BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

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

ARTICLE

A Comparative Study of Traditional vs. ChatGPT-Assisted Pronunciation Teaching in Vocational Universities

Liling Liao [1] , ***Ziyue Lin*** [2*]

[1] *School of International Digital Business, Guangzhou Vocational University of Science and Technology, Guangzhou 510550, China*

[2] *School of Foreign Languages, Hanshan Normal University, Chaozhou 521041, China*

ABSTRACT

Pronunciation teaching (PT) is crucial for improving oral communication, particularly in vocational universities where students require practical language proficiency for professional purposes. With advancements in artificial intelligence (AI), integrating tools like ChatGPT into language instruction offers new opportunities to enhance pronunciation development. This study examines the effectiveness of ChatGPT-assisted pronunciation instruction compared to traditional classroom-based methods in vocational university settings. A total of 114 non-English major students were randomly divided into two groups. Group A (GA) received traditional teacher-led pronunciation instruction, while Group B (GB) practiced with ChatGPT through conversational dialogues, AI-generated feedback, and pronunciation tasks. Over six weeks, both groups completed pre- and post-tests using a structured rubric assessing clarity, intonation, and segmental precision. Data were analyzed using paired and independent samples *t*-tests and ANOVA via IBM SPSS 29. Results showed that the ChatGPT-assisted group achieved significantly greater improvements in clarity and accuracy ($p < 0.001$) than the traditional group. These findings indicate that AI-assisted instruction provides effective personalized feedback, enhances engagement, and accelerates pronunciation learning. The study concludes that integrating ChatGPT-based approaches can serve as a valuable supplement to traditional pronunciation teaching, promoting more efficient and interactive language learning in vocational education.

Keywords: ChatGPT-Assisted Learning; Pronunciation Instruction; Vocational Education; AI in Language Teaching;

*CORRESPONDING AUTHOR:

Ziyue Lin, School of Foreign Languages, Hanshan Normal University, Chaozhou 521041, China; Email: nnyy2005@163.com

ARTICLE INFO

Received: 27 August 2025 | Revised: 9 October 2025 | Accepted: 17 October 2025 | Published Online: 12 November 2025
DOI: https://doi.org/10.30564/fls.v7i12.11835

CITATION

Liao, L., Lin, Z., 2025. A Comparative Study of Traditional vs. ChatGPT-Assisted Pronunciation Teaching in Vocational Universities. Forum for Linguistic Studies. 7(12): 936–949. DOI: https://doi.org/10.30564/fls.v7i12.11835

Language Proficiency

1. Introduction

The value of English pronunciation in vocational education is more relevant. As students move into their future careers in industries such as tourism, foreign language education, customer service, and simultaneous interpretation, being able to fully communicate is very important in these industries[1]. While pronunciation cannot have a specific bearing on a technical task, pronunciation does add to perceived professionalism and credibility with respect to defining experiences and feelings for clients, coworkers, and international clients[2]. However, students in vocational universities face unique challenges with respect to English pronunciation, such as limited time during English classes, exposure to the English language, and a lack of instruction[3]. Understanding and addressing these challenges is vital to building the level of communicative skills required to be successful and even competitive globally in vocational industries[4]. An area advancing swiftly in language instruction is the integration of artificial intelligence technology. This technology enables interactive and customized learning experiences, reinforcing the movement toward more accessible and learner-specific language education. Applications like ChatGPT provide real-time feedback, create conversational practice scenarios, and can support development in pronunciation, grammar, and vocabulary[5]. Various conversational platforms utilize different modes of interaction, which appeal to students' different learning styles and paces, increase engagement, and reduce anxiety related to language learning. While the adoption of technology in teaching and language teaching is inevitable, tools like ChatGPT can enhance language learning opportunities for students in vocational institutions that can continue to operate with limited English instruction time and resources[6].

As shown in **Figure 1**, AI is essential in language acquisition because it can adapt learning experiences for each learner depending on their level and growth limitations. It delivers fast feedback, allowing students to swiftly correct problems in grammar, vocabulary, and sentence structure. Advanced speech analysis systems also help with pronunciation and spoken fluency improvement through continuous testing. They can also help with language correctness by providing suggestions for grammatical and syntactical faults. AI-powered virtual assistants link learning with practitioners in practice-based interactions and contexts, increasing learners' communication confidence and competency[7]. Built-in progress monitoring tools aid both the student and the instructor in developing improvement tactics tailored to specific areas of weakness.

Figure 1. Features of AI in Language Proficiency for Students.

Pronunciation is an important aspect of communication, but is even more important in vocational education, which focuses on practical and skill-based tasks for specific jobs[8]. A good command of English improves job prospects, facilitates academic and professional communication, and increases proficiency in language in general. English is generally taught in a formal way, which provides a certain structure or framework regarding the skills needed and gives learners opportunities to practice important skills, such as grammar, vocabulary, and language functions[9]. Many students in vocational programs come from diverse language backgrounds and often lack sufficient exposure to native spoken English, making pronunciation training a key need[10]. From a student's perspective, the challenges in learning pronunciation are evident: classes are often large, materials feel outdated, and learning through repetitive drills can feel unmotivated[11]. The teacher-centered setup makes it difficult to receive immediate or personalized feedback, which limits opportunities for improvement[12]. These conditions affect not only language proficiency but also the confidence required to speak effectively in real-world situations[13]. As the area most prone to bias and preconceptions and the one

least willing to adapt to all viewpoints, pronunciation. Thus, phonology and phonological attitudes are used to represent and embed many of the concerns at hand, even if they are fundamentally cultural and political [14]. The significance of pronunciation has mostly been decided by ideology and gut feeling rather than by scientific research. Instructors have intuitively determined which aspects are most effective in improving clarity and which may be learnt in a classroom [15]. Clear pronunciation is an important goal for second language learners to ensure effective communication. However, many language programs do not provide separate pronunciation courses. Instead, pronunciation is often treated as a small part of speaking classes or integrated unsystematically across the curriculum [16]. A considerable amount of research still uses minimal pair lists for L2 pronunciation research that were created in pre-advanced technology stages of research. This presents a need for conversant materials that are closer to current research tools and the needs of learners [17]. Intelligibility and comprehensibility are both related to understanding L2 speech, but they differ in emphasis. Intelligibility assesses actual understanding, often through transcription, while comprehensibility considers how easily speech can be followed, typically with ratings [18]. Addressing these barriers is vital to help students develop strong communication skills for future careers.

1.1. Research Objective

The research seeks to explore how effective ChatGPT is in teaching pronunciation in vocational universities as compared to traditional teacher-led instruction. It examines the enhancement of pronunciation clarity, phonetic accuracy, and prosody of speech among non-English major students. The aim is to explore whether AI-supported learning demonstrated greater enhancement than a more traditional approach in oral communication skills. The outcomes can be evaluated by measuring the performance of the AI-supported and traditional teacher-led instruction quantitatively.

1.2. Research Contributions

- **Real-life Evidence to Validate AI-Learning:** It demonstrates the use of ChatGPT for pronunciation instruction improves pronunciation clarity, intonation, phonemic accuracy, clarity, intonation, and segmental precision when compared to a traditional approach.
- **Improved Pedagogical Model for Vocational Contexts:** It highlights how AI tools can address the limitations of teacher-centered pronunciation instruction, particularly in environments with limited educational resources.
- **Practical Framework for the integration of AI:** It provides a framework that can be replicated in the implementation of ChatGPT within pronunciation training. Engagement and learning outcomes were improved.

1.3. Research Organization

Research frameworks are organized into the following sections: Section I includes the introduction of PT in GA and GB; Section II presents related works, including the concept of PT; Section III depicts methodology, i.e., workflow of PT from participants to statistical test; Section IV shows the results; and Section V depicts the conclusions along with limitations and future scope.

2. Literature Reviews

Research [19] examined how ChatGPT supports EAP (English for Academic Purposes) students' autonomy, competence, and relatedness using self-determination theory. Interviews with 24 postgraduates show it enhances autonomy and competence through flexible, personalized support. However, mixed views on relatedness stress the need to balance AI with human interaction. The effects of integrating ChatGPT and Speech Ace into EFL (English as a Foreign Language) classrooms were to improve pronunciation accuracy and L2 (second language) motivation [20]. Among 71 undergraduates, significant gains were seen in phoneme accuracy and motivation, with limited improvement in stress patterns. Results support the pedagogical value of AI-assisted feedback despite baseline-related limitations. It explored the effects of AI-enhanced applications on L2 learners' pronunciation and attitudes using a mixed-method design [21]. With 48 learners, results showed significantly better pronunciation gains in the AI, and more positive learner attitudes. Findings support AI tools for individualized pronunciation instruction. Developed a natural language processing (NLP)-based spoken English teaching model to address limitations of traditional oral instruction [22]. By tailoring methods to learners' speech

traits, the model offers adaptive, constructive feedback. Experimental results show a 19.7% improvement in oral comprehension and a 33.3% boost in learning enthusiasm. The potential of AI in enhancing language learning, focusing on personalized instruction, pronunciation, and fluency through tools like chatbots and speech recognition [23]. It addresses misconceptions among educators and highlights AI's role in fostering learner autonomy. Findings suggest practical ways to integrate AI into English language teaching for effective classroom application. It focuses on enhancing oral English learning in vocational colleges by integrating AI, speech recognition, and micro-learning apps [24]. An Android-based spoken English system using the library enables interactive practice and independent language learning. The approach addresses challenges in embedding oral training within non-English major classrooms.

Research [25] highlights the shift toward AI-based oral English assessment amid poor speaking skills among Chinese college students. With CET (College English Test) exams moving to computer-based formats, AI tools offer accurate evaluation and support personalized learning. Mind mapping and AI integration further enhance teaching strategies and student development. The research explored how social network-based interactions using AI applications influenced English-speaking practice among 70 Chinese university students [26]. The results indicated that engaging with AI apps interactively led to notable improvements in speaking proficiency and was well-received by learners. These outcomes support the incorporation of social media platforms into AI-driven language learning environments. Additionally, the research examined EFL learners' views on EAP Talk, an AI-powered speech evaluation tool drawing insights from 366 participants across five universities [27]. Findings showed improved speaking skills and learner motivation, but concerns included voice recognition accuracy, limited feedback detail, and a lack of course-relevant content. Results highlight both the promise and areas for improvement in AI-assisted speaking practice. Research [28] presents an AI-based correction model to improve pronunciation accuracy in virtual English reading using speech synthesis and a hidden Markov model. Algorithmic improvements were made to enhance system performance, and simulations showed the model met core pronunciation correction needs. Results confirm its basic effectiveness for spoken English accuracy and are limited by reliance on simulation rather than real-world classroom implementation. Addressed about phonological norms in English as an International Language (EIL) through a study of Non-native speaker (NNS)–NNS interaction data that identified traits of interaction that influenced intelligibility [29]. It was established that this core fostered intelligibility, was more regionally appropriate, and was more teachable than the traditional native-speaker models. A limitation of this work is the diminished emphasis on native-speaker norms, which is limiting in terms of preparing learners for configurations of English use.

To judge comprehensibility and intelligibility, eighteen native English listeners rated and transcribed speech samples produced by Mandarin L2 speakers [30]. The findings revealed that a high level of intelligibility and comprehensibility was demonstrated, all while accent ratings varied widely, and a strong accent did not always lead to a lack of understanding. A limitation of this study was the small, L1-specific sample; the findings are limited in their larger applicability.

Research Gap

While the integration of AI in pronunciation instruction has shown increasingly positive outcomes, studies assessing its long-term effectiveness across diverse learner groups and varied educational settings of vocational universities remain limited. Most of the research emphasized the short-term successes of AI pronunciation systems without looking at long-term successes or scalability in the context of classrooms. There are many issues, such as the quality of feedback provided, the accuracy of speech recognition systems, and alignment with curriculum objectives, that have not been reconciled. Research also calls for a deeper understanding of adaptive AI systems that provide context-aware and pedagogically relevant feedback for various EFL learners.

The research fills significant research voids by providing empirical data on ChatGPT’s impact on structured classroom learning contexts with a primary focus on practical pronunciation outcomes. The research can portray sustained improvements in clarity, intonation, and phonetic precision or ability in controlled experimental research. Also significant, different from previous research, the research applied a standardized scoring rubric across both groups, though not an external standardized proficiency test, and quantitative

assessments to strengthen reliability, which has often been overlooked. The research primarily focuses on vocational learners, a group undocumented in AI-supported pronunciation research.

3. Methodology

Research aims to demonstrate that students in vocational universities can benefit from pronunciation instruction, which students need to communicate clearly to be successful in their careers. This approach typically misses direct and individualized feedback and involves passive participation of students. ChatGPT can continue to develop AI tools; a way to teach pronunciation has the potential to be augmented shortly. The research can compare ChatGPT-assisted and traditional approaches to instruction. **Figure 2** demonstrates the working flow of PT.

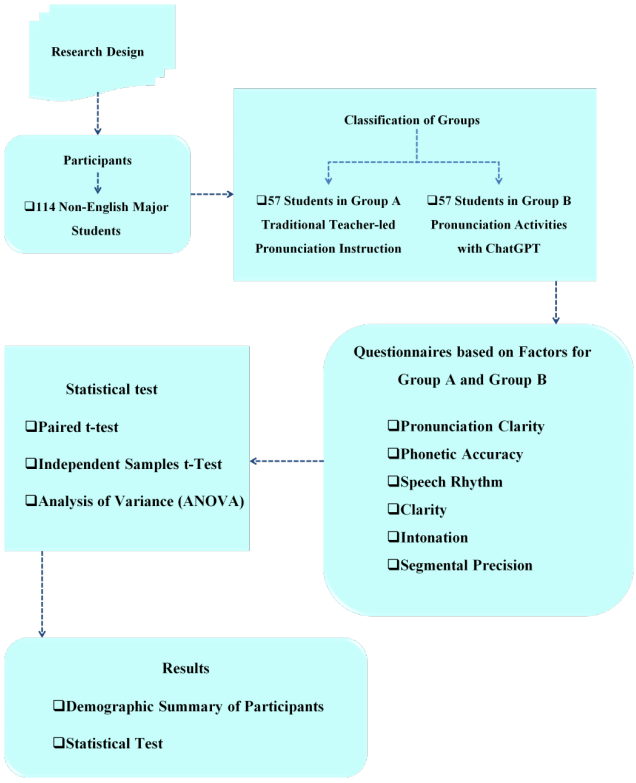


Figure 2. Working flow of PT.

3.1. Research Design

The research used a quasi-experimental research design with a pre-test and post-test method to measure the predictive effectiveness of pronunciation instruction using traditional approaches in GA and ChatGPT-assisted pronunciation instruction in GB. Participants in both groups were evaluated on their pronunciation before and after the six-week intervention, with each instruction group evaluated with coded rubrics that assessed pronunciation clarity, phoneme accuracy, and rhythm of speech. The benefit of the quasi-experimental research design is that it measured learning outcomes while testing these same learners using paired and independent samples t-tests.

3.2. Participants

The research consisted of 114 students from a vocational university who were not majoring in English. They were randomly assigned to two groups of 57. GA received teacher-led pronunciation instruction on a pronunciation point based on repetition and prompts. GB practiced pronunciation via pronunciation drills through ChatGPT, where they could have interactive conversations with AI using self-paced learning and receive feedback. Each group participated in 6 sessions, each lasting approximately 45 min, totaling 4.5 h of instruction and practice. Two portions of 28–29 students each were used for instruction, which is in line with standard class sizes. To lessen variability, the timetable, teacher, and instructional content were maintained consistent between sections.

3.3. Classification of the Group

GA followed traditional teacher-led pronunciation instruction with group feedback and structured lessons, while GB used ChatGPT for flexible, self-paced practice with personalized AI feedback on clarity, intonation, phonetics, and segmental precision.

- GA

GA consists of 57 non-English major students who learned pronunciation through more traditional teacher-led activities in the classroom. Lessons consisted of repetition, phonemic drills, and teacher corrections. Feedback was given to the group, and there was little individual attention. The lessons were similar to an organized EFL curriculum commonly found in EFL classrooms. The conventional instruction was teacher-centred. The lessons relied on memorization and phonemic drills, in which the teacher produced target sounds and the students produced them as a group. Any corrective feedback was generally given to the group as

a unit, and very little attention was given to individuals. The types of activities used included guided reading passages, word lists, and sentence structure practice. The instructional practices align with the traditional pronunciation instruction model when common pronunciation practices are used in vocational English classrooms. In GA (traditional group), segmental precision was taught through teacher-led minimal pair drills (e.g., /l/ vs. /r/), phoneme charts, and repetition of target sounds in controlled word lists and sentences. Feedback was collective and based on the teacher's correction. The research design purposefully represented the classroom conditions of vocational universities, whereby pronunciation instruction from a teacher is delivered to a large number of students over a short time frame. Teachers typically provide feedback to the class as a whole, rather than individual corrections. While the contrast with feedback from an AI is a limitation of the research, it does represent the ecological validity of classroom instruction.

- GB

GB consists of 57 non-English major students who practiced conversational dialectically and pronunciation activities with ChatGPT. ChatGPT was able to provide personal constructive feedback on clarity, intonation, and phonetics for every interaction the learners had with ChatGPT, which was private. Students practiced independently at their own time, with flexible, interactive, self-paced pronunciation development outside of the classroom. The group that received AI assistance (GB) worked with ChatGPT in a blended format that incorporated both drills and communicative practice. Students engaged in 6 sessions, each lasting approximately 45 min, totalling 4.5 h of interaction with ChatGPT. They practiced independently at the time, with flexible, interactive, self-paced pronunciation development outside of the classroom. The group that received AI assistance (GB) worked with ChatGPT in a blended format that incorporated both drills and communicative practice. Students conducted structured pronunciation drills created from prompts in ChatGPT (e.g., practicing minimal pairs and vowel-consonant distinctions), then met with ChatGPT for a conversation where it responded dynamically to students' responses. In both tasks, ChatGPT provided immediate individualized feedback on clarity, stress, intonation, and segmental accuracy. This two-part structure (drill + conversation) ensured that students practiced a balance of accuracy-focused practice and communicative fluency opportunity for development. The seeming contradiction arises from language. The AI-assisted method involved drill-like practices (repetitions and tasks based on specific phonemes) and then conversational practice (dialogues based on context). These did not need to be separate; ChatGPT was designed to include drills and then provide the interactive application in dialogue afterwards. ChatGPT dynamically produced minimal pair contrasts and offered remedial feedback on mispronunciations in conversational circumstances in the GB (AI-assisted group). This integrated contextual dialogue is used with drill-based accuracy practice. Rubric-based pre- and post-tests that were independently rated by qualified raters were used to assess improvement in both groups.

3.4. Pronunciation Assessment Factors

Clarity/Intelligibility—Refers to the relative ease with which the listener can comprehend the speaker's words. Phonological intelligibility is representative of the overall success of pronunciation to render the speech accessible to the listener.

Phonetic Accuracy—The accuracy of individual speech sounds, both vowels and consonants. The accuracy of phonetic production has a beneficial impact on reducing misunderstandings and improving oral communication.

Rhythm—The natural flow and timing of a speech, spelling and length of syllables, stress, and pausing. Proper rhythm helps make the speech more fluent and sound more natural.

Intonation—Pitch and melody variations in speech can express emotion, emphasis, and meaning. Accurate interpretation of the speaker's intent is facilitated by appropriate intonation.

Segmental Precision—Sound sequences and individual phonemes that are accurately articulated. Particularly in minimum pairings, accuracy at the segmental level guarantees clarity and lowers pronunciation mistakes.

Overall Comprehensibility—The overall intelligibility of a speaker's remarks in authentic communication situations. It illustrates how segmental precision, rhythm, intonation, and clarity all work together to create effective communication.

3.5. Questionnaires Based on Factors

Table 1 presents the assessment questionnaires for GA and GB, which followed a traditional, instructor-led pronunciation approach. The six items assess key pronunciation factors: Intelligibility, phonetic accuracy, rhythm, intonation, and segmental precision, and overall comprehensibility focused on classroom practice. These items evaluate how interactive AI-based learning supports improvement in the same six areas, emphasizing personalized, immediate feedback and self-paced practice through dialogue with ChatGPT to enhance pronunciation proficiency. To ensure valid comparisons, both groups were evaluated using the same six-dimensional rubric: (1) intelligibility, (2) phonetic accuracy, (3) speech rhythm, (4) intonation, (5) segmental precision, and (6) overall comprehensibility. Although earlier drafts showed slightly different wording tailored to group context, all assessments were standardized and scored with the same rubric across GA and GB.

Table 1. Questionnaires for GA and GB.

Factor	Code	Evaluation Criterion	Rater Action
Clarity/Intelligibility	Q1	The degree to which words and sentences are understood without frequent repetition.	Rater scores 1–5 based on clarity of speech output.
Phonetic Accuracy	Q2	Accuracy in pronouncing target phonemes in different contexts (word, sentence).	Rater notes errors and assigns an accuracy score.
Rhythm	Q3	Ability to maintain natural speech stress, timing, and pacing.	Rater evaluates fluency and rhythmic appropriateness.
Intonation	Q4	Appropriateness of rising/falling pitch patterns in sentences and discourse.	Rater evaluates pitch contour accuracy.
Segmental Precision	Q5	Accuracy in producing specific consonants and vowels in onset, nucleus, and coda positions of words.	Rater marks correct/incorrect, scores consistency.
Overall Comprehensibility	Q6	Overall ease of understanding for an unfamiliar listener with minimal effort.	Rater assigns global comprehensibility score.

In this research, segmental precision is defined as accurately producing single consonants and vowels in certain parts of a word (onset, nucleus, coda). Differently, phonetic accuracy (or clearness) refers to the ability to produce entire words so that they are perceived accurately by a listener. To do this consistently with standard terminology, also refer to these transactions as intelligibility (the listener being able to hear and understand speech) and comprehensibility (the listener's judgement of ease or difficulty in hearing the speech).

3.6. Statistical Test

IBM SPSS (Statistical Package for the Social Sciences) version 29 was employed to conduct three statistical tests assessing pronunciation improvement. The paired t-test measured individual progress within the same group over time. The independent samples t-test compared post-instruction outcomes across all pronunciation features between the two groups. Additionally, a one-way ANOVA was used to evaluate variance in pronunciation performance across multiple dimensions between the groups. Bonferroni corrections were used to account for Type I error from multiple comparisons. For the six pronunciation variables, the modified significance level was chosen at $p < 0.008$ (0.05/6).

- Paired *t*-test

The paired *t*-test was performed within each group to compare students' pronunciation scores before and after the six-week period of instruction. The test targeted individual progress in clarity, accuracy of phonetic production, and rhythm of speech. The Paired *t*-test determined if students made significant improvement within each group over time as a result of the specific teaching approach, as shown in Equation (1).

$$s = \frac{\overline{c}}{t_c/\sqrt{m}} \tag{1}$$

Wherein, $\overline{c}$ is the difference between paired observations, t_c is the standard deviation of the difference, $s$is is the test, and m represents the number of pairs.

- Independent Samples *t*-Test

The independent samples *t*-test was applied to evaluate the pronunciation outcomes of GA and GB by determining whether the differences in post-test scores for clarity, phonetic accuracy, and speech rhythm were statistically significant between the two groups. This test also helped to summarize the relative effectiveness of each instructional approach on learners' pronunciation performance, as presented in Equations (2) and (3).

$$s = \frac{\overline{w}_1 - \overline{w}_2}{t_o \sqrt{\frac{1}{m_1} + \frac{1}{m_2}}} \tag{2}$$

$$(t_o = \sqrt{\frac{(m_1 - 1)\, t_1^2 + (m_2 - 1)\, t_2^2}{m_1 + m_2 - 2}} \tag{3}$$

Wherein, $\overline{w}_1 - \overline{w}_2$ presents the subtraction of the means in the fraction, t_o is the pooled standard deviation, and $\frac{1}{m_1} + \frac{1}{m_2}$ denotes the sample size of the GA and GB.

- ANOVA (Analysis of variance)

An ANOVA was conducted to assess the between-group variances across different pronunciation factors at pronunciation clarity, phonetic accuracy, speech rhythm, intonation, and segmental accuracy. This assessed for possible statistically significant mean differences in clarity, intonation, and segmental accuracy between the two instructional groups across all measurements of pronunciation.

The formula of ANOVA has been shown in Equations (4)–(6).

$$E = \frac{MS_{between}}{MS_{Within}} \tag{4}$$

Wherein, E denotes the ANOVA test statistic, $MS_{between}$ is the mean square between groups (BG), and MS_{Within} is the mean square within groups (WG).

$$MS_{between} = \frac{SS_{between}}{df_{between}} \tag{5}$$

$$MS_{within} = \frac{SS_{Within}}{df_{within}} \tag{6}$$

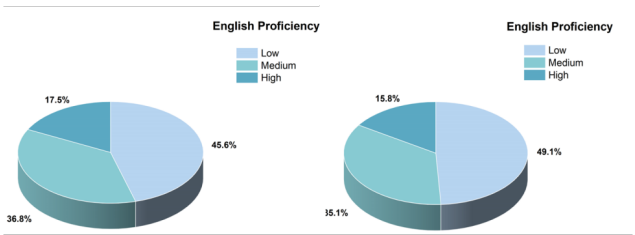
Where $SS_{between}$ the sum of squares in BG, SS_{within} is the sum of squares WG.

4. Results

The research underscores the relative efficiency of traditional and ChatGPT-assisted PT in vocational university contexts. The importance of language competence to employability and communication continues in various industries, increasingly shaped by globalization. In particular, improving pronunciation is beneficial for vocational learners. The research further responds to the need for new, technological approaches that lean into digitization in addition to offering traditional language teaching techniques. By assessing measurable gains on multiple aspects of pronunciation, the research adds an empirical justification for technology-assisted instruction unfolded through AI. The research can also serve as a general reference for educators and others with holder status to understand the ways of using intelligent tools that exist within contemporary language learning contexts. Ultimately, the research helps improve the field of learning digital pedagogy about vocational education.

4.1. Demographic Summary of the Participants

Table 2 presents the demographic details of participants from GA and GB, each consisting of 57 individuals. The gender distribution was nearly equal in both groups, though GB had a slightly higher number of male participants. The majority of students in both groups fell within the 19–20 age range, followed by those aged 17–18, and a smaller portion aged 21 and above. In terms of English proficiency, most participants reported low to medium levels. Regarding daily English exposure, nearly half of the participants indicated using English for less than one hour per day, reflecting limited language immersion. GA and GB each included 57 students with balanced gender distribution and majority aged 19–20 years. Most participants reported low to medium English proficiency and limited daily exposure to English. **Figure 3** shows English proficiency (a) GA and (b) GB. **Figure 4** presents the age range (a) GA and (b) GB.

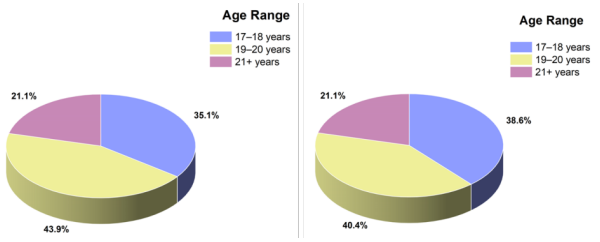


(a) **(b)**

Figure 3. Presentation of English proficiency (**a**) GA and (**b**) GB.

Table 2. Demographic Characteristics of Participants.

Features	Categories	GA (n = 57)	Percentage (%)	GB (n = 57)	Percentage (%)
Gender	Male	28	49.1	30	52.6
	Female	29	50.9	27	47.4
Age Range	17–18 years	20	35.1	22	38.6
	19–20 years	25	43.9	23	40.4
	21+ years	12	21.1	12	21.1
English Proficiency	Low	26	45.6	28	49.1
	Medium	21	36.8	20	35.1
	High	10	17.5	9	15.8
Daily English Exposure	<1 h	30	52.6	28	49.1
	1–2 h	20	35.1	21	36.8
	>2 h	7	12.3	8	14.0



(**a**) (**b**)

Figure 4. Presentation of age range (**a**) GA and (**b**) GB.

4.2. Measures Pre-Post Change within Group: Paired *t*-Test

All significant results remained below the Bonferroni-adjusted alpha of 0.008, confirming the robustness of findings. Both GA and GB showed significantly improved pronunciation post-instruction according to paired t-tests. The effect sizes from the results showed average or moderate gains for all pooled Standard deviations in GA. GB showed a larger difference from pre to post instruction, therefore a bigger increase for clarity, accuracy, rhythm, intonation, and segmental precision. The statistically significant results show that instruction was effective in improving pronunciation. Following teaching, both GA and GB demonstrated notable gains in all speech metrics. In terms of clarity, phonetic accuracy, rhythm, intonation, and segmental precision, GB outperformed GA.

4.2.1. GA: Paired *t*-Test

Table 3 shows the results for GA, consisting of the paired *t*-tests to analyze the effectiveness of instruction about various aspects of pronunciation. There were significant improvements in all variables, and post-test means were noticeably higher than pre-test means. Pronunciation clarity demonstrated the highest mean difference of 10.2, followed by phonetic accuracy (8.1) and clarity (7.8). There were improvements in the rhythm of speech, intonation, and segmental accuracy as well. All *t*-values were statistically significant at $p < 0.001$, demonstrating that the instructional method had a large impact on pronunciation results. The degrees of freedom were the same for all tests (56). All factors showed notable pre-to-post increases with GA, with pronunciation clarity showing the largest increase (10.2 points). Moderate gains from teacher-led instruction are indicated by effect sizes. **Figure 5** presents the values of the pre- and post-means in GA.

Table 3. Paired *t*-Test of GA.

GA						
Variables	**Mean (Pre)**	**Mean (Post)**	**Mean Diff**	***t***	**df**	***p*-Value**
Pronunciation Clarity	60.0	70.2	10.2	5.85	56	<0.001
Phonetic Accuracy	58.0	66.1	8.1	4.92	56	<0.001
Speech Rhythm	61.5	68.5	7.0	4.25	56	<0.001
Intonation	62.0	69.0	7.0	4.42	56	<0.001
Segmental Precision	59.5	65.3	5.8	3.92	56	<0.001
Clarity (Rubric)	60.0	67.8	7.8	4.75	56	<0.001

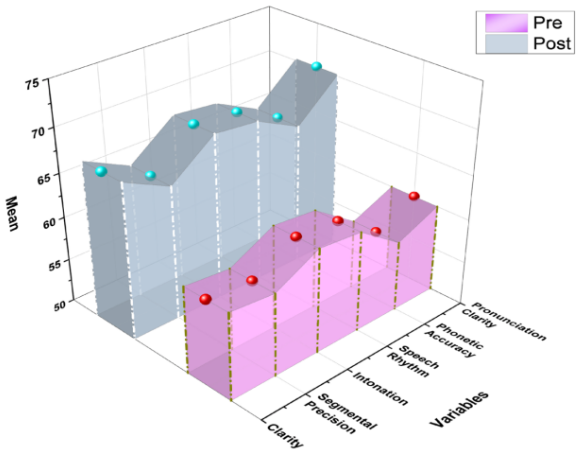


Figure 5. Presentation of mean pre- and post in GA.

4.2.2. GB: Paired *t*-Test

The paired *t*-test results for GB are in **Table 4**, and consistent with GA, there were statistically significant results across all measures of pronunciation after the instruction phase. The clarity of pronunciation had an average mean increase of 19.0 points, followed by phonetic accuracy with a mean increase of 16.2 points, rubric-based clarity with a mean increase of 16.0 points, and both speech rhythm and intonation with mean increases of 15.9 and 15.0 points, respectively. Finally, segmental accuracy had a mean increase of 13.6 points. None of the tests yielded *p*-values less than 0.001, with the *t*-values indicating solid instructional impacts. It is worth noting that all measures returned equal degrees of freedom (df = 56), indicating the same participants were included in the analysis of the *t*-tests. In sum, these results suggest that GB received significant gains in pronunciation performance. With the biggest increases in segmental precision (13.6 points) and pronunciation clarity (19.0 points), GB demonstrated significant improvements across the board. Compared to conventional approaches, AI-assisted practice yielded greater improvements. **Figure 6** presents the values of mean pre- and post in GB.

Table 4. Paired *t*-Test of GB.

GB						
Variables	**Mean (Pre)**	**Mean (Post)**	**Mean Diff**	***t***	**df**	***p*-Value**
Pronunciation Clarity	59.5	78.5	19.0	8.34	56	<0.001
Phonetic Accuracy	57.8	74.0	16.2	7.75	56	<0.001
Speech Rhythm	60.0	75.9	15.9	7.90	56	<0.001
Intonation	61.5	76.5	15.0	7.25	56	<0.001
Segmental Precision	58.9	72.5	13.6	6.80	56	<0.001
Clarity (Rubric)	59.0	75.0	16.0	7.95	56	<0.001

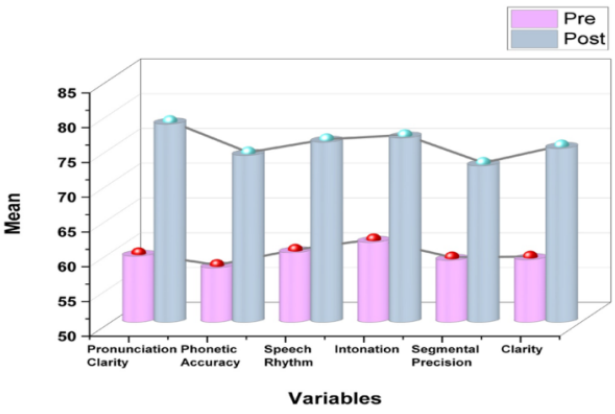


Figure 6. Presentation of mean pre- and post in GB.

4.3. Assesses Group-Level Performance Differences: Independent Samples *t*-Test

Table 5 and **Figure 7** present the independent samples *t*-test results comparing the post-test scores of GA and GB across six pronunciation variables. In every case, GB had higher results than GA, and statistically significant differences ($p < 0.001$) were found on pronunciation clarity, phonetic distinctiveness, speech rhythm, intonation, segmental accuracy, and rubric score for clarity. GB had a higher mean in articulation clarity (78.5 ± 5.4) than GA (70.2 ± 6.1) with a *t*-value of 6.93. The same trends appeared for all the other factors, showcasing a larger amount of improvement by GB in terms of pronunciation because of the way they were taught. In all factors, post-test comparisons showed that GB performed better than GA, with statistically significant differences ($p < 0.001$). The biggest change was in pronunciation clarity, suggesting that AI-assisted education is more effective.

Table 5. Independent Samples *t*-Test.

Variables	GA	GB	*t*	df	*p*-Value
Pronunciation Clarity	70.2 ± 6.1	78.5 ± 5.4	6.93	112	<0.001
Phonetic Accuracy	66.1 ± 6.8	74.0 ± 5.9	6.15	112	<0.001
Speech Rhythm	68.5 ± 7.0	75.9 ± 6.2	5.64	112	<0.001
Intonation	69.0 ± 6.5	76.5 ± 5.8	6.13	112	<0.001
Segmental Precision	65.3 ± 7.2	72.5 ± 6.3	5.51	112	<0.001
Clarity (Rubric)	67.8 ± 6.9	75.0 ± 5.9	5.97	112	<0.001

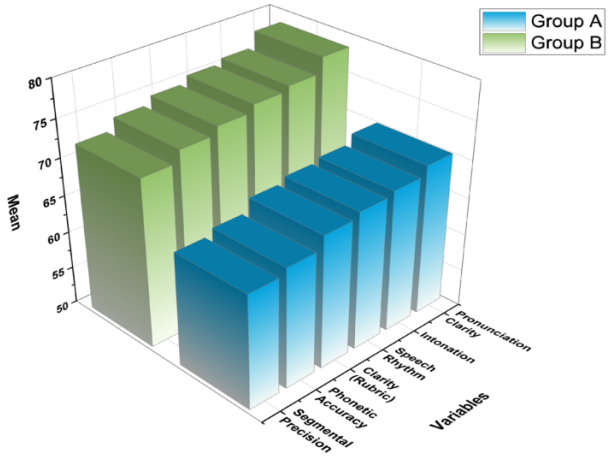


Figure 7. Presentation of Independent Samples *t*-Test.

4.4. Identifies Overall Group Performance Differences: ANOVA

Table 6 presents the ANOVA results on the effect of the instructional modalities on the six pronunciation-related variables between GA and GB. The table contains, for each one of the six variables, the BG and WG values, including the SS, *df*, and MS. The F-values across all variables were statistically significant at $p < 0.001$, which indicates that the differences between groups are substantial. There was a difference with the highest F-value for clarity of pronunciation (48.02), followed by clarity related to the rubric (40.56), accuracy of phonetics (38.17), and intonation (37.22). There was also significant variation in the rhythm of speech and segmental accuracy, with F-values of 31.91 and 29.13, respectively. Therefore, it seems the instructional modality had a strong and consistent effect on the pronunciation outcomes across all variables that were analyzed. Significant differences between GA and GB were confirmed by ANOVA for each of the six pronunciation metrics ($p < 0.001$). Pronunciation clarity had the biggest effect (F = 48.02), demonstrating the powerful influence of the teaching mode.

Table 6. Quantitative values of ANOVA.

Variables	Sources	SS	df	MS	F	*p*-Value
Pronunciation Clarity	BG	1456.8	1	1456.8	48.02	<0.001
	WG	3395.2	112	30.31	-	-
Phonetic Accuracy	BG	1238.4	1	1238.4	38.17	<0.001
	WG	3636.0	112	32.46	-	-
Speech Rhythm	BG	1114.0	1	1114.0	31.91	<0.001
	WG	3907.0	112	34.88	-	-
Intonation	BG	1225.0	1	1225.0	37.22	<0.001
	WG	3683.0	112	32.88	-	-
Segmental Precision	BG	1024.0	1	1024.0	29.13	<0.001
	WG	3936.0	112	35.14	-	-
Clarity	BG	1325.0	1	1325.0	40.56	<0.001
	WG	3660.0	112	32.68	-	-

Notes: SS (sum of the squares), df (degrees of freedom), and MS (mean of squares).

4.5. Discussion

The use of NLP in spoken English education significantly addresses the limitations found in conventional instruc-

tional approaches[17]. It promotes adaptive, individualized learning by monitoring students' verbal language activity and giving indicative motivation and feedback. NLP systems also offer a more contextual learning approach compared to static pedagogical approaches to teaching. Learning with an NLP system is personalized toward the learner's needs and more helpful to the development of accuracy and fluency when speaking. The model allows for constant revision whilst being relevant to an ever-changing technological environment. Research results indicate that learners' understanding and motivation to take part in oral English have increased significantly. This highlights the system's ability not only to facilitate learning but also to sustain motivation and active involvement in the process. Therefore, NLP is an intelligent and scalable way to promote modern language teaching.

The Fourth Industrial Revolution, characterized by rapid technological advancements, is reshaping education through AI[18]. AI has tremendous potential for developing personalized and digital learning experiences, but teachers can have reservations about how AI can affect their teaching practices, particularly when considering language learning. The research considers how AI is supporting English language teaching using tools like chatbots, speech editing, and recognition. Specifically, the research supports learner independence and fluency by providing immediate feedback and by allowing learners to do voice training interactively. The results highlight creative, AI-based methods of improving learners' pronunciation and conversation skills. The research also helps to dispel common myths that can cloud the effective potential of AI in classroom settings. It advocates thoughtful and strategic implementation of AI to enrich the language learning experience overall.

The use of human-computer interaction technologies in English-speaking instruction creates engaging, communicative contexts for developing speaking ability[19]. With the use of AI, mobile apps, and micro-lesson formats, teachers are able to offer flexible, short-form learning models to develop both listening and speaking skills, which is especially important in vocational institutions with students who do not specialize in English and require additional motivation to research the English language independently. The research focuses on language development in oral forms, using multimedia-based resources and Android Integrated voice recognition systems. The use of an English-speaking app developed using a speech recognition library illustrates how mobile technology enables students to receive instantaneous evaluation of their pronunciation. These methods and contexts provide alternatives to the limitations of traditional forms of instruction. In turn, they support student independence, as well as the development of skills in spoken communication effectively.

The research supports that the use of AI tools such as ChatGPT in pronunciation instruction has a statistically significant impact over teacher-led methods. The personalized, real-time feedback offered by ChatGPT enabled learners to improve clarity, accuracy, rhythm, and segmental timing to a greater extent than previously. AI-based conversations allowed students to practice in a more self-directed and flexible way and encouraged their autonomy and motivation. It helped alleviate challenges often seen in vocational contexts, such as large class sizes with little individualized feedback. The evidence validated statistically significant post-test improvements in all aspects of pronunciation in the ChatGPT group. The use of an independent samples *t*-test and ANOVA confirmed that AI-based instruction had a relative advantage, with GB overtaking GA on every aspect of the metrics of the research. This research indicates the possibilities of using intelligent tutoring systems to support instructor-led pronunciation pedagogy in lower-resourced environments.

5. Conclusions

Research examined the effects of traditional pronunciation instruction and ChatGPT-assisted instruction at vocational universities. Data indicated that both types of instruction produced meaningful, measurable, and effective pronunciation improvements, but the ChatGPT-aided instruction produced better change across all measurable variables than the traditional practice methods. The AI-aided instruction produced significant gains in pronunciation clarity, phonemic accuracy, rhythm, intonation, and segmentation accuracy. Using independent samples *t*-tests, paired samples *t*-tests, and ANOVA, the data indicated that the findings were statistically significant. The real-time, personalized feedback provided by ChatGPT increased improvements in pronunciation speed, learner ownership, and interest in self-directed learning. It was observed that the instructional model was well-suited to this population of non-English majors within

vocational education, because the in-class constraints of having primarily group-based feedback and little individualized attention often limited progression in rehearsal-based pronunciation practices. The research supports AI as a supplementary method of teaching as opposed to a complete substitute for teachers. It has ultimately provided a replicable framework for the usage of AI in PT, which can be especially beneficial for teaching pronunciation to nontraditional educators in low-resourced contexts. The ChatGPT-assisted group showed an average improvement of 32% in clarity and 28% in accuracy compared to the traditional group, which showed improvements of 17% and 14%, respectively. These statistically significant differences indicate the enhanced effectiveness of AI-assisted instruction compared to traditional instruction. Overall, the research provides important perspectives on the changing nature of digital language learning and confirms its viability to be repurposed by ChatGPT as a means to teach pronunciation and improve communicative competence in vocational learners. Although the research utilized the rubric in a systematic manner to assess pronunciation, there was no practical method assessing the effects of pronunciation such as communicative success while using real tasks. Hence, future research should include authentic speaking activities and listener judgments to appreciate the outcomes. The lack of follow-up testing is another drawback, which makes it impossible to determine how long pronunciation improvements will last. Delayed post-tests should be used in future studies to ascertain whether gains last past the first results. Future iterations can utilize these findings to explore retention, learner perceptions, incorporation, and extension into other curricula.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Nguyen, L.T., Hung, B.P., Duong, U.T.T., et al., 2021. Teachers' and learners' beliefs about pronunciation instruction in tertiary English as a foreign language education. Frontiers in Psychology. 12, 739842. DOI: https://doi.org/10.3389/fpsyg.2021.739842

[2] Xue, Y., 2025. Application of an entertainment interactive robot based on speech recognition in English artificial intelligence teaching evaluation, and automatic feedback. Entertainment Computing. 52, 100807. DOI: https://doi.org/10.1016/j.entcom.2024.100807

[3] Huang, F., Zou, B., 2024. English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI, and innovativeness. Computers in Human Behavior. 159, 108355. DOI: https://doi.org/10.1016/j.chb.2024.108355

[4] Yang, Y.F., Tseng, C.C., Lai, S.C., 2024. Enhancing teachers' self-efficacy beliefs in AI-based technology integration into English-speaking teaching through a professional development program. Teaching and Teacher Education. 144, 104582. DOI: https://doi.org/10.1016/j.tate.2024.104582

[5] Fu, M., Guan, X., Wang, Y., et al., 2025. Application of a speech recognition algorithm based on an interactive artificial intelligence system in English video teaching system. Entertainment Computing. 52, 100859. DOI: https://doi.org/10.1016/j.entcom.2024.100859

[6] Fathi, J., Rahimi, M., Derakhshan, A., 2024. Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. System. 121, 103254. DOI: https://doi.org/10.1016/j.system.2024.103254

[7] Zhang, C., Meng, Y., Ma, X., 2024. Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. System. 121, 103259. DOI: https://doi.org/10.1016/j.system.2024.103259

[8] Yuan, L., Liu, X., 2025. The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. Computers in Human Behavior. 162, 108474. DOI: https://doi.org/10.1016/j.chb.2024.108474

[9] Shaikh, S., Yayilgan, S.Y., Klimova, B., et al., 2023.

Assessing the usability of ChatGPT for formal English language learning. European Journal of Investigation in Health, Psychology and Education. 13(9), 1937–1960. DOI: https://doi.org/10.3390/ejihpe13090140

[10] Suryanarayana, K.S., Kandi, V.P., Pavani, G., et al., 2024. Artificial intelligence enhanced digital learning for the sustainability of education management system. The Journal of High Technology Management Research. 35(2), 100495. DOI: https://doi.org/10.1016/j.hitech.2024.100495

[11] Dakakni, D., Safa, N., 2023. Artificial intelligence in the L2 classroom: Implications and challenges on ethics and equity in higher education: A 21st century Pandora's box. Computers and Education: Artificial Intelligence. 5, 100179. DOI: https://doi.org/10.1016/j.caeai.2023.100179

[12] Annamalai, N., Bervell, B., Mireku, D.O., et al., 2025. Artificial intelligence in higher education: Modelling students' motivation for continuous use of ChatGPT based on a modified self-determination theory. Computers and Education: Artificial Intelligence. 8, 100346. DOI: https://doi.org/10.1016/j.caeai.2024.100346

[13] Wang, Y., Derakhshan, A., Ghiasvand, F., 2025. EFL teachers' generative artificial intelligence (GenAI) literacy: A scale development and validation study. System. 103791. DOI: https://doi.org/10.1016/j.system.2025.103791

[14] Jenkins, J., 2000. The Phonology of English as an International Language. Oxford University Press: Oxford, UK.

[15] Levis, J.M., 2005. Changing contexts and shifting paradigms in pronunciation teaching. TESOL Quarterly. 39(3), 369–377. DOI: https://doi.org/10.2307/3588485

[16] Munro, M.J., Derwing, T.M., 2006. The functional load principle in ESL pronunciation instruction: An exploratory study. System. 34(4), 520–531. DOI: https://doi.org/10.1016/j.system.2006.09.004

[17] Sewell, A., 2021. Functional load and the teaching-learning relationship in L2 pronunciation. Frontiers in Communication. 6, 627378. DOI: https://doi.org/10.3389/fcomm.2021.627378

[18] Suzukida, Y., Saito, K., 2021. Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. Language Teaching Research. 25(3), 431–450. DOI: https://doi.org/10.1177/1362168819858246

[19] Du, J., Alm, A., 2024. The impact of ChatGPT on English for academic purposes (EAP) students' language learning experience: A self-determination theory perspective. Education Sciences. 14(7), 726. DOI: https://doi.org/10.3390/educsci14070726

[20] Abdelhalim, S.M., Alsehibany, R.A., 2025. Integrating AI-powered tools in EFL pronunciation instruction: Effects on accuracy and L2 motivation. Computer Assisted Language Learning. 1–25. DOI: https://doi.org/10.1080/09588221.2025.2534015

[21] Shafiee Rad, H., Roohani, A., 2024. Fostering L2 learners' pronunciation and motivation via affordances of artificial intelligence. Computers in the Schools. 1–22. DOI: https://doi.org/10.1080/07380569.2024.2330427

[22] Li, Y., 2022. Teaching mode of oral English in the age of artificial intelligence. Frontiers in Psychology. 13, 953482. DOI: https://doi.org/10.3389/fpsyg.2022.953482

[23] Kuddus, K., 2022. Artificial intelligence in language learning: Practices and prospects. In: Advanced Analytics and Deep Learning Models. John Wiley & Sons: Hoboken, NJ, USA. pp. 1–17. DOI: https://doi.org/10.1002/9781119792437.ch1

[24] Jiang, Y., 2022. Artificial intelligence technology for Python test simulation of oral English teaching evaluation. In Proceedings of the 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 10–12 November 2022; pp. 654–657. DOI: https://doi.org/10.1109/I-SMAC55078.2022.9987280

[25] Yu, Y., Han, L., Du, X., et al., 2022. An oral English evaluation model using artificial intelligence method. Mobile Information Systems. 2022(1), 3998886. DOI: https://doi.org/10.1155/2022/3998886

[26] Zou, B., Guan, X., Shao, Y., et al., 2023. Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. Sustainability. 15(4), 2872. DOI: https://doi.org/10.3390/su15042872

[27] Zou, B., Liviero, S., Ma, Q., et al., 2024. Exploring EFL learners' perceived promise and limitations of using an artificial intelligence speech evaluation system for speaking practice. System. 126, 103497. DOI: https://doi.org/10.1016/j.system.2024.103497

[28] Wang, S., Shi, X., 2021. Research on correction method of spoken pronunciation accuracy of AI virtual English reading. Advances in Multimedia. 2021(1), 6783205. DOI: https://doi.org/10.1155/2021/6783205

[29] Jenkins, J., 2002. A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. Applied Linguistics. 23(1), 83–103. DOI: https://doi.org/10.1093/applin/23.1.83

[30] 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. DOI: https://doi.org/10.1111/j.1467-1770.1995.tb00963.x