Al Bahith Journal for Social Sciences
Homepage: https://journals.uokerbala.edu.iq/index.php/bjh

Teaching English Pronunciation through AI-Assisted Feedback: A Quasi-Experimental Study Integrating Automated and Human Evaluation within a Structured Pedagogical Framework

Mustafa Hussein Noor AL kashaa
Ministry of Education, Holy Karbala Education Directorate,
Nahr al-Alqami Intermediate School for Boys, Karbala, Iraq
Esraa Mohammed Ali Jaber
Information Technology Center, University of Kerbala, Karbala, Iraq

Abstract in English

This quasi-experimental study investigates the pedagogical effectiveness of AI-assisted pronunciation feedback in improving EFL learners' intelligibility and segmental accuracy within a structured instructional framework. The study integrates automated speech analytics with human evaluation to ensure both technological precision and pedagogical validity.Eighty-four intermediate Arabic-speaking university students participated in an eight-week intervention. Participants were assigned to either an experimental group receiving AI-supported real-time diagnostic feedback integrated into guided classroom instruction, or a control group receiving conventional teacher-led drills with delayed corrective feedback. Pretest and posttest measures included intelligibility ratings provided by trained evaluators and automated segmental accuracy scores validated against expert phonetic assessment.

Analysis of covariance revealed statistically significant improvements in the AI-assisted group for both intelligibility ($\eta^2 = .11$, $p < .01$) and segmental accuracy ($\eta^2 = .14$, $p < .001$). Engagement analytics further indicated a positive relationship between sustained practice cycles and pronunciation gains. While findings suggest that AI-Assisted feedback can enhance individualized and scalable pronunciation training when pedagogically mediated, the results are context-bound to intermediate Arabic-speaking learners and short-term intervention conditions. The study highlights the importance of structured teacher mediation, critical evaluation of AI feedback systems, and cautious generalization of outcomes.

Paper Info

Keywords

artificial intelligence, pronunciation instruction, computer-assisted language learning (CALL), intelligibility, segmental accuracy, deliberate practice, quasi-experiment

doi: https://doi.org/10.63797/bjh.

Introduction

Teaching pronunciation remains one of the most challenging aspects of second language acquisition (SLA). Although pronunciation constitutes a central component of communicative competence, it is frequently marginalized in English as a Foreign Language (EFL) classrooms in favor of grammar and vocabulary instruction (Isaacs & Trofimovich, 2020; Saito, 2021). Contemporary pronunciation pedagogy increasingly emphasizes intelligibility rather than native-like accent attainment, given its direct relationship to listener comprehension and communicative success (Derwing & Munro, 2019; Munro & Derwing, 2020). Research in second language speech also emphasizes that pronunciation plays a crucial role in shaping listeners' perceptions of comprehensibility and communicative effectiveness in spoken interaction (Isaacs & Trofimovich, 2020).

Effective pronunciation instruction requires timely, individualized corrective feedback. Research on oral corrective feedback highlights the importance of immediacy, specificity, and repetition in facilitating phonological restructuring (Li, 2022; Lyster et al., 2019). However, in many institutional contexts characterized by large class sizes and limited instructional time, teachers face structural constraints that restrict opportunities for sustained, individualized phonological guidance.

Recent advances in Computer-Assisted Pronunciation Training (CAPT) and AI-driven feedback systems have sought to address these limitations. Modern systems employ automatic speech recognition (ASR) and acoustic modeling techniques to generate real-time diagnostic feedback on learner speech (Kartushina et al., 2022; Liakin et al., 2023). Such systems are capable of detecting segmental deviations and, in some cases, identifying stress and rhythm irregularities (Wang & Xu, 2020). The provision of immediate, item-level feedback aligns with principles of deliberate practice and noticing in SLA, which emphasize repeated, focused engagement with target forms (Li, 2022).

Nevertheless, the pedagogical integration of AI-Assisted feedback remains insufficiently theorized. While several studies report measurable improvements in phoneme production accuracy (McCrocklin, 2019; Neri et al., 2022), concerns persist regarding the interpretability of automated scoring, potential algorithmic bias, and limited contextual sensitivity to communicative intent (Evers & Chen, 2023; Isaacs et al., 2021). Moreover, many studies rely primarily on automated metrics without parallel human validation (Saito et al., 2022).

In light of these considerations, the present study investigates the effectiveness of AI-Assisted pronunciation feedback within a structured pedagogical framework that integrates teacher mediation, automated analytics, and human perceptual evaluation. The study compares AI-supported real-time feedback with conventional delayed teacher feedback in improving intelligibility and segmental accuracy among intermediate Arabic-speaking EFL learners.

In addition to examining the effectiveness of AI-Assisted pronunciation feedback, the study contributes to the literature by integrating automated speech analytics with human perceptual evaluation within a structured

pedagogical framework.

1.1 Conceptual Pedagogical Integration Framework (New Section)

To operationalize the integration of AI-Assisted feedback within classroom instruction, the study adopts a four-stage pedagogical model:

1. Targeted phonological focus aligned with curricular objectives.
2. Guided AI-supported diagnostic practice.
3. Teacher-mediated consolidation of recurrent errors.
4. Reflective performance monitoring.

This framework conceptualizes AI as a structured practice amplifier rather than an autonomous instructional agent, ensuring pedagogical coherence and preventing overreliance on automated scoring.

2. Literature Review

2.1 Pronunciation Instruction and Corrective Feedback

Pronunciation instruction has undergone a notable shift in recent decades from an emphasis on native-like accent attainment toward an intelligibility-oriented approach that prioritizes effective communication. Within this perspective, the primary goal of pronunciation teaching is not to eliminate a learner's accent but to ensure that speech is understandable to listeners in communicative contexts (Levis, 2018; Munro & Derwing, 2020). Research in second language acquisition consistently highlights the importance of pronunciation as a key component of communicative competence, influencing listener comprehension, speaker confidence, and overall interactional effectiveness (Derwing & Munro, 2019).

Corrective feedback plays a central role in the development of pronunciation skills. Studies on second language speech learning suggest that learners benefit from feedback that is timely, explicit, and targeted to specific phonological features (Li, 2022; Lyster et al., 2019). Immediate feedback, in particular, has been shown to facilitate phonological restructuring by enabling learners to notice discrepancies between their production and the target pronunciation. However, providing individualized pronunciation feedback in traditional classroom settings is often challenging due to time constraints, large class sizes, and the need for teachers to divide attention among multiple learners.

These limitations have prompted increasing interest in technological solutions that can supplement classroom instruction by providing individualized pronunciation feedback outside the constraints of teacher-led interaction.

2.2 Computer-Assisted Pronunciation Training (CAPT)

Computer-Assisted Pronunciation Training (CAPT) has emerged as a promising approach for supporting pronunciation learning through technology-enhanced practice. Early CAPT systems relied primarily on rule-based detection mechanisms that evaluated learner pronunciation using predefined phonetic parameters. More recent systems employ automatic speech recognition (ASR) and machine learning algorithms to analyze learner speech and generate diagnostic feedback (Kartushina et al., 2022;

Neri et al., 2022).
Research on CAPT systems indicates that technology-mediated pronunciation practice can increase opportunities for repetition, promote learner autonomy, and support individualized learning pathways (McCrocklin, 2019; Xie et al., 2023). These systems allow learners to practice pronunciation tasks multiple times while receiving immediate feedback on specific phonological features.
Despite these advantages, the effectiveness of CAPT systems has produced mixed results across different learning contexts. Some studies report improvements in phoneme production accuracy, while others suggest that gains measured through automated acoustic analysis do not always correspond to improvements in listener-based comprehensibility (Isaacs et al., 2021; Saito et al., 2022). These findings highlight the importance of combining automated feedback with human evaluation in order to ensure that improvements detected by technological systems reflect meaningful improvements in communicative pronunciation.

2.3 AI-Assisted Pronunciation Feedback and Pedagogical Validity

Recent advances in artificial intelligence have significantly expanded the capabilities of pronunciation training systems used in language learning environments. AI-driven platforms now integrate speech recognition technologies, acoustic modeling, and data-driven feedback mechanisms to provide learners with detailed diagnostic information about their speech production (Liakin et al., 2023; Zhang & Liu, 2024).
One of the key advantages of AI-Assisted assisted pronunciation training lies in its ability to provide immediate and individualized feedback. This feedback may include phoneme-level accuracy scores, visual representations of speech patterns, and automated prompts encouraging learners to repeat mispronounced segments. From a theoretical perspective, such feedback mechanisms align with principles of deliberate practice, which emphasize repeated performance, immediate feedback, and gradual refinement of skills (Li, 2022).
However, the pedagogical validity of AI-Assisted feedback remains an area of ongoing discussion. While AI systems can efficiently analyze acoustic features of speech, they may not fully capture contextual aspects of communication, including pragmatic intent, discourse-level meaning, or affective dimensions of spoken interaction (Evers & Chen, 2023). In addition, concerns have been raised regarding the potential for algorithmic bias or misrecognition when automated systems process speech produced by learners with diverse linguistic backgrounds (Lee et al., 2024).
For these reasons, several scholars recommend hybrid approaches that combine automated speech analysis with human perceptual evaluation in order to ensure that pronunciation improvements are both technically measurable and communicatively meaningful (Saito et al., 2022).

2.4 Intelligibility and Segmental Accuracy in Pronunciation Assessment

Within contemporary pronunciation research, intelligibility is widely recognized as a central construct in evaluating pronunciation performance.

Intelligibility refers to the degree to which a speaker's speech can be understood by a listener and is considered a more realistic and pedagogically meaningful objective than native-like pronunciation (Levis, 2018; Munro & Derwing, 2020).

Segmental accuracy, which refers to the accurate production of individual consonant and vowel sounds, also plays a critical role in pronunciation development. Certain phonological contrasts carry a high functional load, meaning that errors in these sounds may significantly affect listener comprehension (Saito et al., 2022). As a result, pronunciation training often focuses on improving the production of these high-impact phonological contrasts.

However, pronunciation competence extends beyond segmental production. Suprasegmental features such as stress patterns, rhythm, and prosodic fluency also contribute to comprehensibility and naturalness in spoken language (Wang & Xu, 2020). Because automated systems often prioritize measurable acoustic features, research designs must clearly define which dimensions of pronunciation are being assessed and acknowledge those that fall beyond the scope of the measurement tools used.

2.5 Research Gaps and Rationale for the Present Study

Although previous research has demonstrated the potential of AI-Assisted pronunciation training, several limitations remain in the existing literature. Many studies involve relatively short intervention periods, homogeneous learner samples, or exclusive reliance on automated scoring metrics without parallel human evaluation (Isaacs et al., 2021; Saito, 2021). These limitations highlight the need for more comprehensive research designs that integrate both technological assessment tools and human perceptual evaluation.

In addition, the pedagogical integration of AI-Assisted feedback within classroom instruction is often insufficiently specified. While technological capabilities are frequently discussed, fewer studies provide explicit instructional frameworks explaining how AI tools can be integrated into teacher-mediated learning environments.

The present study addresses these gaps by examining the effectiveness of AI-Assisted pronunciation feedback within a structured instructional framework that combines automated speech analysis, human perceptual evaluation, and guided classroom practice. By comparing AI-supported pronunciation training with conventional teacher-led instruction, the study seeks to provide empirical evidence regarding the potential role of AI-Assisted feedback in supporting pronunciation development among EFL learners.

3. Methodology

This section describes the research design, participant characteristics, instructional procedures, data collection instruments, and statistical analysis methods used in the study. The methodological framework was designed to evaluate the effectiveness of AI-Assisted pronunciation feedback compared with conventional teacher-led pronunciation instruction.

3.1 Research Design

The study employed a quasi-experimental pretest–posttest design involving two instructional conditions: an experimental group receiving AI-Assisted pronunciation training and a control group receiving conventional teacher-led pronunciation instruction. This design allowed for the comparison of learning outcomes while controlling for baseline differences in pronunciation performance.

Although participants were assigned to intact classroom groups rather than through full randomization, both groups received identical instructional content, contact hours, and learning objectives. Pretest scores were used as covariates during statistical analysis in order to control for initial differences between the groups.

The intervention lasted eight weeks and was integrated into the regular instructional schedule of the participants' English courses.

3.2 Participants

The participants consisted of 84 undergraduate students enrolled in intermediate-level English courses at a public university in Iraq. All participants were native speakers of Arabic and had completed at least six years of formal English instruction prior to participating in the study. Participants were divided into two groups:

- Experimental group (n = 42) – received AI-Assisted pronunciation training.
- Control group (n = 42) – received traditional teacher-led pronunciation instruction.

Placement test results confirmed that there were no statistically significant differences between the two groups prior to the intervention ($p > .10$). The participants ranged in age from 18 to 23 years and were enrolled in similar academic programs.

Participants were selected through convenience sampling from intact English language classes available during the semester.

While the linguistic homogeneity of the sample helped control for cross-linguistic variation in pronunciation learning, it may also limit the generalizability of the findings to other learner populations.

3.3 Instructional Procedures

Both groups participated in an eight-week instructional program consisting of two 90-minute sessions per week.

During the intervention period, both groups practiced identical pronunciation targets, including selected English consonant and vowel contrasts that commonly present difficulties for Arabic-speaking learners.

The control group practiced pronunciation using conventional classroom activities such as teacher modeling, repetition drills, and delayed corrective feedback during class discussion.

The experimental group participated in one weekly session of AI-Assisted pronunciation training. During these sessions, learners practiced

pronunciation tasks using an AI-based speech analysis system that provided real-time diagnostic feedback on pronunciation accuracy. The instructor supervised the AI-Assisted sessions and helped learners interpret the feedback results when necessary. The AI system therefore functioned as a supplementary practice tool within teacher-mediated instruction rather than as a replacement for classroom teaching.

3.4 AI Tool Description

The AI-Assisted pronunciation training system used in this study was based on automatic speech recognition (ASR) technology combined with phoneme-level acoustic analysis. The system analyzed learner speech input and generated immediate feedback on pronunciation accuracy.
The feedback provided by the system included several components:

- phoneme-level pronunciation scores
- visual waveform representations of speech output
- color-coded indicators highlighting pronunciation accuracy
- repetition prompts encouraging learners to reattempt mispronounced segments

Pronunciation scores were generated on a 0–100 scale based on acoustic similarity between learner speech and reference pronunciation models. To examine the reliability of automated scoring, segmental accuracy scores generated by the AI system were compared with expert phonetic ratings. The correlation between automated scores and expert ratings was r = .79, indicating substantial agreement between the two evaluation methods. Although the system was capable of detecting certain suprasegmental patterns, the present study focused primarily on segmental pronunciation features and intelligibility outcomes.

3.5 Deliberate Practice Protocol

To encourage sustained engagement with pronunciation practice, learners in the experimental group followed a structured repetition protocol.
For each pronunciation task, learners were required to either:

- achieve a minimum pronunciation score of 80/100, or
- complete three practice attempts before moving to the next task.

This procedure was designed to encourage iterative practice and reduce the likelihood that learners would move through the training tasks without attempting to refine their pronunciation.

3.6 Data Collection Instruments

Three sources of data were used to evaluate pronunciation development during the study.

Intelligibility Assessment

Participants completed read-aloud passages and picture-based oral production tasks during both the pretest and posttest stages. Six trained

evaluators independently rated intelligibility using a nine-point scale. Inter-rater reliability analysis produced an intraclass correlation coefficient of ICC = .86, indicating strong agreement among raters.

Segmental Accuracy Scores

Segmental pronunciation accuracy was measured using automated scores generated by the AI system. These scores were validated through comparison with expert phonetic judgments as described in Section 3.4.

Learner Engagement Metrics

The AI system recorded engagement data including:

- total practice duration
- number of repetition attempts
- frequency of same-session self-corrections

These metrics were used to explore possible relationships between learner engagement and pronunciation improvement.

3.7 Statistical Analysis

Data analysis was conducted using Analysis of Covariance (ANCOVA). Posttest scores for intelligibility and segmental accuracy were used as dependent variables, while pretest scores were entered as covariates in order to control for baseline differences between the groups.

Effect sizes were calculated using partial eta squared (η^2), allowing for the interpretation of the magnitude of instructional effects in addition to statistical significance.

Descriptive statistics were calculated to summarize group performance, and correlation analysis was used to examine relationships between learner engagement variables and pronunciation gains.

3.8 Ethical Considerations

All participants were informed about the purpose of the study prior to data collection. Participation was voluntary, and students were allowed to withdraw from the study at any time without academic penalty.

Audio recordings collected during the study were anonymized and stored securely in accordance with institutional ethical guidelines. The research procedures complied with standard ethical practices for educational research involving human participants.

4. Results and Analysis

This section presents the statistical findings of the study, comparing the outcomes of the AI-Assisted pronunciation training with those of conventional teacher-led instruction. The analysis includes descriptive statistics, inferential statistical testing using ANCOVA, and an examination of learner engagement patterns recorded through the AI system.

4.1 Descriptive Statistics

Both groups demonstrated improvement from pretest to posttest; however, the magnitude of improvement differed between instructional conditions. Descriptive statistics for intelligibility and segmental accuracy are presented in Table 1.

The results indicate that learners in the AI-Assisted group achieved greater mean gains in both intelligibility and segmental accuracy compared to the control group. While both instructional approaches resulted in improvement, the AI-supported condition produced more pronounced gains over the eight-week intervention period.

Table 1
Descriptive Statistics for Intelligibility and Segmental Accuracy (Pretest–Posttest)

Measure	Group	Pretest Mean (SD)	Posttest Mean (SD)	Mean Gain
Intelligibility (1–9 scale)	AI Feedback	4.86 (0.77)	6.31 (0.83)	+1.45
	Control	4.83 (0.81)	5.59 (0.88)	+0.76
Segmental Accuracy (0–100)	AI Feedback	61.8 (7.5)	73.9 (8.1)	+12.1
	Control	62.1 (7.2)	68.5 (7.7)	+6.4

As shown in Table 1, the AI-Assisted group demonstrated nearly double the improvement in intelligibility compared to the control group. A similar pattern was observed in segmental accuracy, where the experimental group achieved a mean gain of 12.1 points compared to 6.4 points in the control condition.
These descriptive patterns suggest that real-time diagnostic feedback may have provided learners with additional opportunities for iterative phonological refinement during practice.

4.2 Inferential Analysis

To determine whether the observed differences between groups were statistically significant, an Analysis of Covariance (ANCOVA) was conducted. Posttest scores for intelligibility and segmental accuracy served as dependent variables, while pretest scores were entered as covariates in order to control for baseline differences.
The ANCOVA results revealed statistically significant differences between the two instructional conditions.
For intelligibility, the analysis produced the following result:
$F(1,81) = 10.13$, $p = .002$, $\eta^2 = .11$
This effect size indicates that approximately 11% of the variance in posttest intelligibility scores can be attributed to the instructional condition after controlling for pretest performance.
For segmental accuracy, the ANCOVA yielded the following result:
$F(1,81) = 13.07$, $p < .001$, $\eta^2 = .14$
This effect size suggests that approximately 14% of the variance in posttest segmental accuracy scores is associated with the type of pronunciation

feedback provided during instruction.
These values indicate moderate instructional effects, suggesting that the AI-Assisted intervention accounted for a meaningful proportion of variance in pronunciation improvement.
These results indicate that the AI-Assisted training condition produced statistically significant improvements in both intelligibility and segmental accuracy when compared with conventional teacher-led feedback.
A visual representation of the comparative gains between the two groups is provided in Figure 1, which illustrates the differences in mean improvement across instructional conditions.

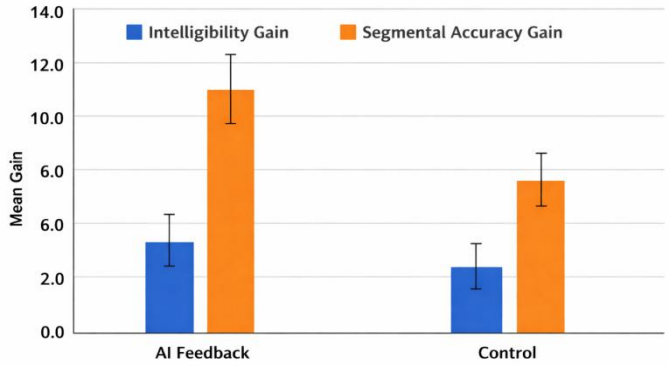

Figure 1
Comparison of Mean Pronunciation Gains between AI-Assisted and Control Groups

4.3 Learner Engagement Patterns

In addition to performance outcomes, engagement data collected from the AI platform were analyzed to explore patterns of learner interaction during the intervention period.
System logs recorded several indicators of learner engagement, including total practice time, number of repetition attempts, and frequency of same-session self-corrections. Correlation analysis indicated that higher engagement levels were associated with greater pronunciation gains. Specifically, the relationship between practice frequency and pronunciation improvement produced correlation coefficients ranging between $r = .33$ and $r = .36$ ($p < .05$). Learners who completed more than three practice cycles per week tended to demonstrate greater improvements in segmental accuracy.
These findings suggest that the effectiveness of AI-Assisted pronunciation training may depend not only on the availability of real-time feedback but also on the degree to which learners actively engage with iterative practice opportunities.

4.4 Summary of Results

Overall, the results indicate that AI-Assisted pronunciation feedback integrated within structured instructional sessions produced greater

improvements in intelligibility and segmental accuracy than conventional delayed feedback.

However, the results should be interpreted within the specific scope of the assessment measures used in this study. The improvements observed reflect targeted phonological gains measured through intelligibility ratings and segmental accuracy scores rather than broader dimensions of pronunciation performance such as prosodic fluency or communicative rhythm.

5. Discussion

The present study examined the effectiveness of AI-Assisted pronunciation feedback when integrated within a structured instructional framework. The findings indicate that learners who engaged in AI-supported real-time diagnostic practice demonstrated greater improvements in intelligibility and segmental accuracy than those who received conventional teacher-led pronunciation instruction. These results suggest that immediate feedback and repeated practice opportunities provided through AI systems may facilitate more focused phonological development during instructional activities.

The observed improvement can be interpreted in light of theories of deliberate practice and corrective feedback in second language acquisition. Deliberate practice emphasizes repeated engagement with specific performance targets combined with immediate feedback and opportunities for refinement (Li, 2022). In the AI-Assisted condition, learners were able to repeat pronunciation tasks multiple times while receiving instant diagnostic feedback on mispronounced phonemes. This iterative practice cycle may have enabled learners to identify discrepancies between their own speech production and the target forms more efficiently.

The findings also align with research suggesting that immediate corrective feedback plays an important role in phonological learning (Lyster et al., 2019). Unlike traditional classroom feedback, which is often delayed due to time constraints and the need to address multiple learners simultaneously, AI-Assisted systems provide individualized feedback in real time. This immediacy may allow learners to adjust articulatory patterns while the target pronunciation remains cognitively salient.

In addition, the positive relationship observed between learner engagement and pronunciation improvement suggests that the effectiveness of AI-Assisted pronunciation training is closely related to the degree of learner interaction with the system. Learners who engaged more frequently in repetition cycles and self-correction attempts demonstrated greater gains in segmental accuracy. Similar patterns have been reported in studies of computer-assisted pronunciation training, where increased practice frequency is associated with improved phonological outcomes (Zhao & Chen, 2021).

However, the results should be interpreted with caution. The improvements observed in this study were limited to intelligibility and segmental accuracy measured during structured tasks. Pronunciation competence involves additional dimensions, including prosodic fluency, speech rhythm, and broader communicative effectiveness (Munro & Derwing, 2020). Since these aspects were not directly measured in the present study, the results

should not be interpreted as evidence of comprehensive pronunciation development.

Another important consideration concerns the pedagogical role of teachers in AI-supported learning environments. The findings of this study do not suggest that AI systems replace teacher instruction. Instead, the results support a supplementary instructional model in which AI-Assisted feedback extends opportunities for individualized pronunciation practice while teachers continue to provide contextual explanation, communicative guidance, and motivational support. Teacher mediation remains essential in interpreting feedback, prioritizing learning targets, and integrating pronunciation practice within broader communicative objectives.

Furthermore, although automated scoring demonstrated substantial correspondence with expert phonetic ratings in this study, AI-based pronunciation evaluation systems remain constrained by algorithmic design. Issues such as misrecognition of accented speech and limited contextual interpretation have been discussed in the literature (Evers & Chen, 2023). For this reason, hybrid approaches that combine automated analytics with human perceptual evaluation may provide a more balanced framework for pronunciation assessment (Saito et al., 2022).

Overall, the results of this study contribute to a growing body of research indicating that AI-Assisted pronunciation tools can support targeted phonological practice when integrated within pedagogically guided instructional contexts. However, their effectiveness appears to depend on structured implementation, active learner engagement, and continued teacher involvement rather than technological availability alone.

6. Conclusion

This study examined the instructional effectiveness of AI-Assisted pronunciation feedback within a structured learning environment for EFL learners. By comparing AI-supported real-time diagnostic feedback with conventional teacher-led pronunciation instruction, the study aimed to evaluate whether technology-mediated practice could contribute to measurable improvements in pronunciation performance.

The findings indicate that learners who engaged in AI-Assisted pronunciation training demonstrated greater improvements in intelligibility and segmental accuracy than those who received traditional delayed feedback during classroom instruction. These results suggest that real-time diagnostic feedback combined with repeated practice opportunities may facilitate more focused phonological development during pronunciation training.

One possible explanation for this improvement lies in the immediacy and consistency of automated feedback. AI-based systems allow learners to receive individualized corrective feedback instantly, enabling multiple attempts at pronunciation refinement during practice sessions. This repeated interaction with targeted phonological features may support learners in identifying and correcting articulation errors more effectively than traditional classroom feedback alone.

However, the results of the present study should be interpreted within the specific context in which the research was conducted. The findings reflect

pronunciation development among intermediate Arabic-speaking learners within an eight-week instructional period and therefore cannot be generalized automatically to learners with different linguistic backgrounds, proficiency levels, or instructional settings.

In addition, the improvements identified in this study relate specifically to intelligibility and segmental accuracy. Pronunciation competence includes additional dimensions such as prosodic fluency, speech rhythm, and broader communicative effectiveness, which were not directly measured in the present investigation.

Importantly, the results do not suggest that AI systems replace the pedagogical role of teachers. Instead, the findings support an instructional model in which AI-Assisted feedback functions as a supplementary learning tool that expands opportunities for individualized practice while teacher mediation continues to guide instruction, contextualize feedback, and support learner motivation.

Overall, the study provides empirical evidence that AI-Assisted pronunciation feedback can contribute to targeted pronunciation improvement when integrated within structured instructional practice. These findings highlight the potential value of combining technological feedback systems with pedagogically guided instruction in order to enhance pronunciation training in EFL contexts.

7. Pedagogical Implications

The findings of this study offer several pedagogical considerations for integrating AI-Assisted pronunciation tools into English language instruction.

First, AI-Assisted feedback systems may be most effective when implemented within a blended instructional model that combines teacher guidance with technology-supported practice. Rather than replacing classroom instruction, AI tools can extend opportunities for individualized pronunciation practice beyond the limited time available in traditional classroom settings. By providing immediate feedback on pronunciation attempts, these systems allow learners to engage in repeated practice cycles that may support the refinement of specific phonological targets.

Second, pronunciation instruction should prioritize intelligibility-oriented objectives. Research in pronunciation pedagogy has emphasized that intelligibility plays a central role in effective communication and should therefore be a primary goal of pronunciation instruction (Levis, 2018; Munro & Derwing, 2020). AI-Assisted systems can support this objective by enabling learners to practice high-functional-load phonological contrasts that have a significant impact on listener comprehension.

Third, learners may benefit from explicit training on how to interpret automated pronunciation feedback. AI systems typically present numerical scores, visual indicators, or diagnostic cues that learners must interpret in order to adjust their pronunciation. Without appropriate guidance, learners may focus excessively on numerical performance scores rather than on meaningful phonological improvement. Teacher support can therefore help learners understand feedback results and apply them constructively during pronunciation practice.

Fourth, teacher mediation remains an essential component of effective pronunciation instruction. While AI systems can provide detailed diagnostic feedback, teachers continue to play a crucial role in contextualizing pronunciation practice within communicative activities, prioritizing instructional targets, and supporting learner motivation. In this sense, AI tools should be viewed as instructional supplements that enhance practice opportunities rather than as substitutes for pedagogical expertise.
Finally, practical implementation considerations should also be taken into account. Successful integration of AI-Assisted pronunciation tools requires access to appropriate technological infrastructure, including reliable internet connectivity, compatible devices, and institutional support for digital learning resources. In addition, teacher training may be necessary to ensure that instructors can effectively interpret AI-generated analytics and integrate them into instructional practice.

8. Limitations and Future Research

Although the present study produced statistically significant results, several limitations should be acknowledged.
First, the study employed a quasi-experimental design using intact classroom groups rather than fully randomized participant assignment. While statistical controls were applied to account for baseline differences between groups, this design may limit the extent to which causal conclusions can be generalized.
Second, the participant sample consisted exclusively of intermediate Arabic-speaking university learners from a single educational context. Although this sample allowed for controlled investigation within a specific instructional setting, it limits the generalizability of the findings to learners from different linguistic backgrounds, proficiency levels, or educational environments. Future research should therefore examine AI-Assisted pronunciation training across more diverse learner populations.
Third, the duration of the instructional intervention was limited to eight weeks. While the results indicate measurable improvements during this period, longer-term studies are needed to determine whether pronunciation gains achieved through AI-Assisted practice are retained over time and transferred to spontaneous communicative contexts.
Fourth, the assessment measures used in this study focused primarily on intelligibility and segmental accuracy. Although these constructs represent important aspects of pronunciation performance, other dimensions of spoken language production, such as prosodic fluency, stress patterns, rhythm, and overall comprehensibility, were not directly evaluated. Future studies could incorporate a broader range of pronunciation measures in order to provide a more comprehensive understanding of pronunciation development.
Finally, while automated pronunciation scores were validated through comparison with expert phonetic ratings in this study, further research is needed to examine the accuracy and reliability of AI-based pronunciation assessment systems across different accents and learner populations. Continued investigation of hybrid evaluation approaches that combine automated analytics with human perceptual judgment may help strengthen

the validity of pronunciation research in technology-enhanced learning environments (Isaacs et al., 2021; Saito et al., 2022).
Addressing these limitations in future research may provide deeper insight into the conditions under which AI-Assisted pronunciation tools can most effectively support language learning.

References

Derwing, T. M., & Munro, M. J. (2019). The interface of teaching and research: What type of pronunciation instruction should L2 learners receive? Language Teaching, 52(2), 253–264. https://doi.org/10.1017/S026144481800051X

Evers, V., & Chen, F. (2023). Explainable AI for language learning: Design principles for transparent feedback. ReCALL, 35(2), 145–164. https://doi.org/10.1017/S0958344023000093

Isaacs, T., & Trofimovich, P. (2020). Second language speech: From research to practice. Annual Review of Applied Linguistics, 40, 161–187. https://doi.org/10.1017/S0267190520000051

Isaacs, T., Saito, K., & Li, Z. (2021). Validating automated measures of L2 speech: Implications for assessment. Language Testing, 38(3), 403–429. https://doi.org/10.1177/0265532220969976

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2022). The effects of feedback on second language pronunciation learning: A meta-analysis. Studies in Second Language Acquisition, 44(1), 1–28. https://doi.org/10.1017/S0272263121000104

Lee, J., Saito, K., & Kang, S. (2024). Automated feedback in L2 pronunciation: Pedagogical effectiveness and validity considerations. System, 120, 102979. https://doi.org/10.1016/j.system.2024.102979

Levis, J. (2018). Intelligibility, oral communication, and the teaching of pronunciation. Cambridge University Press.

Li, S. (2022). Corrective feedback in L2 speech: Timing, explicitness, and durability. Language Learning, 72(S1), 65–98. https://doi.org/10.1111/lang.12488

Liakin, D., Cardoso, W., & Li, P. (2023). Intelligent CALL and pronunciation training: Opportunities and challenges. Language Learning & Technology, 27(2), 1–21. https://doi.org/10.10125/73489

Lyster, R., Saito, K., & Sato, M. (2019). Oral corrective feedback in L2 classrooms. TESOL Quarterly, 53(1), 5–34. https://doi.org/10.1002/tesq.439

McCrocklin, S. (2019). ASR-based pronunciation training and learner autonomy. Journal of Second Language Pronunciation, 5(1), 98–118. https://doi.org/10.1075/jslp.18012.mcc

Munro, M. J., & Derwing, T. M. (2020). Intelligibility in research and practice. Language Teaching, 53(3), 343–356. https://doi.org/10.1017/S0261444820000103

Neri, A., Cucchiarini, C., & Strik, H. (2022). CAPT revisited: Twenty years of research. Computer Assisted Language Learning, 35(7), 1315–1347. https://doi.org/10.1080/09588221.2020.1825091

Saito, K. (2021). What characterizes explicit pronunciation instruction? A synthesis of 100+ intervention studies. TESOL Quarterly, 55(3), 866–900. https://doi.org/10.1002/tesq.3038

Saito, K., Trofimovich, P., & Isaacs, T. (2022). What predicts comprehensibility? Segmental and prosodic contributions revisited. Applied Linguistics, 43(5), 817–842. https://doi.org/10.1093/applin/amaa063

Wang, Y., & Xu, Y. (2020). Prosody-driven CAPT: Detecting and correcting lexical stress errors. Speech Communication, 121, 1–12. https://doi.org/10.1016/j.specom.2020.04.004

Xie, Q., Li, W., & Zhang, Y. (2023). Integrating ASR feedback into mobile-assisted pronunciation learning. ReCALL, 35(3), 289–308. https://doi.org/10.1017/S0958344023000214

Zhang, J., & Liu, S. (2024). From detection to explanation: Improving learner uptake with interpretable AI feedback. Computer Assisted Language Learning, 37(5), 901–924. https://doi.org/10.1080/09588221.2023.2187654

Zhao, T., & Chen, S. (2021). Learner engagement in CAPT: The role of immediate feedback. Language Learning & Technology, 25(3), 40–62. https://doi.org/10.10125/73402

Abstract in Arabic

تستقصي هذه الدراسة شبه التجريبية الفاعلية التربوية للتغذية الراجعة في النطق المدعومة بالذكاء الاصطناعي في تحسين وضوح النطق ودقة المقاطع الصوتية لدى متعلمي اللغة الإنجليزية كلغة أجنبية، وذلك ضمن إطار تعليمي منظم. وتجمع الدراسة بين التحليل الآلي للكلام والتقويم البشري من أجل ضمان كلٍّ من الدقة التقنية والمصداقية التربوية.

شارك في الدراسة أربعة وثمانون طالبًا جامعيًا من متعلمي اللغة الإنجليزية متوسطي المستوى من الناطقين بالعربية، وذلك ضمن برنامج تدخلي استمر ثمانية أسابيع. وقد وُزِّع المشاركون إلى مجموعتين: مجموعة تجريبية تلقت تغذية راجعة تشخيصية فورية مدعومة بالذكاء الاصطناعي ومتكاملة مع أنشطة تعليمية موجهة داخل الصف، ومجموعة ضابطة تلقت تدريبات نطق تقليدية يقودها المعلم مع تغذية راجعة تصحيحية مؤجلة.

وشملت أدوات القياس في الاختبارين القبلي والبعدي تقديرات لوضوح النطق قدمها مقيمون مدربون، إضافة إلى درجات آلية لدقة المقاطع الصوتية جرى التحقق من صحتها من خلال مقارنتها بتقويم صوتي خبير.
وكشفت نتائج تحليل التباين المصاحب عن تحسينات ذات دلالة إحصائية في المجموعة التي استخدمت التغذية الراجعة المدعومة بالذكاء الاصطناعي في كلٍّ من وضوح النطق ($\eta^2 = .11$، $p < .01$) ودقة المقاطع الصوتية ($\eta^2 = .14$، $p < .001$). كما أظهرت تحليلات التفاعل مع النظام وجود علاقة إيجابية بين استمرارية دورات التدريب والمكاسب المتحققة في النطق.
وعلى الرغم من أن النتائج تشير إلى أن التغذية الراجعة المدعومة بالذكاء الاصطناعي يمكن أن تعزز تدريب النطق الفردي والقابل للتوسع عند توظيفها ضمن إطار تربوي موجه، فإن هذه النتائج تبقى مرتبطة بسياق متعلمي اللغة الإنجليزية متوسطي المستوى من الناطقين بالعربية وبظروف تدخل تعليمي قصير المدى. وتؤكد الدراسة أهمية الإرشاد المنظم من قبل المعلم، والتقييم النقدي لأنظمة التغذية الراجعة المعتمدة على الذكاء الاصطناعي، وضرورة توخي الحذر عند تعميم النتائج.