

International Journal of Learning, Teaching and Educational Research
 Vol. 25, No. 9, pp. 400-423, September 2026
<https://doi.org/10.26803/ijlter.25.9.16>
 Received Jul 3, 2026; Revised Aug 18, 2026; Accepted Aug 30, 2026

AI or not to AI in Resource-Constrained Contexts? Exploring the Impact of AI-Mediated Feedback on the EFL Speaking Proficiency of Bangladeshi Undergraduate Learners

MD. Nahid Hasan* 

University of Frontier Technology, Bangladesh
 Kaliakair, Gazipur-1750, Bangladesh

Khadija Akter Onee , **Shakila Akter** , **Nujhat Aslam Nieon** 

Bangladesh University of Professionals
 Mirpur Cantonment, Dhaka-1216, Bangladesh

Pragga Paromita Bose 

University of Barishal
 Karnakathi, Barishal-8254, Bangladesh

Zannatul Ferdushie 

University of Frontier Technology, Bangladesh
 Kaliakair, Gazipur-1750, Bangladesh

Md. Ismaile Hossain 

Global University Bangladesh
 Natullahbad, Barisal, Bangladesh

Citation :

Hasan, M. N., Onee, K. A., Akter, S., Nieon, N. A., Bose, P. P., Ferdushie, Z., & Hossain, M. I. (2026). AI or not to AI in Resource-Constrained Contexts? Exploring the Impact of AI-Mediated Feedback on the EFL Speaking Proficiency of Bangladeshi Undergraduate Learners. *International Journal of Learning, Teaching and Educational Research*, 25(9), 400-423. <https://doi.org/10.26803/ijlter.25.9.16>

Abstract. Developing English-speaking proficiency remains challenging in Bangladesh because teacher-centred instruction, large class sizes, and limited opportunities for individualised feedback constrain the speaking practice of learners. Although AI-mediated feedback has shown promise for supporting language learning, its effectiveness in resource-constrained English as a Foreign Language (EFL) context remains underexplored. This study investigated the impact of SmallTalk2Me, an AI-mediated speaking platform, on the speaking proficiency, fluency, and perceptions of Bangladeshi tertiary-level EFL learners. Using a one-group pretest-posttest pre-experimental design, 28 undergraduate students completed IELTS speaking tests before and after a three-week intervention, while speaking fluency was measured using words per minute (WPM) and the perceptions of learners were examined through a

*Corresponding author: MD. Nahid Hasan; nahid0001@uftb.ac.bd

post-intervention survey. The findings showed significant improvements in IELTS speaking scores (5.75–6.38, $p < .001$, $d = 0.93$) and speaking fluency (84.50–93.10 WPM, $p < .001$, $d = 1.15$), alongside positive perceptions of AI-mediated feedback. However, variation in learning outcomes, concerns regarding feedback reliability, and the lack of human interaction suggest that AI-mediated feedback is most effective when integrated with contextualised learning opportunities and teacher support. The study contributes empirical evidence from a resource-constrained EFL context and highlights the potential of AI as a complementary tool for speaking development rather than a replacement for teachers.

Keywords: AI-generated feedback; EFL speaking; speaking proficiency; learner perception; Bangladesh

1. Introduction

Artificial intelligence (AI) has brought about significant changes in language learning (Gyawali & Mehandroo, 2022). This development has prompted a broader reconsideration of how language teaching and learning can be conceptualised. Through AI, students can receive tailored feedback based on their individual needs and weaknesses, which makes learning a second or foreign language more effective (Betel, 2023). AI-based feedback is particularly important in the development of speaking skills, which are generally considered complex productive skills that involves multiple subskills, including fluency, pronunciation, vocabulary and grammar (Lestari, 2019). In addition, increasing the confidence of students, reducing anxiety, and having a positive mental state also play important roles in improving speaking skills (Murtiningsih et al., 2023).

Various AI-based platforms can analyse the fluency, accuracy, pitch, rhythm, sentence structure and tone of students (Kristiawan et al., 2024). With advances in automatic speech recognition (ASR), natural language processing (NLP), and machine learning (ML), AI has been able to provide very fast, learner-centred speech feedback (Zhou et al., 2025). Conversely, in traditional classrooms, teachers cannot give equal attention to all students owing to the large class size and time constraints (Cruz et al., 2023). AI can help to alleviate these limitations to some extent (Çelik et al., 2022). AI-mediated platforms are particularly useful for developing speaking skills, as regular practice and quick corrective feedback are crucial for speaking improvement (Qiao & Zhao, 2023). In addition, in AI-based environments, students can practise repeatedly without fear of negative evaluation from others, which helps to create a more comfortable, confident and self-reliant learning environment for them (Nhật et al., 2025).

SmallTalk2Me, an AI-mediated platform designed to enhance speaking performance, has been examined previously for its potential to improve the speaking skills of English as a foreign language (EFL) learners in the Saudi context (Farooqi, 2025). To generate feedback on the speaking performance of users, this platform considers various aspects of speaking, including pronunciation, grammar and vocabulary usage (Manggasih et al., 2023). Existing research on AI-generated feedback for EFL speaking performance asserts its positive effects,

including instant feedback (Estaji et al., 2026). But the reliability of such feedback for varying speaking contexts and accents still needs to be explored and reducing communication to a level where human-to-human interactions do not exist is still a matter of concern (Moussalli & Cardoso, 2019). Recent research has started to investigate the effectiveness and challenges of feedback generated by SmallTalk2Me in resource-constrained contexts (Wicaksono et al., 2025). A study by Permatasari et al. (2025) examined the perceptions of students of SmallTalk2Me as an AI-powered conversational assistant for speaking practice; however, evidence on its effectiveness in resource-constrained EFL contexts remains limited.

In many low- and limited-resource countries, such as Bangladesh, English language learning facilities are inadequate to meet the needs of learners (Islam & -E-Jannat, 2025). In Bangladesh, learners have very limited opportunities to receive personalised support to improve EFL speaking proficiency, and overall communication opportunities in the language classroom remain very constrained, with mostly lecture-based, traditional classes being practised (Islam & Roy, 2024). Although English is highly significant for securing jobs and serves as a medium of instruction at the higher education level in Bangladesh, many students struggle to communicate in English (Mridha & Muniruzzaman, 2020).

With the advancement in technology, in highly resourced areas, learners are capable of digitalising the language learning process, but in resource-constrained countries, the scope for using technology for language learning remains comparatively more challenging owing to the unavailability of stable internet bandwidth and the persisting digital divide (Okoye et al., 2022), despite a vast population having access to gadgets like mobile phones (Seraj et al., 2020). Therefore, it is significant to conduct empirical research in such resource-constrained contexts to shed light on the impact of AI tools such as SmallTalk2Me on the EFL speaking proficiency of learners.

Also, the effects of the tool in measuring fluency specifically in experimental settings are underexplored (Al-husban, 2025). Therefore, the present study attempts to shed light on the role of SmallTalk2Me in improving the speaking proficiency of Bangladeshi tertiary-level EFL learners through an experimental study incorporating pre- and posttests to understand its effects on learners' speaking performance and to provide insights into the perspectives of learners on it through perception analysis. In doing so, it seeks to generate context-sensitive insights into the role of AI-assisted speaking practice in limited-resource educational environments.

Accordingly, the study addresses the following research questions (RQ):

- RQ1:** What is the impact of AI-mediated feedback provided by SmallTalk2Me on the EFL speaking proficiency of tertiary-level learners in Bangladesh?
- RQ2:** To what extent does AI-generated feedback influence the fluency development of Bangladeshi EFL learners?
- RQ3:** How do Bangladeshi EFL learners perceive the effectiveness, usability and pedagogical role of SmallTalk2Me in supporting speaking development?

2. Literature Review

2.1 AI as a Platform for Learning Feedback

AI has expanded opportunities for immediate, individualised, and data-driven feedback in language learning through the use of advanced technology (Evenddy, 2024). Compared with exclusively teacher-mediated feedback, AI-supported systems can provide more immediate and individualised feedback, allowing learners to receive timely responses to their linguistic needs and to engage in repeated opportunities for practice and revision (Evenddy, 2024; Wang & Han, 2022). However, the pedagogical value of AI-mediated feedback depends on how these technologies are integrated into the learning process and instructional context (Evenddy, 2024; Kasneci et al., 2023).

Despite these affordances, AI-mediated feedback has limitations, particularly in interpreting context-dependent, pragmatic, social and cultural dimensions of communication that may require human judgement and lived experience (Nazaretsky et al., 2025). Moreover, its effectiveness may vary according to the capacity of learners to engage with AI, their learning ecology, and the instructional support available to them (Salameh, 2024). Thus, AI-mediated feedback should be understood not as a universally effective intervention, but as a context-dependent pedagogical resource whose value is shaped by learner, instructional and contextual factors.

2.2 AI-Generated Feedback for Speaking Development

AI-generated instant feedback on English speaking has increasingly gained attention in EFL education (Nguyen, 2024). These systems can evaluate multiple dimensions of speaking, including pronunciation, fluency, grammar, vocabulary, lexical choice and speech rate, thereby providing learners with opportunities for repeated and self-directed practice (Nguyen, 2024). Such repeated opportunities for language production can be interpreted through the output hypothesis of Swain (1985), which proposes that producing language encourages learners to notice gaps between their current linguistic knowledge and their intended expression, possibly promoting language development.

Empirical studies generally have reported positive effects of AI-generated feedback on EFL speaking, particularly in observable aspects of oral performance such as pronunciation, fluency, speech rate, pause frequency and response time (Loewen et al., 2020). These findings suggest that AI-mediated feedback can facilitate repeated speaking practice and provide learners with immediate information about aspects of their oral performance (Qiao & Zhao, 2023). However, the extent to which such benefits can be achieved may vary across educational contexts. Therefore, although AI-generated feedback shows promise for EFL speaking development, further empirical investigation across diverse learning environments is needed to establish the consistency and scope of its effectiveness.

2.3 Learners' Perceptions of AI-Mediated Feedback

The perceptions of learners play a significant role in their engagement with AI in language learning. While AI can offer instant, personalised feedback, concerns remain about its grasp of pragmatic meaning, discourse context and the

sociocultural aspects of communication, as well as its capacity to provide genuine interaction and teacher support (Cogo et al., 2024). Therefore, whether learners accept AI depends not only on technical performance but also on trust, instructional context and opportunities for human interaction.

The responses of learners to AI-generated feedback can be influenced by affective and feedback-related factors (Bhar, 2026). Lower levels of anxiety and higher levels of self-confidence can support language learning, as outlined in the affective filter hypothesis of Krashen (1982). According to Hattie and Timperley (2007), feedback is useful when learners can understand and apply it to improve future work. The pedagogical potential of AI feedback is as much about the feedback itself as about the perceptions, trust, and interaction of learners with it (Liu et al., 2024). The technology acceptance model (TAM) also helps to explain the adoption of AI-enabled technologies by highlighting perceived usefulness and perceived ease of use (Davis, 1989).

Recent studies indicate that these expectations can affect the attitudes of learners towards AI-assisted speaking practice (Zou et al., 2023). However, the relationship between positive attitudes towards AI feedback and improvements in speaking skills remains poorly established, especially in low-resource EFL settings (Al-husban, 2025). This gap indicates that it is important to examine what learners actually think and do regarding their speaking performance, rather than assuming that positive perceptions are evidence of the effectiveness of the learning.

2.4 AI-Supported Speaking Development Across Contexts

Evidence from diverse educational contexts demonstrates both shared and context-specific benefits of AI-assisted speaking instruction. Research from China has shown consistently that core speech features, such as pronunciation, fluency and speech rate, can be improved using ASR systems (Yan & Singh, 2026), whereas in different EFL contexts, the focus is on reducing speaking anxiety and increasing willingness to communicate (Bhar, 2026). However, the studies also emphasise learner autonomy, self-regulated learning and individual speaking practice (Viberg et al., 2020). These differences suggest that the pedagogical effectiveness of AI-mediated feedback depends on various factors, including the demographics of learners and specific learning ecologies.

The efficacy of a particular AI-mediated technology does not rely solely on how well it is built; rather, it depends on how a particular learner integrates the technology into learning habits, how the learning environment facilitates the incorporation of AI-mediated learning into its regular pedagogical activities and how much knowledge a learner has about embarking on and continuing learning through AI (Kim et al., 2022). This scenario is relevant especially to Bangladesh because EFL classrooms remain overcrowded, teachers traditionally dominate classes, student participation remains limited and assessing student performance and catering to individual needs are not up to the mark, which affects EFL speaking skills (Islam & Roy, 2024). Therefore, findings from technologically advanced settings cannot be applied to limited-resource contexts like Bangladesh.

2.5 SmallTalk2Me as an AI-Powered Speaking Platform

With the emergence of ASR, NLP and ML, AI-enabled language learning platforms have proliferated, and SmallTalk2Me is among the most prominent, specifically designed to improve EFL skills. It offers a range of learning features, including practice for everyday conversations and interviews, as well as an IELTS practice platform. This AI platform assesses different subskills in speaking to help learners become proficient (Manggiasih et al., 2023). The platform can reduce the need always to approach a language teacher to monitor learning. Instead, SmallTalk2Me helps the completely different speaking tasks of learners regularly and provides immediate feedback based on structured rubrics. Learners can check the learning graph on their own (Manggiasih et al., 2023). Although some studies report that SmallTalk2Me is effective in improving ESL speaking skills (Farooqi, 2025), its effectiveness in resource-limited contexts, such as Bangladesh, has not yet been studied and requires further research.

2.6 Research Gap

Although interest in researching AI-facilitated EFL skills enhancement is growing, most studies are conducted in developed settings where the learning ecology is rich in digital equipment. However, how AI-facilitated educational technology would work in limited-resource contexts like Bangladesh remains under-researched. Because findings from developed nations may not generalise to resource-constrained settings, this area remains underexplored (Qiao & Zhao, 2023). Also, few studies have assessed the role of AI in enhancing the speaking proficiency of learners using IELTS rubrics, with a focus on both overall speaking performance and specific measures of fluency, followed by an analysis of learner perceptions. Existing research has focused largely on overall speaking proficiency and the perceptions of learners of AI-supported learning (Nguyen, 2024).

A further gap concerns the limitations of AI-mediated feedback. Issues such as inconsistent feedback, limited sensitivity to sociocultural context and difficulties in evaluating pragmatic and higher-order communicative features remain insufficiently examined (Yan & Singh, 2026). Although SmallTalk2Me has potential as an AI-supported speaking practice platform, empirical evidence of its effectiveness in resource-constrained EFL settings remains limited (Wicaksono et al., 2025). Accordingly, the present study addresses these gaps by examining whether AI-mediated feedback through SmallTalk2Me can improve the speaking performance and fluency of Bangladeshi tertiary-level EFL learners, while also exploring the perceptions of learners of its effectiveness, usability and pedagogical value.

3. Methodology

3.1 Research Design

The study used a one-group pretest-posttest pre-experimental design, which is an exploratory intervention design typically used when random assignment is not possible or when there is no control group used in an educational study (Creswell & Creswell, 2017). The design allowed within-participant changes in speaking to be studied by comparing pretest and posttest scores after the AI-mediated intervention. The repeated measures design was used with the same subjects

tested in the same conditions to assess changes over time and to minimise the influence of individual differences on the results (Field, 2018). In addition, a post-intervention questionnaire was administered to examine the perceptions of participants of the AI-mediated learning experience. As the study did not include a control group, the findings should be interpreted as changes associated with the intervention rather than evidence of causal effects. Moreover, the design is subject to potential threats to internal validity, including maturation, testing effects and history (Shadish et al., 2002).

Figure 1 presents an overview of the study procedure, illustrating the sequence of the IELTS pretest, AI-mediated speaking intervention, fluency practice sessions, IELTS posttest and post-intervention questionnaire administered to participants.

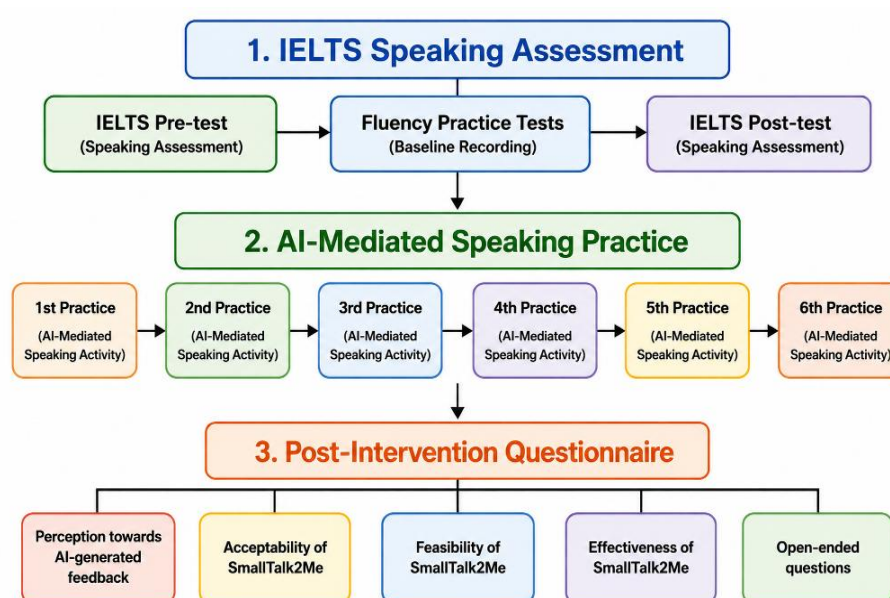


Figure 1: Overview of the study design and data collection procedure

3.2 Research Setting, Participants, and Sampling

The study was conducted at a public university in Bangladesh, where limited opportunities for individualised English-speaking practice and timely formative feedback provided an appropriate context for evaluating AI-mediated feedback in resource-constrained EFL speaking instruction (Seraj et al., 2021).

The participants were 28 undergraduate English major students (64% fifth semester- and 36% sixth-semester students). Participants were recruited using criterion purposive sampling, a non-probability sampling technique appropriate for selecting individuals who meet predefined eligibility criteria aligned with the research objectives (Etikan, 2016). To ensure participant comparability, only students who (a) were enrolled in the undergraduate English programme; (b) were in the fifth or sixth semester; (c) had a cumulative GPA between 3.40 and 3.60; (d) had access to a smartphone or computer with a stable internet connection; and (e) agreed to complete the three-week intervention were included.

The final sample comprised 28 participants, all of whom completed the intervention and both assessments. The sample size was deemed adequate for an exploratory one-group pretest–posttest pre-experimental study, focusing on preliminary changes in the individual rather than at the population level. In pilot and exploratory intervention studies with repeated measures designs, similar sample sizes are considered acceptable (Julious, 2005). However, the limited number of participants in the studies reduces the validity of the conclusions drawn, and future studies should feature larger and more diverse samples, as well as controlled experimental designs.

3.3 Intervention Procedure

The intervention was conducted over three weeks using the AI-powered speaking platform SmallTalk2Me. Participants were oriented to the study procedures before the intervention and were provided with premium access to the platform. Following the pretest, they completed six AI-mediated speaking activities from the *How to Make Great Small Talk* module, submitting the AI-generated performance report after each activity. After completing the intervention, participants completed the posttest and a perception questionnaire. All participants completed the same sequence of activities, received identical instructions and followed the intervention over the same three-week period to ensure procedural uniformity.

3.4 IELTS Speaking Pretest and Posttest

Participants completed an IELTS speaking pretest before the intervention and an IELTS speaking posttest after the intervention using the SmallTalk2Me assessment platform. Before the pretest, each participant received a general introduction to the assessment process and the SmallTalk2Me platform. The orientation familiarised participants with the assessment process and how to navigate the platform, but no practice was provided on the actual test prompts to minimise testing bias and maintain the integrity of the pretest. The pretests and posttests followed the same format as the IELTS speaking test, with three sections: Part 1 (introduction and interview), Part 2 (individual long turn), and Part 3 (two-way discussion).

The platform generated pretest and posttest IELTS speaking items in a different but standardised manner to reduce practice effects while still ensuring that the items measured the same constructs in each test. Assessments were administered in the same IELTS speaking format, and all responses were scored according to the same criteria and using the same AI scoring system, which ensured consistent assessment conditions for all participants and across measurement occasions. The AI-generated reports provided overall IELTS speaking band scores and analytic feedback. Words per minute (WPM) was also used as an objective measure of speech fluency, since IELTS band scores do not correlate directly with speaking rate.

3.5 Questionnaire Development and Validation

A 25-item questionnaire based on a 5-point Likert scale and a 3-item open-ended question was used to determine the perception of participants. The questionnaire structure and scale-development procedures were informed by Petrić and Czár

(2003), and Raoofi et al. (2017). These studies were consulted solely to guide item organisation and scale development. Their writing-strategy constructs and items were not adopted, as the present study focuses on oral fluency and EFL speaking proficiency.

Content validity was ensured by content expert review by an expert in technology-mediated language teaching. During pilot testing for clarity and understanding, the questionnaire was modified as necessary. The last instrument showed excellent internal consistency (Cronbach's $\alpha = .932$).

3.6 Data Collection Procedure

The data were obtained from three sources: (a) IELTS speaking pretest and posttest scores; (b) WPM recorded during the first and sixth AI-mediated speaking tasks; and (c) responses participants to the post-intervention questionnaire.

3.7 Data Analysis

Data were analysed using IBM SPSS Statistics (Version 27). Prior to inferential analysis, the assumption of normality was assessed using the Shapiro-Wilk test (Shapiro & Wilk, 1965), which confirmed that all outcome variables satisfied the assumption of normality (all $p > .05$; Appendix 1), thereby supporting the use of parametric analyses (Field, 2018). Descriptive statistics (mean, standard deviation, mean difference and percentage difference) were used to summarise speaking performance and questionnaire responses.

Changes in IELTS speaking scores and speaking fluency were examined using paired-samples *t*-tests, as the study compared pretest and posttest scores obtained from the same participants (Pallant, 2020). Statistical significance was set at $p < .05$. Closed-ended questionnaire responses were analysed descriptively, whereas responses to the open-ended questions were analysed using the six-phase reflexive thematic analysis of Braun and Clarke (2006).

3.8 Ethical Considerations

Ethical approval was obtained from the relevant institutional authority before data collection. Participation was voluntary, and all participants provided informed consent before taking part in the study. They were informed about the purpose and procedures of the study and their right to withdraw at any time without penalty. All collected data were anonymised to protect the privacy of participants and were used solely for the purposes of this research.

4. Results

4.1 The Band Score of IELTS, Pretest and Posttest Results

A paired-samples *t*-test (Table 1) showed a statistically significant improvement in IELTS speaking scores from the pretest ($M = 5.75$, $SD = 0.65$) to the posttest ($M = 6.38$, $SD = 0.67$), $t(27) = 4.92$, $p < .001$, Cohen's $d = 0.93$, indicating a large effect.

Table 1: Comparison of IELTS speaking pretest and posttest scores

Test Phase	<i>n</i>	<i>M</i>	<i>SD</i>	Mean Difference	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Pretest	28	5.75	0.65					
				0.63	4.92	27	<.001	0.93
Posttest	28	6.38	0.67					

Note. *df* = 27. *p* <.001 indicates a statistically significant difference. Cohen's *d* = 0.93 indicates a large effect

4.2 Comparison Between the First and Sixth Practice Test in Terms of Fluency

Participants also demonstrated a significant improvement in speaking fluency (Table 2) from Practice 1 to Practice 6, $t(27) = 6.11$, $p <.001$, Cohen's $d = 1.15$, indicating a large effect.

Table 2: Comparison of fluency based on first and sixth practice test scores

Practice Stage	<i>n</i>	<i>M</i> (WPM)	<i>SD</i>	Mean Difference	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Practice 1	28	84.50	15.20					
				8.60	6.11	27	<.001	1.15
Practice 6	28	93.10	16.10					

Note. *df* = 27. *p* <.001 indicates a statistically significant improvement. Cohen's $d = 1.15$ indicates a large effect

The subsequent analysis (Figure 2) presents the percentages of student scores that increase, remain unchanged or decrease from pretest to posttest and from the first to the sixth fluency test.

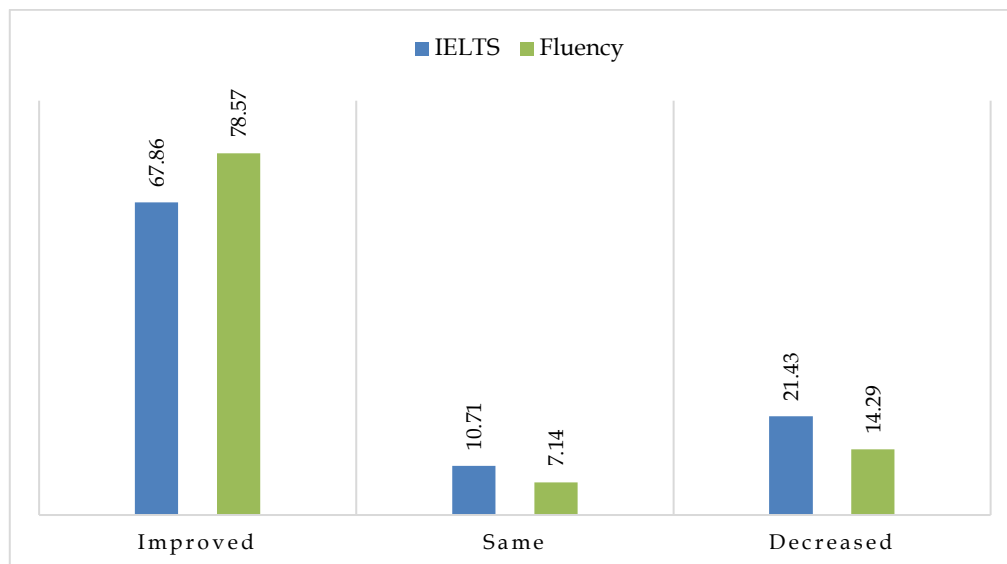


Figure 2: Distribution of participant performance change in IELTS speaking scores and fluency

Figure 2 presents the distribution of changes in the IELTS speaking scores and speaking fluency of participants between the pretest and posttest assessments. Overall, 67.86% of the participants improved their IELTS band scores in the posttest, whereas 10.71% showed no change and 21.43% demonstrated a slight decrease in their scores. A similar pattern was observed for speaking fluency, with

78.57% of participants showing improvement, 7.14% remaining unchanged and 14.29% showing a decrease in performance. Overall, the descriptive results indicate that a majority of participants exhibited higher IELTS speaking scores and speaking fluency in the posttest than in the pretest.

4.3 Perceptions and Experiences of Learners of AI-Generated Feedback

The questionnaire results (Table 3) indicated that 89.28% of the participants had prior experience using AI-based tools, whereas 10.71% reported no prior experience. The overall questionnaire mean score was 3.67, indicating generally positive perceptions of AI-mediated speaking practice.

Table 3: Quantitative results of the close-ended questionnaire items

Questions	Yes%	No%					
1. Prior experience using AI tools.	89.28	10.71					
Perception Of Feedback Generated By AI	Strongly Agree %	Agree %	Neutral %	Disagree %	Strongly Disagree %	Mean	Std. Deviation
2. My overall speaking proficiency was enhanced by the assistance provided by the AI-mediated platform.	14.3	71.4	10.7	0	3.6	3.93	.766
3. Feedback received from AI was comprehensible and clear.	3.6	67.9	25	0	3.6	3.68	.723
4. I could identify errors related to pronunciation and vocabulary with the help of AI.	3.6	60.7	32.1	0	3.6	3.61	.737
5. The AI-generated feedback was specific to the tasks and assignments.	3.6	75	21.4	0	0	3.82	.476
6. I was motivated to improve my English skills by the feedback I received from the AI platform.	7.1	64.3	25	0	3.6	3.71	.763
7. The feedback motivated me to experiment with increasingly diverse speaking techniques.	3.6	60.7	32.1	0	3.6	3.61	.737

8.	Getting AI feedback helped lessen my worry about errors.	7.1	64.3	25	0	3.6	3.71	.763
9.	I relied on the AI feedback to be correct.	3.6	42.9	42.9	10.7	0	3.25	.799
10.	The AI accurately detected my mistakes most of the time.	7.1	53.6	28.6	10.7	0	3.57	.790
Overall Mean of Perception							3.65	
Acceptability of SmallTalk2ME								
11.	This SmallTalk2Me task, designed specifically for IELTS speaking, is beneficial for improving my speaking skills for the test.	14.3	67.9	10.7	0	7.1	3.82	.945
12.	I am eager to integrate AI-mediated speaking practice activities on IELTS to prepare for the IELTS exam.	3.6	71.4	17.9	0	7.1	3.64	.870
13.	I recommend this AI-mediated task and feedback to others preparing to take the IELTS speaking test.	14.3	53.6	28.6	3.6	0	3.79	.738
Overall Mean of Acceptability							3.75	
Feasibility of SmallTalk2Me.								
14.	SmallTalk2Me is a user-friendly platform that is easy to navigate.	7.1	67.9	17.9	0	7.1	3.68	.905
15.	I felt confident using SmallTalk2Me on my own.	3.6	71.4	14.3	3.6	7.1	3.61	.916
16.	The activities for IELTS in SmallTalk2Me are practical and direct.	3.6	71.4	17.9	0	7.1	3.64	.870
17.	I can incorporate the speaking tasks on IELTS	3.6	64.3	25	0	7.1	3.57	.879

from SmallTalk2Me into my study schedule based on my preferred timing.								
Overall Mean of Feasibility								3.63
Effectiveness of SmallTalk2Me								
18.	SmallTalk2Me's AI-driven feedback assists me in tackling my speaking difficulties.	0	60.7	28.6	7.1	3.6	3.46	.793
19.	I believe that practising speaking with SmallTalk2Me will significantly enhance my test performance.	7.1	57.1	28.6	0	7.1	3.57	.920
20.	The feedback provided by SmallTalk2Me is without vagueness and is direct, making it understandable.	7.1	67.9	14.3	7.1	3.6	3.68	.863
21.	The IELTS speaking tasks on SmallTalk2Me help me focus on specific speaking skills that need further development.	3.6	75	14.3	3.6	3.6	3.71	.763
22.	Using this activity and feedback system has boosted my confidence in my IELTS speaking skills.	7.1	57.1	35.7	0	0	3.71	.600
Overall Mean of Effectiveness								3.63
Overall Mean of the Questionnaire								3.67

For the perception dimension, the overall mean score was 3.65. The highest-rated item was Q2 ($M = 3.93$, $SD = .766$), followed by Q5 ($M = 3.82$, $SD = .476$). Items related to feedback clarity, error identification and motivation (Q3, Q4, and Q7) also received relatively positive ratings ($M = 3.61$ – 3.68). By contrast, Q9 received the lowest mean score ($M = 3.25$, $SD = .799$), indicating comparatively lower ratings for trust in AI-generated feedback.

The acceptability dimension recorded the highest overall mean ($M = 3.75$) among the four dimensions. Participants gave the highest rating to Q11 ($M = 3.82$, $SD = .945$), while Q13 ($M = 3.79$, $SD = .738$) and Q12 ($M = 3.64$, $SD = .870$) also received positive ratings.

The feasibility dimension had an overall mean of 3.63. Q14 received the highest mean score ($M = 3.68$, $SD = .905$), followed by Q16 ($M = 3.64$, $SD = .870$) and Q15 ($M = 3.61$, $SD = .916$). Q17 received the lowest rating in this dimension ($M = 3.57$, $SD = .879$).

The effectiveness dimension also produced an overall mean of 3.63. The highest mean scores were observed for Q21 and Q22 (both $M = 3.71$), whereas Q18 received the lowest mean score ($M = 3.46$, $SD = .793$).

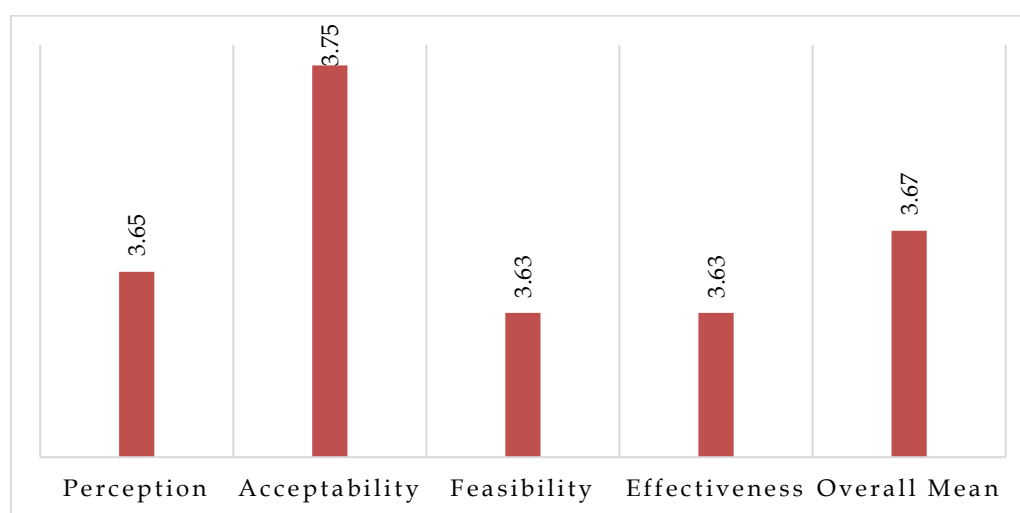


Figure 3: Comparison of perceptions, acceptability, feasibility and effectiveness

Overall (Figure 3), acceptability recorded the highest mean score ($M = 3.75$), followed by perception ($M = 3.65$), whereas feasibility and effectiveness both recorded mean scores of 3.63. These descriptive findings indicate generally positive participant responses across all four dimensions, with comparatively lower ratings for trust in AI-generated feedback and some aspects of long-term use and perceived effectiveness.

4.4 Open-ended Questions

To gain deeper insights into the experiences of learners beyond the quantitative findings, participants were asked to respond to three open-ended questions:

1. What aspect of the AI-generated feedback did you find most helpful?
2. What improvements would you recommend for the SmallTalk2Me platform?
3. How has AI feedback impacted your learning experience compared to traditional methods?

Table 4: Thematic summary of learner responses to open-ended questions

Theme	Key Findings	Example Responses
Speed and immediacy of feedback	Students viewed the quick and immediate response of AI-generated feedback positively, as it helped them understand and correct their mistakes quickly while practising speaking.	'Answers faster', 'instant feedback', 'quick response'
Clarity and personalisation	Participants preferred feedback that was clear, specific and tailored to their individual needs. Many said that the system identified mistakes clearly and also showed them ways to improve.	'Clear correction', 'specific feedback', 'shows how to improve'
Language skill development	According to students, the platform helped them to improve their speaking skills, grammar accuracy, listening comprehension and confidence. Some participants also noted that the feedback highlighted both strengths and weaknesses, not just weaknesses.	'Improved speaking ability', 'improved grammar', 'shows strengths and weaknesses'
Flexibility and learner autonomy	Many participants perceived that AI-mediated feedback offered greater flexibility, as there were no time or place restrictions. Their affective filters were lowered, as they felt less anxious and more motivated to practise speaking.	'Practise anytime', 'less fear of mistakes', 'more confidence'
Suggested pedagogical improvements	Participants suggested adding more detailed explanations, additional examples, a transparent scoring system, and more speaking topics and realistic communication scenarios.	'More examples', 'clearer explanation', 'more speaking topics'
Technical and user experience improvements	Participants reported that the free features of SmallTalk2Me were limited and that the paid features would be easier to use if they were less expensive. They also noted that the platform could be more user-friendly and that the pronunciation assessment needs to be more effective. The system for monitoring student progress should be better organised, and mistakes should be identified more quickly and in a variety of ways.	'Better pronunciation scoring', 'more user-friendly navigation', 'quicker and diverse feedback'

Qualitative analysis of the open-ended responses indicated generally positive perceptions of SmallTalk2Me. Participants frequently described the platform as useful for providing immediate, clear and personalised feedback during speaking practice. Many reported that the AI-generated feedback supported the improvement of their English-speaking skills, increased their interest in language learning and made them feel more comfortable during speaking practice. Several participants also described the platform as convenient and learner-centred because it identified speaking errors quickly and provided individualised suggestions based on their performance.

Participants also identified several areas for improvement. Common suggestions included improving the accuracy of pronunciation assessment, providing clearer explanations of the scoring system, enhancing the usability of the platform and offering more personalised learning materials. Overall, the qualitative findings indicated generally positive perceptions of SmallTalk2Me while highlighting several aspects that participants believed could further improve the platform.

5. Discussion

The present study investigated the role of SmallTalk2Me in enhancing the English proficiency of learners, with a specific focus on their speaking rate and on providing insight into the perspectives of learners on using this AI platform to develop their EFL speaking skills. The findings indicate that AI-mediated feedback affects the proficiency of learners positively by providing opportunities for individual practice sessions and immediate feedback. These findings align with results from previous studies on AI-mediated EFL learning (Qiao & Zhao, 2023).

The findings of the study indicate the significance of acknowledging the role of individual learning ecologies, facilities and obstacles while exploring the pedagogical effect of an AI tool (Salameh, 2024). Because the present study was conducted using an in-group experimental method with a limited number of participants who were exposed to the AI platform for a short period, the results of the study cannot be generalised to other settings and the study does not claim to have established any cause-and-effect relationship.

5.1 Improving Speaking Performance and Fluency

The results show that the exposure of learners to the SmallTalk2Me platform during the intervention period helped them to improve their English proficiency, as measured using the IELTS format. Their fluency was also improved. This result is similar to past studies exploring the role of AI in improving EFL speaking skills (Zou et al., 2023). The findings indicate that, because the affective filters of learners were low, they were motivated to keep practising speaking through the AI platform (Qiao & Zhao, 2023), suggesting the potential benefits of incorporating AI in limited-resource contexts like Bangladesh. The findings of the present study can also be related to the output hypothesis of Swain (1985), since learners had ample opportunities to practise speaking and to be assessed on speaking tasks.

This procedure may have facilitated monitoring their output and making necessary edits. Individual variances were noticed in the extent of speaking improvement from AI-mediated feedback, indicating that learner demographics, particularly the ability to use feedback, affective filters while using the digital platform level, and perseverance, influenced how it is integrated and utilised for productivity impact (Bashori et al., 2020). However, the findings are indicators, not evidence of a strict cause-and-effect relationship, because the study involved only one group, the sample was small and the experiment was short.

5.2 Affective Aspects: Reduced Anxiety and Enhanced Motivation

The findings suggest that participants experienced lower speaking anxiety and greater motivation during the SmallTalk2Me intervention. This pattern is consistent with the affective filter hypothesis of Krashen (1982), which proposes that lower anxiety and greater confidence may create more favourable conditions for language learning (Bashori et al., 2021). Participants also reported that the absence of peer and teacher judgement encouraged more frequent speaking practice, consistent with previous research on AI-supported and autonomous language learning (Ding & Yusof, 2025; Wu et al., 2025). However, given the small sample and the pretest-posttest design of the study, these findings should be interpreted as suggestive rather than conclusive evidence of a causal effect. In the Bangladeshi EFL context, where opportunities for individualised oral practice may be limited (Islam & Stapa, 2021), AI-mediated speaking practice may therefore offer a supportive environment for reducing affective barriers and encouraging oral participation.

5.3 Learner Perception: The Gap Between Acceptance and Trust

Despite using AI for language learning, EFL learners still expressed concerns about the accuracy of AI-generated feedback. This finding is broadly consistent with previous research highlighting limitations in AI-mediated speaking tools, particularly in recognising diverse pronunciations, interpreting pragmatic meaning and reproducing authentic human interaction (Manggiasih et al., 2023). In the present sample, the findings suggest that, in this resource-constrained Bangladeshi EFL context, the trust of learners in AI-mediated feedback may be influenced not only by perceived technical performance but also by the contextual conditions surrounding its use.

Consistent with feedback theory, effective feedback involves helping learners to understand their current performance and identify how they can improve subsequent performance (Hattie & Timperley, 2007). Accordingly, concerns about the credibility of AI-generated feedback may have influenced the willingness of learners to rely on and engage with such feedback in the present study. These findings suggest that trust may be an important consideration in the effective and sustained use of AI-mediated feedback, although this interpretation should be treated cautiously given the small sample and one-group experiment.

5.4 Feasibility and Contextual Limitations in Bangladesh

Participants were largely positive about the feasibility of SmallTalk2Me for speaking practice but found it challenging to use regularly and to navigate the platform. These findings suggest that the pedagogical potential of AI-based language learning depends not only on technological capabilities but also on the digital readiness and contextual support of learners (Akter et al., 2026). The results of the present study contradict those from more technologically advanced educational environments, suggesting that further issues remain when implementing AI-supported speaking practice in resource-limited settings. The challenges that participants faced in integrating the learning environment into their routines and following its instructions mirror the importance of learner and learning environment readiness, as highlighted by previous research (Otero et al., 2023). These findings underscore the importance of aligning AI affordances with

the preparedness and contextual support of learners when integrating AI-mediated speaking practice in resource-constrained EFL settings.

5.5 AI Versus Human Feedback: Moving Towards a Hybrid Approach

While SmallTalk2Me was appreciated as a system that provided customised and timely feedback, participants noted that it did not offer any human interaction, emotional support or instructor guidance. This finding is consistent with research emphasising the complementary, rather than substitutive, role of AI in language learning (Alhalangy & AbdAlgane, 2023). The findings indicate that participant responses suggest the potential value of combining AI-supported practice and immediate feedback with teacher guidance and human interaction. This blended approach can offer the benefits of both AI and human interaction to some extent, since AI may not be able to provide the interpersonal touch and engagement that a person can. From a sociocultural perspective, the findings also reiterate the significance of human mediation in facilitating the speaking development of learners (Lantolf, 2006).

5.6 Theoretical Implications

The findings suggest that the effectiveness of AI-mediated EFL speaking practice may depend not only on the technological affordances of the tool but also on the affective states of learners, engagement with AI-mediated practice and the surrounding learning context. The results can be explained by the affective filter hypothesis (Krashen, 1982), output hypothesis (Swain, 1985), and feedback framework (Hattie & Timperley, 2007). On the affective side, the emotional reactions of students can influence their motivation to use AI for speaking practice, and the output hypothesis emphasises the importance of language output in language learning. The feedback framework also highlights that feedback, when given at the right time and with the right information, can facilitate learning. Taken together, these perspectives suggest that AI-mediated speaking development may be understood better as an interaction between technological affordances, learner affect, productive engagement and feedback rather than as a function of the technology alone.

5.7 Practical Implications

The results indicate that AI and human instruction should not be considered rivals in developing speaking ability in EFL but rather as complementary techniques. SmallTalk2Me can also provide opportunities for repeated practice and immediate feedback (Farooqi, 2025), while teachers can provide interpersonal interaction, encouragement and contextualised guidance. However, combining the responsiveness and accessibility of AI with the social and pedagogical strengths of teachers may be useful. At the same time, pronunciation variation and pragmatic interpretation constraints point to the need for teacher oversight to ensure that using AI for speaking practice in EFL classrooms is effective (Al-Shallakh, 2024; Manggiasih et al., 2023).

6. Conclusion

This study examined the use of SmallTalk2Me as an AI-mediated feedback platform to develop English-speaking skills among tertiary-level EFL learners from Bangladesh. The findings indicate that participants improved in speaking performance and fluency after the three-week intervention and they generally viewed AI-mediated feedback positively, particularly for its immediacy, personalisation, and motivational value. However, learning outcomes varied and lower trust in AI-generated feedback suggests that the educational value of AI will vary depending on the digital readiness, engagement and teaching context of learners. The study adds to the emerging body of research on AI-supported language learning by offering empirical data from the EFL context in the Global South, where language resources are limited.

The results suggest that a blended teaching model combining AI feedback with the pedagogical knowledge, contextual support and socio-emotional support of teachers can enhance the teaching process. The findings should be considered in light of the limitations of the study, including a one-group pretest–posttest pre-experimental design, no control group, a small sample, a single-institution study and a short intervention period. The study, therefore, cannot draw any firm causal inferences. Future studies could use larger, more varied samples; include control groups; explore longer-term effects of AI-mediated feedback; and continue examining the relationship between digital literacy, feedback literacy, trust in AI and language learning outcomes in various contexts.

7. Conflict of Interest

The authors declare that there is no conflict of interest associated with the publication of this manuscript. The research was conducted independently, without any commercial or financial relationships that could be construed as a potential conflict of interest.

8. Acknowledgements

The authors acknowledge the use of ChatGPT (OpenAI) and Grammarly during the preparation of this manuscript. These tools were used to support language refinement, including grammar, translation, clarity, readability, coherence and sentence fluency. All final decisions and substantive intellectual contributions were made and verified by the authors.

9. References

- Akter, M., Akter, M., & Ali, Md. M. (2026). Impact of digital learning on primary education in Bangladesh. *Global Mainstream Journal of Arts Literature History & Education*, 2(1). <https://doi.org/10.62304/alhe.v4i01.245>
- Alhalangy, A., & AbdAlgane, M. (2023). Exploring the impact of AI on The EFL context: A case study of Saudi universities. *Journal of Intercultural Communication*, 23(2), 41–49. <https://doi.org/10.36923/jicc.v23i2.125>
- Al-husban, N. (2025). The impact of AI- assisted language learning tools on augmenting university EFL students' speaking skills in Jordan. *Journal of Applied Learning & Teaching*, 8(1), 116–127. <https://doi.org/10.37074/jalt.2025.8.1.13>
- Al-Shallakh, M. A. (2024). Embedding artificial intelligent applications in higher educational institutions to improve students' pronunciation performance. *Theory*

- and Practice in Language Studies*, 14(6), 1897–1906. <https://doi.org/10.17507/tpsls.1406.31>
- Bashori, M., Van Hout, R., Strik, H., & Cucchiarini, C. (2020). Web-based language learning and speaking anxiety. *Computer Assisted Language Learning*, 35(5-6), 1058–1089. <https://doi.org/10.1080/09588221.2020.1770293>
- Bashori, M., Van Hout, R., Strik, H., & Cucchiarini, C. (2021). Effects of ASR-based websites on EFL learners' vocabulary, speaking anxiety, and language enjoyment. *Data Archiving and Networked Services (DANS)*, 99, 102496–102496. <https://doi.org/10.1016/j.system.2021.102496>
- Betal, A. K. (2023). Enhancing second language acquisition through artificial intelligence (AI): Current insights and future directions. *Journal for Research Scholars and Professionals of English Language Teaching*, 7(39). <https://doi.org/10.54850/jrspelt.7.39.003>
- Bhar, S. K. (2026). Artificial intelligence in EFL speaking instruction: A systematic review of pedagogical design, affective conditions and instructional input. *Encyclopedia*, 6(4), 74–74. <https://doi.org/10.3390/encyclopedia6040074>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Çelik, İ., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research *TechTrends*, 66(4), 616–630. <https://doi.org/10.1007/s11528-022-00715-y>
- Cogo, A., Patsko, L., & Szoke, J. (2024). Generative artificial intelligence and ELT. *ELT Journal*, 78(4), 373–377. <https://doi.org/10.1093/elt/ccae051>
- Creswell, J. W., & Creswell, J.D. (2017). *Research design Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Cruz, R., de Lourdes M., Villanueva, M. S., & Pluvinet, D. (2023). Mediation in autonomous practices among novice language teachers while discussing conflicts in their teaching practice. *Colombian Applied Linguistics Journal*, 25(2), 44–57. <https://doi.org/10.14483/22487085.18736>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of Information Technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Ding, D., & Yusof, A. M. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. *Humanities and Social Sciences Communications*, 12(1). <https://doi.org/10.1057/s41599-025-05550-z>
- Estaji, M., Mirfendereski, Y., & Marian, K. S. (2026). Artificial intelligence as a catalyst for transformation: EFL learners' perceptions of AI-powered language tools in delivering effective oral and written corrective feedback. *Research and Practice in Technology Enhanced Learning*, 22, 14–14. <https://doi.org/10.58459/rptel.2027.22014>
- Etikan, İ., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1–1. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Evenddy, S. S. (2024). Investigating AI's automated feedback in English language learning. *FLIP Foreign Language Instruction Probe*, 3(1), 76–87. <https://doi.org/10.54213/flip.v3i1.401>
- Farooqi, S.-H. (2025). Efficacy of AI-generated feedback by SmallTalk2Me for improving speaking skill of Saudi EFL learners. *Forum for Linguistic Studies*, 7(3), 714–772. <https://doi.org/10.30564/fls.v7i3.8294>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. Sage.

- Gyawali, Y. P., & Mehandroo, M. (2022). Artificial intelligence in English language teaching: Navigating the future with emerging perspectives. *Journal of Language and Linguistics in Society*, 26, 21–27. <https://doi.org/10.55529/jlls.26.21.27>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Islam, M. S., & Stapa, M. (2021). Students' low proficiency in spoken English in private universities in Bangladesh: Reasons and remedies. *Language Testing in Asia*, 11(1), 22. <https://doi.org/10.1186/s40468-021-00139-0>
- Islam, Md. S., & Roy, S. (2024). Challenges in developing learners' English-speaking skills at the tertiary level in the EFL context: Teachers' and learners' perceptions. In *Research Square*. <https://doi.org/10.21203/rs.3.rs-4491944/v1>
- Islam, S., & -E-Jannat, B. (2025). English language learning challenges in Bangladeshi secondary schools. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(4), 450–454. <https://doi.org/10.62225/2583049x.2025.5.4.4613>
- Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics*, 4(4), 287–291. <https://doi.org/10.1002/pst.185>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274–102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kim, J., Lee, H.-K., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27(5), 6069–6104. <https://doi.org/10.1007/s10639-021-10831-6>
- Krashen, S. (1982). *Principles and practice in second language acquisition*. 168–168. <https://doi.org/10.2307/328293>
- Kristiawan, D., Bashar, K., & Pradana, D. A. (2024). Artificial Intelligence in English language learning: A systematic review of AI tools, applications, and pedagogical outcomes. *The Art of Teaching English as a Foreign Language*, 5(2), 207–218. <https://doi.org/10.36663/tatefl.v5i2.912>
- Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28(1), 67–109. <https://doi.org/10.1017/s0272263106060037>
- Lestari, N. (2019). Improving the speaking skill by vlog (video blog) as learning media: The EFL students' perspective. *International Journal of Academic Research in Business and Social Sciences*, 9(1). <https://doi.org/10.6007/ijarbss/v9-i1/5490>
- Liu, G., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, 38(7), 1632–1660. <https://doi.org/10.1080/09588221.2024.2310288>
- Loewen, S., Isbell, D. R., & Sporn, Z. (2020). The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability. *Foreign Language Annals*, 53(2), 209–233. <https://doi.org/10.1111/flan.12454>
- Manggiasih, L. A., Loreana, Y. R., Azizah, A., & Nurjati, N. (2023). Strengths and limitations of SmallTalk2Me App in English language proficiency evaluation. *Tell Teaching of English Language and Literature Journal*, 11(2). <https://doi.org/10.30651/tell.v11i2.19560>
- Moussalli, S., & Cardoso, W. (2019). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, 33(8), 865–890. <https://doi.org/10.1080/09588221.2019.1595664>

- Mridha, Md. M., & Muniruzzaman, S. Md. (2020). Developing English speaking skills: Barriers faced by the Bangladeshi EFL learners. *Englisia Journal of Language Education and Humanities*, 7(2), 118–118. <https://doi.org/10.22373/ej.v7i2.6257>
- Murtiningsih, S. R., Darmanto, F. M. I., & Hidayatulloh, S. M. M. (2023). Students' psychological state degree and their speaking ability: To what extent is the link? *Voices of English Language Education Society*, 7(2), 210–220. <https://doi.org/10.29408/veles.v7i2.14815>
- Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2025). The critical role of trust in adopting AI-powered educational technology for learning: An instrument for measuring student perceptions. *Computers and Education Artificial Intelligence*, 8, 100368–100368. <https://doi.org/10.1016/j.caeai.2025.100368>
- Nguyen, H. A. (2024). Harnessing AI-based tools for enhancing English speaking proficiency: Impacts, challenges, and long-term engagement. *International Journal of AI in Language Education*, 1(2), 18–29. <https://doi.org/10.54855/ijaile.24122>
- Nhật, N. Q., Vũ, L. H., & Trúc, N. T. T. (2025). The impact of AI-supported speaking practice on EFL learners' confidence in Vietnam. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(5), 384–393. <https://doi.org/10.62225/2583049x.2025.5.5.4924>
- Okoye, K., Haruna, H., Arrona-Palacios, A., Quintero, H. N., Ortega, L. O. P., Sanchez, A. L., Ortiz, E. A., Escamilla, J., & Hosseini, S. (2022). Impact of digital technologies upon teaching and learning in higher education in Latin America: An outlook on the reach, barriers, and bottlenecks. *Education and Information Technologies*, 28(2), 2291–2360. <https://doi.org/10.1007/s10639-022-11214-1>
- Otero, L. C., Catalá, A., Fernández-Morante, C., Taboada, M., López, B. C., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, 10(1). <https://doi.org/10.1186/s40594-023-00418-7>
- Pallant, J. (2020). *SPSS Survival Manual*. Routledge. <https://doi.org/10.4324/9781003117452>
- Permatasari, I. R., Wulandari, R. M., Palupi, D., Faizin, H. A., & Adji, B. P. W. (2025). Examining Students' Perceptions of a Smalltalk2Me powered conversation assistant in the speaking class. *Educational Dynamics International Journal of Education and Social Sciences*, 2(4), 16–23. <https://doi.org/10.70062/educationaldynamics.v2i4.243>
- Petrić, B., & Czár, B. (2003). Validating a writing strategy questionnaire. *System*, 31(2), 187–215. [https://doi.org/10.1016/s0346-251x\(03\)00020-4](https://doi.org/10.1016/s0346-251x(03)00020-4)
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Raoofi, S., Miri, A., Gharibi, J., & Malaki, B. (2017). Assessing and validating a writing strategy scale for undergraduate students. *Journal of Language Teaching and Research*, 8(3), 624–624. <https://doi.org/10.17507/jltr.0803.23>
- Salameh, W. A. K. (2024). The dual role of AI in personalized learning: Enhancements and hindrances for diverse learners. *International Journal of Science and Research Archive*, 13(2), 62–67. <https://doi.org/10.30574/ijsra.2024.13.2.2057>
- Seraj, P. M. I., Habil, H., & Hasan, Md. K. (2020). Investigating EFL learners' readiness of using smartphone and problems for learning oral English communication skills at tertiary level in an EFL Context. *International Journal of Interactive Mobile Technologies*, 14(14), 148–148. <https://doi.org/10.3991/ijim.v14i14.14759>
- Seraj, P. M. I., Habil, H., & Hasan, Md. K. (2021). Investigating the problems of teaching oral English communication skills in an EFL context at the tertiary level. *International Journal of Instruction*, 14(2), 501–516. <https://doi.org/10.29333/iji.2021.14228a>

- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Quasi-experiments: interrupted time-series designs. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, 1, 171–205. Houghton-Mifflin.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*, 15, 165–179.
- Viberg, O., Wasson, B., & Kukulska-Hulme, A. (2020). Mobile-assisted language learning through learning analytics for self-regulated learning (MALLAS): A conceptual framework. *Australasian Journal of Educational Technology*, 36(6), 34–52. <https://doi.org/10.14742/ajet.6494>
- Wang, Z., & Han, F. (2022). The effects of teacher feedback and automated feedback on cognitive and psychological aspects of foreign language writing: Mixed-methods research. *Frontiers in Psychology*, 13, 909802–909802. <https://doi.org/10.3389/fpsyg.2022.909802>
- Wicaksono, C., Afidah, N., Ma'arif, I. B., & Agustina, U. W. (2025). The effect of SmallTalk2Me in English instruction at MTs Pesantren Al-Amin Sooko Mojokerto. *Application: Applied Science in Learning Research*, 4(3), 101–107. <https://doi.org/10.32764/application.v4i3.5578>
- Wu, T., Hapsari, I. P., & Huang, Y. (2025). Effects of incorporating AI chatbots into think-pair-share activities on EFL speaking anxiety, language enjoyment, and speaking performance. *Computer Assisted Language Learning*, 39, 1408–1446. <https://doi.org/10.1080/09588221.2025.2478271>
- Yan, H., & Singh, M. K. S. (2026). The impact of AI-mediated instruction on speaking proficiency, enjoyment, anxiety, and emotional engagement: A mixed-methods approach. *Humanities and Social Sciences Communications*, 13(1). <https://doi.org/10.1057/s41599-026-06705-2>
- Zhou, Y., He, Y., Wang, X., Song, L., Liu, X., & Liu, W. (2025). AI-based real-time feedback system for improving English oral proficiency. 349–353. <https://doi.org/10.1109/icdel65868.2025.11193512>
- Zou, B., Lyu, Q., Han, Y., Li, Z., & Zhang, W. (2023). Exploring students' acceptance of an artificial intelligence speech evaluation program for EFL speaking practice: An application of the integrated model of technology acceptance. *Computer Assisted Language Learning*, 38, 1366–1391. <https://doi.org/10.1080/09588221.2023.2278608>

Appendix 1

Table A1: Shapiro-Wilk Tests of Normality for the Outcome Variables

Outcome Variable	<i>n</i>	Shapiro-Wilk (<i>W</i>)	<i>p</i>	Interpretation
Pretest	28	0.973	.671	Assumption of normality met
Posttest	28	0.968	.512	Assumption of normality met
WPM – Practice 1	28	0.961	.387	Assumption of normality met
WPM – Practice 6	28	0.976	.742	Assumption of normality met

Note. Normality was assessed using the Shapiro-Wilk test because of the relatively small sample size ($n = 28$) (Shapiro & Wilk, 1965; Field, 2018). None of the variables showed a statistically significant departure from normality ($p > .05$), indicating that the assumption of normality was satisfied. Accordingly, paired samples *t*-tests were considered appropriate for subsequent analyses.