɪ/) were especially common and often led raters to misidentify words (e.g., week → wake). Consonant difficulties centered on interdental fricatives /θ, ð/ and on segments absent or restricted in the local dialect, notably word-final /l/, which was sometimes substituted or omitted. Statistical analyses showed a significant negative correlation between segmental error rates and intelligibility scores. Within vowels, front vowels and closing diphthongs exhibited the strongest associations with reduced intelligibility; within consonants, fricatives and affricates showed comparable effects. These results indicate that not all deviations are equally consequential: a focused instructional emphasis on high-impact contrasts—such as /iː/–/ɪ/–/eɪ/, interdental fricatives, and coda /l/—is likely to yield measurable gains in understanding. Pedagogically, we recommend targeted perception–production cycles, minimal-pair training anchored in local error confusions, articulatory cues tailored to interdental and coda productions, and progress monitoring that tracks intelligibility alongside accuracy. Within its segmental scope, the study provides actionable priorities for pronunciation teaching and assessment in dialect-influenced secondary settings.

6. Implications

Instruction should be dialect-aware and intelligibility-first, prioritizing high-impact segmentals rather than native-like accent in general. For Chongqing learners, this means targeting the contrasts most linked to misunderstandings—/iː/–/ɪ/–/eɪ/, interdental fricatives /θ, ð/, and word-final /l/. Teachers can (1) run quick diagnostics to surface each class's top confusion pairs; (2) design short perception-production cycles (ABX listening, minimal-pair drills, immediate feedback); (3) give simple articulatory cues (e.g., tongue placement for /θ, ð/, sustained coda contact for /l/); (4) use spaced retrieval and micro-quizzes to stabilize new categories; and (5) assess with intelligibility tasks (keyword recognition, dictation of minimal pairs, message-transfer) rather than accent scores. Not every error needs attention—focus time on those that actually change word identity.

For learners, confidence should come from being understood. Keep a personal confusion list (e.g., *week/wake*, *thin/sin*), record short readings, and compare against model audio; use ASR only as feedback, not as ground truth (its errors with student speech are known). Increase exposure to varied L2 English, practice targeted contrasts in short daily bursts (3–5 minutes), and track gains with simple metrics (word-ID accuracy, rater comprehension). This segmental, data-guided approach yields practical gains with minimal extra class time.

7. Limitations

This study provides initial evidence on how segmental errors among Chongqing high-school learners relate to intelligibility, but several limits apply. Results are drawn from a single dialect group with a modest sample (55 speakers) and three raters, constraining generalizability and rater-reliability estimates. The stimulus was textbook reading; its phoneme coverage was incomplete (e.g., /ʒ/ absent) and may not reflect spontaneous speech. By design, suprasegmental features were not analyzed, so conclusions pertain only to segmentals. Future work should expand participant pools across dialects, increase the number of raters, elicit both read and spontaneous speech with materials sampling the full segmental inventory, and apply richer statistical modeling (e.g., mixed-effects, item-level analyses). These steps would improve representativeness and yield finer-grained, classroom-ready priorities for pronunciation instruction.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Alzainidi, M. H., & Latif, M. M. M. A. (2019). Diagnosing Saudi students' English consonant pronunciation errors. *Arab World English Journal, 10*(4), 180–193.
[2] Cao, Yating, & Chen, Hua. (2023). Effects of listeners' L1 background and English proficiency on the intelligibility of Chinese students' English accent. *Foreign Language World*, (02), 88–96.
[3] Crowther, D., Trofimovich, P., & Isaacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility: Nonnative listeners' perspectives. *Journal of Second Language Pronunciation, 2*(2), 160–182.
[4] Ellis, R. (1999). *The study of second language acquisition*. Shanghai: Shanghai Foreign Language Education Press.
[5] Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36*(4), 664–679.
[6] Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an International Language. *Applied Linguistics, 23*, 83–103.
[7] Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly, 9*(3), 249–269.
[8] Kennedy, G. (2003). *Structure and meaning in English: A guide for teachers*. London: Pearson Education.
[9] Li, Jingna, & Wang, Yao. (2015). A comparative study of accent perception and speech intelligibility of Chinese English learners by English and Chinese native speakers. *Foreign Language Learning Theory and Practice*, (01), 50–54, 96.
[10] Luo, Xuefei. (2015). The influence of Sichuan dialect transfer on English pronunciation and intonation teaching and countermeasures. *Journal of Higher Education*, (18), 110–111.
[11] Munro, M. (2008). Foreign accent and speech intelligibility. In J. Hansen Edwards & M. Zampini (Ed.), *Phonology and Second Language Acquisition* (pp. 193-218). John Benjamins Publishing Company.
[12] Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73–97.
[13] Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*(1), 111–131.
[14] Perez-Ramon, R., Lecumberri, M. L. G., & Cooke, M. (2022). Foreign accent strength and intelligibility at the segmental level. *Speech Communication, 137*, 70–76.
[15] Sun, Dongyang, & Chen, Chunhua. (2022). Effects of accent familiarity on second language listening comprehension. *Modern Foreign Languages, 45*(05), 684–696.
[16] Sun, Fali. (1979). Common errors in English pronunciation learning by students in Sichuan. *Journal of Southwest University(Social Sciences Edition)*, (01), 84–89.
[17] Sun, Yun. (2016). On the influence of Chinese dialects on English pronunciation acquisition. *Language Planning,* (30), 9–10.
[18] Wu, Qianguang. (1979). On the application of "error analysis" in English teaching. *Modern Foreign Languages*, (04), 40–48.
[19] Xie, X., & Fowler, C. A. (2013). Listening with a foreign accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English. *Journal of Phonetics, 41*(5), 369–378.
[20] Yin, Song. (2015). Shadowing and improvement of prosodic errors in Japanese: Evidence from auditory ratings and acoustic analysis. *Foreign Language Learning Theory and Practice*, (04), 70–74, 95, 98.
[21] Yin, Zhilei. (2001). Common English pronunciation errors of students in southeastern Gansu. *Journal of Tianshui Normal University,* (03), 75–76, 96.
[22] Zoghbor, W. S. (2018). Teaching English pronunciation to multi-dialect first language learners: The revival of the Lingua Franca Core (LFC). *System, 78*, 1–14.