



RESEARCH ARTICLE



DEVELOPING ENGLISH SPEAKING SKILLS THROUGH AI-ASSISTED PRONUNCIATION INSTRUCTION: A PILOT STUDY IN MULTILINGUAL INDIAN SECONDARY CLASSROOMS

Parul Kathuria^{1*}, Sanjay K. Jha²

^{1*}Research Scholar, Amity University, Haryana.

²Director, Amity School of Liberal Arts, Amity University, Haryana.

Email: Parulkathuria57@gmail.com ; skjha@ggn.amity.edu

DOI: <https://doi.org/10.54513/JOELL.2026.13118>

ABSTRACT



Article history:

Received on: 24-02-2026

Accepted on: 25-03-2026

Available online: 30-03-2026

The pilot study aims to explore the possibility of using AI-supported pedagogy in multilingual secondary classrooms in India to facilitate the development of English-speaking competence, such as pronunciation, conversational fluency, and speaking confidence, given large class sizes, the lack of individualized oral feedback, and learners' pronunciation anxiety.

A single group, pre-test, post-test quantitative phase (embedded mixed method design) was followed by a qualitative phase of data collection. The intervention involved one teacher-led, human-in-the-loop program deployment for 100 students in Class 9 across four intact classes in three schools, using four AI speaking tools (ELSA Speak, ChatGPT Voice Mode, Speechace, and Microsoft Speaker Coach). Classroom observations, semi-structured teacher interviews, student focus groups, reflective journals, and pre- and post-speaking assessments were used to gather data, which were analyzed using Braun and Clarke's (2006) thematic analysis and descriptive and paired-sample comparisons.

This study provides pilot evidence of intelligibility in the context of classroom-based, multilingual Indian secondary schools through a multi-tool AI speaking module, a principle that differs from native-speaker imitation, in contrast to previous studies that focused on higher education or single-tool laboratory settings.

Keywords: *AI-assisted pronunciation training, Multilingual secondary classrooms, Teacher-led AI integration, Speaking anxiety, Indian English language teaching (ELT).*

Author(s) retain the copyright of this article. Copyright© 2026 VEDA Publications.

Author(s) agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License



1. INTRODUCTION

The context of English language learning in India is highly multilingual, where learners' native language use is shaped by the context in which they learn English as an additional language. Within this context, students encounter unique challenges in developing pronunciation, sound systems, stress patterns, rhythm, and oral fluency, as well as in accessing opportunities to hear 'real' English outside the classroom. Anxiety among many students about public speaking in front of peers, unequal access to personalized feedback, and large class sizes add to these challenges. This discourages many students from using their speaking skills, even when they are grammatically correct and have good vocabulary.

With the recent advancements in Artificial Intelligence (AI), language learning and pronunciation help are now possible. Oral rehearsal with AI can also be more discreet than in a traditional classroom setting: AI can provide immediate feedback, offer multiple passes for students to practice, and provide a more comfortable environment for oral rehearsal, helping students recognize sound production issues, improve pacing and fluency, and develop proficiency without the pressure of a high-stakes performance. The pedagogical potential of such tools, however, relies not only on their technical potential but also on the practical feasibility of using them in a multilingual school context with limited resources.

Interest in using AI tools for language learning has been rising. However, the research conducted so far focuses on higher education institutions, experimental designs, or short-term trials, which may not be representative of Indian secondary classrooms. The impact on teachers and students at

the school level, the nature of teachers' roles when using AI, and the practical and implementation issues are insufficiently explored. In particular, evidence from multilingual classrooms in India is needed, in which pronunciation support should serve to make the language intelligible and to foster confidence in speaking it, without stunting the growth of local languages and identities.

Although there have been recent developments in the research on AI-assisted language learning in multilingual developing-country contexts, little attention has been paid to pronunciation instruction at the school level. Previous studies have focused on higher education settings, lab settings, or a single intervention with AI speaking models, and the feasibility of implementing a multi-tool, teacher-supervised AI speaking model has been largely overlooked. This is an issue because multilingual secondary classrooms pose pedagogical, technological, and sociolinguistic challenges that university-based studies cannot fully capture. In view of the above, the current pilot study seeks to collect context-specific evidence to inform future large-scale pronunciation teaching using AI in Indian schools.

This study uses a pilot feasibility study to assess the feasibility of implementing a multi-tool AI speaking module in a secondary school's English learning environment. The intervention was a teacher-led approach that incorporated ELSA Speak, ChatGPT Voice Mode, Speechace, and Microsoft Speaker Coach to aid pronunciation, conversation fluency, and speech confidence. The study does not aim to prove effectiveness but rather to discuss the acceptability of using AI pronunciation support in a



multilingual school setting, as well as the difficulties in implementing it in practice.

2. Research Questions and Objectives

2.1 Research Questions

- RQ1. Is it possible to implement the use of AI in pronunciation teaching in multilingual secondary English classrooms in India?
- RQ2. What is teachers' and students' acceptability of, and perceived usefulness for, the AI speaking tools?
- RQ3. What are the barriers to the implementation of the intervention in the classroom?
- RQ4. What initial changes do you notice in pronunciation, fluency, and speaking anxiety?

2.2 Objectives

- Discuss the possibility of introducing a multi-tool AI speaking module in the classroom and its suitability.
- Discuss the importance of AI for teachers and students in pronunciation practice, fluency, and speaking self-confidence.
- Identify the pedagogical, technical, and contextual challenges that may arise during implementation.
- Gather preliminary descriptive and comparative results regarding pronunciation change, fluency, and speaking anxiety.

3. Literature Review

3.1 AI and Computer-Assisted Pronunciation Training (CAPT) in Language Education

AI technology in language learning has been associated with more positive language learning environments (Belda-Medina & Calvo-Ferrer, 2022), in which pronunciation practice is no longer a "drill and practice" process but rather a flexible, AI-driven, computer-assisted language learning (ICALL) environment (Albedah, 2025). In traditional pronunciation classes, students used and learned a large number of static audio-visual materials, and the assessment focused only on a limited amount of information on the correctness of pronunciation (Vančová, 2023). Today's AI pronunciation devices, on the other hand, rely on advanced speech-recognition technology to analyze students' pronunciation errors and provide real-time feedback, enabling students to repeat, rehearse, and speak while receiving personalized guidance on their pronunciation needs.

The advent of generative AI and conversational interfaces has further extended these uses. ChatGPT Voice Mode is a technology that enables students to engage in continuous, context-specific conversations in voice, similar to natural speech (Panagiotidis, 2024; Yan et al., 2025). The research findings show that these speech situations created with AI technology can be used to enhance learner pronunciation and fluency, provide opportunities for repeated oral practice in an individualized and low-stakes environment, and contribute to increased learner confidence and motivation, especially in the context of repeated oral practice (Ningrum et al., 2026; Xodabande et al., 2025).



3.2 Multilingualism and Phonological Challenges in Indian ELT

English pronunciation development is a challenge in India because Indians are multilingual. It is well known that learners' L1 phonological and rhythmic characteristics interfere with their L2 production (Winson et al., 2023). Such a transfer can influence the way consonants are produced, the placement of stress, the rhythm, and connected speech in Northern Indian contexts (where Indo-Aryan languages like Hindi, Punjabi, and Haryanvi are spoken), as in the confusion between /v/ and /w/ and the substitution of fricative sounds and the influence of syllable-timed rhythm on stress-timed English.

The pronunciation features they possess can lead to communicative difficulties, which, over time, can cause them to hesitate to speak and be less willing to participate in verbal communication (Crowther & Loewen, 2025). The challenge is exacerbated in larger classrooms, where teachers may be unable to provide frequent and individualized pronunciation feedback, and in classrooms with many teachers who have not undergone training or who may not have enough time to focus on pronunciation to address fine-grained pronunciation errors (Ekizer, 2025). If they are not regularly exposed to authentic spoken English outside the classroom, these L1 speech patterns may become fossilized.

Multilingualism should, however, not only be considered as a restriction on learning English. With the emergence of contemporary pedagogy, knowledge of learners' L1 is being treated as a useful tool for translating, transcodifying, and comparing learners' phonological knowledge across languages (Winson et al., 2023). With increasing access to

digital technology, there is a growing opportunity to develop AI pronunciation tools that encourage linguistic diversity and context sensitivity rather than emulating native-speaker pronunciation.

3.3 The Nativeness Principle versus the Intelligibility Principle

One of the main issues in L2 pronunciation research is whether pronunciation teaching should focus on native-speaker pronunciation or on effective, intelligible communication. Until recently, the Nativeness Principle had shaped pronunciation pedagogy, with native-speaker pronunciation (e.g., Received Pronunciation or General American English) serving as the ideal pronunciation model for students to aim for, while it was recommended that they minimize traces of their L1 pronunciation (Levis, 2020).

More recent studies challenge the viability of native norms in multilingual contexts. A fully 'native-like' accent is not considered to be practical or necessary for most SLA learners, especially when they engage in communication in various language settings in the world beyond their native language (Levis, 2020; Derwing & Munro, 2015). However, undue attention to accent reduction can lead to excessive anxiety on the part of learners, lower speaking confidence, and implicit devaluation of multilingual identities.

This has encouraged a movement towards the Intelligibility Principle, a form of communication that focuses on clear, comprehensible, and successful communication rather than imitating an accent. From this perspective, pronunciation teaching serves a function in communicative effectiveness without



students losing their regional or linguistic identity, which is particularly relevant in multilingual Indian schools, where language diversity is an integral part of learners' sociocultural identity. AI pronunciation tools are therefore likely to be pedagogically better if they are focused on intelligibility and confidence in communicating, than if they are based on native-speaker norms.

3.4 Theoretical framework: Communicative competence, input, and feedback.

This research will draw on the complementary theories of second language acquisition (SLA) and communicative language teaching to examine the potential benefits of AI-based pronunciation support in multilingual ELT classrooms.

Communicative competence places a greater emphasis on the acquisition of the language than simply grammatical competence, and communicative competence is defined as the ability to participate effectively in real communication, that is, the ability to manage discourse, to be appropriate, and to pronounce correctly (Celce-Murcia et al., 2010). AI-powered speaking tools can enhance communicative competence by providing low-stakes opportunities for learners to practice pronunciation, pacing, and interactive speaking.

According to Krashen's Input Hypothesis (1982), language acquisition occurs when Input is presented meaningfully and comprehensibly, and at a level just above the learner's current level of acquisition. Multilingual classrooms can provide 24/7 access to authentic spoken English and personalized input using AI-powered speaking platforms, which may not have been available otherwise. Conversely, Swain's

(1985) Output Hypothesis highlights that language production (including the ability to self-correct and engage in active language production) enables the refinement of accuracy and fluency, and AI pronunciation tools can facilitate repeated output and immediate feedback.

The Affective Filter Hypothesis of Krashen also provides insights for the current study, suggesting that language learning may be hindered by anxiety, fear of embarrassment, and low self-confidence. In the case of affective barriers, they could be minimized by providing AI-mediated speaking environments that offer private, non-judgemental practice space, thereby boosting learners' willingness to engage in communication (Xodabande et al., 2025; Ningrum et al., 2026).

Finally, the study is based on Vygotsky's (1978) Sociocultural Theory and the concept of the Zone of Proximal Development (ZPD), focusing on guided interaction and mediation in learning. In this teacher-led approach to design, AI tools serve as a means of mediation rather than a replacement for the teacher, who still plays a key role in contextualizing feedback, facilitating tasks, and interpreting and applying AI-generated corrections.

4. Methodology

4.1 Research Design

This study employed a pilot embedded mixed-methods design (QUAL+quan), which involves a single group pre-test–post-test quantitative phase and a qualitative phase involving data collection in the classroom to test the feasibility and class acceptability of the use of AI-mediated instruction to assist English pronunciation, conversational fluency,

and speaking confidence in multilingual Indian secondary classrooms. Following the advice on pilot and feasibility research (Arain et al., 2010), the study was not designed to generate statistically generalizable claims about effectiveness based on an

appropriately sized sample. However, it was designed to test study procedures, estimate classroom acceptance, and identify real-world issues for a larger, appropriately sized follow-up study.

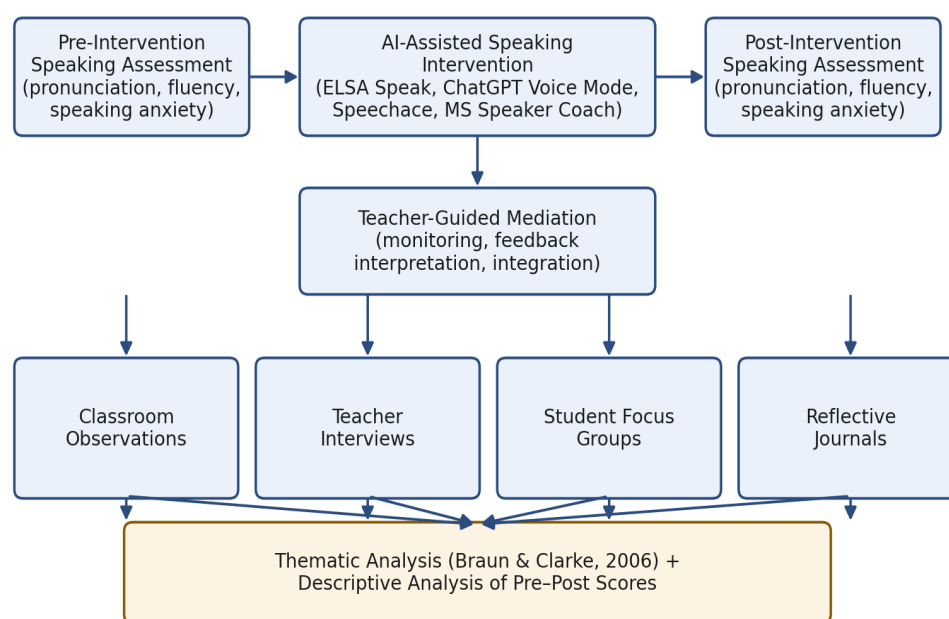


Figure 1. Procedural Flow of the Embedded Mixed-Methods Pilot Design

4.2 Setting and Participants

The study took place in the secondary school English classrooms in a multilingual educational environment in India. 39 English teachers for class 9 (228 students) were recruited from 4 intact classes across 3 secondary schools and participated in the study. All students who were engaged in the intervention had access to the basic digital tools (Internet connection and smartphones or tablets). Given the recommendations of pilot and feasibility studies, the sample was context-specific. It would be prudent to proceed with context-specific generalizations before moving on to population-level generalizations, which

would be sought in a larger, well-powered follow-up study (Arain et al., 2010).

4.3 Intervention

The intervention used a multi-tool AI speaking module (ELSA Speak, ChatGPT Voice Mode, Speechace, and Microsoft Speaker Coach). Segmental pronunciation practice was conducted using ELSA Speak; conversational speech with ChatGPT Voice Mode; formative speech feedback was given using Speechace; and public-speaking rehearsal was done with Speaker Coach from Microsoft. The teacher was the main actor throughout the intervention: guiding students, interpreting AI-generated information, and



implementing the tools as part of their daily instruction.

classroom routine, not in place of classroom

Table 1. AI Speaking Tools Used in the Intervention and Their Pedagogical Function

Tool	Primary Function	Target Speaking Sub-Skill
ELSA Speak	Segmental pronunciation drills with automated phoneme-level feedback	Pronunciation accuracy
ChatGPT Voice Mode	Open-ended, context-sensitive spoken conversation practice	Conversational fluency
Speechace	Formative, criterion-referenced speech scoring and feedback	Pronunciation and fluency
Microsoft Speaker Coach	Real-time feedback on pacing, pitch, and filler words during rehearsal	Public speaking confidence

4.4 Procedure

The intervention was delivered during the regular English program as part of a short pilot. Students engaged in a set of speaking activities facilitated by AI for a designated amount of time per day, and teachers supervised and managed the use of AI and adapted activities based on student needs. This protocol aimed to mirror the real-life classroom environment closely and to determine the tools' viability in that setting.

4.5 Data Collection

The feasibility of implementation and participants' classroom experiences in developing English-speaking skills were captured using multiple

qualitative and quantitative sources. Qualitative data consisted of non-participant classroom observations of the AI-supported speaking sessions, semi-structured teacher interviews, semi-structured student focus group discussions, and student reflective journals kept throughout the intervention. The quantitative data included a brief pre-/post speaking assessment that encompassed three outcome areas: segmental pronunciation accuracy (ELSA Speak and Speechace applications, 0-100 scale), conversational fluency (short structured speaking task, rated 1-5 scale), and self-reported speaking anxiety (before and after speaking assessment using a Likert scale, rated 1-5 scale).



Table 2. Data Collection Instruments and Their Purpose

Instrument	Data Type	Purpose / Timing
Classroom observations	Qualitative	Document real-time classroom use, engagement, and implementation issues; conducted throughout the intervention
Semi-structured teacher interviews	Qualitative	Capture teacher perceptions of feasibility, role change, and barriers; conducted after the intervention.
Student focus groups	Qualitative	Explore learner acceptability, confidence, and anxiety; conducted post-intervention.
Reflective journals	Qualitative	Track students' ongoing reflections on speaking practice; maintained across the intervention period.
Pronunciation/fluency assessment (ELSA Speak, Speechace)	Quantitative	Measure pronunciation accuracy and fluency; administered pre- and post-intervention
Speaking-anxiety questionnaire	Quantitative	Measure self-reported speaking anxiety; administered pre- and post-intervention

4.6 Data Analysis

Thematic analysis (Braun and Clarke, 2006) was used to analyze qualitative data from interviews, focus group discussions, observations, and journals to identify themes related to feasibility, acceptability, teacher role change, learner confidence, and implementation barriers. Paired-samples t-tests with Cohen's d (Cohen's d) as standardized effect size were used to analyze quantitative pre- and post-scores (N = 100), both descriptively (means, standard deviations, data ranges) and by comparing pre- and post-scores. The results from this pilot study are only quantitative. They may thus be used as an exploratory indicator to provide the direction and

magnitude of change. Still, this study is not representative of the intervention's causal effectiveness because it was not designed as a single-group study with a control condition and random assignment (Arain et al., 2010).

4.7 Trustworthiness and Rigor

Steps were taken to enhance the study's rigor. To validate the findings, data triangulation was undertaken, and reflective journals and classroom observations were conducted to provide context for interpreting the overall trends in the quantitative data. Similar to the research best practices in the pilot phase, caution has been given in reporting



findings rather than overemphasizing their certainty (Arain et al., 2010).

4.8 Ethical Considerations

Participation was voluntary. Informed consent was obtained after a briefing by all teachers and students (and parents/guardians, as appropriate) on the aims of the study, confidentiality procedures, and the right to withdraw. All data were anonymized and only used for research.

5. Results

5.1 Feasibility and Implementation

The pilot intervention was easily and reasonably adapted to the normal classroom routine. The opportunity was provided for the teachers to present the tools in guided sections, and students could perform the speaking activities to a significant degree. Implementation depended on the availability of digital resources, classroom space, and institutional support.

During the intervention, students' comfort with using AI in speaking activities was observed to increase. Teachers indicated that they did not need to correct pronunciation as often and as repeatedly in class because of fast, automatic feedback.

5.2 Participant Response

The overall feedback received from staff and pupils was positive. The teachers indicated that the tools helped them correct pronunciation more easily and concentrate more on the communicative aspects of

teaching and the higher-level aspects of feedback. The students reported that the AI practice environment was less stressful than speaking aloud in class and that it helped them feel more confident rehearsing in front of others.

5.3 Practical Barriers

The pilot revealed several implementation challenges. Stable Internet access was not available or reliable on some devices, and on some students' devices, and automatic speech recognition was not available or reliable, making delivery more challenging. The tools were also less effective at identifying regional accents, causing some learners to be frustrated and to need further teacher-mediated feedback. The findings indicate that this intervention has potential and needs specific adaptations for classroom-wide implementation.

5.4 Quantitative Outcomes: Pronunciation, Fluency, and Anxiety

The descriptive statistics for the pre- and post-intervention outcome measures (N = 100) are shown in Table 3.

**3. Descriptive Statistics for Pre- and Post-Intervention Outcomes (N = 100)**

Outcome Measure	Pre M	Pre SD	Post M	Post SD	Min–Max (Pre)	Min–Max (Post)
Pronunciation accuracy (0–100)	60.54	9.36	69.88	10.32	45–75	50–90
Conversational fluency (1–5)	2.97	0.49	3.56	0.53	2.0–3.8	2.4–4.6
Speaking anxiety (1–5)	3.91	0.56	3.16	0.65	2.9–4.8	1.8–4.4

The paired-samples comparisons (Table 4) showed statistically significant differences for all three outcomes. Pronunciation accuracy increased by an average of 9.34 points (SD = 3.92), $t(99) = 23.83$, $p < .001$, a very large effect (Cohen's $d_z = 2.38$). Conversational fluency increased by an average of 0.58 points on the 5-point scale (SD = 0.23), $t(99) = 25.64$, $p < .001$, $d_z = 2.56$. Self-reported speaking

anxiety decreased by an average of 0.76 points (SD = 0.28), $t(99) = -26.68$, $p < .001$, $d_z = -2.67$. All 100 students who participated improved in all three areas in the expected direction: 100% improvement in pronunciation, fluency, and stress level reduction, with qualitative reports of increased comfort and confidence in oral practice.

Table 4. Paired-Sample Comparisons of Pre- and Post-Intervention Outcomes (N = 100)

Outcome Measure	Mean Difference (Post–Pre)	SD of Difference	t (df = 99)	p	Cohen's d_z
Pronunciation accuracy	+9.34	3.92	23.83	< .001	2.38
Conversational fluency	+0.58	0.23	25.64	< .001	2.56
Speaking anxiety	–0.76	0.28	–26.68	< .001	–2.67

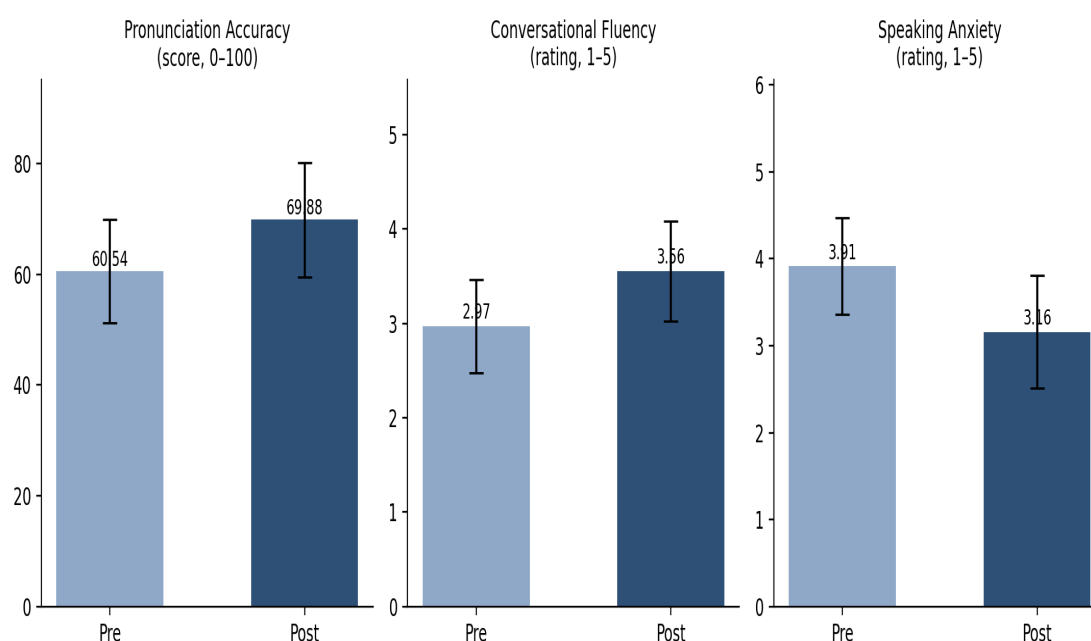


Figure 2. Pre- and Post-Intervention Means ($\pm$ SD) for Pronunciation Accuracy, Conversational Fluency, and Speaking Anxiety (N = 100)

The results presented should be taken with a grain of salt. The design is a single-group, non-randomized pilot with no control or comparison condition, so other factors that could have contributed to the gains seen in this classroom cannot be ruled out, such as practice and testing effects, familiarisation with the assessment tasks, maturation, and the general motivational effect of having students complete a novel classroom activity. The size and uniformity of the change across the 100 students, however, provide a strong initial indication that this warrants further investigation in a more rigorously designed, well-powered subsequent study.

5.5 Summary of Findings

Overall, the results indicate that AI can be used in pronunciation instruction in multilingual secondary classrooms in India to augment teacher-led oral practice and confidence-building exercises. The pilot

also revealed technical and contextual challenges, such as internet instability, the uneven distribution of devices among users, and the lower accuracy of regional accents, which will require further improvements before a wider roll-out.

6. Discussion

Based on this pilot study, it can be concluded that AI-supported pronunciation can be used in multilingual secondary classrooms and has pedagogical value. The AI tools were most useful as supports for the classroom rather than replacements for teacher-led instruction, allowing for more private practice, feedback, and rehearsal of target structures. Teachers and students felt more self-confident and less anxious, and teachers said their role had shifted from a corrector to a more facilitative one. The quantitative pre-post data also support these qualitative patterns, as all 100 students who



participated in this study demonstrated large increases in pronunciation accuracy and fluency and a decrease in speaking anxiety. These are encouraging results from a single-cohort, single-site pilot study with no control group; further research in a multi-site, controlled pilot study is required to generate more definitive conclusions about the effectiveness of the intervention.

The results are in line with the overall hypothesis that pronunciation development can be fostered through a low-pressure opportunities to listen, repeat, and make corrections. The feedback and teacher mediation in this study appeared to be effective factors in making the speaking environment more supportive and conducive to a greater willingness to communicate. In a multilingual context in which the aim of teaching is not native language but intelligibility and confidence, this pilot's interpretation holds: AI tools are best used to complement purposeful pedagogy, not replace it.

The study also pinpointed some major hurdles to broader implementation. Internet connectivity problems made it difficult to deliver smooth classes, and device availability affected their smoothness. At the same time, automatic speech recognition was sometimes inaccurate when processing regional accents, resulting in a learner not being completely confident in the feedback provided. The issues suggest that the effectiveness of AI-based pronunciation teaching depends on the software's capabilities, teachers' readiness and sensitivity to language diversity, and school infrastructure.

Overall, this pilot provides some initial proof of feasibility, acceptability, and potential in the classroom, and a better understanding of the

conditions that will help to make scaling up successful. The main finding of this study is that the use of AI in pronunciation instruction is a teacher-led, context-specific approach suitable for the Indian secondary education context and requires further research with a larger number of outcome measures and a longer intervention period.

6.1 Practical Implications

The findings indicate that artificial intelligence-based pronunciation tools can be effectively integrated into regular classroom teaching to provide more opportunities for individual practice and feedback on pronunciation and to foster students' confidence in speaking when using a teacher-led method. To ensure students are all provided with equitable access to devices and connectivity, to support teachers in developing structured training programs to use AI pronunciation tools in their classrooms, and to ensure the digital infrastructure is available, the effective implementation of AI pronunciation tools in schools will need to be supported.

7. Limitations

The following are limitations of this pilot study. The context-specific, purposively recruited sample limits the generalizability of this intervention's results to other multilingual contexts; the short time window of this intervention precludes claims regarding longer-term effects. These factors had a practical impact on implementation, including inconsistent internet access, device availability, and the accuracy of automatic speech recognition (ASR) in local accents. The statistically significant quantitative results should be viewed as descriptive and exploratory patterns of change, rather than as evidence to test the



effectiveness of the intervention, because the study was not a true experimental study with a control group and employed a single-group pre-post design.

8. CONCLUSION

The results of this exploratory pilot study suggest that AI-based pronunciation instruction is feasible and acceptable in multilingual secondary English classrooms in India. In a teacher-led HITM design, the intervention was associated with decreases in speaking anxiety and improvements in both pronunciation accuracy and conversational fluency among all 100 students who participated in the study. The results indicate that AI tools can help move towards a more facilitative teaching role and assist students in becoming more confident in oral communication, despite infrastructure concerns and challenges with accented speech recognition. Based on the empirical evidence provided in this study, the integration of technology in language instruction with contextually responsive and teacher-mediated approaches seems to hold real promise, awaiting further validation in larger, longitudinal, and controlled studies.

REFERENCES

- Albedah, F. (2025). Artificial intelligence in language education: A systematic review of multilingual applications, large language models, and emerging challenges. *Language Teaching Research Quarterly*, 49, 247–268. <https://doi.org/10.32038/ltrq.2025.49.13>
- Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Medical Research Methodology*, 10, Article 67. <https://doi.org/10.1186/1471-2288-10-67>
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), Article 8427. <https://doi.org/10.3390/app12178427>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.
- Crowther, D., & Loewen, S. (2025). Instructed second language acquisition and second language pronunciation. *Language Learning*, 75(S1), 30–54. <https://doi.org/10.1111/lang.12716>
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins. <https://doi.org/10.1075/lll.42>
- Ekizer, F. N. (2025). Exploring the impact of artificial intelligence on English language teaching: A meta-analysis. *Acta Psychologica*, 260, Article 105649. <https://doi.org/10.1016/j.actpsy.2025.105649>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Levis, J. M. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310–328. <https://doi.org/10.1075/jslp.20050.lev>
- Ningrum, A., Kembaren, F. R. W., & Dalimunte, M. (2026). The role of ChatGPT's voice conversation feature in enhancing EFL freshmen students' speaking confidence from an affective perspective. *ELTIN Journal*, 14(1), 167–180. <https://doi.org/10.22460/eltin.v14i1.p167-180>
- Panagiotidis, P. (2024). LLM-based chatbots in language learning. *European Journal of Education*, 7(1), 102–123.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Vančová, H. (2023). AI and AI-powered tools for pronunciation training. *Journal of Language and Cultural Education*, 11(3), 12–24. <https://doi.org/10.2478/jolace-2023-0022>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Winston, V. R. V., Arunkumar, V., & Rao, D. P. (2023). Exploring the landscape of teaching and learning English as a second language in India: A systematic review of challenges, opportunities, and prospects for future development. *Assyfa Learning Journal*, 1(2), 104–111. <https://doi.org/10.61650/alj.v1i2.123>



Xodabande, I., Shiri, S., & Zohrabi, M. (2025). Exploring the impacts of an AI-driven instructional intervention on Iranian EFL learners' pronunciation skill development. *Discover Education*, 4, Article 307.

<https://doi.org/10.1007/s44217-025-00782-2>

Yan, W., Li, B., & Lowell, V. L. (2025). Integrating artificial intelligence and extended reality in language education: A systematic literature review (2017–2024). *Education Sciences*, 15(8), Article 1066.

<https://doi.org/10.3390/educsci15081066>