ARTICLE

AI Applications for EFL Learners: Enhancing Speaking Performance and Reducing Anxiety with Gender-Based Analysis

Iwan Fauzi , *Rudi Hartono* * , *Dwi Rukmini* , *Hendi Pratama*

English Language Education Department, Semarang State University, Semarang, Central Java 50229, Indonesia

ABSTRACT

This study explores the effectiveness of AI-based applications in enhancing speaking performance and reducing speaking anxiety among EFL learners, with a focus on gender differences. It examined three main areas: the impact of AI tools on students' speaking anxiety, their effect on speaking performance, and how these outcomes vary by gender. A quantitative approach was used through a quasi-experimental design involving pre- and post-tests and questionnaires. The intervention lasted six weeks and included two intact second-year English major classes at an Indonesian university, with 46 participants—22 males and 24 females. Findings showed a statistically significant reduction in students' speaking anxiety and a notable improvement in speaking performance after using AI-based speaking platforms. These results suggest that AI tools are effective in supporting English-speaking practice for EFL learners. Regarding gender, female students exhibited slightly higher anxiety levels, though not significantly different from males. However, females performed significantly better in speaking tasks than their male peers. This indicates that gender may play a role in how learners engage with AI-assisted language learning tools. The findings have important implications for integrating AI in language education, particularly in fostering more equitable and effective speaking instruction. Nonetheless, due to the relatively small sample size and the use of a quasi-experimental design with only two intact classes, the generalizability of these results is limited. Future studies with larger and more diverse populations are recommended to validate and expand on these findings.

Keywords: AI-Based Applications; Speaking Anxiety; Speaking Performance; Gender Differences; EFL Learners

CORRESPONDING AUTHOR:

Rudi Hartono, English Language Education Department, Semarang State University, Semarang, Central Java 50229, Indonesia; Email: rudi.hartono@mail.unnes.ac.id

1. Introduction

Foreign language speaking anxiety (FLSA) significantly hampers learners' progress in acquiring English proficiency, particularly in speaking performance. This anxiety involves feelings of nervousness, fear, and apprehension when using a foreign language, and is rooted in psychological concerns—such as fear of making mistakes, embarrassment, and negative evaluation by peers or teachers—as well as increased self-consciousness in social interaction [1–3]. Linguistic limitations, including insufficient vocabulary, inaccurate grammar, and pronunciation difficulties, further exacerbate this anxiety, making speaking tasks even more challenging. Consequently, learners may develop avoidance behaviors and engage passively in class, which can negatively impact their overall language development [2]. Considering this deep impact, addressing the FLSA issue is essential, not only for enhancing speaking fluency but also for maintaining a positive attitude toward language learning itself, in addition to improving learners' intrinsic motivation and self-confidence [4].

Various technologies under Computer-Assisted Language Learning (CALL) have been introduced to support speaking practice, yet their effectiveness remains limited due to several inherent shortcomings. Traditional CALL systems, though innovative in their time, often lack personalization, real-time interactivity, and adaptability, which can result in reduced engagement and limited speaking improvement [5,6]. These systems typically do not offer nuanced, individualized feedback or simulate realistic conversational contexts, leaving learners underprepared for spontaneous real-world communication. Moreover, the absence of artificial intelligence (AI) in this technology restricts its capacity to respond dynamically to the differences of individual learners and developing linguistic needs, thereby limiting their potential to address speaking anxiety and foster speaking performance.

In contrast, the integration of AI-based digital technology into language education marks a significant transformation, offers a practical solution for reducing FLSA effectively [7–9], and improves the speaking performances of English as a foreign language (EFL) learners [10]. AI-powered platforms leverage machine learning and natural language processing (NLP) to deliver instant, personalized feedback, adaptive practice tasks, and simulations of interactive dialogue [11,12]. These applications are specifically designed to create a supportive, conducive, and low-pressure environment to sustained practice, which enables EFL learners to build their progressive self-confidence in their speaking skills. By replicating real-world communication scenarios, learners become more prepared for spontaneous speech, which contributes to reduced anxiety and improved fluency [13,14]. Furthermore, AI's adaptive capabilities allow instruction to be tailored to each learner's specific strengths and weaknesses, supporting autonomy, communicative competence, and overall language development. Despite these benefits, fewer studies have explored how gender affects the engagement and effectiveness of this platform, particularly in the context of EFL learners.

Studies related to gender roles in the use of AI-based applications for speaking practice among EFL learners have resulted in valuable insights and implications. The differences in how male and female learners engage with AI-driven tools can significantly influence the effectiveness of the application and user experience as a whole, particularly in the way of managing speaking anxiety. Evidence shows that such applications can significantly reduce anxiety among EFL learners, with gender acting as a moderating factor. Female learners, for instance, tend to report higher levels of anxiety in language learning, which can be alleviated through technological support that offers personalized, non-judgmental environments [1,15]. Additionally, the use of AI-based tools has led to substantial improvement in speaking fluency, especially in pronunciation and interactive speaking performance [16,17].

In the Indonesian context, however, research focusing specifically on speaking performance in AI-based applications from a gender perspective remains limited. While existing studies have explored speaking proficiency in relation to gender as an independent variable, few have addressed how gender interacts with AI-enhanced language learning [18–20]. Considering the disruptive impact of FLSA, the adaptive potential of AI to improve speaking skills while reducing anxiety, and the influence of gender on learning engagement, this study aims to investigate how AI-based applications can support learners of diverse genders in reducing speaking anxiety and enhancing their speaking proficiency. To achieve this aim, the following

research questions are proposed:

1. How does the use of AI-based applications impact students' speaking anxiety?
2. How does the use of AI-based applications impact students' speaking performance?
3. Do the effects of AI-based applications on speaking anxiety and performance vary by gender?

1.1. AI-Based Applications in Language Learning

Artificial Intelligence (AI) has become a transformative force in language education, enabling the development of intelligent applications that enhance the effectiveness of language acquisition. These AI-based tools have evolved significantly, progressing from simple systems to highly sophisticated platforms that improve various aspects of language learning, such as conversational practice, pronunciation, personalized instruction, and real-time feedback.

AI-based applications in language learning are generally classified into four categories: *chatbots*, *speech recognition tools*, *virtual tutors*, and *pronunciation assessors*. Chatbots act as conversational agents that simulate human-like interactions, offering learners an interactive platform for real-time dialogue practice, thereby improving their communicative competence [21]. Speech recognition tools focus on interpreting spoken language and providing instant feedback on pronunciation and fluency, which is essential for real-time feedback, significantly improving students' spoken performance by identifying and correcting mispronunciations as they occur [16]. Virtual tutors leverage AI to personalize the learning experience by adapting lesson plans based on individual progress, offering a level of instruction that often exceeds traditional classroom methods. This personalized approach allows students to progress at their own pace, resulting in more effective learning outcomes and improved language retention [22]. Pronunciation assessors, using advanced algorithms, have been shown to be effective in diagnosing pronunciation errors in EFL contexts, guiding learners toward more accurate pronunciation through targeted feedback [23]. These tools together integrate interactive dialogue, immediate feedback, personalized lessons, and pronunciation analysis, creating a comprehensive and effective language learning experience.

The evolution of AI-based language learning tools has been marked by the transition from basic text-based systems to advanced AI technologies incorporating *deep learning*, *neural networks*, and *Natural Language Processing (NLP)*. The integration of deep learning has enhanced the accuracy of language comprehension and interaction, allowing for more nuanced and context-aware exchanges [24]. NLP, in particular, has significantly improved the development of conversational agents and virtual tutors by enabling more fluent and natural language interactions [25]. Additionally, advancements in *speech synthesis* and *real-time recognition* technologies have created more immersive and interactive learning experiences, providing immediate feedback that reinforces pronunciation and fluency development [26]. While these studies highlight the functionality of AI tools individually, limited research critically evaluates their comparative effectiveness across diverse learner profiles—particularly in relation to gender—which suggests a gap in understanding how different learner demographics engage with various AI modalities. As AI technologies continue to advance, these applications are expected to play an even more central role in language education, offering increasingly effective and accessible solutions for learners around the world.

1.2. The Role of AI-Driven Applications on EFL's Speaking Performance and Speaking Anxiety

Recent research has demonstrated the significant role of AI-driven applications in enhancing EFL learners' speaking performance and alleviating speaking anxiety. These applications have shown facilitating improvements in pronunciation accuracy, fluency, and overall speaking proficiency. For instance, studies demonstrated that EFL learners using AI-driven language apps experienced notable improvements in speaking proficiency, as measured by pre- and post-tests [27,28]. These AI tools enable real-time feedback and provide a non-judgmental, interactive environment that promotes confidence and fluency. Furthermore, other studies revealed improvements in fluency and accuracy, along with increased learner autonomy and confidence, highlighting the effectiveness of AI in fostering an enriched, supportive learning environment [8,29].

In addition to improvements in speaking performance,

AI-driven tools have shown promise in reducing speaking anxiety. Reduction in anxiety among EFL learners using AI-based language tutors was indicated, with anxiety declining [30]. Continuous interaction with AI chatbots has been shown to lead to a marked decrease in anxiety levels among students [31]. These findings are further supported by research highlighting that the personalized feedback and non-judgmental nature of AI tools allow learners to practice more freely, thereby reducing the fear and pressure typically experienced in traditional classroom settings [8,32]. Additionally, professional speakers using an AI coaching app have reported increased confidence and decreased public speaking anxiety [33].

While this body of literature affirms the dual benefits of AI—enhancing speaking skills and reducing anxiety—it tends to treat these two outcomes in isolation. Few studies examine the interaction between decreased anxiety and improved performance within the same context, leaving the relationship between affective and cognitive gains underexplored. Moreover, limited attention has been given to whether such gains differ based on learner characteristics such as gender or prior technological familiarity, which could significantly mediate outcomes.

1.3. Gender Differences in the Use of AI Applications

AI-based applications in language learning effectively enhance speaking performance and reduce anxiety, but gender disparities influence their impact. Males typically engage more with technology, leading to better learning outcomes and increased motivation [34–36]. In contrast, females benefit from the interactive and personalized features of AI, which help them overcome challenges in engagement [37]. Therefore, AI tools should cater to both competitive and collaborative learning styles.

AI also boosts confidence in speaking foreign languages by offering a private, non-judgmental environment. This fosters greater confidence from females [38], while males benefit from consistent feedback and practice without the stress of peer or teacher evaluation [39,40]. Moreover, self-efficacy plays a role, as females often show lower self-efficacy in language learning [41]. However, AI applications that provide tailored and positive feedback can boost their confidence and motivation [34]. Although several studies highlight gendered preferences or responses to AI learning environments, the findings are often inconsistent or context-dependent. Some report stronger outcomes for males due to greater technological familiarity, while others suggest females benefit more due to reduced anxiety in private settings. These mixed findings underscore the need for comparative, within-group studies that directly evaluate gender as a moderating factor in AI-enhanced learning. Particularly in the Indonesian context, empirical data on this topic remains scarce, suggesting a need for culturally and contextually relevant investigations. To conclude, gender differences impact the efficacy of AI-based language learning apps that improve speaking abilities and lower anxiety. More research is necessary to fully comprehend these gender dynamics, especially in the Indonesian setting.

2. Materials and Methods

2.1. Design

This study employed quantitative research to provide a measurement of the effectiveness of AI-based applications in enhancing speaking performance and reducing speaking anxiety. A quasi-experimental design with pre-test and post-test measures was used to evaluate the impact of the intervention. Two AI applications—*ELSA* and *SmallTalk2Me*—were deliberately selected to support the design of the experiment in this research. The researcher used these two AI applications so that the result of the experiment would not be dependent on one AI application only, and on coincidental results. In addition, the consideration of using two different AI applications for this experiment is because both applications are familiar to participants. In addition, both AI apps were equipped with Automatic Speech Recognition (ASR) technology, which can be used by users as instant feedback and self-evaluation progress. Therefore, *ELSA* and *SmallTalk2Me* had considerable benefits for the experiment in improving speaking performance and reducing speaking anxiety. Moreover, the quantitative approach was also performed to examine the influence of gender on speaking performance and speaking anxiety across male and female participants, and pre- and post-interventions as well.

2.2. Instrument and Participants

Three instruments were used for this research: a test and a questionnaire. First, two tests were used to measure the improvements in speaking performance: pre- and post-test. Pre-test and post-test were administered to evaluate students' speaking performance. The test of speaking skills has been provided in each AI platform of the *ELSA* and *SmallTalk2me* applications. Each application provides an AI-based IELTS test of speaking, rated by AI technology. Both AI applications evaluated participants' speaking proficiency using IELTS-equivalent scores. The IELTS speaking results are reported on a 9-band scale, ranging from 1 (the lowest) to 9 (the highest). Band scores of 1, 2, 3, and 4 correspond to the descriptors "non-user," "intermittent user," "extremely limited user," and "limited user," respectively. In contrast, scores of 5, 6, 7, 8, and 9 represent "modest," "competent," "good," "very good," and "expert" levels of English-speaking competence [42]. Both AI applications assessed and analyzed participants' speech based on the following criteria: fluency and coherence, vocabulary, grammatical range and accuracy, and pronunciation. Although no human raters were employed to triangulate AI-generated scores, the decision was based on the consistency, efficiency, and objectivity provided by the automated scoring systems. Given the standardized evaluation criteria and the use of two independent AI platforms, this approach offered a practical and replicable method of assessing speaking performance within the scope and constraints of the study.

To measure speaking anxiety, the researchers adopted the Foreign Language Classroom Anxiety Scale (FLCAS) questionnaire, originally developed by Horwitz et al [2]. The original FLCAS consists of 33 items measured on a five-point Likert scale, ranging from "strongly agree" to "strongly disagree." For this study, 20 items specifically related to speaking anxiety were adopted from the original scale, forming in this research what is referred to as the Foreign Language Speaking Anxiety (FLSA) questionnaire. This version was used in the pre-intervention phase to assess participants' initial levels of speaking anxiety (indicated in **Appendix A**). Following this, the same 20 items were adapted to align with the context of AI-based English-speaking practice (see **Appendix A**). The adapted FLSA questionnaire was then used in the post-intervention phase to evaluate any changes in speaking anxiety after the use of AI applications.

The participants of this study were taken from two intact classes who enrolled in the second year of the English study program of a university in Indonesia. There were 46 students involved in the whole research process, characterized by 22 males and 24 females. Initially, researchers purposively assigned 48 students, which is a balanced proportion between both genders, since this study was viewed from a gender perspective. However, until the final stage of this study was completed, two male participants did not complete the entire research process, starting from pre- to post-interventions. The participants were aged between 18 and 20 years, had intermediate (B1) to upper-intermediate (B2) levels of English proficiency as indicated by their responses on the FLSA questionnaire, and of course they were native Indonesian speakers.

The selection of participants from intact classes was based on logistical and institutional considerations, ensuring the feasibility of implementing the intervention in a natural classroom setting. While this approach may introduce selection bias and limit randomization, it reflects real-world classroom conditions. Therefore, although random assignment would have strengthened the internal validity by minimizing potential confounding variables, the use of intact classes allowed the intervention to be carried out more effectively and ethically within the existing academic structure.

2.3. Procedure

This research was conducted over six weeks and followed a structured sequence to examine the impact of AI-based applications on English-speaking anxiety and performance. In the first week, before the intervention, participants completed a pre-intervention assessment. This included the administration of the FLSA questionnaire to measure their initial levels of speaking anxiety. In addition, a pre-test based on the IELTS speaking test was used to evaluate participants' baseline speaking proficiency. In this experiment, researchers used the *Smalltalk2Me* app for testing participants' speaking skills since it was considered more convenient by the participants, although they had also tried the *ELSA* speaking test of IELTS. Following the initial assessments, the intervention phase was carried out

over the next four weeks, from week two through week five. Each week the *ELSA* app was used in the first half, and the *Smalltalk2Me* was used in the second half of the week. During this period, participants engaged in regular English-speaking practice using those two AI-based applications specifically designed to support language learning. These tools provided interactive and personalized speaking tasks aimed at enhancing fluency, accuracy, and confidence in spoken English. Finally, in the sixth week, post-intervention assessments were conducted to measure any changes resulting from the use of AI-based speaking tools. Participants completed the FLSA questionnaire once again to identify any reduction in speaking anxiety. Additionally, they took a post-test modelled on the IELTS of AI speaking format provided by the *Smalltalk2Me* app to determine improvements in their speaking skills. This structured procedure allowed for a comprehensive evaluation of the effectiveness of AI-assisted speaking practice in reducing anxiety and enhancing performance in English speaking. A more specific description of materials and activities in the research procedure is given in **Appendix B Table A1**.

2.4. Data Analysis

The first objective of this study's data analysis is to evaluate the effectiveness of AI-based applications in improving the speaking performance of EFL learners (RQ-1). Quantitative data were obtained from IELTS speaking scores before and after the intervention. Although these scores are often considered ordinal, in this study they were treated as interval data. This is because the AI scoring system used provides consistent and detailed ratings that reflect small, meaningful differences in performance. The system is designed to apply the scoring rubric evenly across all test-takers, making the score differences more like equal units. This approach is also supported by previous studies that treat IELTS scores as interval when using reliable scoring systems [43,44]. Since the data are paired and not assumed to be normally distributed, the Wilcoxon

signed-rank test was used to check whether there was a significant improvement in speaking scores after the intervention.

The second objective was to assess the effect of AI-based tools on reducing students' speaking anxiety (RQ-2). Speaking anxiety was measured using Foreign Language Speaking Anxiety (FLSA) scores, collected before and after the intervention. The Wilcoxon signed-rank test was also used for this comparison. For the third objective (RQ-3), which explores whether gender plays a role in how students respond to AI-based speaking tools, a Multivariate Analysis of Covariance (MANCOVA) was used. This analysis examined the impact of gender on both speaking performance and speaking anxiety, while controlling for students' baseline (pre-intervention) scores.

3. Results

The three research questions proposed in this study are answered separately in three sections of the following description of the research results.

3.1. The Effect of the Use of AI-based Applications on Students' Speaking Anxiety

There were 20-statement items of speaking anxiety adapted from 33 items of the foreign language classroom anxiety scale (FLCAS) to measure learners' speaking anxiety (FLSA) before and after they were treated with an AI-based application for English speaking practice. Those statements in the questionnaire were scaled with a 5-point Likert scale (**Table 1**). The reliability of the questionnaire items before intervention was high, with Cronbach's Alpha value = 0.79. The average score of FLSA in the pre-intervention was M = 3.59; SD = 0.41, which means that this score belongs to relatively anxious. Of the 20 items, four items have a score of 4 or more than 4, which belong to the seriously anxious; and five items have a score of more than 3.5 and below 4, which belong to the mildly anxious.

Table 1. Descriptive Statistics of FLSA Scores.

FLSA scores	N	Mean	Std. Deviation	Minimum	Maximum
Pre-questionnaire	46	3.59	0.41	2.25	4.65
Post-questionnaire	46	3.36	0.56	1.45	4.10

Four items reflecting learners seriously anxious are items 3, 5, 6, and 20. The highest mean score was 4.37 for item 6, which indicated that *learners can get so nervous when they suddenly forget ideas that they want to speak.* Following closely, both items 5 and 20 had a mean score of 4.07. Item 5 informed that *learners started to panic when they had to speak without preparation in the speaking class*, while item 20 revealed that *learners got nervous when the lecturer asked questions in the speaking class, which they hadn't prepared in advance.* The lowest mean score of 4.00 was for item 3, where *learners trembled when they knew that they were going to be called on in the speaking class*.

The other situation of mildly provoking learners' speaking anxiety before the intervention of using an AI-based application was indicated by five items in the questionnaire (items 4, 7, 14, 15, and 19), which had scores of more than 3.5 and below 4. The highest mean score was 3.80 for item 7, where *learners felt less confident when they were speaking in front of the class*. Closely following this, with a mean score of 3.76, was item 4, which indicated that *learners kept thinking that the other students were better at the English language than they are*, and item 15, where *learners got nervous and confused when they were speaking in their speaking class*. Item 14, with a mean score of 3.70, reflected how *learners felt very self-conscious about speaking the foreign language in front of other students*. Last, item 19, with the lowest score of 3.61, indicated that *learners got trembling because of forgetting things they knew while speaking*.

After intervention of speaking practice using AI website platforms driven by *ELSA* and *SmallTalk2Me* applications, learners' speaking anxiety was remeasured to find out the effect of the intervention. The reliability of 20 items in the post-questionnaire was still very high, with Cronbach's Alpha value = 0.91. Evaluating from the post-intervention, nine items that had FLSA scores higher than 3.5 in the pre-intervention indicated a decrease after the intervention. However, three new items scored higher compared with the pre-intervention, such as items 11, 13, and 17. For instance, item 11 informs *learners get nervous when an AI app evaluates every mistake they made during speaking practice*, with a mean score of 3.13, which was previously 3.11. Then, item 13 with a mean score of 3.72, which was previously 2.76, indicates that *learners always feel that they are better, not just entering the speaking class when using AI apps.* Last, item 17 showed *learners felt overwhelmed by the rules given in the feedback of AI apps regarding grammar and pronunciation that they had to learn during speaking practice,* with the mean score of 3.61, which relatively increased from the previous score of 3.45.

In short, while the data highlighted an increase in speaking anxiety after using AI-based applications, the increase in new scores and participants' perception did not impact the overall effectiveness of AI applications in reducing learners' speaking anxiety. To examine the difference between the results of pre- and post-questionnaires on FLSA, a Wilcoxon signed-rank test was used (see **Table 2**).

Table 2. Wilcoxon Test on FLSA.

		N	Mean rank	Sum of rank	Z	Asymp. Sig. (2-tailed)
FLSA Pre-score / Post-score	Negative Ranks	28	26.95	754.50	–2.678	0.007
	Positive Ranks	17	16.50	280.50		
	Ties	1				
	Total	46				

The test revealed a significant difference, with 28 participants showing a decrease in FLSA scores (negative ranks) and 17 participants showing an increase (positive ranks). The negative ranks had a higher mean rank (26.95) compared to the positive ranks (16.50), suggesting that FLSA post-questionnaire scores were generally lower than the pre-questionnaire ones. One participant showed no change between the two conditions. As a result, these findings suggest that speaking anxiety scores after intervention (*M* = 3.36; *SD* = 0.56) were significantly lower than the scores before intervention (*M* = 3.59; *SD* = 0.41) which means speaking practice using AI website platforms is effective in reducing speaking anxiety (*Z* = –2.678; *p* = 0.007).

3.2. The Effect of the Use of AI-based Applications on Students' Speaking Performance

An overview of the speaking scores before and after the intervention of using AI-based applications is provided in **Table 3**. The average score for the pre-test (M = 6.59; SD = 0.82) indicates a moderate level of performance among the participants. The scores ranged from 5.00 to 8.50, showing a spread of abilities. After the intervention, the post-test average score increased with a slightly lower standard deviation (M = 7.09; SD = 0.76), suggesting that the majority of participants showed improvement in their speaking performance. The range of scores remained the same, between 5.00 and 8.50, indicating that while scores improved, they did not fluctuate drastically. This suggests a general enhancement in performance across the group following the use of AI applications.

Table 3. Descriptive Statistics of Speaking Performance Scores.

Speaking Scores	N	Mean	Std. Deviation	Minimum	Maximum
Pre-test	46	6.59	0.82	5.00	8.50
Post-test	46	7.09	0.76	5.00	8.50

Figure 1 shows that the speaking performance scores of pre-test and post-test demonstrate a notable improvement in performance following the use of AI applications for speaking practice as an intervention.
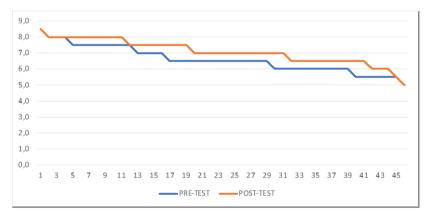

Figure 1. Speaking Performance Scores Before and After the Intervention.

The post-test scores show a clear upward shift compared to the pre-test scores, indicating that the use of AI applications contributed to enhancing IELTS speaking proficiency. Specifically, the post-test scores are more concentrated in the higher bands (7.0 and 8.0), whereas the pre-test scores are more widely distributed across the lower and middle ranges. This shift suggests that the AI-based intervention effectively helped participants achieve higher scores in the post-test. Additionally, the post-test scores display a more consistent performance, with fewer fluctuations, highlighting the positive impact of the AI applications on test-taker performance.

To examine the difference between the results of the pre-test and post-test on speaking performance, a Wilcoxon signed-rank test was used (see **Table 4**). The test revealed there was a statistically significant difference between the pre-test and post-test scores of speaking performances (Z = –3.335; p = 0.001). The results indicate that there was a notable change in speaking performance after the intervention. Specifically, eight participants had lower post-test scores than pre-test scores, which were ranked with a mean of 20.69, while 32 participants showed improvement, with their post-test scores ranked with a mean of 20.45. Additionally, six participants had no change in their scores, resulting in ties. In short, the Wilcoxon test strongly supports the use of AI applications, leading to a significant improvement in students' speaking performance.

Table 4. Wilcoxon Statistic on Speaking Performance.

		N	Mean rank	Sum of rank	Z	Asymp. Sig. (2-tailed)
	Negative Ranks	8	20.69	165.50		
IELTS scores of pre-test / post-test	Positive Ranks	32	20.45	654.50	–3.335	0.001
	Ties	6				
	Total	46				

The Effects of Gender on Speaking Anxiety and Speaking Performance in Using AI Apps

The study included 22 male participants (48%) and 24 female participants (52%). **Figure 2** depicts the data of post-intervention mean scores in speaking anxiety (FLSA) and speaking performance with regard to gender. Female participants had a bit higher FLSA mean score (*M* = 3.50; *SD* = 0.51) compared to male participants (*M* = 3.22; *SD* = 0.59). Meanwhile, in speaking performance mean scores after intervention, females performed better (*M* = 7.29; *SD* = 0.64) than males (M = 6.86; *SD* = 0.82). This indicates that there is a difference score between males and females in speaking anxiety, with females experiencing slightly higher than males, but it has no significant effect (*F* = 2.220; *p* = 0.121). However, on speaking performance, gender has a significant effect (*F* = 12.925; *p* < 0.001), with females performing better than males. In conclusion, there are clear gender differences in the speaking performance of using AI-driven applications for English-speaking practice, but not in speaking anxiety.
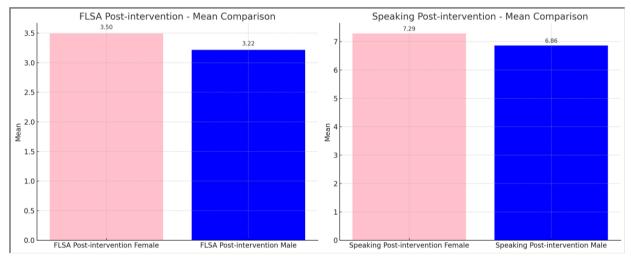

Figure 2. Description of FLSA and Speaking Scores After Intervention with Regard to Gender.

With regard to intervention, the MANCOVA test of time (AI apps intervention) in **Table 5** indicates the AI-based intervention had a very large effect on reducing speaking anxiety (*F* = 564.046; *p* < 0.001; *Partial Eta Squared* = 0.982). This is similar to speaking performance, indicating the intervention also had a very large effect (*F* = 1119.847; *p* < 0.001; *Partial Eta Squared* = 0.991). Values of both *Partial Eta Squared* reveal 98.2% of the variance in speaking anxiety scores is explained by the intervention, suggesting that the AI intervention has a very strong impact on reducing anxiety. Also, 99.1% of the variance in speaking performance scores can be attributed to the AI intervention. As a result, the AI-based intervention was highly effective in both reducing speaking anxiety and improving speaking performance, with extremely large effect sizes for both outcomes.

Table 5. MANCOVA Test Main Effects of Intervention on FLSA and Speaking Performance.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Model	FLSA post-intervention	524.609	4	131.152	564.046	<0.001	0.982
	Speaking post-intervention	2314.300	4	578.575	1119.847	<0.001	0.991

With regard to gender, MANCOVA test (**Table 6**) is used to examine the interaction between gender and time (post-intervention) if the effect of the AI-based intervention on speaking anxiety (post-FLSA) and speaking performance (post-speaking performance) differs between male and female participants.

Table 6. MANCOVA Test Between-Subject Effects of Gender and Speaking Anxiety.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Gender	FLSA post-intervention	1.032	2	0.516	2.220	0.121	0.096
	Speaking post-intervention	13.355	2	6.678	12.925	<0.001	0.381

In terms of speaking anxiety, the result shows a non-significant interaction between gender and the AI-based intervention for speaking anxiety ($F = 2.220$; $p = 0.121$). The effect of the intervention on reducing speaking anxiety is not significantly different between male and female participants. It means that the AI-based apps did not show any different anxiety levels between males and females in terms of their alleviation after intervention. However, the result is highly significant between gender and the AI-based intervention for speaking performance ($F = 12.9215$; $p < 0.001$). The effect of the intervention on improving speaking performance is significantly different between male and female participants. Thus, there is a significant interaction between gender and time for post-speaking performance, but not for post-speaking anxiety. To conclude, the reduction in speaking anxiety contributed to the improvement in speaking performance. However, the reduction in speaking anxiety did not significantly differ between males and females (albeit males experience lower anxiety than females), while the improvement in speaking performance significantly varies between both, where female students tend to perform better than male students.

4. Discussion

The results of this study show a significant reduction in speaking anxiety among EFL learners after the use of AI-based applications (RQ-1). The findings revealed a decrease in anxiety scores after the intervention, albeit some increase in anxiety related to AI judging mistakes and giving feedback, the overall effect was a notable decrease in speaking anxiety. This confirms several previous studies that highlight AI's potential to reduce anxiety in language learners through offering a non-judgmental, personalized learning environment. Confirming the present study's results, AI-assisted speaking platforms were found to help lower anxiety by offering personalized feedback without the judgment that students might feel in a traditional classroom setting [32]. Other research has also supported the role of AI in reducing speaking anxiety. Learners using AI tools for language practice exhibited a marked reduction in speaking anxiety, particularly during formal assessments [45]. Similarly, students practicing with AI-powered speaking tools showed more confidence, which contributed to reduced anxiety levels due to their ability to practice in a low-pressure environment [9]. Comparable effects were reported with the use of VoiceThread, an AI-based voice-conferencing tool, which was seen as effective in lowering foreign language anxiety by allowing students to engage in language tasks in a more relaxed and less stressful setting [46].

Although AI apps can significantly reduce speaking anxiety, AI tools can still be noted to cause apprehension in students who use them for speaking practice, as several students experienced it, as shown by the results of this

study. For instance, students' nervousness increased when using AI due to the immediate evaluation of their performance. This condition is similar to findings showing that EFL learners experience speaking anxiety when engaging in AI-powered language tasks outside the classroom [31]. It exhibited higher anxiety levels than those participating in traditional speaking tasks. On the other hand, the findings of this study contrast with previous research, where AI use in speaking class settings was found to contribute to anxiety, unlike the preference for independent practice with AI apps at home [31]. This is to say, technological limitations contribute to increased anxiety. Last, the immediate and sometimes critical feedback from AI can also make students feel overwhelmed and insecure. This is reinforced by findings of increased anxiety related to AI feedback occurring when students are evaluated by technology [47]. It must be acknowledged that while AI generally helps alleviate anxiety, there may be specific cases where learners perceive AI feedback as insecure or tend to make themselves uneasy.

The second research question (RQ-2) explored the impact of AI applications on students' speaking performance. AI speaking practice applications offer insightful feedback for students' skill development. The study's findings showed that students' interactions with AI app features were particularly beneficial for improving their speaking skills, including speech rate, pronunciation, grammar, and vocabulary accuracy. This is to confirm the positive impact of the utilization of AI-assisted applications in the improvement of English skills, such as pronunciation, grammar accuracy, vocabulary improvement, and fluency [17,27,48,49]. Further studies lend additional support to the findings, showing that AI-based applications led to significant improvements in students' oral proficiency, particularly by enhancing their fluency and pronunciation through continuous feedback [9]. Similarly, studies have concluded that AI technologies, specifically speech recognition systems, help students refine their speaking skills by providing real-time corrections and encouraging consistent practice [50]. Additionally, AI applications tailored to individual learning styles have been shown to increase speaking performance by allowing learners to engage at their own pace, fostering better retention and fluency [51].

In addition, as this study revealed, AI helps students concentrate on refining their speaking skills since they are free of stress and provided a private practice environment with AI. This confirms previous findings that practice without the stress of peer or teacher evaluation and interaction with AI-personalized features leads to better learning outcomes and increased motivation [17,22,40]. In terms of real-time feedback, AI-based language learning tools are highly effective in helping students overcome speaking mistakes and significantly boost their performance using instant feedback [49]. However, some challenges using AI tools, especially regarding user adaptation to AI feedback, may slow down improvement for certain students [17]. These mixed findings suggest that while AI applications are generally beneficial, they may require further customization to cater to diverse learner needs.

In examining gender differences (RQ-3), the study found that the effect of AI applications on improving speaking performance significantly differs between male and female participants. This aligns with the observation that female students exhibited greater enhancement in oral skill levels compared to male students, suggesting that gender may influence the effectiveness of technology-assisted language learning, particularly concerning improvements in speaking skills [38]. Moreover, the study indicates that female participants showed slightly better speaking performance, reinforcing the idea that gender plays a role in the effectiveness of mobile-assisted language learning tools, also known as digital technology supported with AI [52]. Additionally, female participants reported more positive engagement in speaking practice with AI conversational systems than males, indicating that the AI-driven system successfully maintained students' interest in practicing English speaking with regard to gender [53]. These findings collectively suggest that gender influences both the effectiveness and engagement with AI-based language learning tools, particularly in improving speaking skills. These results raise important pedagogical implications for future development and use of AI-supported learning tools. For example, AI language apps may benefit from incorporating gender-sensitive motivational strategies or feedback mechanisms to ensure more personalized and equitable learning experiences. Female students' higher engagement and performance might be linked to their responsiveness to socially supportive or collaborative features, suggesting

the value of designing AI interfaces that adapt to different user preferences or learning behaviors. Conversely, further exploration is needed to understand how to boost motivation and performance among male learners through AI-driven interventions.

Still in RQ-3, studies of using AI-driven tools and their impact on speaking anxiety have resulted in several findings regardless of gender [9,32,45,46,54]. If studies were concerning gender, they did not specifically use an AI app as the intervention to measure its impact on speaking anxiety [55–59]. Only a few studies that examine the effect of AI tools on speaking anxiety with regard to gender as the affecting factor [60–62]. For instance, a study investigated gender differences in need satisfaction of self-regulated learning (SRL) through an AI chatbot experiment in the K-12 setting, indicating that girls perceived more need for support than boys in AI learning environments [60]. Then, it has been discussed that gender is a variable influencing factors affecting the intention to use ChatGPT, an AI-based technology application. Findings showed that female teachers experienced higher anxiety than male teachers when using ChatGPT for academic tasks [61]. This is to imply that the use of AI tools has a real influence on the level of anxiety of its users with regard to gender. However, speech-enabled corrective feedback (SECF) technology was utilized to enhance speaking skills rather than to examine EFL learners' anxiety and confidence in English-speaking proficiency. The results showed a clear reduction in foreign language anxiety and an increase in confidence, with no significant gender differences [62]. The above-mentioned studies support the findings of this research, confirming that AI-based speaking apps help reduce speaking anxiety, with no significant difference in anxiety reduction between male and female participants, although females tended to feel slightly more anxious when using the apps. In light of these findings, AI developers and educators may consider integrating adaptive feedback or emotional support features that respond differently based on users' anxiety levels, potentially varying by gender. Tailoring AI feedback mechanisms may help learners feel more supported and reduce disengagement during speaking practice. This personalization could also contribute to reducing the gender gap in language learning outcomes observed in AI-mediated environments.

5. Conclusions

Based on the findings, this study concludes that the integration of AI-based speaking platforms significantly reduces students' speaking anxiety and enhances their speaking performance. The post-intervention scores demonstrated a marked decrease in anxiety and a notable improvement in oral proficiency, confirming the efficacy of AI tools in language learning environments. Furthermore, the gender-based analysis revealed that while anxiety levels were slightly higher in female participants, this difference was not statistically significant. However, in terms of performance, females outperformed males, suggesting that gender may play a role in how learners benefit from AI-driven speaking practice.

Despite these promising results, this study is not without limitations. The absence of a true experimental design—namely, the lack of a control group—limits the extent to which causal claims can be made. Moreover, the exclusive use of AI-generated speaking scores may not fully capture the nuances of human communication skills. Involving human evaluators, such as English teachers or instructors, would have enriched the assessment and offered a more comprehensive perspective on speaking performance. These constraints of research methodology and test evaluation should be addressed in future research to validate and extend the current findings.

Future research could explore the longitudinal effects of sustained AI use on speaking proficiency and anxiety reduction, potentially using mixed-methods designs that combine quantitative outcomes with qualitative insights into learners' experiences. Additionally, cross-cultural replications would help determine the generalizability of these findings across different educational and sociocultural contexts. Such directions would not only strengthen the empirical foundation but also guide more inclusive and effective implementation of AI technologies in language education.

The implications of this study are twofold. Pedagogically, the results advocate for the incorporation of AI speaking platforms into language curricula to support learners' confidence and competence in speaking. The statistically significant performance improvement suggests that such tools are effective in creating interactive and low-pressure environments for practice. Additionally, the observed gender differences in performance signal the need for more personalized and gender-sensitive learning

interventions that accommodate diverse learner characteristics. Developers of AI tools should consider adaptive features that respond to learners' individual emotional and performance profiles, while educators are encouraged to integrate these tools in ways that support differentiated instruction. Finally, for educators, these findings provide valuable practical implications. AI-based speaking platforms can be strategically embedded in curriculum design to foster active participation and reduce language anxiety. When aligned with pedagogical goals and learner diversity, these tools hold transformative potential for enhancing language instruction in a technology-enhanced classroom.

Author Contributions

Conceptualization, I.F. and R.H.; methodology, D.R.; software, I.F.; validation, R.H., D.R. and H.P.; formal analysis, I.F.; investigation, I.F.; resources, I.F.; data curation, I.F.; writing—original draft preparation, I.F.; writing—review and editing, I.F., R.H., D.R.; visualization, I.F.; supervision, R.H., D.R., H.P.; project administration, I.F.. All authors have read and agreed to the published version of the manuscript.

Funding

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Palangka Raya University, Indonesia (protocol code IRB-2024-001 and 09 August 2024).

Informed Consent Statement

Patient consent was waived due to minimal risk and the use of routine educational practices that do not necessitate formal consent, in accordance with institutional guidelines.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to privacy and ethical considerations, the data are not publicly accessible.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

Appendix A

Before you answer this questionnaire, please answer some of the questions about yourself below *(This is only given once in the pre-intervention questionnaire)*.

Your gender:	Your age:	Your English level:			
☐Male ☐Female	…… years old	☐Elementary (A2)	☐Intermediate (B1)	☐Upper-intermediate (B2)	☐Advanced (C1)

Foreign Language Speaking Anxiety (FLSA) Questionnaire

This questionnaire is prepared to gather information about the level of English anxiety you experience when you speak English in your speaking class. After reading each statement, put a check mark (ü) on the circle that you think fits best. There are no true or false answers to the statements in this questionnaire.

FLSA Questionnaire in Pre-intervention	**FLSA Questionnaire in Post-intervention**
1. I never feel quite sure of myself when I am speaking in my English-speaking class.	1. I never feel confident in myself when I speak English using an AI apps.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
2. I am worried about making mistakes when in English speaking class.	2. I am worried about making mistakes when using AI apps in English speaking class.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
3. I tremble when I know that I'm going to be called on in English speaking class.	3. I tremble when I know that I'm going to be called on in English speaking class by using AI apps.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
4. I keep thinking that the other students are better at English than I am.	4. I keep thinking that the other students are better at English than I am in the speaking class using AI applications.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
5. I start to panic when I have to speak without preparation in English speaking class.	5. I start to panic when I have to speak in AI apps without preparation in English speaking class.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
6. In speaking class, I can get so nervous if I forget things I know.	6. In speaking class using AI apps, I can get so nervous if I forget things I know.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
7. I am not confident to volunteer for speaking in front of the class.	7. I am not confident to volunteer for speaking in front of the AI apps.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
8. I feel nervous when I speak English with the lecturer of the speaking class.	8. I feel nervous when I speak English with the AI apps as the speaking tutor in the speaking class.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
9. I get upset when I don't understand what the teacher is correcting me when I am speaking.	9. I get upset when I don't understand what the AI apps correct me when they give feedback.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree
10. I get nervous and confused when my lecturer ask me to answer his questions in the speaking class.	10. I get nervous and confused when AI apps assign me a topic in random for speaking practice.
○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree	○Strongly agree ○Agree ○Neutral ○Disagree ○Strongly disagree

11. I am afraid that the other students will laugh at me when I speak in English.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

12. Even if I am well prepared for the speaking class, I feel anxious about it.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

13. I often feel like not going to my speaking class.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

14. I feel very self-conscious about speaking English in front of other students in the speaking class.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

15. I was nervous and confused when lecturer asked me to speak up in the speaking class.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

16. I get nervous when I don't understand every word spoken by the lecturer in the speaking class.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

17. I feel overwhelmed by the number of rules I have to learn in the speaking class.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

18. I am afraid that my lecturer is ready to correct every mistake I make in the speaking class.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

19. I immediately get trembling during practice in the speaking class when I suddenly do not know what I have to speak.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

20. I get nervous when my lecturer asks questions in the speaking class which I haven't prepared in advance for my answers.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

11. I am afraid when AI apps evaluate my speaking skills during speaking practice.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

12. Even if I am well prepared for the speaking class using AI apps, I feel anxious about it.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

13. I often feel like not going to my speaking class when the class is using AI apps.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

14. I feel very self-conscious about speaking English in front of AI apps.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

15. I am nervous and confused when AI apps ask me to start speaking based on the topic they select.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

16. I get nervous when I don't understand every command written by the AI apps to prompt me in the speaking practice.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

17. I feel overwhelmed by the number of rules (grammar & pronunciation) provided by AI apps that I have to learn during the practice.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

18. I am afraid that AI apps would correct every mistake I make in the platform.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

19. I immediately get trembling during practice using AI apps when I suddenly do not know what I have to speak.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

20. I get nervous when AI apps ask questions that I haven't prepared in advance for my answers.

○Strongly agree ○Agree ○Neutral
○Disagree ○Strongly disagree

Appendix B

Table A1. Procedure of Research Intervention.

Time (in-week)	Types of class activities	Description of Practice Using AI apps
Week-1 (first half)	Pre-questionnaire of anxiety (FLSA)	• Students are asked to answer questionnaire regarding with speaking anxiety before intervention.
Week-1 (second half)	Pre-test of speaking (IELTS)	• Students are asked to test independently with IELTS speaking test simulator provided by the AI app before intervention. Then report their results by sharing links of test score given by the app.
Week-2 (first half)	Warming up speaking practice with *ELSA* Speech Analyzer	• To warm up with a mini recording in addition to familiarise themself with Speech Analyser, students are given a series of picture, and they are asked to speak with a series of narrative picture orderly. • To perform their speech, students are given several questions from the series of picture as a hint to complete the task of their speaking such as: • Who are the two characters in the story? What competition were they taking part in? By nature, who is faster at running? Did the two main characters run continuously throughout the competition? If not, who did what during the race? Who is the winner? Is the result surprising? What lesson do you pick up from the story? • • After students complete their speech, the App will give AI-based evaluation in the aspects of pronunciation, intonation, fluency, grammar, and vocabulary.
Week-2 (second half)	Warming up speaking practice by a spontaneous speech with *SmallTalk2Me* Speech Analyzer	• This activity is for the warming up with the second AI app. Students speak for at least one minute to receive useful feedback from the speech analyser AI feature from the app. • Here students can practice their spontaneous speech with assigned topic by researcher and get feedback from the application such as grammar, vocabulary, and fluency.
Week-3 (first half)	Speaking practice with sound natural in English by *ELSA*	• There are three topics of speaking practice available that can be used by free users of *ELSA* namely *food, shopping,* and *work.* The AI of the application will ask students to perform their speech where each topic has five questions to hint the speech. • FOOD: *What is your favourite food? Do you prefer to eat at home or eat out? What foods do you like to snack on? What is your favourite dessert? Do you prefer sweet or salty foods?* • SHOPPING: *Do you enjoy shopping? What things do you like shopping for? Do you like shopping on the internet? Do you have a favourite store? Do you think men and women shop differently?* • WORK: *What work do you do? Why did you choose to do that type of work? Do you remember your first day of work? Do you enjoy your job? Would you like to change your job in the future?*
Week-3 (second half)	Speaking practice by shadowing with *SmallTalk2Me*	• A menu in the general speaking course provided by app will allow students to speak on imitating native speaker's pronunciation. • In this activity, students are asked to watch video first, and read the transcript out loud repeating intonation and pronunciation. Then, the app will give feedback about students' intonation and pronunciation.
Week-4 (first half)	Speaking practice with interview through *ELSA*	• Mock Interview-1 • Interview questions set about student self-introduction. In this section, students will be asked a list of commonly asked questions in an interview as in the following. • First, tell me about yourself; Tell me about your current position and responsibilities? How did you hear about this position? Tell me why you are interested in this position. What can you bring to the company? • Mock Interview-2 • Interview questions set about personal interests and skillsets. In this section, students will be asked a list of commonly asked questions in an interview as in the following. • Tell me about a time you demonstrated leadership skills. What do you find most interesting about our company? What is your ideal working environment? How do you deal with the pressure of stressful situations? Describe the latest project you worked on. • In this mock interview, students can opt one of two topics of the interview.

Table A1. *Cont.*

Time (in-week)	Types of class activities	Description of Practice Using AI apps
Week-4 (second half)	Speaking practice for boosting vocabulary with *SmallTalk2Me*	• Students practice for pronunciation by choosing 1 of 15 topics provided the app. The selected topic is generated by artificial intelligence. • First, students read the text out loud paying attention to the keywords. Second, students should answer three questions using key words provided by the app. Then, students tell a story that includes the key words given by the app by speaking at least 1 minute. Last, the app will give feedback based on students' speech.
Week-5 (first half)	Speaking practice by impromptu speech with *ELSA*	• Here students can practice their spontaneous speech with assigned topic by researcher and get feedback from the application. • Students are asked to record their voice to get speech analyser feedback.
Week-5 (second half)	Speaking practice with *SmallTalk2Me* to simulate English speaking level	• The app will ask students several questions to hint their speaking such as: *how are you today? where do you live? pronounce the quote from the given picture 1 to 5; read the text provided by app.* • Then, students are asked to speak for 2 to 3 minutes by telling their daily routines. This practice section is prepared by prompting questions based photos given by app, such as *where and when the photo was taken; what / who is in the photo; what is happening; why you keep the photo in your album; what is so special about this photo; and what emotions does it bring to you.* • Last, the app will give feedback of pronunciation, grammar, vocabulary, and fluency levels based on students' speech.
Week-6 (first half)	Post-questionnaire of anxiety (FLSA)	• Students are asked to answer questionnaire regarding with speaking anxiety after intervention.
Week-6 (second half)	Post-test of speaking (IELTS)	• Students are asked to test independently with IELTS speaking test simulator provided by the AI app after intervention. Then report their results by sharing links of test score given by the app.

References

[1] Gregersen, T., Horwitz, E.K., 2002. Language Learning and Perfectionism: Anxious and Non-Anxious Language Learners' Reactions to Their Own Oral Performance. The Modern Language Journal. 86(4), 562–570. DOI: https://doi.org/10.1111/1540-4781.00161

[2] Horwitz, E.K., Horwitz, M.B., Cope, J., 1986. Foreign Language Classroom Anxiety. The Modern Language Journal. 70(2), 125–132. DOI: https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

[3] MacIntyre, P.D., 2017. An overview of language anxiety research and trends in its development. In: Gkonou, C., Daubney, M., Dewaele, J. (eds.). New Insights into Language Anxiety: Theory, Research And Educational Implications. Multilingual Matters: Bristol, UK. pp. 11–30. DOI: https://doi.org/10.21832/9781783097722-003

[4] Brown, H.D., 2007. Principles of Language Learning And Teaching. Pearson Education: New York, USA.

[5] Ozturk, N., 2013. Using CALL in Language Teaching and Learning, in consideration of its Strengths and Limitations. Academia Medicine. 3(1), 36–41.

[6] Chapelle, C.A., Voss, E., 2008. Utilizing technology in language assessment. In: Hornberger, N. (ed.). Encyclopedia of Language and Education, 2nd ed. Springer: New York, USA. pp. 2263–2274. DOI: https://doi.org/10.1007/978-0-387-30424-3_172

[7] Qin, P., Zhu, Z., Yamashita, N., et al., 2025. AI-Based Speaking Assistant: Supporting Non-Native Speakers' Speaking in Real-Time Multilingual Communication. arXiv:2505.01678v1. DOI: https://doi.org/10.48550/arXiv.2505.01678

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

[9] Zou, B., Guan, X., Shao, Y., et al., 2023. Supporting Speaking Practice by Social Network-Based Interaction in Artificial Intelligence (AI)-Assisted Language Learning. Sustainability. 15(4), 2872. DOI: https://doi.org/10.3390/su15042872

[10] 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, 1255594. DOI: https://doi.org/10.3389/fpsyg.2023.1255594

[11] Li, Y., Chen, C-Y., Yu, D., et al., 2022. Using Chatbots to Teach Languages. In Proceedings of the 9th ACM Conference on Learning Scale, New York, USA, (1–3 June 2022); pp. 451–455. DOI: https://doi.org/10.1145/3491140.3528329

[12] Sheriffdeen, K., 2024. Enhancing Personalized Learning: The Role of Adaptive AI-Driven Natural Language Processing in Real-Time Feedback for Online Education. Journal of Advanced Research in Artificial Intelligence & Its Applications. 2(1), 76–83. DOI: https://doi.org/10.5281/zenodo.14292153

[13] Hsiang, O., Huang, L., 2024. Enhancing English Oral Communication Skills through Virtual Reality: A Study on Anxiety Reduction and Authentic Learning. English Language Teaching. 17(5), 1–11. DOI: https://doi.org/10.5539/elt.v17n5p1

[14] Kaplan-Rakowski, R., Gruber, A., 2023. The impact of high-immersion virtual reality on foreign language anxiety. Smart Learning Environments. 10(1), 46. DOI: https://doi.org/10.1186/s40561-023-00263-9

[15] Koul, R., Roy, L., Kaewkuekool, S., et al., 2009. Multiple goal orientations and foreign language anxiety. System. 37(4), 676–688. DOI: https://doi.org/10.1016/j.system.2009.09.011

[16] Dja'far, V.H., Hamidah, F.N., 2024. Improving English Pronunciation Skills through AI-Based Speech Recognition Technology. Ethical Lingua: Journal of Language Teaching and Literature. 11(2), 565–572. DOI: https://doi.org/10.30605/25409190.747

[17] 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. DOI: https://doi.org/10.54855/ijAILE.24122

[18] Dashela, T., Mustika, Y., 2022. Gender Differences in English Language Learning. Jurnal Ilmiah Spectral. 8(2), 92–105. DOI: https://doi.org/10.47255/f8aa1v48

[19] Erdiana, N., Bahri, Y.S., Akhmal, C.N., 2019. Male vs. female EFL students: Who is better in speaking skill? Studies in English Language and Education. 6(1), 131–140. DOI: https://doi.org/10.24815/siele.v6i1.13024

[20] Erviona, L., Arsyad, S., 2022. Gender Differences and their Impacts on Students' Performance in Speaking Ability. Journal of English for Specific Purposes in Indonesia. 1(1), 43–49.

[21] Wang, Y., Xue, L., 2024. Using AI-driven chatbots to foster Chinese EFL students' academic engagement: An intervention study. Computers in Human Behavior. 159, 108353. DOI: https://doi.org/10.1016/j.chb.2024.108353

[22] Mohammadkarimi, E., 2024. Exploring the Use of Artificial Intelligence in Promoting English Language Pronunciation Skills. LLT Journal: A Journal on Language and Language Teaching. 27(1), 98–115. DOI: https://doi.org/10.24071/llt.v27i1.8151

[23] 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. DOI: https://doi.org/10.4018/IJCALLT.2021070105

[24] Hongli, F., 2025. A Deep Learning-Based Blended Teaching Model for Enhancing English Proficiency in English Education. International Journal of Research and Innovation in Social Science. 8(3s), 6247–6266.

[25] Ye, J., Wang, S., Zou, D., et al., 2025. Position: LLMs Can be Good Tutors in Foreign Language Education. Computation and Language. arXiv preprint arXiv:2502.05467. DOI: https://doi.org/10.48550/arXiv.2502.05467

[26] Kang, B.O., Jeon, H.B., Lee, Y.K., 2024. AI-based language tutoring systems with end-to-end automatic speech recognition and proficiency evaluation. ETRI Journal. 46(1), 48–58. DOI: https://doi.org/10.4218/ETRIJ.2023-0322

[27] Baharloo, A., Baghi, A.M., 2024. The Impact of AI-Assisted Learning on EFL Speaking Skills: A Mixed-Methods Study in the Iranian Context. Technology Assisted Language Education. 2(4), 69–96. DOI: https://doi.org/10.22126/TALE.2025.11299.1070

[28] Khalaf, M., Makhlouf, 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), 36–57. DOI: https://doi.org/10.5296/ijele.v9i2.18782

[29] Pertiwi, P.C., Segoh, D., Rohmadhani, A., 2025. Artificial Intelligence in Academic Environments: Reducing or Increasing Foreign Language Anxiety? Inspiring: English Education Journal. 8(1), 100–131. DOI: https://doi.org/10.35905/inspiring.v8i1.13009

[30] Hapsari, I.P., Wu, T.T., 2022. AI Chatbots Learning Model in English Speaking Skill: Alleviating Speaking Anxiety, Boosting Enjoyment, and Fostering Critical Thinking. In: Huang, Y.M., Cheng, S.C., Barroso, J., et al. (eds.). Innovative Technologies and Learning. Springer: Cham, Switzerland. pp. 444–453.

DOI: https://doi.org/10.1007/978-3-031-15273-3_49

[31] Ballıdağ, M., Aydın, S., 2025. A comparison of the effects of AI-based chatbots and peer interactions on speaking anxiety among EFL learners. Future in Educational Research. 3(2), 224–238. DOI: https://doi.org/10.1002/fer3.70005

[32] El Shazly, R., 2021. Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. Expert Systems. 38(3), e12667. DOI: https://doi.org/10.1111/exsy.12667

[33] Jr., F.V.G., Pascual, N.F., Sybingco, M.E., et al., 2024. Addressing Public Speaking Anxiety with an AI Speech Coach. In Proceeding of the 32nd International-al Conference on Computers in Education, Quezon City, Philippines, 25-29 November 2024; 1–9. DOI: https://doi.org/10.58459/ICCE.2024.4955

[34] Møgelvang, A., Bjelland, C., Grassini, S., et al., 2024. Gender Differences in the Use of Generative Artificial Intelligence Chatbots in Higher Education: Characteristics and Consequences. Education Sciences. 14(12), 1363. DOI: https://doi.org/10.3390/educsci14121363

[35] Huang, F., Peng, D., Teo, T., 2025. AI Affordances and EFL Learners' Speaking Engagement: The Moderating Roles of Gender and Learner Type. European Journal of Education. 60(1), e70041. DOI: https://doi.org/10.1111/ejed.70041

[36] Fihris, Alfianika, N., Nasikhin, N., 2024. Differences in Male and Female Responses to Artificial Intelligence Integration for Education Faculty: Study of Thailand International Students at Islamic Universities in Indonesia. EL-HIKMAH: Jurnal Kajian Dan Penelitian Pendidikan Islam. 18(1), 31–60. DOI: https://doi.org/10.20414/elhikmah.v18i1.10037

[37] Kim, K., Kwon, K., 2024. Designing an Inclusive Artificial Intelligence (AI) Curriculum for Elementary Students to Address Gender Differences With Collaborative and Tangible Approaches. Journal of Educational Computing Research. 62(7), 1617–1644. DOI: https://doi.org/10.1177/07356331241271059

[38] Harb, J., Bakar, N.A., Krish, P., 2014. Gender differences in attitudes towards learning oral skills using technology. Education and Information Technology. 19(4), 805–816. DOI: https://doi.org/10.1007/s10639-013-9253-0

[39] MacIntyre, P.D., Noels, K.A., Clément, R., 1997. Biases in self-ratings of second language proficiency: The role of language anxiety. Language Learning. 47(2), 265–287. DOI: https://doi.org/10.1111/0023-8333.81997008

[40] Maurat, J.I.C., Isip, E.V., Lumabas, A.G.R., 2024. A Comparative Study of Gender Differences in the Utilization and Effectiveness of AI-Assisted Learning Tools in Programming Among University Students. In Proceedings of the 2024 International Conference on Artificial Intelligence and Teacher Education, Beijing China, 12–14 October 2014; pp. 30–34. DOI: https://doi.org/10.1145/3702386.3702392

[41] Csizér, K., Albert, Á., 2024. Gender-Related Differences in the Effects of Motivation, Self-Efficacy, and Emotions on Autonomous Use of Technology in Second Language Learning. Asia-Pacific Education Researcher. 33(4), 819–828. DOI: https://doi.org/10.1007/s40299-023-00808

[42] IELTS, 2025. Understanding IELTS test scores. Available from: https://ielts.org/take-a-test/your-results/ielts-scoring-in-detail (cited 23 May 2025).

[43] Hashemi, A., Daneshfar, S.A., 2018. A review of the IELTS test: Focus on validity, reliability, and washback. Indonesian Journal of English Language Teaching and Applied Linguistics. 3(1), 39–52. DOI: https://doi.org/10.21093/ijeltal.v3i1.123

[44] Gagen, T., Faez, F., 2024. The predictive validity of IELTS scores: a meta-analysis. Higher Education Research and Development. 43(4), 873–888. DOI: https://doi.org/10.1080/07294360.2023.2280700

[45] Hayashi, K., Sato, T., 2024. The Effectiveness of ChatGPT in Enhancing English Language Proficiency and Reducing Second Language Anxiety (L2). In Proceeding of the WorldCALL 2023, Chiang Mai, Thailand, 9–13 November 2023; pp. 201–208. DOI: https://doi.org/10.22492/issn.2759-1182.2023.23

[46] Ataiefar, F., Sadighi, F., 2017. Lowering Foreign Language Anxiety through Technology: A Case of Iranian EFL Sophomore Students. English Literature and Language Review. 3(4), 23–34.

[47] Wardat, S., Akour, M., 2025. AI-Powered Language Learning Tools and Their Impact on EFL Students' Speaking Anxiety in Jordanian Universities. Journal of Posthumanism. 5(3), 756–772. DOI: https://doi.org/10.63332/joph.v5i3.785

[48] Hồng, N.X., Thuý, L.T., Khôi, Đ.Đ., et al., 2024. Students' Attitudes toward Utilizing AI-Based Technologies to Their Speaking Proficiency: A Case Study. Tạp Chí Khoa Học. 21(5), 957–970. DOI: https://doi.org/10.54607/hcmue.js.21.5.4270 (in Vietnamese)

[49] Qassrawi, R.M., ElMashharawi, A., Itmeizeh, M., et al., 2024. AI-Powered Applications for Improving EFL Students' Speaking Proficiency in Higher Education. Forum for Linguistic Studies. 6(5), 535–549. DOI: https://doi.org/10.30564/fls.v6i5.6966

[50] Tajik, A., 2025. Integrating AI-Driven Emotional Intelligence in Language Learning Platforms to Improve

English Speaking Skills through Real-Time Adaptive Feedback. Research Square. 1–41. Preprints.org. DOI: https://doi.org/10.21203/rs.3.rs-5919944/v1

[51] Dakhil, T.A., Karimi, F., Abbas, R., et al., 2025. The Effect of Artificial Intelligence (AI)-Mediated Speaking Assessment on Speaking Performance and Willingness to Communicate of Iraqi EFL Learners. International Journal of Language Testing. 15(2), 1–18.

[52] Tareh, F., Bijani, H., 2024. EFL Learners' Speaking Enhancement through WhatsApp Messages: A Case of Learners' Gender and Affection. International Journal of Research in English Education. 9(2), 46–59.

[53] Ericsson, E., Johansson, S., 2023. English speaking practice with conversational AI: Lower secondary students' educational experiences over time. Computers and Education: Artificial Intelligence. 5, 100164. DOI: https://doi.org/10.1016/j.caeai.2023.100164

[54] 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

[55] Mahmoodzadeh, M., 2012. Investigating Foreign Language Speaking Anxiety within the EFL Learner's Interlanguage System: The Case of Iranian Learners. Journal of Language Teaching and Research. 3(3), 466–476. DOI: https://doi.org/10.4304/jltr.3.3.466-476

[56] Öztürk, G., Gürbüz, N., 2013. The Impact of Gender on Foreign Language Speaking Anxiety and Motivation. Procedia - Social and Behavioral Sciences. 70, 654–665. DOI: https://doi.org/10.1016/j.sbspro.2013.01.106

[57] Çağatay, S., 2015. Examining EFL Students' Foreign Language Speaking Anxiety: The Case at a Turkish State University. Procedia - Social and Behavioral Sciences. 199, 648–656. DOI: https://doi.org/10.1016/j.sbspro.2015.07.594

[58] Rafek, M.B., Ramli, N.H.L.B., Iksan, H.B., et al., 2014. Gender and Language: Communication Apprehension in Second Language Learning. Procedia-Social and Behavioral Sciences. 123, 90–96. DOI: https://doi.org/10.1016/j.sbspro.2014.01.1401

[59] Fauzi, I., Asi, N., 2023. Examining Speaking Anxiety of Indonesian Learners of English: A Case of University Students. Indonesian Journal of English Language Teaching and Applied Linguistics. 7(2), 263–275. DOI: https://doi.org/10.21093/ijeltal.v7i2.1338

[60] Xia, Q., Chiu, T.K.F., Chai, C.S., 2023. The moderating effects of gender and need satisfaction on self-regulated learning through Artificial Intelligence (AI). Education and Information Technology. 28(7), 8691–8713. DOI: https://doi.org/10.1007/s10639-022-11547-x

[61] Bolívar-Cruz, A., Verano-Tacoronte, D., 2025. Is Anxiety Affecting the Adoption of ChatGPT in University Teaching? A Gender Perspective. Technology, Knowledge and Learning. 30(2), 1–20. DOI: https://doi.org/10.1007/s10758-025-09830-0

[62] Shadiev, R., Feng, Y., Zhussupova, R., et al., 2024. Effects of speech-enabled corrective feedback technology on EFL speaking skills, anxiety and confidence. Computer Assisted Language Learning. 1–37. DOI: https://doi.org/10.1080/09588221.2024.2430761