

The Use of Gemini AI in Teaching Speaking

"Gemini AI as a Speaking Coach: Transforming EFL Speaking Practice"

Lukman Arif Rachman
Universitas Negeri Semarang
lukmanarifrachman@stutdent.unnes.ac.id

Abstract

This study investigates the effectiveness of Gemini AI in teaching speaking skills to second-grade senior high school students in Bandung. The study aims to reveal the use of Gemini AI in teaching speaking abilities. The participants consisted of 24 students and one English teacher from one of the private high schools in Bandung; most of the students possessed an intermediate level of speaking proficiency. The study used Classroom action research for three three-cycle implementations. Each cycle consisted of planning, implementation, and evaluation, aiming to get betterment in implementing Gemini AI in teaching speaking. The instruments used in the study were an observation sheet and a researcher's journal in order to know the steps and the techniques used in implementing Gemini AI for teaching speaking in every cycle. Findings from this research demonstrate that Gemini AI can be used as a tool in speaking lessons, providing interactive, real-time feedback and practice opportunities. Even though the implementation faced some obstacles, there were always some improvements in every cycle, which led the teacher and students to make Gemini AI as an effective digital tool for enhancing students' speaking skills. The significance of the study indicates that integrating AI in teaching speaking can provide individualized and interesting language practice. The use of artificial intelligence in teaching speaking also can enhance conventional teaching strategies. This study contributes to the growing body of literature on AI in education and provides practical implications for English language teaching, especially in enhancing students' communicative competence.

Keywords: Gemini AI, Classroom Action Research, Speaking, EFL, Teaching

1. Introduction

Speaking, the means of expression and communication by oral is an integral part for students to succeed in life, it is definitely fundamental. It is a dialogical oral process creating meaningful signs and making meaning in the sign system production, reception, process of information. Good speaking skills accompany almost all the long term social, emotional and academic achievements. In English teaching, especially speaking, has been a non-stop focus of The Ministry of National Education in Indonesia like curricular revision to fulfill global demand (in creating opportunities to learn). Unfortunately, most of English teachers have not always delivered what was needed by students. The students in Indonesia run into all kinds of problems in learning English speaking such as inappropriate vocabulary, insufficient mastery of grammar, wrong pronunciation and lack of confidence (Jon et al., 2021). This is a big obstacle; mostly there are not enough practice environments, particularly in terms of partners for a real-life conversation that makes fluency and communication skills more practical.

Artificial Intelligence (AI) such as Gemini AI is an innovative way to solve some of these teaching and learning challenges when it comes to teaching speaking skills (McGehee, 2024). By imitating real-life conversations, Gemini AI can offer personalized speaking practice creating personalized feedback in the

moment like pronunciation and grammar feedback that real-time data feedback is adaptive and adaptive amount of consultation (feedback) on top end all for one-users. As Marpaung (2024) said, Gemini AI is at least zero pressure learning space for students, thus will lead to less anxiety, and the other thing that is increased is the participation for more objectiveness, ultimately improving fluency; vocabulary confidence. In addition, the role a Gemini AI can play as patient, always-on assistant that is not time-constrained guides with examples and corrects step by step, quizzes students, creates additional learning opportunities enriched by traditional classroom time through a safe space for students to question without fear of judgement. Research related to AI in English Language Teaching (ELT) has examined tools, such as "Little Language Lessons" developed by Gemini for personalized vocabulary and slang practice through interactive games in Severe Educational Blinding studies. Another research documented lesson plans with the use of Gemini to deliver text-based explanations, develop exercises and supportive answers in real-time, in particular beneficial in low-resource environments.

Due to the integration of AI in teaching and learning methodologies in this era of education, The study on the use of Gemini AI on speaking within today's educational scene must be conducted. Alhusaini (2024) says rapid development of AI technologies has redefined the potential for personalized and tailored language learning experiences. To be clear, more empirical studies on the use of various AI strategies when employing tools like Gemini in distinct educational settings are needed here. This study is to find out the actual implementation of Gemini AI for teaching speaking for the students in second grade of senior high school. The implementation was done in three cycles so as to evaluate the effect of introducing Gemini AI into teaching speaking.

2. Literature Review

2.1. The Foundational Role of Speaking in Language Acquisition and the Emergence of AI in ELT

It forms the basis of speaking, which is at the center of language proficiency since communication takes place (Soans, 2022). Having very often been considered the most crucial language skill, communicative function between social/professional interactions sets in directly. The capacity for effective verbal communication to think, feeling and idea is core to creating relationships and promoting cultural literacy. But as in language learning Rao (2019) said, 'speaking is generally the second skill you learn after listening and is still considered very important for language acquisition. Another way that speaks to the importance of speaking is that native speakers of a language are often said to be 'speaking' that language.

The progression of teaching speaking approaches is quite impressive. Drills and rote memorization are more a part of the traditional methods. Abdullaev (2022) states that present-day language teaching is more communicative oriented targeting face-to-face work and communication based on real-life situations. The emphasis moved from an intransigent accuracy, to meaningful use and competence putting language to work well and appropriately in a wide range of contexts. This new age view acknowledges the need of learners of language to engage in authentic communication for effective learning.

Teachers are key agents in supporting an effective learning process for students, it is a place teachers aim for students to speak. Good teaching is about systematically supporting students to engage in integrated and structured activities that build knowledge of the specific knowledge, skills and strategies required for various types of interaction (Assel, 2014). This is about giving those opportunities and explicitly teaching the components of speaking (pronunciation, fluency and language in context) including when and where to

speak. Rezeki and Dalimunte (2024) concludes that the aim of teaching speaking is to promote efficient communication which allows learners to make themselves understood using their current language abilities.

There could be many beneficial sides of using AI for teaching speaking. Wu (2024) states that AI is used at personal learners based on needs and level of each student by adapting from previous releases. Then, the ability to give instantaneous feedback is really essential for learning and self-correcting. The fluidity many AI tools embody fosters a responsive classroom environment that is more engaging for students. Additionally, AI provides a great medium for practicing speaking (especially if in the form of automated performance recording) outside the land of bricks; without the time and again pressures of a classroom. For those students who do not want to speak in front of their peers, AI creates a space where they can gain confidence virtually.

In the last few years, Artificial Intelligence (AI) has advanced really fast and is also getting widespread, ranging from training to the specific case of English Language Teaching. The relentless advancement of AI, particularly in the space of generative models has garnered a good deal of interest as this technology has the potential to forever change language acquisition pedagogy with language learning. The emergence of more sophisticated AI models and tools presents new avenues for solving the existing challenges in language education as well as for delivering more personalized, adaptive learning. The implications of these developments being to improve the learning experience and revolutionize how languages are taught and learned.

2.2. Gemini AI: Capabilities and Potential Applications in English Language Teaching

Gemini AI is a leading large language model (LLM) from Google and its family of multimodal AI models (Imran and Almusharraf, 2024). The models are designed to take and comprehend different type of data like audio, images, software code, text and video. Because of that Saeidnia (2024) informs that Gemini uses a transformer model architecture where it is an artificial intelligence network framework invented by Google and it is fed with large quantities of multilingual or multimodal data. While previous models focus on text and/or image-text pairs, Gemini hosts interleaved data types in an amortizing manner so that later model processing can take the far more powerful and context-aware approaches.

The Gemini family includes several versions tailored for different devices and tasks, such as Nano, Pro, Ultra, and Flash. The smallest and most efficient is the Gemini nano for off-device on mobile devices for on-device tasks. For a wide range of applications, Gemini Pro is the best performing and scalable platform. Working closely with Intel; the large-sized (Ultra) version intended for heavy tasks that need advanced reasoning. Gemini Flash, which is a speed and efficiency focused lightweight version more recently also came out. The wide spectrum of models also highlights the versatility and the ability to fit into longer tail use cases / platforms, which is so great about Gemini.

One of the main features of Gemini AI is its generative ability. Imran and Almusharraf (2024) argue that “by recognizing and resampling the patterns learnt from its big training data, Gemini can synthesize new content such as human-like text, images or even pieces of code. This ability to generate text in context, as well as working with multimodal recognition places Gemini as an intriguing tool for developing experienced, personalized learning experiences in educational settings.

3. Methodology

Following in line with solving specific educational context problems and promoting a culture of improvement which is typical at the time slot, this study used Classroom Action Research (CAR) as its methodological foundation. Linear Planning-Evaluation in response to investigate the dynamic integration of new technology (Gemini AI) operating in speaking instruction based on CAR cyclical process of planning, implementation & evaluation would help prepare responsive pedagogical interventions (Burns, 2019). This study recruited one English teacher and 24 students in a senior high school in Bandung, purposively selected because of their accessibility, meaning that students were intermediate in speaking level reflecting a relevant context for examining the effect of AI on enhancing conversational depth.

Data were collected systematically in three cycles, each cycle replicating the phases of CAR planning, implementation and evaluation. Teacher and researcher planned speaking activities with integrated aids of Gemini AI in the planning stage where the learning outcomes and rubrics for assessment were identified. The implementation phase was the use of Gemini AI in direction of speaking lessons, student-centered communication, coach-guided feedback and also giving personal attention from AI. It also offered us some ideas for the next cycle to optimally adapt and improve students' speaking performance. As such, the repetitive nature helped in tuning Gemini AI as per requirement where it could be used to enhance speaking ability.

Data collection was the main purpose of using an observation sheet and researcher journal. By using the observation sheet students' engagement and interactions with Gemini AI, as well language development visible in their speaking across classes was more systematically recorded by standard speaking assessment criteria (fluency, pronunciation, grammar, vocabulary and interactional proficiency) in place. The researcher's journal was a detailed and insightful account of the teaching-learning process with Gemini AI; it also acted as a reflective tool wherein qualitative data (teacher comments, novel problems, unexpected insights and contextual data) were collected. Both empowers tools to ensure full understanding of the implementation and its influence on students' speaking learning.

4. Result and Discussion

4.1. Results: Implementation and Outcomes of Gemini AI in Speaking Instruction

4.1.1. Overview of the Action Research Cycles

This research investigated the integration of Gemini AI into teaching speaking through a classroom action research methodology. The classroom action research is characterized by that it systemizes, reflectively and iteratively approaches that aim to improve practice with a more refined strategy on the basis findings found immediate use. This research followed a cyclical process as observations and evaluation from phase one informed the planning and implementation of phase two directly. The main purpose was to explore how Gemini AI can be used strategically for students' speaking skills, especially for their engagement and conversational eloquence. This exemplified a pedagogic refinement over time as the study proceeded through three cycles of notional reflective practice, where insights from each cycle informed subsequent ones sequentially. This iterative nature of the method was a big strength that allowed us to tweak

pedagogical tools on the go based on empirical findings. In the specific use of a technology like artificial intelligence when integrated with dynamic and complex human systems (for instance, language learning) such adaptability is particularly important.

Table 1: Summary of Action Research Cycles and Outcomes

Cycle Number	Key Intervention/ Strategy	Planning (Brief)	Implementation (Brief)	Evaluation/Observed Outcomes	Rationale for Next Cycle
1	Individual AI Interaction	Target 5 min individual talk on self/post-graduation.	Students talked individually with Gemini AI.	Significant "silent mode"; no student reached 5 mins, many <2 mins.	Need for more active engagement and sustained conversation.
2	Pair Work with AI	Shift to pair work for longer, more active conversations.	Students talked in pairs with Gemini AI.	Longer conversations achieved; still "difficult to build" dialogue.	Need for conversational support and initiation strategies.
3	Scaffolded Pair Work with AI	Provide guiding questions to trigger longer, easier talks.	Students used teacher-provided questions in pairs with AI.	Most groups achieved >5 mins; students found it "easier to talk."	Achieved desired outcomes, indicating effective integration.

Cycle 1: Individual AI-Mediated Speaking Practice

Planning: The first phase in the study, Cycle 1 was to investigate the baseline efficacy of Gemini AI as a tool for solo speaking practice. The pedagogical plan was for students to talk one on one with an AI. In order to make the contents pertinent and available, chosen topics were kept simple for conversation, i.e., self and talking to self after leaving school. The teacher and researcher were very specific to the students. Students have to keep Gemini AI up and running, for at least 5 minutes. Primary data-capture related work in this cycle was information excavation from the transcriptions turned out by Gemini AI of conversations.

Implementation: In Cycle 1, implementation phase when students interacted individually with the bots of Gemini AI on their given topics. The role of a teacher was more direct and as an observer while guiding from the back with limited intervention during the AI's actual conversation.

Evaluation: Cycle evaluation found real problems with implementation in Cycle 1. The biggest two, first being a ton of "silent mode" from the kids indicating long sections of not engaging, hesitation or "nothing to say"; and second, none of the students could hold a 5-minute AI Chat. In reality we saw many students ending their conversations in less than 2 mins.

These observations suggest that while AI can really help create a 'no-stress' chatting environment, kind of taking the edge off social anxiety when talking to a machine. Moreover, it doesn't automatically fix all the other big challenges that come with keeping a conversation going. The short chats and moments of silence

show that students might be feeling overwhelmed or unsure of what to say next, or maybe just not confident enough when students are dealing with AI on their own. Students might find it hard to come up with things to say, ask follow-up questions, or handle the AI's responses in an open-ended chat. This points to a key difference such as even though AI might help reduce feelings of anxiety, it doesn't automatically give students the skills or strategies they need to have smooth, sustained conversations. That initial difficulty shows that, despite AI being a pretty powerful tool, it doesn't eliminate all the obstacles to produce language. Instead, it is necessary to incorporate specific teaching strategies to really get the most out of it, especially for learners who aren't yet confident or fluent speakers.

Cycle 2: Pair Work with AI-Mediated Speaking Practice

Planning: After noticing some areas to improve from Cycle 1, next planning was decided to switch things up a bit for Cycle 2. Instead of students working alone, they were asked to team up in pairs when talking with Gemini AI. The main idea was to encourage longer, more lively conversations by using the advantages of working together. To keep things fair and see how this change made a difference, the teacher and researcher kept the discussion topics the same as before like talking about "self" and what they think about after graduating from school.

Implementation: In Cycle 2, students teamed up in pairs and worked together during their chats with Gemini AI. They collaborated on coming up with questions, reacting to the AI's responses, and keeping the conversation going. The goal was to support each other and make it easier for everyone to stay engaged without feeling overwhelmed or exhausted trying to carry the dialogue alone.

Evaluation: For the evaluation, it can be seen that some good progress in Cycle 2 like students talking in pairs were able to have longer chats with the AI than those in Cycle 1 when they worked alone. This lines up well with ideas from collaborative learning and sociocultural theories, which tell us that social interaction boosts language skills. Working together helps build confidence, supports each other, and keeps the conversation flowing, all of which are key for improving speaking skills. The fact that conversations were longer in pairs shows that having a social aspect really helped lighten the load and provided immediate support, making the process less challenging.

However, students were still not able to initiate and keep conversations with the AI. It tells us that talking to a partner helps in making interactions sustainable but not students were able of how to formulate good questions or how to ask questions with AI. Coming up with new content and making the dialog make sense remained the hardest part, partner sharing this task. At times, students would actually talk more amongst themselves about what to say back vs. talking to the actual AI. Perhaps the AI's adaptive replies do not suffice in the sense that students require even more assistance to take charge of leading conversations.

Cycle 3: Scaffolded Pair Work with AI-Mediated Speaking Practice

Planning: To tackle the ongoing challenge of students struggling to hold conversations, the teacher made a key change in Cycle 3. This time, she provided specific questions to guide students when they interacted with Gemini AI. The goal was pretty simple: make conversations longer and easier by giving students a clear starting point and consistent prompts for chatting with the AI.

Implementation: During this final phase, students still worked in pairs, but now they used the questions the teacher supplied to jumpstart and keep their conversations going with Gemini AI. This approach gave them a clear structure, which made it much easier for them to think of questions and topics on their own, reducing the mental effort normally involved.

Evaluation: The results from Cycle 3 were definitely the best. Most groups managed to chat for over 5 minutes, which was even longer than what we aimed for back in Cycle 1. Students also said it was noticeably easier to talk with the AI, mainly because they had those questions ready to go to keep the conversation flowing.

This big jump in both conversation length and how easy it felt shows just how powerful scaffolding can be in learning a language. By providing discussion prompts or conversational frames, it could cut down the mental workload of students to build context talking with AI. Instead, they can focus more on actually speaking and being fluent. This aligns really well with Vygotsky's idea of the Zone of Proximal Development, where the right kind of support helps students reach higher levels than they could on their own.

This cycle is a good example of how humans and AI can work together effectively. The teacher acts as the expert, offering the necessary scaffolding, while Gemini AI creates a space for real conversation. It shows that the goal isn't to replace teachers with AI but to support them. AI takes on some of the back-and-forth, freeing teachers to focus on bigger-picture stuff like designing better prompts, keeping track of progress, and giving more targeted feedback. The teacher's skill in designing those prompts was key to making this work—this isn't just about using a tool, but about thoughtful, strategic lesson planning.

4.2. Discussion: Interpreting the Use of Gemini AI in Speaking Instruction

4.2.1. Synthesis of Key Findings Across Cycles

Going through those three cycles of action research really tells a story about how teaching with Gemini AI for speaking practice has gotten better over time. It started with students just talking to the AI by themselves, then moved to students in pairs working with the AI together, and finally added in structured support to make collaboration even more effective. What's clear across all these phases is that while AI has some big perks like letting students practice anytime and providing a low-pressure environment simply using AI without thoughtful teaching strategies doesn't cut it. To truly help learners, really need to design the lessons carefully, making sure to include social interaction and clear support structures. Early on, it was obvious that just letting students talk to the AI on their own wasn't enough to really improve their speaking skills. The noticeable progress in how long students could hold conversations and how naturally they spoke grew each time by added a layer of support moving from solo work to pair work, and then to structured prompts. This shows that AI's usefulness depends a lot on how it's embedded into a well-thought-out learning plan, not just on the technology itself. It challenges the idea that AI alone can fix everything and reminds us how important good teaching methods still are if want to get the most out of it.

4.2.2. The Evolving Role of Interaction and Scaffolding

A. Analyzing the shift from individual to collaborative AI interaction

In the first cycle, we noticed some pretty common struggles things like lots of silence and conversations that barely went beyond a few words. This really showed that just using AI alone, without any guidance, makes it tough for learners to get comfortable practicing speaking. It kind of goes against the idea that AI is supposed to create a stress-free space for chat (Jiang, 2024). Sure, AI can help reduce anxiety around talking to people, but these results suggest that other mental challenges like figuring out how to kick off a conversation and keep it going are still pretty challenging.

Then, when switched to pair work in Cycle 2, things really took a turn for the better. Conversations got much longer, which is a huge win. This lines up with a bunch of research on working together and how social learning works. Moran & Iza (2024) talk about how activities done with partners help build confidence, support each other, and keep the talking flowing, all of which are super important for becoming more fluent. Vygotsky's idea of the Zone of Proximal Development (ZPD) ties in here, showing how buddies create a safe space where learners can push past their limits, sharing the mental load and getting immediate feedback from someone right there with them (Govindasamy & Shah, 2020).

B. Discussing the effectiveness of teacher-provided prompts as a scaffolding technique

Even though working in pairs has clear positives, students in Cycle 2 still found it tough to chat, meaning they still needed more help from teacher. So, in Cycle 3, the students got some questions to lead them, which was a big help. Vasquez (2025) notes that scaffolding, like Bruner and Vygotsky talked about, is about giving short-term help so students can do things they couldn't do alone. In our study, the questions acted as prompts that made it easier for students to come up with things to say, so they could focus on speaking the language. Cycle 3 went great while chats were longer and easier which shows that this kind of help really works when learning languages with AI.

The study shows that AI, group work, and teacher support go well together for language speaking practice. AI works best when it's part of a supportive setup that considers students' social and mental needs, not just used by itself. Going from just AI, to AI with friends, and then to AI with friends and teacher help, each step really boosted how well students learned. It looks like AI helps to make people smarter, but really it is good teaching that makes AI more useful. This backs up the idea that AI is a strong tool, but teachers need to know how to use it well, and it won't replace good teaching.

4.2.3. Gemini AI as a Tool for Speaking Practice

A. Exploring the benefits of AI for on-demand, interactive language practice

Gemini AI, as a large language model (LLM), is undoubtedly an extremely useful tool for language learning since it has advanced conversational abilities, has easy accessibility, and can offer interactive, real-time practice. It is no wonder that it can provide activities that simulate human conversations, which enables learners to practice speaking a language at high frequency without the constant presence of human instructors. The "no-stress communication environment" that is provided by AI is very good for second language learners, who often think that speaking is the hardest language skill. In this way, the use of AI in the second language learning process can be very beneficial because it can help to reduce anxiety and increase the willingness to participate. Besides this, the latest Gemini features, for example, "contextually adaptive responses" and its capacity to "explain how you get there" ' point to a bright future of further

development of sophisticated pedagogical applications, like highly customizable learning experiences and multimodal interactions.

B. Addressing the challenges encountered and how they were mitigated

On the other hand, the study highlighted some initial challenges, which are the problem of "silent mode" and students' insistence that it was hard to start and keep the conversations going. The main reason for these difficulties is the cognitive load notion, which states that learners may have been overburdened by the fact that on the one hand they had to create content and form grammatically correct utterances and at the same time they had the need to manage the interaction with the AI. The iterative nature of the action research was very helpful in lessening the impact of such difficulties. The implementation of pair work in Cycle 2 not only successfully lessened the burden on the individual but also created a space that allowed for mutual support, and the teacher-provided prompts in Cycle 3 directly tackled the problem of initiating the conversation because they resulted in less cognitive load and thus the teacher's adaptive strategies in response to the emergent difficulties were manifested clearly.

The research clearly shows that AI can significantly lower people's affective barriers like speaking anxiety but it is not guaranteed that it will solve cognitive or strategic barriers forever. The "silent mode" feature first and then the "difficulty building conversation" result in the study, even though AI is "no-stress" environment, clearly indicates this difference. It suggests that, besides AI being a non-judgmental practice space, learners should definitely have or be taught conversational strategies (for example topic initiation, turn-taking, elaboration) if they want to be able to make use of this environment effectively. Methods such as external scaffolding, which was successfully done in Cycle 3, were definitely important in closing this gap and thus allowing students to use AI's features more effectively.

4.2.4. Theoretical Implications

a. Connecting findings to established pedagogical theories of language acquisition:

This research contains solid evidence for the interactionist approach to language acquisition, which is the work of Lev Vygotsky's sociocultural theory. The noteworthy gains witnessed when pupils moved from performing tasks individually to working in pairs and further on to conducting pair work with scaffolding indicate the indispensability of social interaction and cooperative learning for language development. The idea of the Zone of Proximal Development (ZPD) finds its best expression in the teacher's offering of support through directing questions, thus allowing students to do conversational tasks, which they have not been able to do independently, at first.

Moreover, the first deficiencies of unregulated solo AI communication in Cycle 1, which in some ways may be likened to the behaviorist model of stimulus-response are shown in that such a model is not the correct one for language practice in that it does not give room for the establishment of a sustained and meaningful conversation. Then the accomplishment of the various cycles that took into account social interaction and cognitive scaffolding portrays a roadmap that is more socially and cognitively inclined for language learning. Such approaches focus on the learners' active construction of meaning and the strategic interactions rather than simple imitation or unguided practice. The case is a concrete demonstration of how the integration of AI into education can be reconceived through the reinterpretation of traditional

pedagogical theories and not as a disruptive force that requires completely new theoretical perspectives. Furthermore, it is a reaffirmation of the perpetually timely presence of Vygotsky and Bruner's ideas in the epoch of AI.

b. Practical Implications for Language Educators

The findings of this classroom action research study offer several actionable recommendations for language educators seeking to integrate conversational AI effectively into their speaking curricula:

- **Prioritize Structured, Collaborative AI Interaction:** Educators should move beyond simply asking students to interact individually with AI. Instead, they should design activities that leverage the benefits of *pair or group work* with AI. This fosters a supportive learning environment, reduces individual pressure, and demonstrably promotes longer, more active conversations.
- **Implement Teacher-Designed Scaffolding:** It is crucial for teachers to provide *explicit scaffolding* in the form of guiding questions, sentence starters, or topic prompts. This strategic intervention reduces cognitive load, helps students initiate and sustain conversations, and allows them to focus more effectively on language production and fluency.
- **Leverage AI for Fluency, Retain Teacher for Accuracy:** While AI is an excellent tool for *on-demand practice* and *fluency development* in a low-stress environment, teachers must maintain their critical role in providing deeper, nuanced linguistic feedback. This includes addressing complex grammar, syntax, and pragmatic errors, given AI's known limitations in identifying subtle mistakes or occasionally fabricating errors. The teacher's analysis of AI-generated transcripts remains a vital component of the learning process.
- **Explore AI for Personalized Learning:** Educators should explore Gemini's potential for *personalized learning*. Its ability to offer tailored responses, create custom quizzes, or provide step-by-step problem-solving guidance can be adapted to meet individual student needs and learning styles.
- **AI as an Augmentation, not a Replacement:** This study reinforces that AI should be viewed as a tool to *augment teaching* and learning, promoting "intelligence amplification" and potentially decreasing certain aspects of teacher workload (e.g., providing an accessible conversational partner). However, it does not diminish the essential and irreplaceable role of the human educator in pedagogical design, nuanced feedback, and fostering a holistic learning environment.

The study provides a clear roadmap for educators on how to move beyond basic AI tool adoption to pedagogically sound integration. This approach transforms AI from a mere novelty into a strategic instructional asset. It shifts the focus from a broad question of "can AI teach?" to a more precise and actionable inquiry: "how can teachers effectively *use* AI to teach?"

4.2.5. Limitations of the Study and Directions for Future Research

a. Limitations

As an action research study, the results are by nature limited to the particular classroom context, student demographics, and the teacher's unique implementation. Hence, it is quite unlikely that the findings apply generally to other educational settings or student populations without conducting additional research. The

research explored the immediate effects over three different cycles and hence gave interesting insights about short-term gains. Nevertheless, the long-term effect of AI-supported speaking practice on continuous language ability, keeping the learned material, and the autonomy of the learner is largely unknown. Gemini AI, although conversational interaction, it is far from providing deep, accurate linguistic feedback (e.g., the nuances of grammar, syntax, and semantics) because it has limited capabilities in this respect. The study mostly depended on the teacher's analysis of the transcripts which were AI-generated to ascertain the performance. At the same time, this is a strength that highlights the future area for AI development where the implementation of more sophisticated automated feedback is possible. The observations here that are "difficulty to build conversation" and "silent mode" strongly indicate that increased cognitive load was present; however, the study did not use some quantitative measures to really assess cognitive load.

b. Directions for Future Research

Building upon these findings, a number of directions for new areas of research become apparent. One of the possibilities which future research could pursue is the inquiry of the success rate of the different kinds of scaffolding (e.g., verbal, procedural, visual) and the best coalescing of the AI with different learner profiles. Longitudinal studies are imperative to measure the lasting influence of AI-supported speaking exercises on overall language proficiency, motivation, and the formation of learner autonomy during the time of continuous involvement. Research can also be aimed at investigating effective models for teacher training when using AI in language classrooms, regarding pedagogical design, prompt engineering, and exploiting AI as a tool for "intelligence amplification". Besides, comparative studies can be undertaken for assessing the effectiveness of speaking practice using the AI versus traditional talk activities or other technological interventions. Conducting deeper qualitative and quantitative research on student perceptions of AI interaction, including their experience of cognitive load and anxiety, would be of great importance. Moreover, regarding Gemini's multimodal features (audio, video, images, text) as a resource to improve speaking practice beyond text-based conversations offers a potential area to be explored in the future. The recognition of these as both limitations and directions for the future of the study is critical for its academic integrity and also its axiom that the study is only the beginning of an ever-expanding field. It emphasizes the constant progression of the AI technology and the improvements of its practical usage in pedagogy, thus allowing for more research to be done instead of claiming it's the only answer.

5. Conclusions

The use of Gemini AI in teaching speaking that is shown through this classroom action research study is a clear trail of effective pedagogical integration. First of all, students had individual and unguided sessions with AI, which were not enough for continued speaking practice and thus, short conversations and student hesitation occurred. Moving to pair work has notably improved conversation length, thus, indicating the indispensable role of collaborative learning and social interaction in language development. Nevertheless, there were still some problems in starting and developing conversations, which means that more assistance was needed. A very indispensable source of the students' overcoming of the cognitive load and focusing on longer and more fluent conversations with AI was teacher-provided guiding questions, which as a kind of support, they could use as a vehicle of scaffolding.

The results corroborate that conversational AI uniquely carves out opportunities like on-demand and low-stress practice; however, its real magic in language education comes when placed in a well-thought-out pedagogical framework that brings in established principles of collaborative learning and scaffolding. The study puts forward solid evidence for a synergistic model of human-computer collaboration, where the teacher's expertise in instructional design perfectly complements the AI's interactive capabilities. Essentially, after the teacher has carefully chosen and strategically implemented Gemini AI as a tool to augment teaching and learning and as a facilitator of better speaking practice, it can no longer be viewed as a standalone agent replacing human intervention in teaching.

References

- Abdullaev, Z. (2022, June 10). *Importance of teaching speaking skill*. Proceedings of the International Scientific-Practical Conference on Translation Studies and Pragmalinguistic Development Tendencies: Problems, Solutions and Achievements. Uzbekistan State World Languages University. <https://www.researchgate.net/publication/373901543>
- Alhusaiyan, E. (2025). A systematic review of current trends in artificial intelligence in foreign language learning. *Saudi Journal of Language Studies*, 5(1), 1–16. <https://doi.org/10.1108/SJLS-07-2024-0039>
- Assylbayeva, A. (2014). Teaching speaking skills to adult EFL learners. *Kazakh-American Free University Academic Journal*, (6). <http://www.kafu-academic-journal.info/journal/6/>
- Burns, A. (2019). Action research in English language teaching: Contributions and recent developments. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. [halaman tidak tersedia]). Springer. https://doi.org/10.1007/978-3-319-58542-0_52-1
- Govindasamy, M., & Shah, P. M. (2020). Students' perceptions on collaborative speaking tasks in ESL classrooms. *Creative Education*, 11(11), 2280–2292. <https://doi.org/10.4236/ce.2020.1111167>
- Imran, M., & Almusharraf, N. (2024). Google Gemini as a next generation AI educational tool: A review of emerging educational technology. *Smart Learning Environments*, 11(22). <https://doi.org/10.1186/s40561-024-00310-z>
- Jiang, H. (2024). Applications and research gaps of LLM-based English-as-a-foreign-language education. *Journal of Education, Humanities and Social Sciences EMSS*, 39, 429–433.
- Jon, R. B., Embong, R., Purnama, B., & Wadi, A. S. (2021). The challenges of English language teaching in Indonesia. *International Journal of English and Applied Linguistics (IJEAL)*, 1(3), 158–160. <https://doi.org/10.31098/ijeal.v1n1.1157>
- McGehee, N. (2024). *AI in education: Student usage in online learning*. Michigan Virtual Learning Research Institute. <https://michiganvirtual.org/research/publications/ai-in-education-student-usage-in-online-learning/>
- Morán, J. I. A., & Iza, M. F. A. (2024). The cooperative learning in the English speaking – fluency development. *Revista Científica de Innovación Educativa y Sociedad Actual "ALCON"*, 4(2), 81–89.

- Rao, P. S. (2019). The importance of speaking skills in English classrooms. *Alford Council of International English & Literature Journal (ACIELJ)*, 2(2), 6–7. <https://www.researchgate.net/publication/334283040>
- Rezeki, P., & Dalimunte, M. (2024). Exploring English teachers' difficulties in teaching speaking. *Inspiring: English Education Journal*, 7(1), 34–48. <https://doi.org/10.35905/inspiring.v7i1.8793>
- Saeidnia, H. R. (2024). Gemini is multimodal. *Information Matters*. <https://ssrn.com/abstract=4703935>
- Vasquez, V. (2025, May 27). Scaffolding techniques for English language learners: Part 1. *Collaborative Classroom*. <https://www.collaborativeclassroom.org/blog/scaffolding-techniques-english-language-learners-part-1/>
- Wu, L. (2024). AI-based writing tools: Empowering students to achieve writing success. *Advances in Educational Technology and Psychology*, 8(2), 40–50. <https://doi.org/10.23977/aetp.2024.080206>