



SCAFFOLDING EFL SPEAKING PRACTICE: UNIVERSITY STUDENTS' ENGAGEMENT WITH GEMINI AI FEEDBACK

Arda¹, Anne Ratna Suminar², Yustika Nur Fajriah³

Program Studi Pendidikan Bahasa Inggris, Institut Pendidikan Indonesia Garut (IPI Garut),
Indonesia¹²³

e-mail: ardarakana@gmail.com; anners@institutpendidikan.ac.id;

yustikanurfajriah@institutpendidikan.ac.id

Diterima: 07/08/2026; Direvisi: 14/08/2026; Diterbitkan: 17/08/2026

ABSTRAK

Perkembangan kecerdasan buatan generatif dalam pembelajaran bahasa membuka peluang baru untuk menyediakan umpan balik lisan yang cepat dan personal, tetapi pemanfaatannya sebagai perancah digital dalam praktik komunikasi interaktif masih memerlukan kajian lebih lanjut. Penelitian ini bertujuan mengeksplorasi keterlibatan mahasiswa Program Studi Pendidikan Bahasa Inggris dengan umpan balik berbasis Kecerdasan Buatan (AI) generatif sebagai mekanisme perancah digital (*digital scaffolding*) selama praktik komunikasi lisan interaktif. Penelitian ini menggunakan studi kasus kualitatif yang melibatkan tiga mahasiswa EFL tingkat menengah di Institut Pendidikan Indonesia Garut, Jawa Barat. Data dikumpulkan melalui interaksi lisan dalam beberapa sesi dan wawancara retrospektif semi-terstruktur, kemudian dianalisis menggunakan analisis tematik enam tahap Braun dan Clarke. Penelitian ini menggunakan Teori Sosiokultural Vygotsky dan Hipotesis *Noticing* Schmidt sebagai landasan teoretis untuk memahami proses perancahan, kesadaran linguistik, dan perkembangan interaksi lisan mahasiswa. Hasil penelitian menunjukkan bahwa Gemini AI berfungsi sebagai perancah afektif dan kognitif dalam praktik berbicara. Secara afektif, interaksi dengan AI menciptakan ruang praktik dengan tingkat kecemasan yang lebih rendah dan minim tekanan sosial sehingga meningkatkan kesiapan mahasiswa untuk berkomunikasi. Secara kognitif, umpan balik yang cepat, jelas, dan objektif membantu mahasiswa mengenali kekuatan serta kesalahan penggunaan bahasa. Secara linguistik, umpan balik waktu nyata mendorong kesadaran terhadap ketidaktepatan *tense* dan penggunaan leksikal, sekaligus membantu mahasiswa memantau artikulasi dan mengatur kecepatan berbicara. Meskipun terdapat kendala berupa pemotongan giliran bicara akibat AI salah menginterpretasikan jeda berpikir sebagai akhir tuturan, mahasiswa menunjukkan agensi digital melalui penggunaan penanda percakapan strategis. Dengan demikian, Gemini AI dapat berfungsi sebagai ruang latihan bertekanan rendah yang menjembatani kesiapan mahasiswa menuju interaksi lisan di dunia nyata.

Kata Kunci: *digital scaffolding; EFL speaking practice; Gemini AI feedback; grammatical accuracy; speaking fluency*

ABSTRACT

The rapid development of generative artificial intelligence in language education offers new opportunities to provide immediate and personalized oral feedback; however, its use as digital scaffolding in interactive communication practice remains underexplored. This study aimed to explore how undergraduate English as a Foreign Language (EFL) students engage with generative AI-mediated feedback as a digital scaffolding mechanism during interactive oral communication practice. This qualitative case study involved three intermediate EFL university



students at Institut Pendidikan Indonesia Garut, West Java. Data were collected through multi-session oral interactions and semi-structured retrospective interviews and analyzed using Braun and Clarke's six-phase thematic analysis. Vygotsky's Sociocultural Theory and Schmidt's Noticing Hypothesis served as theoretical foundations for understanding scaffolding processes, linguistic awareness, and students' oral interaction development. The findings revealed that Gemini AI functioned as both an affective and cognitive scaffold during speaking practice. Affectively, interaction with AI created a low-anxiety practice environment with minimal social pressure, thereby enhancing students' willingness to communicate. Cognitively, immediate, clear, and objective feedback helped students recognize both linguistic strengths and errors. Linguistically, real-time feedback promoted explicit awareness of tense inconsistencies and lexical misuse while encouraging students to monitor articulation and regulate speaking rate. Despite a technological constraint involving premature turn-taking caused by the AI misinterpreting cognitive pauses as the end of an utterance, students demonstrated digital agency by employing strategic conversational fillers to maintain interactional continuity. Thus, Gemini AI can serve as a low-pressure practice environment that bridges students' readiness toward real-world oral interaction.

Keywords: *digital scaffolding; EFL speaking practice; Gemini AI feedback; grammatical accuracy; speaking fluency*

INTRODUCTION

Developing proficient oral communication skills remains one of the most demanding objectives for undergraduate students learning English as a Foreign Language (EFL), particularly in contexts where opportunities for authentic English interaction are limited (Amoah & Yeboah, 2021; Hosni, 2014; Normawati et al., 2023). EFL learners commonly encounter linguistic difficulties, including limited vocabulary, inaccurate pronunciation, and grammatical inconsistencies. These challenges are compounded by affective factors such as foreign language speaking anxiety, low self-efficacy, fear of negative evaluation, and embarrassment, which can reduce students' willingness to communicate (Gobena, 2025; Rahmon, 2024). In conventional classroom settings, peer observation and perceived teacher authority may intensify these concerns, leading some learners to avoid speaking because of the possibility of making public errors (Hanifa, 2018; Putri et al., 2020). The growing availability of generative Artificial Intelligence (AI) therefore creates an urgent need to investigate whether AI-mediated speaking practice can provide learners with accessible, immediate, and psychologically safer opportunities to receive feedback and develop oral communication skills. Understanding how students actually experience and respond to such feedback is particularly important because the pedagogical value of AI depends not only on its technical capabilities but also on how learners engage with and act upon the feedback provided.

Corrective feedback is widely recognized as an important component of second language learning because it helps learners identify linguistic problems, modify interlanguage forms, and improve grammatical accuracy and fluency (Chu, 2011; Khasanah & Fauzi, 2024; Liu & Feng, 2023). Nevertheless, conventional teacher-mediated oral feedback presents considerable challenges in higher education. Limited instructional time, large student populations, and increasing administrative responsibilities make it difficult for instructors to provide immediate, individualized, and continuous feedback during oral communication activities (Asiri & Khadawardi, 2024; Hyland, 2025). Delayed feedback may also reduce its pedagogical value because learners may have difficulty recalling the specific linguistic context



in which an error occurred (Ellis, 2009). These limitations highlight the need for complementary technology-mediated approaches that can provide learners with opportunities for immediate and repeated speaking practice beyond formal classroom instruction. Generative AI, particularly systems capable of multimodal and voice-based interaction, offers one potential solution because learners can engage in conversational practice and receive feedback without being constrained by classroom time or the immediate presence of peers and teachers.

The potential of AI-mediated feedback can be understood through Vygotsky's (1978) Sociocultural Theory. From this perspective, language development occurs through mediated social interaction, in which learners receive assistance that enables them to perform beyond their current independent level within the Zone of Proximal Development (ZPD). Scaffolding refers to temporary and adaptive assistance that supports learners while they gradually develop greater independence in performing a task (Thorne & Lantolf, 2006). In technology-mediated language learning, AI can function as a form of digital mediation by providing prompts, responses, corrections, and guidance during learning activities. Rather than positioning Gemini AI as a direct equivalent of a human More Knowledgeable Other (MKO), this study conceptualizes the system as a digital scaffolding tool through which learners obtain mediated assistance during speaking practice. Its value therefore lies in how its feedback supports learners' attempts to perform oral communication tasks and regulate their subsequent performance.

Schmidt's Noticing Hypothesis provides a complementary perspective for understanding the cognitive processes involved in responding to AI feedback. The hypothesis emphasizes the importance of conscious attention to linguistic features in the process of second language development (Swinnen et al., 1990). During spontaneous communication, learners must simultaneously formulate meaning, retrieve vocabulary, organize grammatical structures, and manage interaction, which may limit their ability to consciously recognize errors in their own production. Immediate feedback can potentially direct learners' attention toward discrepancies between their output and target language forms, thereby creating opportunities for noticing, reflection, and subsequent modification (Ellis et al., 2008; Yaseen et al., 2025). In this study, feedback concerning verb tenses, sentence structures, lexical choices, pronunciation, and speech delivery is therefore examined not simply as corrective information but as a potential trigger for learners' conscious attention to their own linguistic performance.

Recent research has increasingly examined the use of generative AI in foreign language education; however, important empirical gaps remain. Much of the existing literature focuses on automated writing evaluation, quantitative measures of learning outcomes, or general perceptions and acceptance of AI technologies (Hou et al., 2024; S, 2024; Zou et al., 2023). Such research provides valuable evidence regarding the potential of AI but offers less insight into how learners actually experience and respond to voice-based generative AI feedback during interactive speaking practice. In particular, qualitative evidence concerning the affective, cognitive, and behavioral dimensions of university EFL students' engagement with Google's Gemini AI remains limited. The issue is important because speaking practice involves not only linguistic accuracy but also confidence, willingness to communicate, self-monitoring, and the ability to maintain interaction. Moreover, AI-mediated oral interaction may introduce technological constraints, such as inaccurate speech recognition, inappropriate feedback, response delays, or premature turn-taking, which can influence how learners adapt their communication strategies. Within university English Education contexts in West Java, these dimensions have not yet been sufficiently examined.



Accordingly, this study does not seek to determine the effectiveness of Gemini AI through quantitative gain scores. Instead, it investigates how undergraduate EFL students experience, interpret, and respond to Gemini AI-generated feedback during interactive English speaking practice. Drawing on Sociocultural Theory and the Noticing Hypothesis, the study focuses on three interconnected dimensions of engagement: affective responses to AI-mediated practice, cognitive engagement with feedback and linguistic discrepancies, and behavioral responses reflected in learners' subsequent speaking strategies. Particular attention is also given to learners' responses to technological constraints during interaction. The study addresses two research questions: (1) How do university EFL students perceive the affective and cognitive scaffolding provided by Gemini AI-generated feedback during English speaking practice? and (2) How do students engage with and respond to Gemini AI feedback regarding grammatical accuracy, speaking fluency, and associated technological constraints? Through this qualitative inquiry, the study seeks to contribute empirical insight into the role of generative AI as a digital scaffolding mechanism for EFL oral communication and to clarify the conditions under which AI-mediated speaking practice can complement conventional language learning.

METHODS

This study employed a qualitative case study design (Yin, 2018). A case study approach provides a robust empirical framework for conducting an intensive, holistic, and in-depth examination of a contemporary educational phenomenon within its real-world context. Because student perceptions, emotional states, and cognitive evaluations of AI feedback are inherently nuanced, subjective, and context-dependent, a qualitative case study design enabled the researcher to capture detailed descriptions of participant experiences without manipulating the natural practice environment (Creswell, 2012; Lim, 2025).

The research was conducted within the English Education Program at Institut Pendidikan Indonesia (IPI) Garut. The study involved three undergraduate EFL students, designated as Respondent 1 (R1), Respondent 2 (R2), and Respondent 3 (R3). To ensure rich data collection, participants were selected using purposive sampling (Etikan et al., 2016; Palinkas et al., 2016; Tajik et al., 2024) based on four explicit criteria: (1) enrolled as active undergraduate students in the English Education Program, (2) completed foundational coursework in English speaking skills, (3) demonstrated intermediate-level oral English proficiency, and (4) had limited prior exposure to intensive, voice-based AI speaking practice tools.

Data were collected through semi-structured retrospective telephone interviews, supplemented by secondary artifact analysis of participants' Gemini conversation logs and feedback records. The semi-structured interview protocol was framed around the Technology Acceptance Model (TAM) constructs (Davis & Davis, 1989), focusing on Perceived Ease of Use, Perceived Usefulness, and Behavioral Intention, and adapted from validated CALL instruments (Azzahra et al., 2024; Mudawy, 2025; Nasution et al., 2025; Saraswati et al., 2024; Ulfa, 2025; Yasa, 2024).

The research was executed across seven sequential steps: (1) orientation and ethics briefing where participants signed informed consent, (2) application configuration on personal smartphones using the free tier of Google Gemini, (3) role-setting prompting to establish AI persona as a supportive partner focusing on future aspirations, (4) spoken practice sessions lasting 5 to 10 minutes per session across four consecutive practice days, (5) feedback generation prompting focusing on fluency and grammatical accuracy, (6) artifact archiving of



Gemini feedback transcripts and dialogue histories, and (7) retrospective audio-recorded telephone interviews transcribed verbatim.

Interview transcriptions were analyzed using Braun and Clarke's (2006) six-phase thematic analysis framework: (1) familiarization through repeated reading of verbatim transcripts, (2) generating initial codes, (3) searching for themes aligned with theoretical constructs, (4) reviewing themes against primary datasets, (5) defining and naming themes, and (6) producing the final report supported by verbatim extracts.

RESULTS AND DISCUSSION

Results

Affective Scaffolding: Low-Anxiety Practice and Emotional Safety

Analysis of the participant interviews indicated that interaction with Gemini AI created a comfortable and low-anxiety environment that reduced psychological barriers to speaking English. Before using the AI tool, all three participants reported experiencing fear of making mistakes, embarrassment, and concern about peer evaluation during conventional classroom speaking activities. The absence of direct human judgment appeared to create a psychologically safer practice environment, allowing participants to speak more freely and gradually develop confidence. R1 explained that the supportive prompt configuration and constructive responses from Gemini AI encouraged him to speak more:

“I felt comfortable... according to the prompt given to instruct Gemini AI to provide supportive and constructive responses, and yes, Gemini AI indeed gave responsive and constructive feedback. So during my speaking process, response by response from the AI built up my confidence to speak more... speaking anxiety decreased because Gemini is not a human; its responses are always positive.” (R1)

R2 similarly reported that interaction with AI reduced feelings of shame when making speaking errors:

“When I speak with the AI, it feels more relaxed... there is no anxiety, and if I make mistakes, there is no feeling of shame. Basically, when speaking with AI, I feel more relaxed in practicing... I am someone who is afraid of making mistakes and feeling embarrassed, but with the AI, that fear and shame are reduced.” (R2)

R3 described a gradual transition from nervousness during the initial sessions to greater confidence in using English beyond the practice sessions:

“At first, I was nervous and awkward, thinking deeply about how to answer because I had never spoken English fully like this before... But I felt more relaxed with the AI... after using the AI, I became a little bit more courageous, step-by-step, to speak English, even at home with my family.” (R3)

These accounts indicate that the participants experienced a gradual reduction in affective barriers and an increase in confidence throughout the speaking practice sessions.

Feedback Quality: Structural Clarity, Directness, and Balanced Evaluation

Participants consistently characterized Gemini AI feedback as clear, direct, and easy to understand. The feedback was perceived as specific to participants' spoken responses rather than as generic evaluation. Participants also valued the combination of corrective information and positive reinforcement. R1 described the feedback as immediately actionable:

“Overall, it was easy to understand and clear. The AI explained where my mistakes were, such as verb tenses or choosing a more accurate word, clearly and with encouraging



language... It pointed directly to the points, so I immediately knew what should be used.” (R1)

R2 emphasized that the feedback did not focus exclusively on errors but also identified strengths:

“Every feedback given was clear because it went straight to the core... Moreover, at the end, there was overall feedback that showed my strengths. That was what made me excited, because it was not just about my weaknesses, but I was also told what my strengths were and given tips for the future.” (R2)

Similarly, R3 emphasized the contextual specificity of the feedback:

“It was easy to understand because when we finished speaking, it immediately told us, ‘You made a mistake here, since you are telling a past story, you should use past tense.’ It was specific... each answer had different mistakes, so it was not a generic template.” (R3)

The three accounts consistently show that participants perceived the feedback as comprehensible, specific, immediate, and balanced between corrective comments and positive reinforcement.

Perceived Affordances Relative to Human Teacher Feedback

Participants identified several differences between Gemini AI feedback and feedback from human teachers, particularly in terms of response speed, perceived objectivity, and privacy. Although participants acknowledged that teachers could provide more extensive explanations, they considered AI particularly useful for repeated speaking practice and immediate feedback. R1 emphasized the speed and directness of the AI response:

“The main difference is speed; Gemini's response is instant and goes straight to the point. For theory, I feel Gemini is easier to understand because the language is direct and detailed... Gemini is clearly objective because it is a machine, so it is nicer to listen to. With real teachers, human errors or subjectivity can sometimes happen.” (R1)

R2 highlighted the perceived absence of interpersonal judgment and greater privacy when communicating with AI:

“The AI evaluates your ability strictly at that moment. A teacher might look at how you were in the past or how you behave outside the class... Also, in terms of topic privacy, I don't feel embarrassed with the AI. If there is an embarrassing or secret story, I can be completely open with the AI, whereas with a teacher, I would be shy or secretive.” (R2)

R3 viewed AI and teachers as serving complementary functions:

“For practice, it is better with AI because you feel more confident. But if we want to increase deep knowledge, it is better with a teacher... The AI gives instant feedback after we speak, whereas with a teacher, we have to wait, but the teacher can be more detailed if we ask.” (R3)

Thus, participants did not perceive Gemini AI as a replacement for teachers. Rather, they identified its primary advantages in providing accessible, immediate, and relatively private opportunities for independent speaking practice.

Impact on Grammatical Accuracy, Fluency Regulation, and Technical Coping

The multi-session speaking activities increased participants' awareness of linguistic errors that they had not consciously recognized during spontaneous speech. Participants reported that their attention during speaking was initially focused primarily on conveying meaning, while grammatical and lexical inaccuracies often went unnoticed. Feedback from Gemini AI subsequently directed their attention to specific linguistic forms.



R1 primarily identified difficulties with verb tense:

“The feedback really made me aware... Most of my errors were in using tenses, where I should have used past tense verbs, but I used present tense instead.” (R1)

R2 reported that the feedback helped identify pronunciation and lexical problems:

“For example, when I wanted to say ‘snack’, I accidentally mispronounced or used ‘snake’. The AI immediately pointed it out.” (R2)

R3 reported receiving alternative expressions that helped him recognize ways to make his speech more natural:

“It suggested things like, ‘If you want to sound more natural, it is better to say it like this.’ It directly guided me on how to fix it.” (R3)

The participants also reported changes in their regulation of speaking fluency. Feedback concerning speech clarity and pace encouraged them to modify how they delivered their utterances. R1 explained:

“There were parts where Gemini misheard me, and I realized it was because I spoke too fast... I then tried to speak more patiently and clearly.” (R1)

R2 similarly stated:

“The AI explicitly advised me, ‘Please be more relaxed, be more calm.’ That made me realize how to control my speed.” (R2)

Despite these perceived benefits, the participants encountered a technical limitation involving Gemini AI’s turn-taking mechanism. The system sometimes interpreted a cognitive pause as the end of a participant’s utterance and responded before the participant had completed the intended message. R3 described the problem:

“The AI jumped in and answered before I was finished... It felt a bit rushed.” (R3)

Rather than terminating the interaction, participants adapted their communication behavior to accommodate this limitation. R1 reported using explicit conversational markers to signal that he had not finished speaking:

“If I wanted to pause, I could just say ‘hold on a minute’ or ‘hold on a sec’, and the AI would reply ‘oh sure, take your time’ and wait for me to think.” (R1)

This adaptation indicates that participants actively modified their interactional strategies in response to the technical characteristics of the AI system. Overall, the findings demonstrate that participants experienced benefits in affective comfort, feedback comprehension, grammatical awareness, and fluency regulation, while also developing strategies to maintain interaction when technological interruptions occurred.

Discussion

The findings demonstrate that Gemini AI can function as a multidimensional digital scaffold in EFL speaking practice, particularly by supporting learners’ affective readiness, awareness of linguistic forms, and self-regulation of oral production. This finding is consistent with the growing evidence that AI-supported language learning can facilitate learner engagement through personalized and interactive feedback (Wang et al., 2023; Yaseen et al., 2025). Rather than functioning solely as an automated correction tool, Gemini AI in this study became an interactive practice partner that enabled participants to repeatedly produce language, receive immediate responses, identify problems, and modify subsequent performance. This finding is particularly relevant to the persistent difficulties experienced by EFL learners in oral communication, including limited linguistic resources and difficulties in producing English spontaneously (Amoah & Yeboah, 2021).



The first major finding concerns the affective scaffolding provided by Gemini AI. Participants consistently perceived interaction with the AI as less threatening than conventional classroom speaking activities because the interaction reduced fear of peer judgment, embarrassment, and negative evaluation. This finding is important because psychological barriers can substantially influence learners' willingness to participate in English-speaking activities. Amoah and Yeboah (2021) identified speaking difficulties and motivational factors as important dimensions of EFL learners' oral communication experiences. In the present study, the absence of direct interpersonal judgment appeared to create a psychologically safer environment in which participants were willing to experiment with English and tolerate mistakes. This finding is consistent with Wang et al. (2023), who demonstrated that the quality of human-AI interaction is an important factor in AI-supported language learning. Thus, the affective value of Gemini AI appears to derive not simply from the presence of technology but from the particular interactional conditions created by the technology.

The participants' perception that AI interaction reduced embarrassment also has implications for the development of self-efficacy and cognitive engagement. Liang et al. (2023) found that students' interaction with generative AI was associated with learning achievement through the mediating roles of self-efficacy and cognitive engagement. Although the present study did not quantitatively measure these variables, participants' accounts indicate a similar process: reduced anxiety encouraged them to speak more freely, repeated interaction increased their confidence, and greater confidence encouraged continued practice. The qualitative evidence therefore helps explain how interaction with generative AI may support learning beyond the immediate provision of linguistic feedback.

A second important finding concerns the clarity, immediacy, and individualized nature of AI feedback. Participants regarded Gemini AI's feedback as direct and easy to understand because the feedback was connected to their specific spoken production. This finding complements Liu and Feng (2023), who emphasized that learner engagement with corrective feedback is an important component of the effectiveness of feedback in EFL/ESL contexts. Feedback becomes pedagogically meaningful when learners attend to it and use it to modify their subsequent performance. In the present study, participants did not merely receive corrections; they reported becoming aware of particular problems, such as incorrect verb tense, inappropriate lexical choices, and pronunciation errors, and subsequently attempting to modify their speech.

The immediacy of AI feedback is another important contribution of the findings. Ellis (2009) explains that corrective feedback plays an important role in language development, but the effectiveness of feedback is closely related to how learners engage with the information provided. In conventional classroom contexts, teachers may face difficulties providing individualized feedback continuously to every learner. Nazaretsky et al. (2024) similarly demonstrate that students perceive AI-generated and human feedback through different affordances and limitations. The present findings indicate that participants particularly valued AI because it provided immediate responses without requiring them to wait for individual teacher attention. However, participants simultaneously recognized that teachers could provide deeper explanations. This distinction suggests that AI feedback should be understood as complementary to teacher feedback rather than as a complete substitute for it.

The findings regarding grammatical accuracy provide further evidence of the pedagogical potential of AI-mediated corrective feedback. Participants became particularly aware of tense inconsistencies, lexical errors, and inappropriate expressions after receiving



immediate feedback. This observation is consistent with Chu (2011), who demonstrated the relevance of corrective feedback to the development of oral accuracy among English learners. It is also consistent with Liu and Feng (2023), who emphasize the importance of learner engagement in processing corrective feedback. In the present study, Gemini AI made errors that might otherwise have disappeared from learners' attention more salient by immediately identifying them after the speaking activity. The feedback therefore functioned as a trigger for reflection on the relationship between learners' actual production and more appropriate English forms.

This finding is also supported by recent research on AI-generated feedback. Hou et al. (2024) found that learners perceived AI-generated corrective feedback as useful, although its usefulness was related to learner proficiency and attitudes toward the technology. Similarly, S. (2024) reported the potential of Google Gemini-generated feedback to support English writing proficiency. Although that research focused on writing rather than speaking, the present findings extend the discussion to oral communication by showing that generative AI can also provide immediate feedback on transient spoken production. Furthermore, Zou et al. (2023) specifically demonstrated the potential of AI speech-evaluation programs with automatic feedback for developing EFL learners' speaking skills. The present study complements this quantitative-oriented evidence by showing how learners themselves experience and interpret such feedback during repeated interactive speaking practice.

The findings concerning speaking fluency are particularly noteworthy because participants did not equate fluency with speaking rapidly. Instead, interaction with Gemini AI encouraged them to regulate their speaking rate and articulation after recognizing that speech delivered too quickly or unclearly could be misrecognized by the system. This interpretation is consistent with Burns (2019), who conceptualizes speaking proficiency in terms of meaningful, comprehensible, and appropriately delivered oral communication rather than speed alone. Newton (2024) similarly emphasizes the importance of quality and comprehensibility in language production. In the present study, the AI system consequently functioned as an external feedback mechanism that made participants more conscious of how their delivery affected communication. The technological system therefore indirectly encouraged learners to monitor and regulate their oral production.

At the same time, the study identified a significant technological constraint, namely Gemini AI's premature turn-taking when participants paused to formulate their responses. This problem indicates that AI-mediated speaking interaction does not yet fully reproduce the flexible timing of human conversation. Asad et al. (2021) emphasize that effective digital learning requires appropriate techno-pedagogical integration rather than technological adoption alone. From this perspective, the interruption problem observed in the present study demonstrates that the pedagogical usefulness of AI remains partly dependent on the system's capacity to interpret human communicative behavior accurately.

However, participants responded to this limitation by developing strategic coping mechanisms, such as using conversational placeholders to indicate that they had not yet finished speaking. This finding is important because it demonstrates that learners were not passive recipients of technological support. Instead, they developed strategies to adapt their communicative behavior to the characteristics of the AI system. Such behavior can be interpreted as an expression of digital agency, in which learners actively negotiate the affordances and limitations of digital learning environments. Tondeur et al. (2025) highlight the importance of preparing learners and teachers to use digital technologies effectively, while



Yaseen et al. (2025) identify digital literacy as an important factor in determining the impact of adaptive and interactive AI tools on student engagement. The present findings reinforce this perspective by demonstrating that effective AI-supported language learning requires learners to understand not only the language itself but also how to interact strategically with the technology.

The findings also reveal that participants perceived AI and human feedback as complementary rather than interchangeable. Participants appreciated Gemini AI for its speed, privacy, accessibility, and consistent responses, whereas teachers were perceived as more capable of providing deeper explanations and contextual pedagogical guidance. This finding is consistent with Nazaretsky et al. (2024), whose study highlights differences in how students perceive AI-generated and human feedback. The present study therefore challenges an overly simplistic assumption that generative AI should replace teacher feedback. Instead, AI appears to be particularly valuable for repetitive individual practice and immediate corrective feedback, while teachers remain important for interpretation, deeper explanation, social interaction, and pedagogical decision-making.

Finally, participants' intention to continue using Gemini AI for independent speaking practice indicates the potential of generative AI to support learner autonomy and sustained engagement. This finding corresponds with Liang et al. (2023), who identified the relationship between interaction with generative AI, self-efficacy, cognitive engagement, and learning achievement. It is also consistent with Yaseen et al. (2025), who found that personalized feedback and interactive AI tools can contribute to student engagement, particularly when learners possess sufficient digital literacy. Tajik (2025) further highlights the potential of real-time adaptive AI feedback for supporting English speaking development. Nevertheless, the present findings suggest that autonomous AI practice should be positioned as an additional learning opportunity rather than a replacement for authentic human communication.

Overall, the findings answer the first research question by demonstrating that Gemini AI provides affective and cognitive scaffolding through a low-anxiety interactional environment, immediate encouragement, individualized explanations, and accessible corrective feedback. The findings answer the second research question by showing that learners actively used AI feedback to increase awareness of grammatical and lexical problems, regulate speaking rate and articulation, and develop strategies to manage technological interruptions. The contribution of this study therefore lies in demonstrating that the value of generative AI in EFL speaking practice is not limited to automated correction. Its potential emerges from the interaction between immediate feedback, psychological safety, learner engagement, and learners' ability to strategically manage the technology. At the same time, the technological limitations identified in the study reinforce the need to position Gemini AI as a complementary digital scaffold, with human teachers continuing to provide deeper pedagogical guidance and authentic interpersonal communication opportunities.

CONCLUSION

This qualitative case study concludes that Gemini AI can function as an effective digital scaffold for university EFL students during English speaking practice by supporting affective, cognitive, and linguistic dimensions of learning. The AI-mediated environment reduced participants' speaking anxiety and fear of social embarrassment, thereby creating a more comfortable space for repeated oral practice. Participants also perceived the feedback as immediate, clear, individualized, and balanced between correction and positive reinforcement,



which helped them notice grammatical and lexical errors and regulate their speaking rate and articulation. Despite occasional premature turn-taking caused by the AI's interpretation of thinking pauses, participants demonstrated digital agency by using conversational placeholders to maintain interaction. These findings indicate that Gemini AI is best positioned as a complementary, low-pressure preparatory tool that can strengthen learners' confidence, linguistic awareness, and fluency before they engage in authentic human communication.

The study is limited by its qualitative case-study design involving only three intermediate EFL university students and by the use of the available Gemini voice-interaction system, which occasionally disrupted conversational continuity. Therefore, the findings should not be generalized to broader EFL populations without further investigation. Future research should involve larger and more diverse participant groups and employ longitudinal, quantitative, or mixed-methods designs to examine changes in speaking proficiency over time. Comparative research involving different AI models or enhanced voice-interaction systems is also recommended to determine whether improvements in processing, contextual retention, and turn-taking can reduce the technological constraints identified in this study. For EFL practitioners, Gemini AI may be integrated as a supplementary out-of-class speaking practice tool, while teachers should continue to provide deeper linguistic guidance and authentic interpersonal communication opportunities.

REFERENCE

- Amoah, S., & Yeboah, J. (2021). The speaking difficulties of Chinese EFL learners and their motivation towards speaking the English language. *Journal of Language and Linguistic Studies*, 17(1), 56–69. <https://doi.org/10.52462/jlls.4>
- Asad, M. M., Aftab, K., Sherwani, F., Churi, P., Moreno-Guerrero, A. J., & Pourshahian, B. (2021). Techno-pedagogical skills for 21st century digital classrooms: An extensive literature review. *Education Research International*, 2021, 1–12. <https://doi.org/10.1155/2021/8160084>
- Burns, A. (2019). Concepts for teaching speaking in the English language classroom. *LEARN Journal: Language Education and Acquisition Research Network*, 12(1), 1–11. <https://doi.org/10.3126/kvmrj.v6i1.84530>
- Chu, R. (2011). Effects of teacher's corrective feedback on accuracy in the oral English of English-majors college students. *Theory and Practice in Language Studies*, 1(5), 454–459. <https://doi.org/10.4304/tpls.1.5.454-459>
- Davis, F. D., & Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, 1(1), 3–18. <https://doi.org/10.5070/12.v1i1.9054>
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1–4. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Hou, X. L., He, S. Y., & Cui, G. R. X. (2024). Learner use of AI-generated feedback for written corrective feedback in L2 writing: Usefulness, user proficiency and attitude. *ACM International Conference Proceeding Series*, 70–76. <https://doi.org/10.1145/3678726.3678767>
- Liang, J., Wang, L., Luo, J., Yan, Y., & Fan, C. (2023). The relationship between student



- interaction with generative artificial intelligence and learning achievement: Serial mediating roles of self-efficacy and cognitive engagement. *Frontiers in Psychology*, 14, 1285392. <https://doi.org/10.3389/fpsyg.2023.1285392>
- Lim, W. M. (2025). What is qualitative research? An overview and guidelines. *Australasian Marketing Journal*, 33(1), 12–28. <https://doi.org/10.1177/14413582241264619>
- Liu, H., & Feng, M. (2023). The role of learner engagement with corrective feedback in EFL/ESL classrooms. *Frontiers in Psychology*, 14, 1118467. <https://doi.org/10.3389/fpsyg.2023.1118467>
- Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2024). AI or human? Evaluating student feedback perceptions in higher education. *Lecture Notes in Computer Science*, 15159, 284–298. https://doi.org/10.1007/978-3-031-72315-5_20
- Newton, J. M. (2024). Stepping up the quality of English language teaching: How can Nation's four-strands principle help? *VNU Journal of Foreign Studies*, 40(3), 161–171. <https://doi.org/10.63023/2525-2445/jfs.ulis.5201>
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2016). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533–544. <https://doi.org/10.1007/s10488-013-0528-y>
- S, K. (2024). Enhancing writing proficiency through AI-powered feedback: A quasi-experimental study using Google Gemini. *LinguaEducare: Journal of English and Linguistic Studies*, 1(2), 83–96. <https://doi.org/10.63324/h6q1ak58>
- 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. <https://doi.org/10.20944/preprints202501.2205.v1>
- Tondeur, J., Trevisan, O., Howard, S. K., & van Braak, J. (2025). Preparing preservice teachers to teach with digital technologies: An update of effective SQD-strategies. *Computers & Education*, 232, 105262. <https://doi.org/10.1016/j.compedu.2025.105262>
- Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, 194, 104703. <https://doi.org/10.1016/j.compedu.2022.104703>
- Yaseen, H., Mohammad, A. S., Ashal, N., Abusaimh, H., Ali, A., & Sharabati, A. A. A. (2025). The impact of adaptive learning technologies, personalized feedback, and interactive AI tools on student engagement: The moderating role of digital literacy. *Sustainability*, 17(3), 1133. <https://doi.org/10.3390/su17031133>
- Zou, B., Du, Y., Wang, Z., & Chen, J. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open*, 13(3), 1–14. <https://doi.org/10.1177/21582440231193818>