



Enhancing Speaking Proficiency: A Mixed-Methods Study on AI-Assisted Pre-Task Planning for Non-English Major Students

MA Lam My An

Department of Linguistics, Ho Chi Minh university of foreign languages – Information Technology, Vietnam

* Corresponding Author: MA Lam My An

Article Info

ISSN (online): 2583-8261

Volume: 04

Issue: 06

November – December 2025

Received: 09-09-2025

Accepted: 11-10-2025

Published: 03-11-2025

Page No: 41-49

Abstract

The potential of Artificial Intelligence (AI) to aid language acquisition has gained attention as it is increasingly incorporated into education. This study investigates the role of AI in assisting non - English major university students in the pre-task planning stage of speaking activities. Data were collected using a mixed-methods methodology, integrating qualitative information from student interviews and classroom observations with quantitative data from speaking evaluations conducted from control group and experimental group. Results show that when compared to students who employed conventional preparation techniques, those who utilized AI-supported pre-task planning tools performed better in speaking in terms of fluency, coherence, vocabulary, and grammar. Additionally, qualitative data highlight students' positive perceptions of AI tools as supportive, interactive, and personalized learning aids. The study comes to the conclusion that integrating AI into pre-task planning can improve the development of speaking skills and encourage learner autonomy in EFL environments, especially for students who are not majoring in English. Implications for language instruction and technology-enhanced learning are also discussed.

DOI: <https://doi.org/10.54660/IJSSER.2025.4.6.41-49>

Keywords: Artificial Intelligence (AI), Pre-Task Planning, Speaking Skills, Non-English Major Students, Mixed-Methods

1. Introduction

Artificial intelligence's (AI) quick development has created new opportunities to improve language acquisition, especially in circumstances where English is being learned as a foreign language (EFL) (Kuddus, 2022) ^[15]. Learners can receive a multitude of linguistic input and instant, tailored feedback via AI tools, such as generative language models and intelligent tutoring systems. AI's ability to enhance writing and general language acquisition has been extensively studied, but its potential to aid in the development of speaking abilities, particularly during the pre-task planning phase, has received little attention. Pre-task planning, as defined by Ellis (2005) ^[10], involves preparing ideas, vocabulary, and structures before engaging in a communicative task. Such preparation has been shown to increase speaking complexity, correctness, and fluency (Foster and Skehan, 1996; Ortega, 1999) ^[11, 20]. Nonetheless, learners who do not major in English frequently encounter obstacles at this phase, such as a restricted vocabulary, trouble structuring their thoughts, and anxiety when speaking. Positive student perspectives of AI-supported speaking preparation have been emphasized in early research (e.g., Zou *et al.*, 2020; Amin, 2023) ^[30, 2], which have noted enhanced motivation, confidence, and opportunity for interactive practice. However, controlled empirical research assessing real performance improvements from AI-assisted planning is still lacking, especially for non-English majors.

This paper hopefully addresses that gap by investigating how AI-assisted pre-task planning affects the speaking performance of non-English major university students. The study employs a mixed-methods approach, integrating qualitative information from student interviews and classroom observations with quantitative evaluations of vocabulary, grammar, coherence, and fluency. With implications for both EFL teaching and technology-enhanced learning environments, the findings add to the expanding body of research on AI in language education.

2. Literature review

2.1. Artificial Intelligence in Language Education

Artificial Intelligence (AI) in education refers to the application of computational techniques, such as machine learning, natural language processing (NLP), and adaptive systems, to support and enhance teaching and learning processes. Over the past three decades, the field of Artificial Intelligence in Education (AIED) has evolved substantially since its early conceptualizations (O'Shea & Self, 1986)^[22], with rapid advancements attracting interdisciplinary attention from both computer science and education researchers (Chen *et al.*, 2020)^[7]. As Breines and Gallagher (2020)^C note, AI has the potential to transform traditional models of instruction—where students are passive recipients of knowledge—into highly personalized, scalable, and affordable learning environments. This growth reflects not only the technological possibilities of AI but also its strategic role in shaping educational offerings to meet changing learner demands (Moscardini *et al.*, 2022)^[19] and contributing to policy debates on AI literacy in education (Miao & Holmes, 2021)^[18].

Common AI applications in English as a Foreign Language (EFL) instruction include automated feedback systems, intelligent tutoring systems, chatbots, speech recognition, and content recommendation engines. These technologies can analyze learners' performance, detect gaps, and adapt instruction accordingly (Chaudhry & Kazim, 2022; Hwang *et al.*, 2020)^[6, 13]. Personalized learning is a central objective of AI in education, with systems providing targeted support based on learners' status, preferences, and characteristics (Hwang, 2014)^[12]. In the EFL context, AI chatbots have been shown to enhance mastery of grammar and as well as improve oral communication, listening, reading, and argumentative writing skills vocabulary (Wang & Petrina, 2013)^[26]. Moreover, AI-powered conversational agents can offer real-time feedback that boosts both motivation and phonological awareness, leading to measurable improvements in pronunciation (Koka, 2024)^[14].

Research has documented multiple benefits of AI integration in EFL learning including greater learner motivation (Kukulka-Hulme, 2021)^[16] and enhanced feedback quality (Yu & Xu, 2020). By simulating the role of a responsive tutor, AI applications can deliver timely assistance, recommend resources, and sustain learner engagement (El Shazly, 2021)^[9].

2.2. Pre-task Planning in Speaking Development

Pre-task planning (PTP) is widely recognized as a pivotal element in enhancing second language (L2) oral performance, particularly regarding complexity, fluency, and, to a certain extent, accuracy. Early theoretical frameworks, such as those proposed by Skehan (1998)^[23] and Wendel (1997)^[27], posited that planning facilitates greater speech complexity by mitigating the cognitive demands of real-time communication. This allows learners to dedicate more attentional resources to task preparation, leading to more sophisticated linguistic output. Crookes (1989)^[8] first hypothesized that planning would promote more complex and fluent language, although sometimes at the cost of accuracy. This effect has since been linked to the role of PTP in alleviating working memory limitations, thereby enabling learners to better map linguistic forms to meaning and access less-automatized resources (Ellis, 2005)^[10]. Over the past two decades, PTP has become a significant area of research within task-based language teaching, with numerous studies affirming its positive influence on fluency (Abdi, Eslami, & Zahedi, 2012)^[1] and, in some cases, accuracy (Foster & Skehan, 1996)^[11]. The theoretical underpinnings of this phenomenon were further developed by Swain (1985)^[24], who argued that language production involves a deeper level of semantic processing, and by Ortega (2005)^[21], who emphasized that strategic planning allows learners to operate at the upper limits of their interlanguage proficiency without the pressure of time constraints. However, empirical findings are not entirely consistent. For instance, Wigglesworth (1997)^[28] found that planning primarily enhanced accuracy and complexity for high-proficiency learners and in more demanding tasks, whereas Yuan and Ellis (2003)^[29] reported gains in complexity and lexical variety but not in accuracy. The effects of different types of planning have also been explored, with Atai and Nasiri (2017)^[3] observing that strategic planning benefited complexity and fluency in simple tasks, while online and joint planning more often improved accuracy. Collectively, the body of literature suggests that PTP fosters greater linguistic complexity (Wendel, 1997)^[27] and functions as a "problem-solving activity" (Ellis, 2005)^[10] wherein learners strategically select appropriate linguistic

devices to meet communicative goals. Consequently, PTP is considered a valuable pedagogical tool in L2 instruction as it maximizes learners' readiness to engage in communicative tasks, irrespective of their prior knowledge (Bui, 2014) ^[5].

2.3. AI support for Speaking skill

Recent advances in AI have opened new possibilities for enhancing speaking preparation in EFL contexts. New research has looked into using AI to help students before speaking assignments, especially in higher education. Zou *et al.* (2020) ^[30], for instance, examined university students' attitudes towards AI-assisted mobile applications in English for Academic Purposes (EAP) courses. Their findings revealed that students preferred using AI applications to develop speaking ability, with AI-supported feedback serving as a strong motivator. Such tools not only increased learners' confidence but also provided more opportunities for communicative interaction, ultimately fostering speaking skill development.

AI-powered tools, such as automated assessment platforms and intelligent tutoring systems, can provide immediate, customized assistance to help with pre-task planning. These technologies can give immediate content categorization, produce topic-relevant terminology, and suggest sentence patterns and discourse connectors, as noted by Amin (2023) ^[2] and Umar (2024) ^[25]. AI frees up learners to focus more on speaking clearly and smoothly by lowering the cognitive burden associated with idea development and language retrieval. Furthermore, learners can improve their output during the preparation phase by using AI's real-time, adaptive feedback capabilities, which will increase their confidence and preparedness for actual performance.

Although numerous studies have examined AI as a supportive tool in pre-task planning for various language skills, much of this work has centered on writing, reading, or general language development. Little research has specifically investigated AI-assisted pre-task planning in the context of speaking skills, particularly for non-English majors. This gap highlights the need for empirical studies that measure both performance outcomes and learner perceptions to determine the extent to which AI can enhance speaking preparation and delivery.

3. Methodology

3.1. Research design

This study employed a quasi-experimental design with a control group and an experimental group (as full randomization of participants was not feasible within classroom settings) to investigate the effects of AI-assisted pre-task planning on students' speaking performance. Specifically, ChatGPT (version 3.5) is used by all students in this research. The experimental group received structured

AI-based pre-task planning support, while the control group completed the tasks relying on traditional preparation methods (self-planning without AI assistance). Also, a matched-pair design was used to ensure comparability between participants in the two groups.

3.2. Participants

The study involved 46 non-English major students, divided equally into a control group (n = 23; 4 males and 19 females) and an experimental group (n = 23; 5 males and 18 females). They were selected based on the results of an institutional English proficiency test, ensuring intermediate-level equivalence. Students were then matched by proficiency and assigned to either the control group (n=23) or experimental group (n=23). This pairing minimized extraneous variables and enhanced the reliability of the results.

3.3. Materials

The main tool was a picture description task (TOEIC Speaking Part 2 format). Their performance was evaluated using a comprehensive rubric with four distinct criteria: Pronunciation, Vocabulary, Grammar, and Fluency and Coherence was adapted from TOEIC band descriptors from ETS organization. Each criterion was scored on a scale from 0 to 10, resulting in a total possible score of 40 for each participant. Two independent raters evaluated all performances, and inter-rater reliability was calculated. To avoid bias, raters were blind to group assignment and evaluated the recordings in randomized order. For statistical analyses, a final mean score was derived for each participant by averaging the ratings from the two evaluators. These final means served as the dataset for group comparisons. An additional student perception survey (5-point Likert scale and open-ended items, adapted from Likert Scale in Social Sciences Research: Problems and Difficulties (2022) was given to participants to collect qualitative data on their attitudes toward the activities and, for the experimental group, their opinions on the use of AI as a preparatory aid.

3.4. Procedure

The speaking tasks were administered to both groups under identical classroom conditions. The experimental group was given AI-generated support during the pre-task planning stage, which included personalized vocabulary suggestions, grammar reminders, and model responses for inspiration. In contrast, the control group completed the pre-task planning and the speaking task without any additional support. Following the completion of the tasks, all student responses were transcribed. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC, two-way mixed model, absolute agreement)

A qualitative analysis of these transcripts was also conducted to examine specific linguistic features such as the number of

sentences, sentence types (simple, compound, complex).

3.5. Data analysis

Quantitative analysis

The collected data were analyzed using both quantitative and qualitative methods. For the quantitative analysis, descriptive statistics, including mean scores, were used to compare the speaking performance of the control and experimental groups across the five criteria - pronunciation, vocabulary, grammar, and fluency and coherence—on a 0–10 scale, resulting in a maximum total score of 40.

Qualitative

Speaking transcripts were coded for sentence structure (simple, compound, complex, compound-complex) and

vocabulary use. Two coders independently analyzed 20% of the data, reaching 90% agreement. Disagreements were resolved through discussion.

4. Findings and Discussion

4.1. Quantitative findings and discussion

4.1.1. Inter-rater Reliability Analysis

To ensure scoring consistency, inter-rater reliability was examined for both groups. For the control group ($N = 23$), the mean score for Rater 1 was 4.74, while Rater 2's mean was 4.89, yielding a final group mean of 4.85. Reliability analysis indicated excellent agreement for both groups, with $ICC(3,1) = .92$ for the control group and $ICC(3,1) = .91$ for the experimental group. These results confirm that both raters scored consistently across both groups.

Table 1: Scores from Rater 1 and Rater 2 for the Control Group

Student	Rater 1 Average score	Rater 2 Average score	Final mean
Participant 1	5	5	5
Participant 2	4	3.5	3.75
Participant 3	5	5	5
Participant 4	5.5	5.5	5.5
Participant 5	5	5	5
Participant 6	8	7.75	7.88
Participant 7	5.75	6.25	6
Participant 8	4.5	4.25	4.38
Participant 9	6	6.75	6.38
Participant 10	5	5.75	5.38
Participant 11	4	4.5	4.25
Participant 12	5	5.5	5.25
Participant 13	5.5	5.25	5.38
Participant 14	4	4.75	4.38
Participant 15	4	4	4
Participant 16	4	3	3.5
Participant 17	5	5.25	5.12
Participant 18	3	3	3
Participant 19	4	4	4
Participant 20	3	3.25	3.12
Participant 21	4	3.25	3.62
Participant 22	5	4.25	4.62
Participant 23	7	7	7
	Mean score = 4.74	Mean score = 4.89	Mean score = 4.85

The analysis of the experimental group below shows strong consistency between the two raters. For this group ($N = 23$), Rater 1's mean was 5.41 and Rater 2's mean was 5.63, with

a final group mean of 5.52. These results strengthen the reliability of the scoring system across groups.

Table 2: Scores from Rater 1 and Rater 2 for the Experimental Group

Student	Rater 1 Average score	Rater 2 Average score	Final mean
Participant 24	6	6	6
Participant 25	4.75	5	4.875
Participant 26	6.5	6.5	6.5
Participant 27	5.75	5.875	5.8125
Participant 28	4.75	5	4.875
Participant 29	6.5	6.5	6.5
Participant 30	5.75	6.25	6
Participant 31	5.75	5.75	5.75
Participant 32	4.75	5.125	4.9375
Participant 33	5.75	6.25	6
Participant 34	6.25	6.75	6.5
Participant 35	6.25	6.5	6.375

Participant 36	6	5.875	5.9375
Participant 37	4.75	5.875	5.3125
Participant 38	3.75	4.125	3.9375
Participant 39	3	3	3
Participant 40	5	5.125	5.0625
Participant 41	3.75	4.125	3.9375
Participant 42	4.75	4.75	4.75
Participant 43	6.25	6.375	6.3125
Participant 44	5.75	5.625	5.6875
Participant 45	7	6.75	6.875
Participant 46	5.75	6.25	6
	Mean score = 5.41	Mean score = 5.63	Mean score = 5.52

4.1.2. AI in pre-task planning for Speaking

For statistical analysis, the final mean score for each participant was calculated as the average of the two raters' evaluations. These final mean scores were then used in the independent-samples t-test comparing the control and experimental groups.

An independent-samples t-test was conducted to compare the speaking performance of the two groups. The control group had a mean score of $M = 4.85$, $SD = 1.19$, whereas the experimental group achieved a higher mean of $M = 5.52$, $SD = 0.95$. The difference between the groups was statistically

significant, $t(44) = -2.07$, $p = .044$, $d = 0.62$.

These results suggest that AI-assisted pre-task planning positively influenced speaking performance. The experimental group outperformed the control group not only in overall scores but also across individual criteria (pronunciation, vocabulary, fluency & coherence, and grammar). The boxplot below illustrates that the experimental group's scores were more consistently clustered at higher levels, whereas the control group's scores were more widely dispersed.

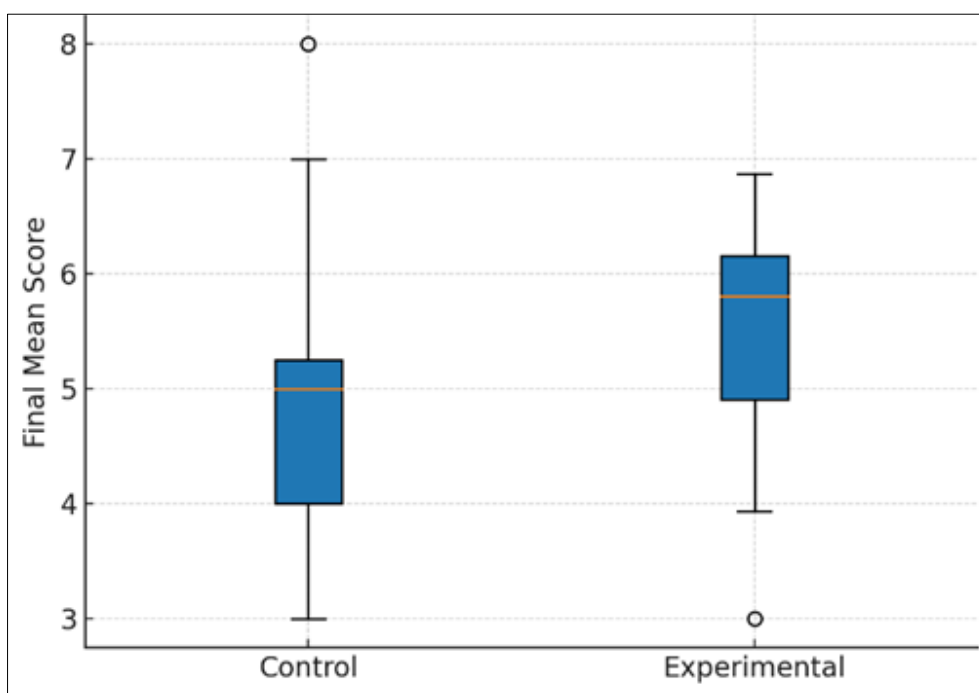


Fig 1: Comparison of speaking scores between control and experimental group

4.2. Qualitative findings and discussion

The qualitative analysis examined students' speaking transcripts to explore differences in sentence complexity, vocabulary level, and overall sentence production between the control group and the experimental group. The aim was to determine how AI-assisted pre-task planning influenced not only the accuracy and fluency of speech, as shown in the quantitative results, but also the quality and richness of language output.

4.2.1. Sentence Complexity

While the total number of sentences produced was relatively similar across the two groups, the content density in the experimental group was higher. Many participants shifted from short, fragmented statements to longer, well-developed sentences.

Analysis revealed a clear distinction in the distribution of sentence types. Students in the control group predominantly relied on simple sentences, with minimal use of compound

and complex structures. For example, several control group participants consistently produced structures such as *"I like reading books"* without extending ideas or linking clauses. In contrast, experimental group participants showed a marked increase in the use of compound and complex sentences. AI-assisted pre-task planning appeared to encourage students to

connect ideas logically using coordinating and subordinating conjunctions. Examples included *"I enjoy reading books because it helps me relax, and I can learn new vocabulary at the same time."* This shift suggests that the pre-task planning tool provided scaffolding that allowed learners to take greater syntactic risks, leading to more varied and cohesive speech.

Table 3: Sentence types across two groups

Pair	Control group	Experimental group	Control group: Sentences	Control group: Sentence Types (Simple/Compound/Complex)	Experimental group: Sentences	Experimental group: Sentence Types (Simple/Compound/Complex)
1	Participant 1	Participant 24	5	3 / 1 / 1	5	3 / 2 / 0
2	Participant 2	Participant 25	3	3 / 0 / 0	4	3 / 1 / 0
3	Participant 3	Participant 26	4	1 / 1 / 2	3	0 / 1 / 2
4	Participant 4	Participant 27	6	4 / 2 / 0	6	3 / 2 / 1
5	Participant 5	Participant 28	3	2 / 0 / 1	3	2 / 1 / 0
6	Participant 6	Participant 29	5	3 / 1 / 1	4	2 / 1 / 1
7	Participant 7	Participant 30	4	2 / 1 / 1	5	2 / 2 / 1
8	Participant 8	Participant 31	4	2 / 1 / 1	3	2 / 0 / 1
9	Participant 9	Participant 32	3	2 / 0 / 1	4	3 / 0 / 1
10	Participant 10	Participant 33	4	3 / 1 / 0	5	3 / 1 / 1
11	Participant 11	Participant 34	3	2 / 1 / 0	5	3 / 1 / 1
12	Participant 12	Participant 35	4	2 / 1 / 1	4	2 / 1 / 1
13	Participant 13	Participant 36	4	2 / 1 / 1	5	3 / 1 / 1
14	Participant 14	Participant 37	3	2 / 0 / 1	3	2 / 0 / 1
15	Participant 15	Participant 38	3	2 / 0 / 1	2	2 / 0 / 0
16	Participant 16	Participant 39	2	2 / 0 / 0	2	2 / 0 / 0
17	Participant 17	Participant 40	3	2 / 1 / 0	3	2 / 1 / 0
18	Participant 18	Participant 41	3	2 / 1 / 0	3	2 / 1 / 0
19	Participant 19	Participant 42	3	2 / 1 / 0	4	3 / 1 / 0
20	Participant 20	Participant 43	4	3 / 1 / 0	4	2 / 1 / 1
21	Participant 21	Participant 44	4	2 / 1 / 1	4	2 / 1 / 1
22	Participant 22	Participant 45	3	2 / 1 / 0	4	3 / 1 / 0
23	Participant 23	Participant 46	3	2 / 0 / 1	4	2 / 1 / 1

In addition to producing more complex sentences, several participants in the experimental group also showed signs of broader lexical choice, although vocabulary differences were not systematically analyzed in this study.

4.2.2. Patterns and Notable Cases

The qualitative findings reinforce the quantitative trends by showing that AI-assisted pre-task planning not only improved speaking scores but also enhanced linguistic sophistication. This aligns with previous studies indicating that structured planning time enables learners to organize ideas, retrieve more varied vocabulary, and attempt complex sentence structures. The data suggests that the observed improvements stem from two main mechanisms: cognitive support, where AI prompts helped students pre-select appropriate vocabulary and sentence patterns, and a confidence boost, as exposure to model responses encouraged learners to take linguistic risks, leading to more diverse and complex output. Overall, the combination of quantitative and qualitative evidence suggests that AI-assisted pre-task planning can positively influence both the form and content of learner speech, leading to more accurate, fluent, and lexically varied performance.

4.2.3. Students' perceptions in using AI to assist speaking skill.

4.2.3.1. Improved Idea Generation and Organization

The higher fluency and coherence scores in the experimental group align with students' reports that AI helped them structure their ideas before speaking. For example, one participant noted, *"It helped me outline what to say, so I didn't forget the main points."* Another added, *"I could arrange my ideas clearly and follow them while speaking."* This suggests that AI-assisted planning directly contributed to more coherent, logically organized speech.

4.2.3.2. Vocabulary Expansion

The vocabulary score gains in the experimental group are consistent with student feedback about learning and applying richer vocabulary. One participant said, *"It gave me new words I don't usually use, like 'sustainable' and 'innovative'."* Another commented, *"I could learn synonyms and avoid repeating the same words."* These findings support Zou *et al.*'s (2020)^[30] claim that AI tools provide topic-related vocabulary input, and further confirm Amin's (2023) suggestion that real-time AI feedback can encourage

learners to incorporate more precise and varied vocabulary in speaking.

4.2.3.3. Increased Confidence in Speaking

Higher fluency scores may also be linked to reduced hesitation, as students felt more prepared. One shared, *"I felt less nervous because I had prepared the sentences with AI beforehand."* Another explained, *"It helped me speak more smoothly without long pauses."* This preparation appears to have minimized speech disruptions and increased overall delivery confidence. This boost in confidence is consistent with Amin's (2023)^[2] and Umar's (2024)^[25] findings that AI-generated feedback reduces learner anxiety and encourages risk-taking in language production.

4.2.3.4. Grammar and Sentence Complexity

The grammar score difference between groups is supported by reports of AI introducing more complex structures. For instance, one student said, *"It showed me how to connect ideas with words like 'although' or 'therefore'."* Another reflected, *"My sentences became longer and more natural after seeing AI examples."* This aligns with Ellis's (2005)^[10] claim that planning supports syntactic complexity, which indirectly strengthens grammar use in performance.

4.2.3.5. Critical Awareness of AI Limitations

Even with higher scores, some students remained cautious about AI-generated language. One observed, *"Sometimes the AI made sentences too formal, so I changed them to sound more natural."* Another warned, *"I still checked the grammar myself because sometimes it was wrong."* This critical awareness indicates that while AI was a useful tool, students retained an active role in editing and personalizing their output.

5. Implication

The findings of this study offer several significant pedagogical and practical implications for English language teaching, particularly within contexts involving non-English major students. First, the significant gains in the experimental group's vocabulary, fluency, and sentence complexity indicate that it is possible to scaffold learners' speaking performance by including AI tools into the pre-task planning stage. To enhance traditional speaking practice, teachers could think about implementing AI-supported preparatory tasks, such as creating topic-specific vocabulary lists, laying out speaking points, and offering real-time linguistic suggestions. Second, the favorable opinions expressed by students in the experimental group suggest that AI can be a significant factor in lowering speaking fear and increasing self-assurance. AI technologies can act as a nonjudgmental preparation partner, enabling students to practice and polish their ideas prior to presenting in class, as anxiety and disorganization of ideas are frequent issues for non-English

major students. Third, the findings of this study show how AI integration can improve strategic and rehearsal planning while reaffirming the significance of these processes as outlined in previous studies (Ellis, 2005; Foster & Skehan, 1996)^[10-11]. Therefore, especially for low-proficiency learners who require extra linguistic and cognitive help, language programs may want to intentionally incorporate AI-based pre-task preparation stages into speaking assignments. Despite its promising findings, the study is not without limitations. First, the sample size (N = 46) was relatively small and drawn from a single institution, which limits the generalizability of the results. Replication with larger and more diverse populations is necessary to strengthen external validity. Second, the study adopted a quasi-experimental design with a relatively short intervention period. It remains unclear whether the observed gains would persist over the long term or transfer to other speaking contexts. Longitudinal research could examine whether AI-assisted planning leads to sustained improvement in communicative competence. Third, while the study focused on quantitative speaking scores and qualitative transcript analysis, other aspects of speaking performance—such as prosody, pragmatics, or interactional competence—were not systematically explored. Future studies could adopt a more comprehensive framework to capture these dimensions. Finally, although inter-rater reliability was high, the scoring process relied on two raters using analytic rubrics. Incorporating automated scoring or triangulating with peer/self-assessments could provide additional perspectives on learner performance.

6. Conclusion

This study set out to explore the role of AI-assisted pre-task planning in enhancing the speaking performance of non-English major students. Through a mixed-methods design, combining speaking test scores and qualitative data from transcripts and student perceptions, the findings showed that AI-supported planning significantly improved fluency, vocabulary use, grammar, and sentence complexity compared to traditional preparation methods. The independent-samples t-test confirmed a statistically significant difference between the experimental and control groups, highlighting AI's measurable impact on speaking performance. Moreover, students reported greater confidence, better idea organization, and richer vocabulary when supported by AI tools. While the results are promising, limitations such as the small sample size, quasi-experimental design, and focus on short-term outcomes suggest that further research is needed to examine long-term effects and test different AI applications across broader learner populations. Nevertheless, the study provides strong evidence that integrating AI into pre-task planning can serve as a valuable pedagogical strategy in EFL classrooms, especially for learners outside English-major programs.

7. References

1. Abdi M, Eslami H, Zahedi Y. The impact of pre-task planning on the fluency and accuracy of Iranian EFL learners' oral performance. *Procedia Soc Behav Sci.* 2012;69:2281-8.
2. Amin MYM. AI and Chat GPT in language teaching: enhancing EFL classroom support and transforming assessment techniques. *Int J High Educ Pedagogies.* 2023;4(4):1-15. doi:10.33422/ijhep.v4i4.554
3. Atai MR, Nasiri M. An investigation into the effects of joint planning on complexity, accuracy, and fluency across task complexity. *J Engl Lang Teach Learn.* 2017;20:49-74.
4. Breines MR, Gallagher M. A return to Teacherbot: rethinking the development of educational technology at the University of Edinburgh. *Teach High Educ.* 2020:1-15. doi:10.1080/13562517.2020.1825373
5. Bui G. Task readiness. In: Skehan P, editor. *Processing perspectives on task performance.* Amsterdam: Benjamins; 2014. p. 63-94.
6. Chaudhry MA, Kazim E. Artificial Intelligence in Education (AIED): a high-level academic and industry note 2021. *AI Ethics.* 2022;2(1):157-65.
7. Chen L, Chen P, Lin Z. Artificial intelligence in education: a review. *IEEE Access.* 2020;8:75264-78.
8. Crookes G. Planning and interlanguage variation. *Stud Second Lang Acquis.* 1989;11(4):367-83.
9. El Shazly R. Effects of artificial intelligence on English speaking anxiety and speaking performance: a case study. *Expert Syst.* 2021;38(3):e12667.
10. Ellis R, editor. *Planning and task performance in a second language.* Amsterdam/Philadelphia: John Benjamins Publishing Company; 2005. p. 3-34.
11. Foster P, Skehan P. The influence of planning and task type on second language performance. *Stud Second Lang Acquis.* 1996;18:299-323.
12. Hwang GJ. Definition, framework, and research issues of smart learning environments—a context-aware ubiquitous learning perspective. *Smart Learn Environ.* 2014;1(1):4.
13. Hwang GJ, Xie H, Wah BW, Gašević D. Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Comput Educ Artif Intell.* 2020;1:100001.
14. Koka NA. An insight into the efficiency of artificial intelligence (AI)-chatbot as digital tutors for enhancing learners' motivation and performance in linguistics courses. *Pak J Life Soc Sci.* 2024;22(2):19812-30.
15. Kuddus K. Artificial intelligence in language learning: practices and prospects. *Adv Anal Deep Learn Models.* 2022:1-17.
16. Kukulska-Hulme A. Moving language teaching and learning from the known to the unknown. In: *Language learning with technology: perspectives from Asia.* Singapore: Springer Nature Singapore; 2021. p. 3-12.
17. Likert scale in social sciences research: problems and difficulties. *FWU J Soc Sci.* 2022:89-101. doi:10.51709/19951272/winter2022/7
18. Miao F, Holmes W. *AI and education: guidance for policy-makers.* Paris: UNESCO; 2021. Available from: <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
19. Moscardini AO, Strachan R, Vlasova T. The role of universities in modern society. *Stud High Educ.* 2022;47(4):812-30.
20. Ortega L. Planning and focus on form in L2 oral performance. *Stud Second Lang Acquis.* 1999;21(1):109-48.
21. Ortega L. What do learners plan? Learner-driven attention to form during pre-task planning. In: Ellis R, editor. *Planning and task performance in a second language.* Philadelphia: John Benjamins; 2005. p. 77-109.
22. O'Shea T, Self J. *Learning and teaching with computers: the artificial intelligence revolution.* Prentice Hall Professional Technical Reference; 1986.
23. Skehan P. *A cognitive approach to language learning.* Oxford: Oxford University Press; 1998.
24. Swain M. Communicative competence: some roles of comprehensible input and comprehensible output in its development. In: Gass S, Madden C, editors. *Input in second language acquisition.* Rowley: Newbury House; 1985. p. 235-52.
25. Umar I, Tech B. Artificial intelligence (AI) in education. *Educ Next.* 2024;1:1-7.
26. Wang YF, Petrina S. Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy. *Editorial Preface.* 2013;4:124-31. doi:10.14569/IJACSA.2013.041117
27. Wendel JN. *Planning and second-language narrative production [dissertation].* Philadelphia: Temple University; 1997.
28. Wigglesworth G. An investigation of planning time and proficiency level on oral test discourse. *Lang Test.* 1997;14(1):85-106.
29. Yuan F, Ellis R. The effects of pre-task planning and on-line planning on fluency, complexity, and accuracy in L2 monologic oral production. *Appl Linguist.* 2003;24:1-27.
30. Zou B, Liviero S, Hao M, Wei C. Artificial intelligence technology for EAP speaking skills: student perceptions of opportunities and challenges. In: *Technology and the psychology of second language learners and users.* 2020. p. 433-63.

How to Cite This Article

An ML. Enhancing speaking proficiency: a mixed-methods study on AI-assisted pre-task planning for non-English major students. *Int J Social Sci Educ Res.* 2025;4(6):41-49. doi:10.54660/IJSSER.2025.4.6.41-49.

Creative Commons (CC) License

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.