International Education Forum, 2026, Volume 4, Issue 6
http://ojs.bbwpublisher.com/index.php/IEF
Online ISSN: 2981-8605
Print ISSN: 3083-4902

A Study on the Effect of AI-assisted Vocational College English Pronunciation Learning on Students' Pronunciation and Oral Proficiency

Chunyan Luo*

Lanzhou college of foreign studies, Lanzhou 730000, Gansu, China

Author to whom correspondence should be addressed.

Abstract: With the digitalization of vocational education, intelligent pronunciation training has been introduced as a new form in the reform of English teaching at vocational colleges. Traditional phonetics teaching has issues, such as delayed feedback, a small training volume, subjective evaluation, and a lack of personalized guidance. To investigate the actual effects of AI on vocational students' second-language pronunciation acquisition, this study selects 42 students majoring in primary school English Education as subjects and employs a one-group pretest-posttest design for a one-semester AI teaching intervention. Praat, professional acoustic software, was used to conduct quantitative comparisons of the pre- and post-intervention speech data at various levels. Through the research, AI-assisted teaching has helped students improve the acoustic features of their pronunciation, correct typical errors, and enhance monotonous intonation and fragmented speech; meanwhile, the accuracy and fluency of their speaking have been significantly improved. Based on objective acoustic data, this study has verified the practical value of AI teaching and reduced the range of subjective evaluation in traditional phonetics classes. It has established a closed-loop teaching model of daily AI training and regular Praat acoustic tests, and provided replicable empirical schemes for the digital reform, quantitative pronunciation evaluation, and personalized phonetic teaching in vocational colleges.

Keywords: Artificial intelligence; Vocational college English; Pronunciation learning; Oral proficiency; Praat; Acoustic parameters

Online publication: July 17, 2026

1. Introduction

1.1. Research background

National Vocational Education Digitalization Policies advocate for intelligent and personalized English classes. Pronunciation is a basic professional quality for graduates of Primary School English Education, but due to large class sizes, delayed teacher feedback, and subjective auditory assessment in vocational phonics

classes, many students' phonetic errors have not been corrected, and their oral expression skills have not been effectively improved.

AI speech platforms provide unlimited practice and instant error correction to compensate for deficiencies in the classroom. Most of the previous studies have used self-reported questionnaires and have not been validated acoustically; therefore, targeted research on vocational students is still needed. Therefore, this study will combine AI training with Praat acoustic analysis to conduct quantitative research on the effects of phonetic teaching.

1.2. Research significance

1.2.1. Theoretical foundation

According to the theory of second-language phonetic acquisition and intelligent education, this paper will investigate the effect of AI on learners' pronunciation prosody through multiple acoustic parameters. Provide objective empirical data for research in intelligent phonics teaching and offer new analytical perspectives on vocational digital pronunciation teaching.

1.2.2. Practical value

This paper establishes an operable teaching mode of AI daily training plus periodic Praat acoustic evaluation to address the problems of homogenization and single evaluation in traditional phonics classes. The research plans can be generalized to vocational colleges to help teachers conduct data-driven, precise phonetic tutoring and enhance students' professional oral literacy.

1.3. Research questions

This paper investigates the application effect of AI-assisted pronunciation learning for vocational college majors of English Education, and through pretest-posttest acoustic data analysis, questionnaire surveys, and semi-structured interviews, explores the following three research questions.

First, can long-term AI-assisted autonomous pronunciation learning improve the acoustic features of students' English pronunciation and effectively correct their typical pronunciation errors, such as vowel deviation, ambiguous consonants, and misplaced word stress? What particular changes will be observed in the main acoustic indicators of vowel formants, syllable duration, and consonant waveforms after intervention?

Secondly, what positive effects will extended training in AI situational oral skills have on students' oral fluency, speech rhythm, and intonation prosody? Can the problems of invalid frequent pauses, unstable speech rate, and flat single intonation be significantly improved in terms of pitch fluctuation and pause distribution data?

Third, based on Praat acoustic quantitative data and students' subjective learning feedback, how to optimize the intelligent pronunciation teaching mode and personalized training system for vocational colleges? What supporting measures should schools, teachers, and students take to compensate for the inherent deficiencies of AI speech platforms in teaching oral prosody?

1.4. Research innovations

First, this study will introduce a new type of research method. Unlike the traditional teaching research mode of subjectivity, which often relies on classroom observation, questionnaires, and interviews, Praat

professional acoustic software is used here for quantitative analysis of multiple measurable acoustic parameters to convert abstract pronunciation quality into visual data and waveforms. The research conclusions are supported by objective acoustic data and thus are more reliable; they have been freed from the bias of a single subjective judgment.

Second, it extends the scope of research and the subjects of study. Most existing research on intelligent phonetic teaching uses general university students or primary and secondary school students as research subjects, and few specific quantitative empirical studies focus on vocational college Primary School English Education majors. Pre-service primary English teachers with weak phonetic foundations were selected as the research subjects, intelligent teaching practice was integrated with acoustic research on second-language pronunciation, and a research gap in quantitative empirical studies on intelligent pronunciation teaching for vocational college students was addressed.

Thirdly, this study establishes a full-featured teaching system of daily autonomous AI training and regular Praat acoustic evaluation, supplemented by teacher-led tutoring. A closed-loop system is used to personalize training based on phonetic diagnosis data, display the learning process through acoustic waveforms and data-driven teaching effect assessment, and is highly suitable for the cognitive rules and phonetic acquisition characteristics of vocational college students with weak English foundations. At the same time, this paper also presents the actual working plan of the hybrid teaching mode, which has good operability and popularization value for vocational education teaching.

2. Literature review

2.1. Research on AI-assisted English pronunciation teaching

With the rapid development of intelligent speech recognition technology in recent years, many Chinese scholars have begun to explore how to apply artificial intelligence in teaching English pronunciation, and they have generally agreed that AI can help address the shortcomings of traditional phonetic classes [1]. Most scholars believe that traditional phonetic teaching is limited by insufficient training time, delayed teacher feedback, and a lack of relevance; AI speech technology can expand the channels for students' after-class autonomous learning, reduce their anxiety in oral expression due to the fear of making mistakes in front of others, and improve the efficiency of daily pronunciation training significantly [2].

Hu and Yan (2026) pointed out that vocational students generally have low self-confidence in speaking English orally, and the group activity of classroom reading makes students with poor pronunciation afraid to speak out loud [1]. An intelligent speech platform can offer students a low-pressure environment to practice speaking repeatedly without fear of embarrassment, and at the same time, the system will promptly point out any pronunciation errors for correction [8]. Ye (2024) carried out an empirical study on an intelligent speech platform to assist oral English teaching, and through questionnaires and classroom observations, found that AI training can enhance students' learning motivation and reduce their fear of speaking [2].

Based on the previous research results at home, most of the existing studies have focused on teaching modes and practical strategies from the perspective of teaching practice, and their teaching effects have been verified through classroom observation, questionnaires, and semi-structured interviews [10]. The above research methods are generally qualitative and subjective analyses. Although the positive effects of artificial intelligence in phonetics education have been widely reported by scholars, few studies have used objective quantitative acoustic data to conduct in-depth research on students' pronunciation of monosyllabic words,

sentence intonation, and speech rhythm [3]. In particular, there are very few acoustic quantitative studies on vocational college students [6]. Most studies have only mentioned that students' pronunciation has improved generally from a subjective perspective and have failed to specify which phonetic indicators have changed significantly or what the range of particular improvements is; therefore, the persuasiveness of the research conclusions is relatively weak [4].

2.2. Applications of Praat in pronunciation teaching

Praat is an excellent commercial program for analyzing sound waves developed by Paul Boersma and David Weenink, and now it is one of the main tools for research in experimental phonetics worldwide [5]. Several main acoustic parameters are automatically extracted, such as fundamental frequency (pitch), syllable duration, vowel formants, speech rate, pause interval, and pronunciation features, and hidden phonetic errors are displayed in visual sound spectrograms, waveform diagrams, and data tables. Praat can be employed in the teaching of English phonetics to collect quantitative data on pronunciation errors that are difficult to perceive with the ear, such as small fluctuations in the opening degree of vowels, an insufficient range of pitch rise and fall, or an overly short duration of final consonants; thus, specific targets for pronunciation correction and sentence intonation training can be set.

Bouchhioua (2024) has systematically studied the combination of Praat acoustic analysis technology and AI pronunciation teaching, indicating that human auditory perception is inherently subjective, different teachers have different standards for judging the same pronunciation sample, and thus, standardized quantitative evaluation of pronunciation can only be achieved by using Praat's acoustic parameters [3]. Chen and Zhang (2022) conducted an empirical study on the acquisition of English intonation based on acoustic parameters and used pitch-fluctuation data to identify the differences in intonation between Chinese ESL learners and native English speakers; they found that acoustic visualization can help learners intuitively recognize their own intonation shortcomings [4].

Currently, Praat is mainly used in linguistic theoretical research and phonetic teaching at undergraduate universities, and its application in daily English teaching in vocational colleges is rare [6]. Most vocational college English teachers have not participated in regular training on experimental phonetics, are unfamiliar with the application of Praat software, and do not know how to interpret acoustic parameters. In addition, there are almost no systematic long-term empirical studies that combine extended AI speech training with periodic Praat acoustic assessments for vocational students; this is the main research entry point and innovation of the present paper [9]. Praat will be used throughout the entire teaching intervention process of this study, and a full set of operable acoustic data collection and analysis procedures for daily teaching at a vocational college will be established.

2.3. Research Commentary

In short, the advantages of AI technology for pronunciation teaching and the scientific objectivity of Praat as an acoustic quantitative analysis tool have been fully verified in the existing domestic and foreign research [5, 10]. However, the present studies on intelligent English phonetic teaching have not yet reached a certain level of development.

First, the research objects are relatively homogeneous. Most of the relevant studies have taken comprehensive university undergraduates or primary and secondary school students as the research subjects,

and special targeted research on vocational college students with weak English foundations and clear professional training goals is lacking [1, 7]. The types of phonetic errors, learning motivation, and degrees of oral anxiety for vocational students are clearly different from those of undergraduates in comprehensive universities, and the research conclusions obtained so far cannot be directly applied to phonetic teaching in vocational colleges [2].

Second, the main kind of research method is qualitative analysis without acoustic quantitative data support. Most research uses questionnaires, interviews, and teacher observations to assess the effects of teaching, but it lacks quantifiable objective data on how students' pronunciation abilities change subtly after the teaching intervention, so the depth and persuasiveness of the research results are relatively low [3–4].

Thirdly, few studies have constructed and verified a hybrid teaching model of intelligent online training and periodic acoustic quantitative assessment. Most research only studies the independent use of AI platforms and fails to explore the advantages of professional acoustic software for post-stage effect testing and personalized teaching diagnosis; therefore, a complete teaching and evaluation closed-loop system has not been established [9].

This paper aims to correct the deficiencies of the above study. Vocational college primary school English education majors were selected as the research subjects, and a new teaching model combining AI teaching intervention with Praat acoustic quantitative analysis was established; longitudinal pretest-posttest empirical research was conducted to extend the existing research system of intelligent English phonetic teaching [7].

3. Research design

3.1. Research subjects

The research subjects are 42 students from an intact class of the Primary School English Education major at Lanzhou College of Foreign Studies. The research selects intact classes for teaching intervention to avoid the disruption of the normal classroom teaching sequence caused by sample group division, thus meeting the ethical standards of school teaching research. All participants have weak foundations in English pronunciation due to their middle school studies, and most commonly exhibit non-standard monosyllabic pronunciation, monotonous flat intonation, and discontinuous, fragmented oral expression. Prior to the start of this study, none of the students had been formally taught English phonetics by professionals; they had not learned about vowel and consonant correction, sentence rhythm and intonation, and the environment for learning English daily, such as learning materials and time, had also been irregular.

This study will use a one-group pretest-posttest design. All 42 students are designated as the only experimental group, and their standardized speech acoustic data before and after one semester of AI teaching intervention are vertically compared to verify the actual teaching effect of AI-assisted pronunciation learning. A one-group pretest-posttest design can rule out the problem of individual differences among groups of students and thus accurately show the changes in phonetic ability of each student under the same teaching intervention.

3.2. Research tools

A three-dimensional integrated tool system was established in this study to collect research evidence through an AI intelligent speech teaching platform, Praat acoustic analysis software, and supplementary subjective survey tools (questionnaires and interview outlines).

3.2.1. AI speech teaching platform

It will be the main medium for all teaching interventions this semester. The several functions of the platform include word phonetic reading, paragraph text recitation, situational oral dialogue simulation, whole-process learning data recording, and personalized differentiated learning task distribution. Throughout the course of the study, students have taken regular in-class training and daily self-study at the end of class on this platform. The platform automatically identifies incorrect phonemes, misplaced stress, abnormal pauses, and flat intonation in students' pronunciation recordings, displays real-time correction prompts, and pushes targeted repeated training exercises based on the classification of each student's errors to provide personalized off-campus supplementary training.

3.2.2. Praat speech analysis software

Praat 6.4, a relatively new and stable professional acoustic analysis program, will be used to process all the sound data uniformly and compare them. In addition to the main evaluation attributes of English pronunciation and oral ability, six core acoustic indicators were selected as the research measurement standards: syllable duration of long and short vowels, average pitch and pitch fluctuation range of sentences, speech rate per minute, frequency of invalid pauses in continuous speech, vowel formant F1/F2 values that reflect vowel opening and tongue position, and overall intonation characteristic curves of declarative and interrogative sentences. Uniform and standardized speech samples are collected from all students before and after the intervention in the same recording environment, and then Praat is used to calculate the mean values of relevant acoustic parameters for each student to quantify the overall changes in their pronunciation accuracy and oral expression ability.

3.2. Three auxiliary survey tools

Self-designed student learning experience questionnaires and semi-structured teacher-student interview outlines are employed as auxiliary research instruments. The main contents of the questionnaire are students' changes in phonetic learning anxiety, learning initiative, recognition of the effect of AI training, and self-assessment of oral ability after one semester. The interview outline selects 8 representative students with various pronunciation levels and 2 English teachers involved in the teaching intervention, collects their subjective feelings, practical difficulties, and optimization suggestions for the AI + Praat teaching mode, and thus forms a research evidence chain together with objective acoustic data.

3.3. Research procedures

The entire research cycle is one full semester (18 weeks), and the four standardized research procedures must be strictly followed: pretest and baseline data collection, long-term AI teaching intervention, posttest speech re-measurement, and integrated data analysis and research conclusion summarization.

In the pretest and baseline data collection stage (Week 1 of the semester), the research group establishes unified speech sample collection standards, sets up a quiet multimedia language laboratory as the fixed recording environment, and uses the same model recording equipment to record standardized speech samples of all 42 students. The contents of the recordings are single-vowel and consonant readings, word-stress readings, short-paragraph recitations, and situational dialogue expressions that cover all phonetic dimensions in the study. All collected speech samples are imported into the Praat software in batches to extract the six main acoustic parameters, and the average values of these parameters for each student are calculated to build

a pretest baseline acoustic database. At the same time, the research group will organize all English teachers participating in the intervention for systematic training on the operation of the AI platform and Praat acoustic analysis, sort out the pretest error data of the whole class, and formulate preliminary personalized training plans for students with different phonetic defect types.

In the teaching intervention period (Weeks 2-16 of the semester), the AI speech platform will be used for daily in-class teaching and self-study after class. Teachers import the pretest Praat acoustic diagnosis data into the platform system, and based on the students' individual pronunciation weaknesses, the platform distributes differentiated learning tasks: students with serious vowel deviation focus on monosyllabic vowel repeated training; students with ambiguous consonants increase terminal consonant, and voiced/voiceless consonant contrast exercises; and students with flat intonation receive targeted sentence rising and falling tone simulation training. Teachers review the students' daily training data on the platform weekly, conduct centralized error explanations and one-on-one targeted tutoring for common phonetic problems shown in acoustic data, and form a hybrid teaching model of online AI autonomous training and offline teacher-led precise tutoring.

After the intervention, the work of collecting posttest speech will be carried out in Week 17 in the same recording environment, with the same equipment and text standards as the pretest. All the new posttest speech samples are imported into Praat for acoustic parameter extraction, and a complete posttest acoustic database is constructed to form a comparative data system with the pretest baseline data. In Week 18, the research group distributes student questionnaires, conducts semi-structured interviews with some teachers and students, collates all the subjective survey data, and finally carries out cross-comparative analysis of pretest-posttest acoustic data and subjective questionnaire/interview data to systematically summarize the research results.

3.4. Research methods

The four complementary research methods are employed in this study to ensure the all-roundness and reliability of the research:

First, research by scholars. The author organizes the domestic and foreign literature on artificial intelligence in foreign language teaching, second-language pronunciation acquisition, Praat acoustic analysis, and vocational college English teaching; reviews the current research progress and limitations; identifies the theoretical basis for all the research, and determines the research breakthrough point of this paper.

Second, a one-group pretest-posttest experimental design. Using a complete class of vocational college students as the sole experimental group, longitudinal comparisons of their speech acoustic data before and after one semester of AI teaching intervention were conducted to observe changes in pronunciation accuracy, oral fluency, and intonation prosody among students, and to verify the actual teaching effect of AI-assisted phonetic learning.

Third, Quantitative Analysis Method for Acoustic Data. Praat software is employed to extract quantifiable acoustic indicators from standardized speech data, quantify minute differences in the students' pronunciation features, transform subjective phonetic judgments into objective numerical comparisons, and conduct all-around data analysis of how well the students have optimized their pronunciation acoustic traits.

Fourth, Questionnaires and semi-structured interviews. Collect students' subjective learning experiences, changes in their learning attitudes, and teachers' practical teaching feedback through questionnaires and

targeted interviews, supplement the objective acoustic data analysis results, and put forward targeted optimization strategies based on the practical difficulties reflected in the subjective feedback.

4. Result

4.1. Comparison of pronunciation acoustic features

After one full semester of systematic AI training and daily personalized error correction, the overall acoustic features of all 42 students' English pronunciation have been significantly improved, and the quantitative changes in core Praat acoustic parameters are clear compared with the pre-test baseline data.

In terms of the pronunciation of monosyllabic vowels, the syllable duration for long and short vowels in students' word and sentence reading is now more uniform and standardized, and the severe problem of inconsistent vowel length before intervention has been effectively resolved. The F1 and F2 formant values, which indicate the position and opening of the vowel tongue, are relatively close to the standard acoustic parameter range of native English speakers; therefore, it can be concluded that students have achieved relatively accurate vowel articulation positions, and the deviation of front and back vowels and high and low vowels has been significantly reduced. Before the intervention, more than 70% of the students had obvious F1/F2 offset errors in /iː/, /ɪ/, /æ/, and /ɑː/; after training, the proportion of students with serious vowel formant deviation dropped to less than 15%.

Typical pronunciation errors of consonants have been corrected, such as unclear consonant sounds, missing final consonants, and confusion between voiced and voiceless consonants. Consonant waveform diagrams obtained by Praat analysis after intervention are clearer and more complete; the vibration waveform of voiced consonants is distinct; and the air-flow mute interval of voiceless consonants meets the norms of phonetics. Before the intervention, most students omitted the pronunciation of terminal plosive consonants /t/, /d/, and /k/ during reading; their consonant waveforms were incomplete. After extended AI real-time correction and repeated focused training, these students can now pronounce terminal consonants properly, and the integrity of their consonant acoustic waveforms has significantly improved.

In general, the real-time error correction function and unlimited repeated training offered by the AI platform help students gradually correct long-standing fixed pronunciation errors caused by mother tongue negative transfer and middle school silent learning, and the overall pronunciation accuracy and standardization level are significantly improved, as fully supported by the objective acoustic parameter data extracted using Praat.

4.2. Comparison of oral fluency and intonation prosody

Based on the quantitative analysis of the three main indicators from Praat, namely speech rate, frequency of invalid pauses, and range of pitch fluctuation, it was found that students' overall oral expression ability improved significantly after receiving AI-led situation-specific oral training, and the two main problems of poor oral fluency and monotonous intonation were effectively solved.

In terms of oral fluency, the students' speech speed per minute is relatively stable and moderate; the number of invalid repeated pauses and stuttering pauses in continuous speech has decreased significantly, and the logical coherence of sentences has improved. Before the intervention, most students read word by word with frequent pauses every two words and had a fluctuating speech rate; after one semester of situational dialogue training on the AI platform, more than 80% of the students could take entire meaning groups as the

unit for continuous expression, and the frequency of meaningless pauses in a 100-word paragraph had been reduced by more than half.

In terms of intonation prosody, as shown in the pitch data, the range of pitch fluctuation of students' sentences is reasonable and layered. Before training, most students' reading pitch curves are almost straight lines with an extremely small fluctuation range and show a typical flat intonation; after intervention, students can use a standard rising tone for general questions and a falling tone for declarative sentences, accurately mark the stress position of key words in sentences, the pitch curve has regular ups and downs, and oral expression is more rhythmic and natural. Extended AI situation-based dialogue simulation and real-time intonation-scoring training help students develop an English language sense and improve the rhythm and intonation of their oral expression at a fundamental level.

4.3. Analysis of students' subjective learning effects and attitudes

Based on the statistics of the recovered valid questionnaires and the sorting records of semi-structured interviews, it can be concluded that the AI-assisted hybrid teaching mode has an overall positive impact on students' phonetic learning situation and learning psychology.

First of all, the fear and anxiety students have about learning English pronunciation have been reduced. In the traditional classroom of collective reading, students with poor pronunciation are often afraid of making mistakes and being laughed at by their classmates, so they do not dare to speak proactively; the AI platform provides an independent, low-pressure learning environment where they can practice repeatedly without psychological pressure, and the learning anxiety of more than 90% of the respondents has visibly decreased.

Secondly, the students' initiative in self-study and independent correction of errors has been promoted. With the help of real-time phonetic error feedback provided by the platform, students can promptly identify their own pronunciation deficiencies without waiting for the teacher's group correction, and after prolonged practice, they gradually learn to self-correct; thus, their self-confidence in speaking has also risen considerably.

Thirdly, most students have a consistent positive view that AI personalized training can accurately address their individual phonetic deficiencies, and the learning results are superior to those achieved by traditional unified collective teaching with teacher demonstration reading. At the same time, a small number of students also mentioned existing shortcomings of the interview, such as a lack of guidance from the platform for free impromptu oral expression, and provided practical directions for the subsequent optimization of the teaching mode.

5. Discussion

5.1. The promotion mechanism of AI-assisted learning for second-language pronunciation acquisition

In light of the second-language pronunciation acquisition theory and the empirical acoustic data obtained in this study, the four core action paths for AI intelligent platforms in promoting vocational college students' second-language pronunciation learning have been established, forming a correction system for phonetic errors.

First of all, AI can overcome the constraints of time and space in traditional classroom training and provide more frequent opportunities for repeated practice to correct pronunciation errors that have occurred

over a long period. Traditional classes have limited time for phonetic practice, and the AI platform allows students to log in for training at any time after class; at the same time, the system will automatically record the duration and type of errors in training to ensure that a sufficient volume of practice is achieved for standard pronunciation consolidation. Sufficient repetition of standardized input can reduce the negative transfer effect of local dialects and middle school incorrect pronunciation habits, thereby providing a foundation for improving phonetic accuracy.

Second, real-time error feedback can be provided promptly for any pronunciation errors to help learners correct them immediately in their second language study. In traditional teaching, teachers can only provide centralized feedback after students finish reading, and the feedback is significantly delayed; the AI platform can mark wrong phonemes, misplaced stress, and abnormal intonation simultaneously during the recording of students' reading, immediately play the standard pronunciation for comparison, and thus allow students to correct mistakes in real time and avoid repeatedly reinforcing incorrect pronunciation through multiple readings.

Thirdly, personalized differentiation of tasks adapts teaching to the various types of phonetic defects for different students. Based on the pretest Praat acoustic diagnosis data of students, categorize their errors, and then push targeted training tasks to address their individual weaknesses; thus, it is no longer necessary to conduct uniform repetitive exercises for the entire class, and the phonetic training has been made more relevant and efficient.

Fourthly, rich standard native audio materials and various situational oral exercises help students develop a natural sense of English and improve oral prosody and sentence rhythm. The large database of standard pronunciation corpora and situative dialogue modules in the AI platform can simulate real-life communication scenarios for students, allowing them to learn how to use all kinds of intonation patterns naturally in speaking rather than merely memorizing words.

5.2. The core value of Praat acoustic evaluation in pronunciation teaching

Traditional pronunciation teaching evaluation relies solely on teachers' subjective human auditory perception and can only identify obvious pronunciation errors, such as completely incorrect phonemes, but fails to observe subtle phonetic deviations in students' reading. Praat professional acoustic analysis software can correct for the defect of subjective judgment in the above step, and its main functions in phonetic teaching include:

First, Praat can capture subtle pronunciation deviations that human hearing cannot detect in objective numerical data and visual sound spectrograms, such as a slight offset of vowel formants, insufficient range in pitch fluctuation, and an overly short pause interval; thus, comprehensive and fine-grained diagnosis of students' pronunciation problems can be carried out.

Second, it changes the mode of evaluation for pronunciation from observation-based subjectivity to data-driven objectivity. Unified acoustic parameter measurement standards eliminate the evaluation deviation caused by different teachers' listening sense, and realize standardized quantitative evaluation of pronunciation quality; thus, reliable, unified, and comparable databases are provided for teachers' teaching adjustment and targeted error correction tutoring.

Thirdly, by comparing the pretest and posttest acoustic parameters longitudinally, it can be seen how much change each student's phonetics has undergone in a semester; at the same time, different parts of

pronunciation can be assessed for improvement precisely, and objective data for grading can be provided to support teachers' decisions rather than relying on subjective feelings.

5.3. Existing problems in the application of AI pronunciation teaching

During the one-semester period of practical teaching intervention, the research group also found some prominent deficiencies and bottlenecks in the joint application of AI speech platforms and Praat acoustic analysis that restrict further improvement in teaching effectiveness:

First, AI platforms perform well in standardized monosyllabic, word, and fixed-text reading training, but are deficient in leading free oral expression and complex contextual prosody. The platform's error recognition logic is based on a fixed standard corpus, and thus cannot flexibly identify reasonable intonation changes in students' personal free expression; only mechanical pronunciation errors can be judged, but no guidance is provided for using flexible rhythm according to the communication context.

Second, a small number of students lack the initiative to learn independently in the after-school AI training. Without offline real-time supervision by teachers, individual students perform the daily training tasks superficially; the frequency and duration of training are low, thus failing to improve their own pronunciation ability effectively, and the desired teaching effect cannot be achieved.

Thirdly, some frontline English teachers are not proficient in operating Praat acoustic analysis software, lack professional knowledge in experimental phonetics to interpret acoustic parameters, and fail to fully utilize the phonetic diagnosis data generated by Praat in daily teaching design and one-on-one targeted tutoring, resulting in the waste of objective acoustic evaluation data resources.

6. Optimization strategies for AI-assisted English pronunciation teaching in vocational colleges

Based on the empirical research results and practical application bottlenecks mentioned above, this study puts forward four targeted operable optimization strategies for the construction of an AI-assisted English pronunciation teaching system in vocational colleges by combining Praat acoustic quantitative evaluation technology.

6.1. Build a hierarchical and personalized intelligent training system

Based on the phonetic diagnosis data of pretest acoustic analysis in Praat, divide the students into three groups according to the type and severity of pronunciation problems: a basic correction group (serious vowel and consonant deviation), a rhythm optimization group (accurate monosyllabic pronunciation but flat intonation and fragmented speech), and a comprehensive promotion group (only minor phonetic defects). Build a hierarchical classified task library on the AI platform, set different training goals, frequencies, and exercise contents for students at different levels, and have the platform automatically push targeted training content based on the acoustic diagnosis labels of the students to achieve precise personalized teaching. Teachers regularly update the task library according to changes in posttest Praat data and adjust the difficulty and focus of training tasks dynamically.

6.2. Strengthen the three-dimensional closed-loop teaching mode

Build a full-featured three-dimensional teaching system that integrates daily AI autonomous training, periodic

Praat acoustic evaluations, and targeted offline teacher guidance. The working mode of the closed-loop system is as follows: Students complete the daily fixed phonetic training and oral dialogue exercises on the AI platform after class; every four weeks, teachers use Praat software to collect speech samples for acoustic parameter analysis, dynamically monitor the learning progress and remaining phonetic deficiencies of the entire class, synthesize quantitative results from acoustic data, and address common phonetic problems in a centralized class session; finally, conduct one-on-one tutoring based on individual defects identified by Praat data to achieve an integrated system of online intelligent training and offline data-driven precise teaching.

6.3. Strengthen skill training for teachers and standardize students' daily learning management

Every semester, schools organize regular, systematic training for English teachers on the two main modules: functional operation of AI intelligent speech platforms and operation and parameter interpretation of Praat acoustic analysis software. Invite experimental phonetics professional teachers to deliver lectures, enhance teachers' digital teaching capabilities and application skills in acoustic data, motivate teachers to use acoustic quantitative data as the foundation for designing daily phonetic teaching, etc. At the same time, set up standardized assessment requirements for students' daily AI after-class practice, link the completion quality and duration of platform training with usual performance scores, establish a weekly training data inspection mechanism, and guide students to develop independent, standardized, and persistent phonetic learning habits to address the problem of insufficient autonomous learning initiative among some students.

6.4. Enrich diversified oral teaching scenarios to compensate for AI platform deficiencies

Based on the basic fixed text reading training modules of the AI platform, expand the diversity of oral teaching scenarios to offset the platform's lack of free expression and contextual prosody guidance. Teachers add offline oral teaching tasks in class hours, such as situational dialogue performance, topic-free retelling, and impromptu speech expression, regularly organize group oral communication activities, and combine the pronunciation accuracy data fed back by AI with teachers' manual evaluations.

7. Conclusion and prospects

7.1. Research conclusions

Based on the empirical data from the one-group pretest-posttest experiment and acoustic quantitative data in Praat, this paper presents the following three main research conclusions for Primary School English Education majors at vocational colleges:

First, extended AI-assisted autonomous pronunciation learning can improve the acoustic features of vocational students' English pronunciation significantly and correct typical phonetic errors such as vowel tongue position deviation, ambiguous consonant pronunciation, and misplaced word stress effectively. Changes in the core acoustic parameters of vowels and syllables, as well as modifications to consonant waveforms, have all been analyzed by Praat to confirm that the optimization effect of AI training has increased monosyllabic pronunciation accuracy.

Second, organized, extended AI situation oral training can help students improve the rhythm and intonation of their speech over time, thus boosting oral fluency. A decrease in the frequency of pauses for invalid reasons and an increase in the range of pitch fluctuations both indicate that students have moved away

from word-by-word mechanical reading and flat single-intonation; thus, their ability to continuously express themselves orally has been all-around improved.

Third, the hybrid teaching model of daily AI regular training and periodic Praat acoustic assessment is highly suitable for the features of English pronunciation teaching in vocational colleges. This way can address the above problems of traditional phonetic teaching, such as a lack of individualized instruction, slow feedback, and single-subject evaluation, and has strong operability and popularization value in vocational college English teaching practice.

7.2. Research limitations

There are certain objective limitations to the scope of this study due to the research environment, and follow-up research will be conducted to address and overcome these deficiencies.

First, the research sample coverage is relatively small. The research subjects are limited to a single class of Primary School English Education majors at one vocational college; the sample size is small, and the research results may not be fully applicable to the pronunciation learning situations of all majors and all levels of vocational college students; therefore, further tests of external validity need to be conducted by expanding the scope of the sample.

Second, the research cycle is relatively short. The entire teaching intervention is for one semester, and therefore, only a short-term improvement in the students' pronunciation skills due to AI training will be achieved. Further long-term follow-up tracking research is needed to examine how the extended, sustained effects of AI-assisted learning on students' phonetic retention ability and professional oral teaching skills manifest after they enter internships.

Third, the indicators of acoustic analysis selected in this paper are limited to basic conventional phonetic parameters. More complex phonetic phenomena, such as sentence stress rhythm, connected speech assimilation, and weak form pronunciation, are not the focus of in-depth analysis, and multi-dimensional acoustic quantitative research on complex oral prosody has yet to be carried out.

Funding

Project category: 2025 Teaching Reform Research Projects of Gansu Provincial College Foreign Language Teaching Research Association
Project name: A Study on the Impact of AI-Assisted English Phonetic Learning in Higher Vocational Education — Taking Pronunciation Training and Oral Expression Ability as Examples
Project Number: 2025DYHYB31

Disclosure statement

The author declares no conflict of interest.

References

[1] Hu WR, Yan XJ, 2026, Empirical Research on AI-enabled Teaching for English Majors in Vocational Colleges. Modern Teaching and Practice, 2(2): 31–38.

[2] Ye BY, 2024, An Empirical Analysis of the Impact of Voice-driven Conversational AI Function on English Speaking Skills of Vocational College Students. Journal of Humanities, Arts and Social Science, 8(11): 2621–2628.

[3] Bouchhioua N, 2024, Fostering the Interplay between Acoustic Phonetics and AI-powered Pronunciation Learning: A Teacher Action Research. Teaching English with Technology, 24(3): 41–67.

[4] Chen YP, Zhang ZY, 2022, An Empirical Research on the Effectiveness and Generalizability of Visual Feedback Paradigm in English Phonetics Learning. Journal of Architectural Education in Institutions of Higher Learning, 31(1): 171–180.

[5] Boersma P, Weenink D, 2024, Praat: Doing Phonetics by Computer (Version 6.4.23), Computer Software.

[6] Yang WG, Zhao X, 2021, Research on the Function of Visual Phonetic Software Praat in Vocational English Phonetics Teaching. 2021 International Conference on Computer Network Security and Software Engineering. IOP Publishing, 2021, 012057.

[7] Luo CY, 2026, A Study on the Path of AI-Assisted Oral English Expression and Pronunciation Training for Vocational College Students. Contemporary Education Frontiers, 4(2): 146–151.

[8] Wu JH, Zhou WT, Cao C, 2024, An Empirical Study on Empowering Oral Teaching with Generative Artificial Intelligence. China Educational Technology, 2024(4): 105–111.

[9] Ma X, 2025, Research on Multimodal Teaching Model of Foreign Language Phonetics Course Based on "Praat + DeepSeek" under the Background of Digital-intelligence Integration. Excellent Papers Online of Chinese Scientific and Technical Papers, 18(4): 375–378.

[10] Chen MT, 2022, Computer-aided Feedback on the Pronunciation of Mandarin Chinese Tones: Using Praat to Promote Multimedia Foreign Language Learning. Computer Assisted Language Learning, 37(1): 1–26.