

ARTICLE

Efficacy of AI-Generated Feedback by SmallTalk2Me for Improving Speaking Skill of Saudi EFL Learners

Shahzad ul Hassan Farooqi 

Department of English Language, College of Education, Majmaah University, Majmaah 11952, Saudi Arabia

ABSTRACT

Feedback is essential in language teaching and learning as it helps learners identify errors, reinforce correct usage, and guide their progress. Traditionally it was the job of a teacher, however, with the advent of Artificial Intelligence (AI) this task has become automated. This study evaluated the efficacy of AI-generated feedback provided by the SmallTalk2Me platform for improving speaking skill among Saudi EFL learners. Forty-four Saudi EFL learners participated in this quasi-experimental study. The data were derived from the IELTS speaking pre-test and post-test, while in between, the participants completed six practice activities over a period of three weeks (two activities each week). Two data sets were compiled: IELTS speaking band scores and proficiency (measured by words spoken per minute). The data were analyzed using paired t-tests to evaluate statistical significance. The study hypothesized that learners would achieve a higher mean score in the IELTS speaking post-test after completing language practices on SmallTalk2Me. This hypothesis was statistically supported, as the difference between the pre-test and post-test mean scores was significant ($p < 0.05$). The learners demonstrated a 12.12% improvement in their IELTS speaking test band scores, while an 11.18% increase was observed in their speaking proficiency between the first and the sixth practice activities. Following the intervention, a 25-item questionnaire was administered to evaluate the participants' perception of AI-generated feedback, the efficacy of SmallTalk2Me, and the overall learning experience. Students found AI-generated feedback effective, motivating, and engaging. The study highlights the efficacy of AI-generated feedback provided by SmallTalk2Me and recommends the use of such AI platforms in the Saudi EFL scenario.

Keywords: Artificial Intelligence; Feedback; IELTS; Saudi EFL Learners; Speaking Skill; SmallTalk2Me

*CORRESPONDING AUTHOR:

Department of English Language, College of Education, Majmaah University, Majmaah 11952, Saudi Arabia; Email: s.farooqi@mu.edu.sa

ARTICLE INFO

Received: 13 December 2024 | Revised: 20 January 2025 | Accepted: 23 January 2025 | Published Online: 14 March 2025

DOI: <https://doi.org/10.30564/fls.v7i3.8294>

CITATION

Farooqi, S.u.H., 2025. Efficacy of AI-Generated Feedback by SmallTalk2Me for Improving Speaking Skill of Saudi EFL Learners. Forum for Linguistic Studies. 7(3): 714–728. DOI: <https://doi.org/10.30564/fls.v7i3.8294>

COPYRIGHT

Copyright © 2025 by the author(s). Published by Bilingual Publishing Group. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License (<https://creativecommons.org/licenses/by-nc/4.0/>).

1. Introduction

Feedback refers to information that is provided to an individual for the improvement of his/her performance. Brown^[1] emphasizes that “Feedback is a fundamental aspect of communicative competence, enabling learners to bridge the gap between their current performance and target proficiency.” Supporting this view Li^[2] highlights feedback’s positive implications for teaching and learning, noting that it not only identifies errors and provides corrective strategies but also enhances learners’ self-awareness and motivation in language learning. Traditionally, the provision of feedback has been a core responsibility of the teacher. However, with recent advances in artificial intelligence (AI), automated feedback systems have emerged as a viable alternative approach. Shi and Aryadoust^[3] observe that automated feedback is gaining popularity in language pedagogy. This development raises critical questions as to how much feedback is effective in developing language skills? Can artificial intelligence supplement and strengthen the teacher’s role or replace it? While experts are grappling with these questions, there is a general consensus that Artificial Intelligence is the essence of human understanding, processed and delivered by computer systems. Russell & Norvig^[4] believe that it functions as a virtual assistant that imitates the human cognitive process.

Numerous online programs now provide automated feedback for language inputs. Among these, *Grammarly*, *ProWritingAid*, and *Linguix* specialize in AI-generated writing feedback, while *Babbel*, *Rosetta Stone*, *Duolingo*, *Speech Evaluation*, and *SmallTalk2Me* are AI speech evaluation programs that provide automated assessment and feedback regarding vocabulary usage, sentence structure, and fluency. As AI technology is continuously evolving, the list of such programs is getting longer and longer every day. Recent research has extensively evaluated efficacy of these tools. Zou et al.^[5] explored AI speech evaluation tools in language learning contexts, specially their effectiveness in developing EFL learners’ speaking skills. Lehman et al.^[6] also note increasing popularity of AI speech evaluation programs in English language education. Studies indicate that advancement of AI in automatic speech recognition and speech evaluation is helping students improve their skills^[7, 8]. Multiple studies confirm their effectiveness for enhancing second and

foreign language learners’ speaking abilities^[9–12]. Simultaneously, both the teachers and the students find these tools user-friendly, unbiased, engaging, and effective in enhancing fluency^[13–15]. Learners feel less anxiety and fear of being error-hunted because they are evaluated by impartial machines rather than their judgmental peers, or critical teachers. Bashori et al.^[16] specially noted reduction of anxiety among language learners due to the use of AI-based feedback systems.

In the Saudi EFL context, speaking skills present significant challenges. Al-Harbi^[17] identifies speech anxiety as a major barrier to effective communication among Saudi learners. Al-Hazmi and Scholfield^[18] attribute these challenges to factors like limited practice and minimal exposure to English outside the classroom. Al-Sibai^[19] identifies lack of vocabulary and fluency in conversation as common problems. Using AI feedback can help improve the language learning process in the Saudi EFL scenario. With this background in mind, the researcher conducted this study by using *SmallTalk2Me*^[20] with an aim to evaluate the efficacy of AI-generated feedback for enhancing the speaking skill of Saudi EFL learners through this platform. *SmallTalk2Me* is an AI-powered interactive self-practice simulator platform that provides online practice and feedback and trains users for IELTS speaking tests and job market interviews. Its automated assessment, vocabulary and sentence structure feedback, and self-graphed visual presentation of results help learners monitor their progress. The platform claims to have helped 100000 people worldwide improve English since its launch in 2022. Very few studies have been conducted to evaluate this platform’s efficacy. Manggiasih et al.^[21] discussed the strengths and limitations of the *SmallTalk2Me* platform for English language proficiency in the Indonesian EFL scenario. Based on interviews with three experienced English teachers, the study identifies several strengths of the platform, including objective and consistent evaluation metrics, immediate feedback, accessibility, and reduced human bias. However, their research also highlights limitations, such as difficulties in accurately assessing diverse accents and speech patterns, and the absence of human interaction. The findings suggest that *SmallTalk2Me* offers valuable tools for language proficiency evaluation.

The lack of research on the effectiveness of *SmallTalk2Me* in the Saudi EFL scenario presents a sig-

nificant research gap, keeping in view the challenges faced by Saudi learners. Hence, this study aims to address this gap by evaluating the efficacy of this platform in enhancing Saudi EFL learners' speaking skills.

Research Questions

- 1- What will be the effect of AI-generated feedback by SmallTalk2Me on the speaking skill of Saudi EFL learners?
- 2- What will be its effect on the fluency of Saudi EFL learners?
- 3- What perception do Saudi EFL learners have regarding SmallTalk2Me and such AI-based platforms and their efficacy for enhancing speaking skill?

2. Literature Review

Communication in the English language has been a daunting challenge for a majority of EFL students worldwide. While language skill development relied heavily on teacher-led instruction, AI is now transforming this paradigm by fostering learner autonomy and self-motivation. A variety of AI tools, gadgets, and online platforms are now available for enhancing fluency and accuracy in speech. Recent studies demonstrate the efficacy of AI tools for strengthening speaking skills, improving pronunciation, and enhancing the confidence level of students. Macias et.al.^[22] synthesized various studies examining AI's potential for improving English language learning.

2.1. Conversational Practice with Chatbots and Pronunciation Tutoring Systems

Garcia et al.^[23] noted that conversational practice with chatbots can improve speaking skills. Their study demonstrated that chatbots can provide realistic conversational experience. However, their study underlined the need to improve the variety of chatbots' responses and make them adapt to different learning levels and styles. This study builds upon earlier research by Li and Chen^[24] who evaluated the efficacy of AI-based pronunciation tutoring systems in enhancing the oral competence of English language students.

2.2. Voice Recognition Systems

Smith and Johnson^[25] conducted experimental research examining the impact of voice recognition systems on pronunciation. Their research findings revealed that students utilizing these systems were better in pronunciation than those who did not. However, the researchers pointed out that these systems had limitations, particularly in the recognition of non-native accents, and in the provision of comprehensive and accurate feedback.

2.3. Virtual Language Assistants

Zhang and Wang^[26] evaluated the efficacy of AI-powered virtual assistants for enhancing speaking fluency and found promising results. By that time this study was conducted, virtual assistant technology had advanced enough to provide immediate and personalized feedback not only in English language but also in multiple languages. For example, Loewen et al.^[27] studied the efficacy of *Duolingo* for learning Turkish as a second language. In a related research, Loewen et al.^[28] studied the impact of *Babbel* on language learning and demonstrated that daily practice of 10 to 15 minutes on the app resulted in improvement between pre and post-test results.

2.4. AI Integration with English Language Teaching

Efforts to integrate AI-driven speech recognition technology into English language teaching started as early as 2016. Wu and Liu^[29] saw a positive impact of AI on students' participation and motivation in speaking activities. However, they were apprehensive of problems like internet speed and lack of personalized feedback for every individual. A similar study was conducted by Warman et al.^[30] wherein AI-powered feedback resulted in improved pronunciation as well as the confidence level of introverted students. Yin and Wei^[31] advanced this field by integrating language acquisition psychology with educational technologies to develop an intelligent learning system for improving college students' listening and speaking skills. AI integration with language instruction was further examined across diverse educational

contexts, like Haong et al.^[32] who noted improvement in vocational students' English pronunciation and Zou et al.^[33] who witnessed development in speaking skills of students of EAP.

2.5. Effect of AI on Saudi EFL Learners' Speaking Skills

Recent studies have shown the potential of AI-based tools in enhancing English speaking skills among Saudi EFL learners. Makhoul's^[34] study at Albaha University investigated the impact of the AI-based mobile app *ELSA Speak* on students' English-speaking proficiency. This pioneering research in the Saudi EFL scenario revealed the efficacy of AI-generated feedback in improving fluency and accuracy. Additionally, the use of AI technology promoted self-regulated and independent learning, enabling students to study on their own at any time and from any location. The study was focused and limited to *ELSA Speak* and promising results of the study called for further investigation into AI usage in the Saudi scenario. Building upon the same foundation, Alsadoon^[35] conducted research on the use of an interactive storytelling chatbot to improve students' communication abilities. Students at the British Council in Riyadh who were EFL low-intermediate level were the focus of the study. A chatbot was introduced as a human-like AI practice tool that could be used at any time to improve English.

With its four features—dictionary, translation, image, and concordance device, this AI tool enabled users to analyze the usage of words and phrases in conversations. According to the study, chatbots were a useful AI tool for enhancing pupils' English-speaking skill. Further contributing to this body of research, Alhalangy, and AbdAlgane^[36] examined the impact of AI on the EFL context within Saudi universities. Their study emphasized the need to increase awareness among teachers and students regarding the integration of AI applications and tools in language education. Most recently, Alsaif^[37] explored the role of AI in developing English language skills in the Saudi EFL context and discussed how AI technologies could provide learners with opportunities to actively engage in speaking activities and receive instant feedback, thereby enhancing their speaking proficiency. These studies collectively underscore the potential of AI-generated feedback in improving speaking skills among Saudi EFL learners, advocating for further integration of AI-based plat-

forms in language education.

2.6. Ways of Providing Feedback

AI-based language learning platforms offer a variety of activities to enhance speaking skills. Platforms like *Duolingo* and *Babbel* offer diverse activities including pronunciation exercises, reading aloud, interactive conversations, and role-play scenarios. Specialized platforms like *SmallTalk2Me* offer targeted preparation for standardized English proficiency tests, including IELTS, while providing structured practice for conversational English and job interviews. These AI-powered platforms provide feedback in a variety of ways. For speaking skills, feedback includes students' scores and band-level assessment, colour-coded graphic visualization, model answers, alternative phrasing suggestions, and syntactic evaluation^[5]. Pronunciation feedback may include phonemic accuracy, syllable stress pattern, and intonation, while fluency assessment may focus on speech rate analysis (words spoken per minute), response latency analysis. As the field of AI-based feedback is ever-expanding, the need to continuously evaluate these platforms and provide empirical evidence about their efficacy is all the more compelling (Burston^[38], Zhang & Zou^[39]). Given the global proliferation of research on AI-generated feedback and speech evaluation systems, this study aims to investigate the efficacy of AI-based feedback by *SmallTalk2Me* for enhancing the speaking skills of Saudi EFL learners.

2.7. SmallTalk2Me

SmallTalk2Me is an AI-driven platform designed to train learners for IELTS tests and job interviews while providing comprehensive language proficiency assessment. While discussing the strengths of this platform, Manggiasih et al.^[21] mention several distinctive features of the platform particularly its standardized evaluation metrics free from any human biases. While discussing platform's AI-generated feedback they highlight the objectivity and consistency of the evaluation process. They appreciate the quickness of the feedback as learners receive an instant evaluation of their weak areas including pronunciation, vocabulary, and fluency. Following **Figure 1** summarizes the strengths of the platform as described by them.

Exploring the Multifaceted Benefits of Smalltalk2me

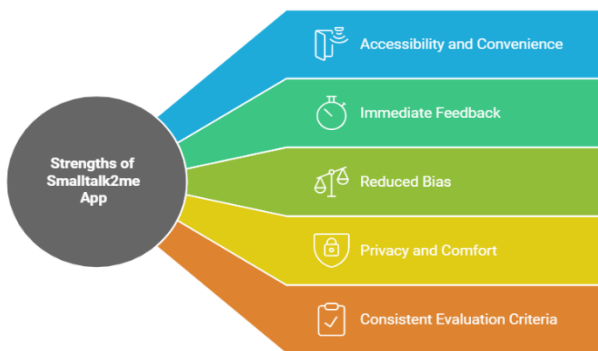


Figure 1. Strengths of SmallTalk2Me.

The platform also allows remote access and flexible usage to a large number of users which is the hallmark of AI-driven language evaluation tools (Gao et al. ^[40]).

However, like other AI-driven tools, SmallTalk2Me has several inherent limitations particularly in speech recognition and pronunciation variations handling which can affect the evaluation accuracy (**Figure 2**). Manggiasih et al. ^[21] observe that although an app can evaluate speech, it cannot replicate the “intricacies of human interaction”. They underline the inability of the AI to quantify the “emotional resonance” because real-world conversations also have non-verbal clues and social and cultural dynamics which are necessary for proper understanding of communication. Besides that, there is an engagement barrier as well. A human instructor’s feedback extends beyond technicalities and fixed format and also incorporates encouragement, motivation, and emotional support, which an online platform lacks because an online AI-platform operates within its technical parameters. Furthermore, Manggiasih et al. ^[21] mentioned technical challenges including system lags, connection interruptions, internet stability issues, and audio disruptions. It is worth noting that these problems are not limited to SmallTalk2Me, rather they are a challenge for every AI-driven online language evaluation platform.

Based on their results, Manggiasih et al. ^[21] advocate a hybrid approach that combines AI’s efficiency with human expertise to address these challenges and enhance the efficacy of assessment process. Ultimately, AI-driven language evaluation platforms like SmallTalk2Me need to achieve an equilibrium between their technical capabilities and pedagogical requirements. Only a balanced approach would

help develop more comprehensive, inclusive, and effective language assessment systems that have the benefit of both artificial as well as human intelligence.

Exploring the Challenges of Smalltalk2me



Figure 2. Weaknesses of SmallTalk2Me.

2.8. Different Styles of Feedback?

SmallTalk2Me provides feedback in multiple formats and categories. Besides different courses about speaking, job simulation interviews, Business English, dialogical practices, and Grammar instructions, it provides structured practices for IELTS. Its IELTS modules are based on a multilayered feedback strategy to achieve high band scores (**Figure 3**).

IELTS Success Pyramid

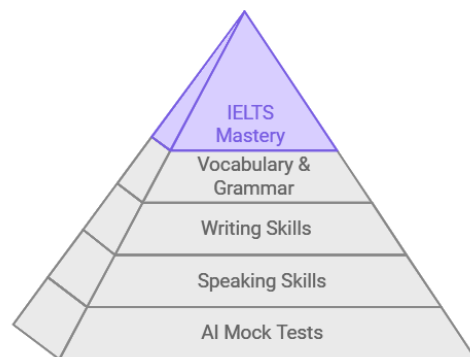


Figure 3. Multi-layered feedback provided by SmallTalk2Me.

In IELTS Speaking tests, SmallTalk2Me first displays the band score, detailing the level of the student, areas of strength, and weaknesses besides a score of how much the native speaker would understand the speakers’ talk. Vocabulary analysis includes the number of frequently used words, rare words, and unique words alongside statistics of vocabulary divided into categories of Beginner (A1), Elementary (A2),

Intermediate (B1), Upper-intermediate (B2), Advanced (C1) and Proficiency (C2) (Figure 4). A “boost your vocabulary” section provides synonyms that may be replaced to improve the writing. For writing tasks, similar multilayered feedback is provided.

However, separate feedback is provided in the speaking test regarding pronunciation and fluency (words spoken in one minute) which reflects the speed of the learners’ talk compared with native American speaker (Figure 5).

The grammar section highlights the mistakes and categorizes them according to grammar rules (Figure 6).

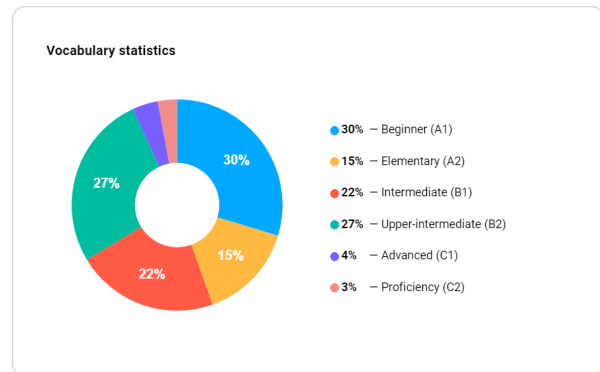


Figure 4. Vocabulary analysis according to A1, A2, B1, B2, C1, C2 levels.

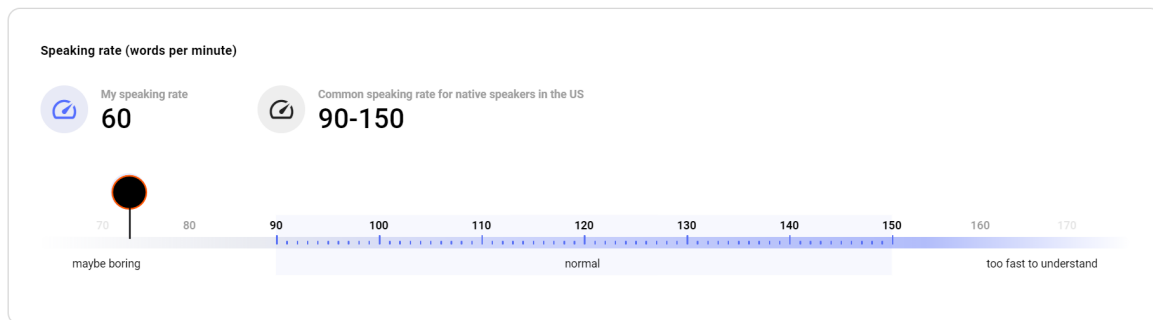


Figure 5. Fluency rate: words spoken per minute.



Figure 6. Grammar feedback.

In this study the researcher employed two key parameters to measure participants’ speaking skills: (1)-compared the band score of IELTS pre and post-tests, and (2)-evaluated fluency rating in the practice tests *How to Make Great Small Talk* provided by SmallTalk2Me to measure the hypothesized improvement in speaking fluency of the participants between 1st and 6th practice session. These two parameters were considered appropriate indicators to measure learners’ speaking skill development in the short period during which

the study was conducted.

3. Methodology

Forty-four male Saudi undergraduate students majoring in English Language participated in this study. These participants belonged to the fourth (82%) and fifth (18%) semesters of English language major at Al-Majmaah University. Using criterion sampling participants were selected on

the basis of their Grade Point Average (GPA) ranging from 3.5 to 4.5 ensuring that the participants had a homogeneous level of language proficiency for standardized pre and post-test administration. Participants created their individual accounts on the SmallTalk2Me platform and completed IELTS speaking practice as a pre-test and post-test, which generated performance reports along with the band score. During the three weeks between IELTS pre and post-tests, students were asked to complete six practice tests using the platform's *How to Make Great Small Talk* module and submit the report to the researcher. These reports were systematically collected through the University's Learning Management System (LMS) and their scores were tabulated and statistically analyzed. For statistical analysis, the researcher used IBM SPSS version 30.0.0.

3.1. Social Validity Scale (Questionnaire)

The third data source was a 25-item questionnaire based on the Social Validity Scale for students (Intervention Rating Profile (IRP-15)^[41]. The instrument was designed to determine participants' perception towards the AI-based feedback and effectiveness of SmallTalk2Me platform for improving their speaking skill. The questionnaire was administered via Google Forms, a convenient and accessible online platform for data collection. While drafting the questionnaire, the researcher sought guidance from established research instruments, including Petrić and Czár's^[42] writing strategies questionnaire, and Raoofi et al.'s^[43] work to ensure that the questionnaire was in line with the contemporary research and well optimized for effective data collection. The following pivotal aspects guided the drafting of the questionnaire:

- 1- Users' perception and overall experience of the AI-generated feedback.
- 2- The acceptability, feasibility, and effectiveness of feedback provided by SmallTalk2Me platform.
- 3- Students' Motivation and willingness to use AI-based platforms in the future.

The questionnaire underwent expert review and standardization to ensure its cultural relevance and appropriateness for the Saudi cultural context. In order to facilitate the understanding and minimize the potential comprehension obstacles, the questionnaire was presented in English

and Arabic languages. The primary objective of the questionnaire was to measure students' social response to the AI-generated feedback by SmallTalk2Me, and any reading comprehension confusion might have affected the validity of the results. A brief pilot study was conducted to identify and address any linguistic or comprehension ambiguities in the questionnaire. Later, Cronbach's alpha was calculated to assess the internal consistency of the questionnaire. The 25-item questionnaire demonstrated an excellent level of reliability, with a Cronbach's alpha of $\alpha = 0.94$, indicating a highly reliable measure. The purpose of utilizing diverse data sources, i.e., pre and post-tests and the questionnaire was to broaden the scope of the study and enhance validity (Zohrabi,^[44]). This comprehensive approach enabled the researcher to present more thorough analysis and results. The data collection process followed a systematic 3-step protocol as shown in the following **Figure 7**.

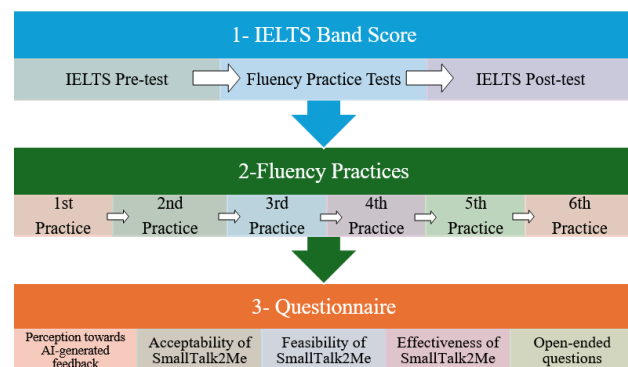


Figure 7. 3-step flow chart of the data collection.

3.2. Delimitations

The quasi-experimental study is based on SmallTalk2Me, which is an AI-driven language training platform. The research design incorporated a pre-test and post-test to evaluate the effectiveness of a three-week-long speaking intervention during which participants engaged in speaking practice twice a week. The study's scope is limited to the students of the English Language Department of Al-Majmaah University Saudi Arabia, and it is specific to the Saudi EFL scenario.

4. Data Analysis

Data collection comprised three sources: IELTS speaking pre and post-tests, first and sixth practice tests of the

SmallTalk2Me portal's *How to Make Great Small Talk* module, and a 25-item questionnaire. Statistical analyses included the calculation of means, standard deviations, mean difference, and percentage differences. A paired sample t-test was conducted to find the statistical significance (p-value) of results. Additionally, descriptive statistics were calculated for the participants' general perceptions towards using the AI-based feedback platform SmallTalk2Me and its efficacy for improving speaking skill.

4.1. IELTS Band Score, Pre and Post-Test Results

The results from paired-sample t-tests to compare the mean band scores of the pre-tests and the post-tests confirmed participants' improvement in IELTS speaking tests.

As shown in **Table 1**, the mean score of the IELTS speaking test increased by 0.70 from the pre-test ($M = 5.86$, $SD = 0.70$) to the post-test ($M = 6.57$, $SD = 0.68$) which is an impressive 12.12% improvement.

The result is statistically significant ($p < 0.05$); hence, we reject the null hypothesis. This indicates that there is a statistically significant difference between the pre-test and the post-test band scores which means that after going through six speaking practices for three weeks, twice a week, the participants improved their band score.

4.2. Fluency Comparison of the First and the Sixth Practice Tests

Here is the statistical comparison of the first and the sixth Practice Tests (**Table 2**).

Table 1. IELTS band score, pre-test and post-test comparison.

| | <i>N</i> | <i>M</i> | <i>SD</i> | <i>t</i> | <i>mean</i> | <i>%Difference</i> | <i>p-Value</i> |
|------------------|----------|----------|-----------|----------|-------------|--------------------|----------------|
| Pre-test | 44 | 5.86 | 0.70 | 5.99 | 0.71 | 12.12 | 0.000000378 |
| Post-Test | 44 | 6.57 | 0.68 | | | | |

$p < 0.05$.

Table 2. Fluency comparison in 1st and 6th Practice.

| | <i>N</i> | <i>M</i> | <i>SD</i> | <i>t</i> | <i>mean</i> | <i>%Difference</i> | <i>p-Value</i> |
|---------------------|----------|----------|-----------|----------|-------------|--------------------|----------------|
| 1st Practice | 44 | 88.20 | 16.25 | 7.38 | 9.86 | 11.18 | 0.0000000364 |
| 6th Practice | 44 | 98.06 | 18.38 | | | | |

$p < 0.05$.

Analysis of fluency score reveals substantial improvement. The fluency score increased from pre-test ($M = 88.20$, $SD = 16.25$) in the first practice to ($M = 98.06$, $SD = 18.38$) in the final practice. The difference in means was recorded 9.86, which reflects an 11.18% improvement in words spoken per minute. Statistical analysis indicates significant difference between the first and the sixth practice performance ($p < 0.05$). This improvement in fluency affected the IELTS post-test performance and we have already seen that the mean band score of the IELTS post-test witnessed a 12.12 percent improvement. Based on this analysis we can say that practice tests provided by the SmallTalk2Me platform and the resultant feedback made a significant impact on the speaking skills of the students.

The following analysis (**Figure 8**) presents a comparison of the percentage of students who improved, remained the same, or decreased their scores in IELTS band score from

pre to post-tests and fluency practice tests from the first to the sixth test.

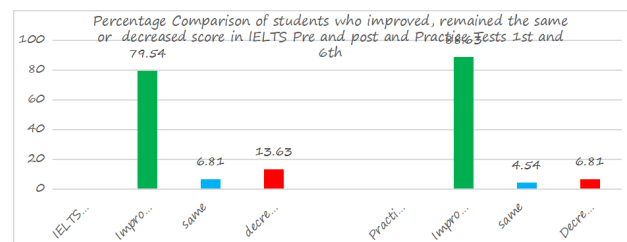


Figure 8. Percentage comparison.

As **Figure 8** illustrates, 79.54% of participants improved their IELTS band score in post-tests, while 6.81% did not witness any change, there are still 13.63% who rather decreased their band score in the post-test. A parallel trend is witnessed in the case of practice tests, with 88.63% of students improving their fluency output while 4% remained

at the same output of words in one minute while 6.81% decreased their performance. Overall, the comparison reveals that the majority of the students improved their speaking performance with more practice. We can further deduce from this analysis that the fluency improvement in the fluency practices led to better results in the IELTS band score post-tests because students' initial IELTS performance (Pre-Test) established the baseline before receiving any fluency training. When they underwent the "treatment" or intervention for fluency improvement on SmallTalk2Me their results improved. The only significant instructional difference between the Pre-Test and the Post-Test was exposure to and use of this platform. After the intervention period students took the IELTS again (Post-Test) and a large majority (79.54% of participants) improved their IELTS band score. Hence, it can be inferred that the intervention contributed to the improvement.

4.3. Questionnaire

The analysis of questionnaire data reveals significant information regarding participants' perception towards AI-generated feedback and usage of the SmallTalk2Me platform. The data shows that majority of them (63.63%) have previous experience of AI tools while 36.36% indicated no previous exposure to such tools. This means growing integration of AI tools in Saudi education as more and more students are turning to AI platforms for improving their language skills.

Participants generally rated the AI-generated feedback and SmallTalk2Me platform positively, as reflected by high mean scores (Overall Mean of the questionnaire = 4.12) suggesting strong approval and acceptability (Table 3).

As the questionnaire was meant to measure four key dimensions: perception towards AI, acceptability, feasibility, and effectiveness of the SmallTalk2Me platform, the mean score of most categories is above 4.12 except the feasibility category with a mean score ($M = 4.07$) (Figure 9). It reflects that the participants have some issues regarding the ease of use of the SmallTalk2Me platform. Under the open-ended question 24 (What improvements would you suggest for the platform SmallTalk2Me?) participants identified some operational problems encountered during language practices on the SmallTalk2Me platform. These functionality related issues need further elaboration which we will discuss in subsequent sections.

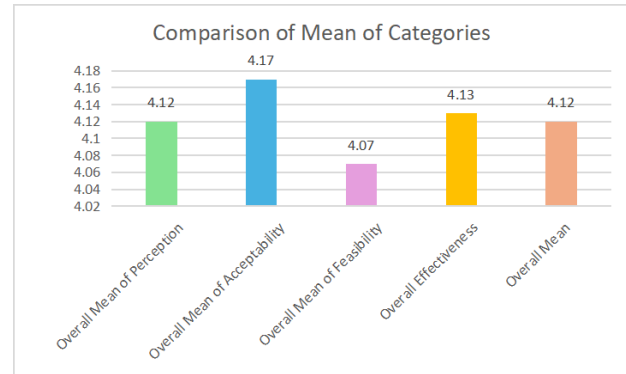


Figure 9. Mean comparison of perception, acceptability, feasibility, and effectiveness.

Regarding their experience with AI-generated feedback, majority of the participants opined under question 2 that it helped to improve their overall speaking skill. The data reveals a clear perception towards feedback's clarity and comprehensibility as is shown by a high mean score ($M = 4.38$, $SD = 0.72$) for question 3 and a similar surging mean score ($M = 4.14$, $SD = 0.96$) for question 20. Responses indicate that the AI feedback was task relevant and helped identify specific errors, with mean scores ($M = 4.04$, $SD = 0.91$) in question 5 and a mean score ($M = 4.09$, $SD = 1.05$) in question 4.

Some psychological aspects of language learning were also discerned in the questionnaire data. Anxiety of being error-hunted is often seen to hampers students' learning, as they become reluctant to receive feedback. However, SmallTalk2Me platform appears to address this problem effectively. Participants positively confirmed that the feedback provided by SmallTalk2Me reduced anxiety about making mistakes (question 8: $M = 4.20$, $SD = 1.04$) and increased motivation to improve (question 6: $M = 4.11$, $SD = 1.10$), besides encouraging them to experiment with more speaking techniques. (question 7: $M = 4.20$, $SD = 0.90$). These findings suggest that AI-generated feedback created a conducive learning environment.

The acceptability of SmallTalk2Me, with the highest mean score ($M = 4.17$), reflects an overwhelmingly positive response from the participants. They also showed a strong commitment to continue practice on the platform in future and recommend it to others.

As discussed previously, the feasibility of SmallTalk2Me with a mean score of ($M = 4.07$), reflects an area where the ease-of-use posed a problem for the Saudi EFL learners. How-

ever, it requires more research in diverse EFL scenarios to determine the ease-of-use related problems encountered by EFL learners from different linguistic communities.

Regarding the effectiveness of SmallTalk2Me (with a mean score of $M = 4.13$), participants believed that the

feedback helped them point out those specific areas of speaking skill where they needed improvement. After taking the IELTS tests and speaking practices on SmallTalk2Me platform, they expressed hoped and confidence to achieve higher scores in actual IELTS examination.

Table 3. Questionnaire data.

| Questions | Yes% | No% | | | | | |
|--|-------------------------|----------------|------------------|-------------------|---------------------------|-------------|---------------------------|
| 1- Previous experience with AI tools | 63.63 | 36.36 | | | | | |
| Perception towards AI-generated feedback | Strongly Agree % | Agree % | Neutral % | Disagree % | Strongly Disagree% | Mean | Standard Deviation |
| 2- The AI feedback helped improve my overall English speaking skills. | 45.45 | 25 | 22.72 | 4.54 | 2.27 | 4.06 | 1.04 |
| 3- The feedback provided by the AI was clear and easy to understand. | 50 | 40.90 | 6.81 | 2.27 | 0 | 4.38 | 0.72 |
| 4- The AI-generated feedback helped me identify specific errors in pronunciation and vocabulary. | 45.45 | 29.54 | 15.90 | 6.81 | 2.27 | 4.09 | 1.05 |
| 5- The feedback was relevant to the tasks or assignments given. | 38.63 | 31.81 | 25 | 4.54 | 0 | 4.04 | 0.91 |
| 6- The feedback motivated me to improve my English skills. | 47.72 | 29.54 | 13.63 | 4.54 | 4.54 | 4.11 | 1.10 |
| 7- I felt encouraged by the feedback to try more and more speaking techniques. | 43.18 | 40.90 | 11.36 | 2.27 | 2.27 | 4.20 | 0.90 |
| 8- Receiving AI feedback reduced my anxiety about making mistakes. | 52.27 | 27.27 | 11.36 | 6.81 | 2.27 | 4.20 | 1.04 |
| 9- I trusted the AI feedback to be accurate. | 38.63 | 29.54 | 20.45 | 4.54 | 6.81 | 3.88 | 1.18 |
| 10- The AI correctly identified my mistakes most of the time. | 43.18 | 34.09 | 20.45 | 0 | 2.27 | 4.15 | 0.91 |
| Overall Mean of Perception | | | | | | 4.12 | |
| Acceptability of SmallTalk2Me | | | | | | | |
| 11- This SmallTalk2Me IELTS speaking activity is an acceptable way to improve my IELTS speaking skills. | 45.45 | 38.63 | 11.36 | 0 | 4.54 | 4.20 | 0.97 |
| 12- I would be willing to use this activity regularly to prepare for the IELTS speaking test. | 38.63 | 38.63 | 13.63 | 9.09 | 0 | 4.06 | 0.94 |
| 13- I would recommend this activity and feedback system to other students preparing for the IELTS speaking test. | 47.72 | 38.63 | 6.81 | 4.54 | 2.27 | 4.25 | 0.94 |
| Overall Mean of Acceptability | | | | | | 4.17 | |
| Feasibility of SmallTalk2Me | | | | | | | |
| 14- The platform SmallTalk2Me was easy to use and navigate. | 47.72 | 31.81 | 9.09 | 9.09 | 2.27 | 4.13 | 1.06 |
| 15- I felt confident in using the platform SmallTalk2Me without external help. | 50 | 20.45 | 18.18 | 9.09 | 2.27 | 4.06 | 1.12 |
| 16- This SmallTalk2Me IELTS speaking activity is practical and easy to implement as part of IELTS preparation. | 47.72 | 31.81 | 15.90 | 2.27 | 2.27 | 4.20 | 0.95 |
| 17- The SmallTalk2Me IELTS speaking activity fits well within my study routine and time constraints. | 47.72 | 20.45 | 11.36 | 15.90 | 4.54 | 3.90 | 1.29 |
| Overall Mean of Feasibility | | | | | | 4.07 | |
| Effectiveness of SmallTalk2Me | | | | | | | |
| 18- The AI-generated feedback provided by SmallTalk2Me is helpful in addressing my speaking challenges. | 39.53 | 39.53 | 13.95 | 4.65 | 2.32 | 4.09 | 0.97 |
| 19- I believe that practicing speaking through SmallTalk2Me will lead to meaningful improvements in my test performance. | 44.18 | 37.20 | 11.62 | 2.32 | 4.65 | 4.13 | 1.03 |
| 20- The feedback provided by SmallTalk2Me is clear and easy to understand. | 41.46 | 41.46 | 9.75 | 4.87 | 2.43 | 4.14 | 0.96 |
| 21- The SmallTalk2Me IELTS speaking activity allows me to work on specific speaking areas where I need improvement.- | 45.23 | 38.09 | 14.28 | 0 | 2.38 | 4.23 | 0.87 |
| 22- I feel more confident about my IELTS speaking abilities after using this activity and feedback system. | 35.71 | 45.23 | 14.28 | 2.38 | 2.38 | 4.09 | 0.90 |
| Overall Mean of Effectiveness | | | | | | 4.13 | |
| Overall Mean of the Questionnaire | | | | | | 4.12 | |

Table 3. *Cont.*

| Open-ended questions |
|---|
| 23- What did you find most helpful about the AI-generated feedback? |
| 24- What improvements would you suggest for the platform SmallTalk2Me? |
| 25- How did AI feedback influence your learning experience compared to traditional methods? |

Analysis of responses to the open-ended questions revealed both significant strengths and areas for improvement in the SmallTalk2Me platform. Answering question 23 the participants highlighted platform's good qualities particularly its "ease of use", "immediate correction of errors", "pinpointing of errors and how to solve them" "learning new vocabulary", "enhancement of performance and speed", because the platform "encourages you to provide a creative answer", and "practicing again and again enriches your skill and confidence". While comparing AI-feedback with traditional human teacher feedback under question 25, participants opined that AI feedback was more personalized and immediate, because the AI-platforms are entirely focused on a single student's performance and not the entire class. One participant pinpointed the availability of more attempts to improve results, which is mostly not available in the traditional method in which a teacher offers only one or very limited attempts to perform a task. Participants also appreciated AI's ability to provide immediate error correction as it saved their time, while a human teacher would take more time to give similar feedback. Moreover, AI-feedback could be generated at any time and any place, which was much easier than face-to-face classes and on-campus meetings with the teacher. According to another participant, the AI-generated feedback took away from him the fear of being error hunted by the teacher.

On the negative side, while answering open-ended question 24, some participants underlined the need for incorporating more tailored exercises based on "the preferences and communication styles of the users". One participant suggested including "translation for people who do not speak English so that it is easy to practice and understand". Another participant demanded "diversity of topics on the platform". Yet another proposed to "bring Improvements to the navigation within the test". Some complained of showing insufficient results in some attempts. A few of them pinpointed technical issues like lag, slow internet connectivity, or sometimes portal kicking them out without posting their results.

However, the most common complaint was that the portal sometimes did not display the band score or the practice test score properly and only displayed a general grade.

5. Discussion

The findings of this study provide empirical evidence that AI-generated feedback through the SmallTalk2Me platform can significantly improve Saudi EFL learners' speaking proficiency. The impact is evidenced through two key quantitative metrics: IELTS speaking band score and speech fluency measurements. Results demonstrate a notable improvement of 12.12% in IELTS speaking band scores with mean scores increasing from 5.86 in the pre-test to 6.57 in the post-test. Similarly, 11.18% improvement was witnessed in fluency (words per minute) from the first to the sixth practice test on SmallTalk2Me platform. These quantitative figures are further supported by the qualitative analysis of the questionnaire that highlights some psychological aspects of AI-generated feedback including reduced anxiety, immediate correction of errors, and sustained motivation for continued practice.

5.1. Improvement in Speaking Proficiency

The statistically significant gains in IELTS band scores align with earlier research showing the potential of AI tools in improving speaking skills among EFL learners. For instance, studies by Zou et al.^[5] and Dizon^[10] reported that AI speech evaluation programs had a positive impact on students' oral performance, highlighting that automated feedback can directly contribute to improved speaking competence. Similarly, Makhoul^[34], in the Saudi EFL context, demonstrated how using the AI-based mobile app *ELSA Speak* led to improvements in fluency and accuracy. The present study's findings not only corroborate these results but also extend them by focusing on a relatively less-explored AI tool, SmallTalk2Me, thus filling a gap in the Saudi EFL

research landscape.

Moreover, the improvement in fluency (mean increase of 9.86 words per minute as seen in **Table 2**) supports prior research on AI-driven speech recognition and feedback. Dai and Wu^[9], Wu and Liu^[29], and Smith and Johnson^[25] highlighted that AI-powered systems encourage repeated practice and enhance learners' speed in delivering spontaneous speech. The current study confirms that engaging with multiple practice tests and receiving immediate, itemized feedback builds learners' confidence, thereby reducing their pauses and increasing their speech rate over time.

5.2. Reduction in Anxiety and Increased Motivation

A notable finding in this study is the reduction in learner anxiety when interacting with an AI-based feedback system. Over half of the participants agreed that receiving AI feedback lessened their fears of being judged or "error hunted." This resonates with Bashori et al.^[16], who found that automated feedback can mitigate learners' anxiety by offering private, objective evaluations without the social pressure of teacher or peer scrutiny. In the Saudi EFL context, where Al-Harbi^[17] and Al-Hazmi & Scholfield^[18] have identified speech anxiety and limited practice opportunities as major obstacles, AI platforms like SmallTalk2Me provide a safe space for learners to experiment with language.

Participants also noted that they felt motivated to keep practicing once they saw immediate feedback about specific areas of pronunciation, fluency, and vocabulary use. These perceptions are consistent with the findings of Ding et al.^[13] and Tai and Chen^[15], who reported that real-time feedback from AI tools encourages learners to speak more frequently and enthusiastically, thus improving overall speaking performance.

5.3. Positive Perceptions of AI Feedback

The questionnaire results revealed an overall mean of 4.12 (on a 5-point Likert scale), indicating a generally favorable view toward AI-generated feedback and the SmallTalk2Me platform. This complements studies by Li^[2] and Lehman et al.^[6], which underscore how automated feedback is becoming popular because it fosters learner autonomy,

allowing them to monitor their progress without constantly depending on the teacher. Moreover, participants' remarks about the quickness and personalization of feedback mirror Garcia et al.^[23], who found that chatbots and interactive AI platforms can provide more targeted, individual-centric language practice compared to traditional, teacher-led settings.

This endorsement of AI also parallels the growing adoption of AI-driven learning tools in Saudi Arabia, as noted by Alsadoon^[35] and Alsaif^[37], who reported a shift in teachers' and students' attitudes toward emerging educational technologies. The participants in the current study not only appreciated the objectivity and immediacy of AI feedback but also demonstrated a willingness to integrate such tools into their routine IELTS preparation.

5.4. Feasibility and Challenges

While the acceptability and effectiveness of SmallTalk2Me scored high, the feasibility dimension received a slightly lower mean (4.07), indicating some challenges. Participants reported occasional technical lags, issues with navigation, and internet connectivity as impediments, which echoes the concerns raised by Manggiasih et al.^[21] about the limitations of AI platforms. Technical glitches are not unique to SmallTalk2Me; rather, they are a broader challenge to many online learning systems, as also pointed out by Wu and Liu^[29] and Burston^[38].

Another relevant issue pertains to accent diversity and the nuanced aspects of oral communication that AI sometimes struggles to capture. Manggiasih et al.^[21] caution that AI-driven evaluation may fail to recognize the full range of dialects and speech idiosyncrasies, reducing the accuracy of feedback. Participants in the present study similarly called for more tailored exercises that accommodate diverse learner backgrounds and preferences.

Lastly, some respondents pointed out that while the platform is beneficial, it cannot replicate the emotional support, encouragement, and rapport-building that a human teacher provides—an observation in line with Manggiasih et al.^[21] Such "human touch" elements remain valuable in sustaining motivation and addressing learners' affective factors, suggesting that a blended or hybrid approach—combining AI's efficiency with teachers' empathy—may be most effective.

5.5. Alignment with the Saudi EFL Context

Previous research on Saudi EFL learners, such as Al-Sibai^[19], Al-Harbi^[17], and Al-Hazmi & Scholfield^[18], has consistently identified speaking anxiety, lack of vocabulary, and insufficient exposure to English as key barriers to language proficiency. The improvement in both fluency and confidence in this study shows that AI feedback can play a crucial role in addressing these issues. Furthermore, the findings align with Makhoul^[34], Alsadoon^[35], and Alhalangy and AbdAlgane^[36], who advocated for a more robust adoption of AI in Saudi universities and language centers. Notably, the current study's participants felt that they could practice any time and gain immediate feedback, thereby compensating for the limited real-life opportunities for English use outside the classroom—an issue highlighted by Al-Hazmi & Scholfield^[18].

5.6. Recommendations for Practice and Future Research

The positive results and generally favorable perceptions suggest that AI-driven platforms like SmallTalk2Me could be a valuable supplement to traditional speaking instruction in the Saudi EFL context. Teachers might consider adopting a hybrid model, where human instructors guide and motivate students while AI tools reinforce practice through individualized, timely feedback. This approach addresses the limitations of purely AI-based learning, such as lack of socio-emotional support and potential technical glitches, while still reaping the benefits of instant correction and autonomous practice.

Future research could focus on long-term impact of AI feedback by investigating whether learners retain their fluency gains when they discontinue using AI tools. Another area of research can be accent and pronunciation diversity focusing on how AI platforms handle a wider range of pronunciations, particularly those from different regions of Saudi Arabia. Teacher Integration and teacher-led AI instruction can be another area of interest to examine the optimal balance between automated and human-mediated feedback. It is also recommended that in-depth case studies are required with learners to understand how AI-driven feedback influences language learning strategies and autonomy over the long term.

6. Conclusions

In light of the above discussion, the significant improvements in IELTS speaking scores and fluency, along with predominantly positive learner perceptions, corroborate existing literature on the benefits of AI-based feedback in language education. Hence, we found the answers to our three research questions. For research question 1, the answer is that AI-generated feedback by platforms like SmallTalk2Me can help students improve their band score as we witnessed that practice on the SmallTalk2Me improved IELTS speaking band score. Regarding research question 2, we saw that students' fluency improved after going through six consecutive practices. For research question 3, we found an overall positive perception of Saudi EFL learners towards AI-generated feedback and the SmallTalk2Me platform. Although some technical and usability challenges were noted, the overall impact of SmallTalk2Me on learners' speaking skills is encouraging. Quantitative and qualitative analyses suggest that the SmallTalk2Me platform has capacity to effectively address traditional challenges in EFL instruction. The study also demonstrates that AI-generated feedback can be a powerful tool for increasing learners' fluency, confidence, and motivation. By aligning these findings with prior studies, this research highlights that careful, blended integration of AI platforms in Saudi EFL scenario can help overcome persistent hurdles in speaking skills development—particularly anxiety, lack of practice opportunities, and insufficient feedback—ultimately leading to greater learner autonomy and improved communicative skills.

Funding

The researcher is indebted to the support provided by the Deanship of Scientific Research, Almajmaah University under the project No: R-2025-1548 for the completion of this research.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

This study was conducted in strict adherence to ethical guidelines to ensure the protection and respect of all participants involved.

Data Availability Statement

The data supporting the findings of this study are available from the authors upon request.

Conflict of interests

The Author declares that there is no conflict of interest.

References

- [1] Brown, H.D., 2014. Principles of language learning and teaching: A course in second language acquisition, 6th ed. Routledge: London, UK.
- [2] Li, S., 2013. The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *The Modern Language Journal*. 97(3), 634–654. DOI: <https://doi.org/10.1111/j.1540-4781.2013.12030.x>
- [3] Shi, H., Aryadoust, V., 2024. A systematic review of AI - based automated written feedback research. *ReCALL*. 36(2), 187–209. DOI: <https://doi.org/10.1017/S0958344023000265>
- [4] Russell, S., Norvig, P., 2016. Artificial intelligence: A modern approach, 3rd ed. Pearson Education: London, UK.
- [5] Zou, B., Du, Y., Wang, Z., et al., 2023. An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open*. 13(3). DOI: <https://doi.org/10.1177/21582440231193818>
- [6] Lehman, B., Gu, L., Zhao, J., et al., 2020. Use of adaptive feedback in an app for English language spontaneous speech. In: Bittencourt, I.I., Cukurova, M., Muldner, K., et al., (Eds.). *Artificial intelligence in education*. Springer International Publishing: Cham, Switzerland. Volume 12163, pp. 309–320. DOI: https://doi.org/10.1007/978-3-030-52237-7_25
- [7] Kannan, J., Munday, P., 2018. New trends in second language learning and teaching through the lens of ICT, networked learning, and artificial intelligence. *Círculo de Lingüística Aplicada a La Comunicación*. 76, 13–30. DOI: <https://doi.org/10.5209/CLAC.62495>
- [8] Natale, S., Cooke, H., 2021. Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces. *Media, Culture & Society*. 43(6), 1000–1016. DOI: <https://doi.org/10.1177/0163443720983295>
- [9] Dai, Y., Wu, Z., 2023. Mobile - assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed - methods study. *Computer Assisted Language Learning*. 36(5–6), 861–884. DOI: <https://doi.org/10.1080/09588221.2021.1952272>
- [10] Dizon, G., 2020. Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*. 24(1), 16–26. DOI: <https://doi.org/10.125/44705>
- [11] Wang, Y.-H., Young, S.S.-C., 2015. Effectiveness of feedback for enhancing English pronunciation in an ASR - based CALL system: The effectiveness of corrective feedback. *Journal of Computer Assisted Learning*. 31(6), 493–504.
- [12] Xiao, W., Park, M., 2021. Using automatic speech recognition to facilitate English pronunciation assessment and learning in an EFL context: Pronunciation error diagnosis and pedagogical implications. *International Journal of Computer - Assisted Language Learning and Teaching*. 11(3), 74–91.
- [13] Ding, S., Liberatore, C., Sonsaat, S., et al., 2019. Golden speaker builder – An interactive tool for pronunciation training. *Speech Communication*. 115, 51–66. DOI: <https://doi.org/10.1016/j.specom.2019.10.005>
- [14] Kan, M.S., Ito, A., 2020. Language cognition and pronunciation training using applications. *Future Internet*. 12(3), 42. DOI: <https://doi.org/10.3390/fi12030042>
- [15] Tai, T.-Y., Chen, H.H.-J., 2023. The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*. 31(3), 1485–1502.
- [16] Bashori, M., van Hout, R., Strik, H., et al., 2022. Web - based language learning and speaking anxiety. *Computer Assisted Language Learning*. 35(5–6), 1058–1089. DOI: <https://doi.org/10.1080/09588221.2020.1770293>
- [17] Al - Harbi, A., 2014. Language learning anxiety among Saudi EFL students: A case study. *Journal of Language and Literature*. 5(1), 20–29. DOI: <https://doi.org/10.1007/s10669-014-9485-0>
- [18] Al - Hazmi, S., Scholfield, P., 2007. The difficulties of learning English as a foreign language in Saudi Arabia. *The Journal of Language and Linguistic Studies*. 3(1), 3–16.
- [19] Al - Sibai, D., 2004. Difficulties in speaking English as a foreign language: A case study of Saudi EFL learners. *Journal of Language Teaching and Research*. 5(3), 123–135. DOI: <https://doi.org/10.4304/jltr.5.3.123-135>
- [20] SmallTalk2Me, 2024. SmallTalk2Me: AI - powered simulator to improve spoken English. Available from: <https://smalltalk2.me/> (cited 1 October 2024).
- [21] Manggiasih, L.A., Loreana, Y.R., Azizah, A., et al., 2023. Strengths and limitations of SmallTalk2Me app

- in English language proficiency evaluation. *TELL: Teaching of English Language and Literature Journal*. 11(2), 134. DOI: <https://doi.org/10.30651/tell.v11i2.19560>
- [22] Macías, A.M.D., Armijos Solano, D.O., Palma Perero, L.M., et al., 2024. The potential of artificial intelligence to improve speaking skills in a second language (English) fluently. *Ciencia Latina Revista Científica Multidisciplinar*. 8(3), 3826–3836. DOI: https://doi.org/10.37811/cl_rcm.v8i3.11592
- [23] García, A., López, M., Martínez, J., 2023. Improving speaking skills through conversational practice with chatbots. *Journal of Language Education and Technology*. 10(2), 45–58.
- [24] Li, M., Chen, J., 2017. The impact of AI - based pronunciation tutoring systems on oral proficiency development. *Language Learning & Technology*. 21(3), 124–142.
- [25] Smith, R., Johnson, L., 2022. Enhancing pronunciation with voice recognition systems. *International Journal of Language Learning*. 8(1), 112–125.
- [26] Zhang, L., Wang, X., 2018. Exploring the use of AI - powered virtual language assistants for speaking practice. *Computer - Assisted Language Learning*. 31(4), 357–376.
- [27] Loewen, S., Crowther, D., Isbell, D.R., et al., 2019. Mobile - assisted language learning: A Duolingo case study. *ReCALL*. 31(3), 293–311. DOI: <https://doi.org/10.1017/S0958344019000065>
- [28] 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. DOI: <https://doi.org/10.1111/flan.12454>
- [29] Wu, Y., Liu, Q., 2016. Integrating AI - driven speech recognition technology into English language teaching: A case study. *Educational Technology Research and Development*. 64(4), 757–776.
- [30] Warman, L.a.D., Erlinda, S., Tashid, T., et al., 2023. Empowering introvert students: How artificial intelligence applications enhance speaking ability. *AI - Ishlah: Jurnal Pendidikan*. 15(4), 4801–4813. DOI: <https://doi.org/10.35445/alishlah.v15i4.4435>
- [31] Yin, X., Wei, Y., 2023. Research on the construction and application of an intelligent learning system to enhance college English listening and speaking. *Journal of Education and Teaching*. 1(4), 39–45. DOI: <https://doi.org/10.59825/jet.2023.1.4.39>
- [32] Hoang, N.T., Han, D.N., Le, D.H., 2023. Exploring chatbot AI in improving vocational students' English pronunciation. *AsiaCALL Online Journal*. 14(2), 140–155. DOI: <https://doi.org/10.54855/acoj.231429>
- [33] Zou, B., Liviero, S., Hao, M., et al., 2020. Artificial intelligence technology for EAP speaking skills: Student perceptions of opportunities and challenges. In *Technology and the Psychology of Second Language Learners and Users*. Palgrave Macmillan: Cham, Switzerland. pp. 433–463. DOI: https://doi.org/10.1007/978-3-030-34212-8_17
- [34] Makhoulouf, M.K.I., 2021. Effect of artificial intelligence - based application on Saudi preparatory - year students' EFL speaking skills at Albaha University. *International Journal of English Language Education*. 9(2), 36. DOI: <https://doi.org/10.5296/ijele.v9i2.18782>
- [35] Alsadoon, R., 2021. 'Chatting with AI Bot: Vocabulary Learning Assistant for Saudi EFL Learners'. *Canadian Center of Science and Education*. 6(14), 1–23.
- [36] Alhalangy, A.G.I., 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. DOI: <https://doi.org/10.36923/jicc.v23i2.125>
- [37] Alsaif, O., 2024. The role of artificial intelligence (AI) in developing English language skills in the Saudi EFL context: An analytical study. *International Journal of Educational Sciences and Arts*, 3(5), 1–10. DOI: <https://doi.org/10.59992/IJESA.2024.v3n5p1>
- [38] Burston, J., 2015. Twenty years of MALL project implementation: A meta - analysis of learning outcomes. *ReCALL*. 27(1), 4–20. DOI: <https://doi.org/10.1017/S0958344014000159>
- [39] Zhang, R., Zou, D., 2022. Types, purposes, and effectiveness of state - of - the - art technologies for second and foreign language learning. *Computer Assisted Language Learning*. 35(4), 696–742.
- [40] Gao, R., Merzdorf, H.E., Anwar, S., et al., 2023. Automatic assessment of text - based responses in post - secondary education: A systematic review. *arXiv*. DOI: <https://arxiv.org/abs/2308.16151>
- [41] Martens, B.K., Witt, J.C., Elliott, S.N., et al., 1985. Teacher judgments concerning the acceptability of school - based interventions. *Professional Psychology: Research and Practice*. 16, 191–198. DOI: <https://doi.org/10.1037/0735-7028.16.2.191>
- [42] Petrić, B., Czárli, B., 2003. Validating a writing strategy questionnaire. *System*. 31(2), 187–215. DOI: [https://doi.org/10.1016/S0346-251X\(03\)00020-4](https://doi.org/10.1016/S0346-251X(03)00020-4)
- [43] Raoofi, S., Miri, A., Gharibi, J., et al., 2017. Assessing and validating a writing strategy scale for undergraduate students. *Journal of Language Teaching and Research*. 8. DOI: <https://doi.org/10.17507/jltr.0803.23>
- [44] Zohrabi, M., 2013. Mixed method research: Instruments, validity, reliability and reporting findings. *Theory and Practice in Language Studies*. 3(2), 254–262. DOI: <https://doi.org/10.4304/tpls.3.2.254-262>