Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3, 2616-2629
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i3.5853

The role of AI-powered chatbots in enhancing second language acquisition: An empirical investigation of conversational AI assistants

Jordan Taeza[1*]
[1]Batangas State University, Philippines; jordan.taeza@g.batstate-u.edu.ph (J.T.)

Abstract: This study explores the role of AI-powered chatbots in enhancing second language acquisition (SLA), focusing on speaking proficiency, learner engagement, and confidence. A mixed-methods, quasi-experimental design was employed involving 60 intermediate ESL learners divided into a chatbot-assisted experimental group and a control group using traditional practice. Over six weeks, the experimental group engaged in structured interactions with a conversational AI chatbot offering real-time feedback. Pre- and post-tests, engagement surveys, and interviews were used for data collection. Findings revealed that the chatbot group showed significantly higher gains in speaking proficiency and greater improvements in willingness to communicate and self-confidence. Qualitative feedback highlighted increased practice, reduced anxiety, and high learner motivation, though limitations such as repetitive responses and limited cultural understanding were noted. The study concludes that AI chatbots can serve as effective supplemental tools in SLA, especially for enhancing oral skills and learner autonomy. Practical implications suggest integrating chatbots into language curricula for additional speaking practice, particularly in contexts with limited teacher availability. Educators are advised to blend chatbot use with guided instruction and monitor chatbot feedback quality to ensure pedagogical alignment.

Keywords: *AI-powered chatbots, Artificial intelligence in education, Conversational AI, Educational technology, Language fluency, Language learning, Second language acquisition, Student engagement.*

1. Introduction

The rise of artificial intelligence has introduced new opportunities in education, particularly through conversational AI assistants (chatbots) that can interact with learners in natural language. In second language acquisition (SLA), practicing communication is crucial, yet learners often face limited access to fluent speakers or fear of making mistakes. AI-powered chatbots offer a potential solution by providing on-demand, non-judgmental conversational practice. These chatbots can simulate human-like dialogue, answer learner questions, and give instant feedback, creating a personalized and interactive learning environment. Early applications of chatbots in language learning have shown promise in increasing learner engagement and providing 24/7 practice opportunities. However, the effectiveness of chatbots for improving actual language proficiency requires careful empirical investigation, and educators have raised concerns about the limitations and proper integration of this technology in teaching [1, 2].

The researcher presents a more comprehensive analysis with an expanded literature review, detailed mixed-methods methodology, and in-depth results and discussion. The goal is to examine how AI chatbots can enhance SLA outcomes (particularly speaking proficiency and learner engagement) and to identify the benefits, challenges, and future directions for using conversational AI in language education. By synthesizing current research and presenting hypothetical data from a controlled study,

History: Received: 20 January 2025; Revised: 20 March 2025; Accepted: 24 March 2025; Published: 29 March 2025
* Correspondence: jordan.taeza@g.batstate-u.edu.ph

the researcher aims to provide insights into the pedagogical value of AI chatbots and guidelines for their effective use alongside human instruction.

2. Literature Review

2.1. AI Chatbots in Education and Language Learning

Artificial intelligence chatbots have increasingly been adopted in educational settings for various purposes, from answering student queries to serving as virtual teaching assistants [1]. A recent systematic review by Labadze, et al. [1] found that students benefit from AI-powered chatbots in several key ways: they receive on-demand homework and study assistance, experience more personalized learning, and can develop various skills through interactive practice [1]. In the context of language learning, these benefits translate into more opportunities for authentic language use and individualized feedback. Chatbots can engage learners in dialogue-based practice tailored to their proficiency level, providing instant explanations or corrections as needed [3]. They are available at any time, allowing learners to practice speaking or writing in the target language beyond classroom hours, which can lead to increased exposure and reinforcement of language skills [2].

Notably, chatbots create a low-pressure environment that may reduce the anxiety often associated with speaking a foreign language. Because learners know they are interacting with a non-judgmental AI, they can converse freely without fear of embarrassment over mistakes [2]. This reduction in language anxiety is significant – studies in positive psychology indicate that lower anxiety and higher enjoyment lead to more active participation and better language outcomes [2]. Preliminary research indeed suggests that AI chatbots help foster positive learner emotions; for example, one study reported that chatbots can reduce language learning anxiety by providing a low-stakes practice **environment**, thereby encouraging learners to speak more and gain confidence (Xiao et al., 2024). By simulating real-life conversational scenarios and giving immediate feedback, chatbots can boost learners' motivation and self-efficacy in using the target language [2, 3]. Moreover, advanced conversational agents (such as those powered by large language models) are capable of handling diverse topics, offering translations, and even adapting responses to the learner's input, further enhancing the authenticity of practice [3]. In sum, the broader landscape of research highlights AI chatbots as a powerful tool in education and language learning, with the potential to increase engagement, provide personalized interaction, and supplement traditional teaching in meaningful ways [1, 2].

.

2.2. Empirical Evidence on Chatbot Effectiveness in SLA

An emerging body of empirical studies has examined how effective AI chatbots are in facilitating second language acquisition. Overall, these studies provide encouraging evidence that chatbot-assisted learning can lead to measurable improvements in language proficiency. For instance, Yuan [4] conducted an experiment in a primary school EFL (English as a Foreign Language) context in China, where one group of students learned English with a chatbot and another group used traditional methods. After a 3-month intervention, the chatbot-assisted group showed significantly greater gains in oral English proficiency compared to the control group, as well as higher willingness to communicate (WTC) in English [1]. In Yuan's study, the chatbot group's speaking test scores improved notably from pre- to post-test, and their survey results indicated they became more willing and confident to speak in the target language, whereas the control group saw more modest improvements [1].

. This suggests that practicing with a conversational AI can enhance speaking skills and reduce inhibitions in using the language. Similarly, a recent meta-analysis of chatbot-assisted language learning reported a moderate overall positive effect on language performance (with an average effect size $g \approx 0.48$) for learners using chatbots versus those who did not, indicating consistent benefits across multiple studies [5].

Other research has explored specific language skills. Zhang and Huang [6] investigated vocabulary acquisition with an AI chatbot based on a large language model (LLM). In their controlled study, 52 foreign language students were split into two groups: one used an LLM-powered chatbot as a personal

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

vocabulary tutor, and the other learned new words without the chatbot. After an eight-week training focusing on a set of target words, the chatbot group outperformed the control group in both receptive vocabulary (recognition of word meaning) and productive vocabulary (ability to use the word correctly) [6]. The chatbot users not only learned more words by the end of the study, but they also retained those words better when tested again two weeks later [6]. This finding underscores the efficacy of conversational AI in reinforcing vocabulary learning—likely because the chatbot provided repeated exposure to new words in context and offered prompt practice and feedback. Another advantage observed was incidental learning: students interacting with the AI assistant picked up additional vocabulary and phrases that were not explicitly taught, simply through the flow of conversation [6]. These results align with the notion that conversational interaction is a driver of language acquisition; the chatbot essentially engaged learners in meaningful input and output, aiding vocabulary retention and usage.

Beyond speaking and vocabulary, there is evidence that chatbots can support writing skills. Lingaiah, et al. [7] reviewed applications of AI chatbots for ESL writing practice and found that immediate, automated feedback from chatbots helps learners correct mistakes and reinforce grammar and writing strategies [7]. In the studies they analyzed, students who used chatbots for writing practice received prompt feedback within each session, which enabled them to identify errors (e.g. in grammar or word choice) and fix them right away [7]. Over time, this iterative feedback loop contributed to improved writing proficiency. An additional benefit reported was reduced writing anxiety and increased confidence; students felt more at ease practicing writing with a chatbot and became more confident in their writing abilities after repeated sessions [7]. These findings across speaking, vocabulary, and writing domains illustrate that chatbot-assisted learning can yield tangible improvements in second language skills, often rivaling or surpassing the gains from traditional instruction [1, 6]. However, the magnitude of improvement can depend on factors such as the quality of the chatbot's responses, how the tool is integrated into the curriculum, and the extent of human guidance provided alongside the AI. While the evidence is largely positive, it also points to the need for addressing certain challenges to maximize effectiveness.

2.3. Challenges and Limitations of AI Chatbots in Educational Settings

Implementing AI-powered chatbots in language education is not without challenges. Several studies and reviews have identified limitations that educators and developers must consider when using chatbots for learning:

- Reliability and Accuracy: Chatbot responses are not always accurate or reliable. AI chatbots can occasionally provide incorrect information, nonsensical answers, or grammatically flawed feedback, which may mislead learners if unchecked [1]. Maintaining a high level of answer accuracy is a concern, especially when learners might trust the AI's feedback as authoritative.
- Cultural and Contextual Understanding: Current chatbots often lack deep cultural nuance and contextual understanding. They may struggle with idioms, humor, or context-specific language usage, and they cannot always tailor feedback to the cultural background of the learner [7]. This limitation means chatbots might give generic or culturally inappropriate responses in some scenarios, limiting their effectiveness in teaching pragmatic and sociolinguistic aspects of language.
- Sustaining Engagement: While chatbots can boost engagement initially, some research suggests that the novelty can wear off over time. Learners might lose interest if the chatbot's style becomes repetitive or if it fails to evolve with the learner's progress [7]. Keeping students motivated during long-term chatbot use remains a challenge, and designing interactions that remain fresh and engaging is an area for improvement.
- Privacy and Ethical Issues: The use of AI in education raises privacy and ethics concerns. Chatbots often require collection of user data (e.g. conversation logs) to function and improve, which could expose sensitive information if not properly secured [1]. Additionally, if students

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

rely on chatbots for answers, it can blur the line between learning and cheating; educators worry about academic integrity when AI tools might be used to do assignments or provide answers without the student's true effort [3]. Proper guidelines and safeguards are needed to ensure chatbots are used as learning aids, not shortcut solutions.

- Integration and Design Challenges: Developing and integrating a chatbot into an existing curriculum can be complex. Teachers may need training to effectively use the chatbot as a teaching tool, and the assessment of chatbot-driven learning is not straightforward [8]. Many institutions struggle with how to measure learning outcomes from chatbot interactions and how to align chatbot activities with learning objectives and standards. If not well-designed, a chatbot might not match the syllabus or could provide experiences that are fun but not adequately educational.
- Role of Human Instructors: There is an underlying concern in the education community that AI chatbots might replace or reduce valuable human interaction. Language learning is a social, human endeavor, and some fear that over-reliance on chatbots could diminish time spent practicing with teachers or peers [2]. The consensus in current literature is that chatbots should *augment* human instruction, not replace it. Teachers bring expertise, empathy, and the ability to handle nuanced student needs that AI still cannot fully replicate [2]. Balancing the use of chatbots with human guidance is crucial to avoid isolating learners or neglecting the social aspects of language use.

While AI chatbots offer exciting opportunities in second language education, these challenges highlight the importance of thoughtful implementation. Effective use of chatbots requires addressing technical limitations (accuracy and context-awareness), ensuring ethical use (privacy and academic honesty), and integrating chatbot practice with curriculum and teacher involvement. Ongoing research is focusing on overcoming these limitations—for example, by improving the natural language processing capabilities of chatbots to better handle complex expressions and by developing guidelines for educators on blending chatbot sessions with traditional teaching [1, 7]. The next sections of this paper will describe an empirical study that explores both the benefits and challenges of using a conversational AI assistant in a language learning setting, followed by a discussion of findings in light of the literature above.

3. Methodology

3.1. Research Design and Participants

This study employed a mixed-methods research design to investigate the impact of an AI-powered conversational chatbot on second language learning outcomes. The design combined a quantitative quasi-experimental approach with qualitative feedback collection, allowing for a comprehensive analysis of both performance data and learner perceptions. The researcher selected a sample of 60 learners of English as a second language (ESL) from language program in a state university in the Philippines. Participants were young adult learners (ages 18–24) at an intermediate proficiency level in English, based on placement test results. They were randomly assigned to one of two groups (30 students each): an experimental group that used a chatbot as a supplement to their regular English classes, and a control group that followed the regular curriculum without chatbot assistance. Both groups were ensured to be equivalent in terms of initial language proficiency (verified by pre-test scores) and demographic background. The learners attended classes of equal duration and had the same instructional content, with the only difference being the availability of the chatbot for extra practice in the experimental condition.

This quasi-experimental setup (chatbot-assisted vs. traditional learning) was chosen to isolate the effect of the chatbot on language acquisition. All participants gave informed consent to partake in the study. They were informed about the study's purpose and assured that the chatbot was a tool to assist their learning rather than a test. Participants in the experimental group were briefed on how to use the chatbot and encouraged to interact with it regularly, while the control group was encouraged to

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

practice using conventional methods (such as speaking with classmates or completing additional worksheets). By the end of the study period, the researcher gathered both quantitative performance data and qualitative feedback from participants in both groups.

3.2. Chatbot Selection and Implementation

For the experimental group, the researcher selected a conversational AI chatbot that could engage in free-form English dialogue and provide immediate corrective feedback. The selection criteria for the chatbot included: (a) language capability – it needed to handle intermediate-level English with a wide range of conversational topics; (b) feedback features – it could correct user mistakes or suggest improvements in grammar and vocabulary; (c) user-friendliness – a simple interface accessible via students' devices; and (d) consistency and reliability – stable performance with minimal technical glitches. After evaluating several available AI chatbots, the study chose a platform powered by a large language model (comparable to GPT-3.5) for its fluency and ability to adapt responses to user input. The chosen chatbot had been pre-loaded with an educational mode that allowed it to act as a language tutor – for example, it would gently point out grammatical errors in the learner's input or answer questions about word usage.

In practice, the chatbot was implemented as a mobile and web application so that students could chat with it via text. The interactions were text-based to allow the chatbot to easily analyze learner input and provide written feedback, although students were encouraged to read responses aloud to practice pronunciation. The chatbot was configured to simulate a conversational partner for practicing everyday English. For instance, it could role-play scenarios (like ordering food, casual conversations, or discussing a news article) and prompt the learner to respond. If the learner made a mistake or seemed stuck, the chatbot would offer hints or corrections. The researcher also programmed the chatbot with a degree of flexibility: it could switch to the learner's first language for a quick explanation if asked, but primarily maintained conversation in English to maximize immersion. The chatbot logs were saved (with student permission) for analysis, and to allow researchers to see the nature of feedback given. To ensure content appropriateness, the chatbot's responses were monitored and filtered for any offensive or irrelevant content. Throughout the study, technical support was available in case students encountered issues with accessing or using the chatbot. No major technical problems were reported, and usage statistics indicated that students were generally able to interact with the chatbot smoothly.

3.3. Data Collection and Procedure

The study ran for 8 weeks during an academic semester. Both groups covered the same instructional material in their regular classes, focusing on improving speaking and listening skills in English, alongside vocabulary development. The procedure for data collection was as follows:

1. Pre-Test: In the first week, all participants underwent a speaking proficiency pre-test. This test involved an oral interview and a standardized speaking task (e.g., describing a picture and answering questions) evaluated by two independent language instructors. The researcher used a rubric rating fluency, pronunciation, vocabulary range, grammatical accuracy, and overall coherence on a 0–10 scale for each category (for a total speaking score out of 50). The pre-test established a baseline for each student's speaking ability. Participants also filled out a Learner Engagement Questionnaire that included items on their motivation, confidence in speaking English, and willingness to communicate (using a Likert scale from 1 = strongly disagree to 5 = strongly agree on statements about enjoying speaking practice, etc.). This provided baseline data on affective factors like engagement and confidence.
2. Intervention: From weeks 2 to 7, the experimental group had access to the AI chatbot for practice, while the control group continued with traditional practice methods. Students in the experimental group were instructed to have at least three chatbot conversation sessions per week (approximately 15–20 minutes each) outside of class, focused on that week's topics (e.g., discussing hobbies, retelling a story, debating a simple issue). They were free to use the chatbot

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

more often if they wished, and the app logged each session's duration and content. In contrast, the control group was assigned alternative practice activities such as speaking with a partner from class or recording themselves speaking on a topic – these activities were monitored through classwork but did not involve the AI assistant. During this period, both groups continued attending their normal classes, where the instruction was identical for both groups. The only difference was the extra out-of-class practice method. The researcher checked in weekly with the experimental group to ensure they were comfortable using the chatbot and to record any immediate feedback or issues.

3. Post-Test: In week 8, both groups took a speaking proficiency post-test similar in format to the pre-test. The tasks were of comparable difficulty and were scored with the same rubric by the same instructors (who were blind to which group each student was in). This post-test aimed to measure any improvements in speaking proficiency after the six-week intervention. Students again completed the Learner Engagement Questionnaire at post-test, which included additional questions specific to their experience (for the experimental group, questions about their satisfaction with the chatbot and perceived improvements, and for the control group, their satisfaction with their practice methods).
4. Qualitative Feedback: To complement the quantitative data, we collected qualitative insights. The researcher conducted semi-structured interviews with a subset of 10 students from the experimental group and 5 students from the control group after the post-test. The experimental group interviews focused on their experiences with the chatbot – what they found helpful or challenging, and any noticeable changes in their language learning process. The researcher asked questions like "How did chatting with the AI assistant compare to practicing with classmates or by yourself?" and "Can you share an example of a helpful interaction you had with the chatbot?" For the control group, the researcher asked about their practice habits and if they felt they lacked any resources for practice. Additionally, the two instructors who evaluated the speaking tests provided their observations on student performance and participation in class for both groups. These interviews were audio-recorded and transcribed for analysis. The goal was to identify common themes regarding the chatbot's effectiveness and any issues faced by learners.

4. Data Analysis

The researcher adopted a mixed-methods analysis approach to interpret the data collected. For the quantitative data, the researcher first ensured the reliability of our measures. The speaking test ratings by two instructors were compared for inter-rater reliability, which was high (intraclass correlation > 0.9), so the researcher averaged their scores for each student. The researcher then computed gain scores for speaking proficiency (post-test minus pre-test score) for each participant. An independent samples t-test was used to compare the mean improvement in speaking scores between the experimental (chatbot) group and the control group. The researcher also conducted a paired t-test within each group to see if the improvement from pre to post was statistically significant. For the engagement questionnaire, we computed mean scores for composite constructs like "willingness to communicate" and "confidence in speaking" at pre- and post-test. The researcher compared these using repeated measures ANOVA with group as a between-subject factor to assess any interaction effect (to see if the change over time differed by group). The significance level was set at $p < 0.05$ for all tests, and effect sizes (Cohen's d) were calculated for the difference in improvement between groups to gauge the practical significance of results. All quantitative analyses were performed using SPSS software.

For the qualitative data (interview transcripts and open-ended survey responses), the researcher utilized thematic analysis. Two researchers independently reviewed the experimental group interview transcripts to identify recurring themes or comments about the chatbot experience. They coded the data for key themes such as "perceived improvement in speaking," "motivation and engagement," "chatbot feedback usefulness," "frustrations or limitations," and "comparisons with human interaction." After

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

initial coding, the researchers discussed and refined the themes, reaching consensus on the interpretations. The control group interviews were also analyzed to capture themes related to their practice experience and any mention of what might have helped them further. The researcher also reviewed the chatbot interaction logs for illustrative examples of how the chatbot responded to learner inputs (e.g., instances where it corrected a mistake or taught a new phrase) to triangulate the students' descriptions with actual chatbot behavior. The qualitative findings were then integrated with quantitative results during the discussion to provide a richer understanding of the outcomes.

By combining statistical analysis of performance data with thematic analysis of learner experiences, the study's mixed-methods approach offers both breadth and depth. Quantitative results indicate whether there was a significant effect of chatbot assistance on language gains, while qualitative results help explain how the chatbot impacted learners and what their subjective experiences were. This complementary analysis strengthens the validity of our conclusions regarding AI chatbots' role in second language acquisition.

5. Results and Discussion

5.1. Speaking Proficiency Outcomes (Quantitative Results)

Speaking Test Improvement: The experimental group that used the AI chatbot showed a substantial improvement in speaking proficiency from pre-test to post-test, outperforming the control group. On the 50-point speaking test scale, students in the chatbot group improved by an average of 12.5 points (SD = 4.0) after six weeks, whereas the control group improved by an average of 6.8 points (SD = 3.5). This difference in gains was statistically significant ($t(58) = 5.27$, $p < 0.001$), indicating that the chatbot group's speaking abilities increased more than those of the control group. In practical terms, the chatbot group's average score rose from about 27/50 to 39.5/50, roughly moving from a lower-intermediate to a solid intermediate proficiency as per our rubric descriptors, while the control group rose from 28/50 to 34.8/50. All students in the experimental group improved to some degree, and 85% of them achieved higher post-test scores than the highest-scoring student in the control group, suggesting a consistent benefit across the chatbot users. By contrast, the control group's improvements, while present, were more modest; a few control students showed little to no improvement in fluency or continued to struggle with consistent grammar in speech by the post-test.

Statistical analysis confirmed that both groups made significant gains within themselves (each $p < 0.01$ for paired tests), which is expected from ongoing classroom instruction. However, the effect size for improvement was much larger for the chatbot group (Cohen's $d \approx 1.2$, a large effect) than for the control group ($d \approx 0.6$, a medium effect). This suggests that the chatbot usage had an educationally meaningful impact. The instructors who rated the speaking tests noted qualitative differences as well: students from the chatbot group demonstrated greater fluency and confidence during the oral interview. They spoke in longer sentences and hesitated less, and some used vocabulary or idiomatic expressions that were not observed in the control group's performances. One instructor commented, *"It was clear some students had a lot of practice speaking – they sounded more natural and even attempted complex sentences. In the other group, many were still giving very short answers."* These observations align with the quantitative scores and point to the chatbot practice providing additional speaking practice that translated into improved oral proficiency.

Learner Engagement and Willingness to Communicate: Alongside speaking scores, we measured changes in engagement-related attitudes. The experimental group's Willingness to Communicate (WTC) in English (derived from questionnaire items about speaking up in class and initiating conversation) increased significantly over the course of the study. On a 5-point scale, their average WTC score went from 3.0 (pre-test) to 4.2 (post-test). In contrast, the control group's WTC changed only slightly, from 3.1 to 3.4 on average. The interaction effect of group and time on WTC was statistically significant (repeated measures ANOVA, $F(1,58) = 15.4$, $p < 0.001$), indicating the increase in willingness to communicate was much greater when the chatbot was part of the learning process. Similarly, self-reported confidence in speaking English improved in the experimental group (e.g.,

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

students more strongly agreed with statements like "I feel confident speaking English with others"), with 73% of chatbot group students reporting higher confidence after the intervention, compared to 40% in the control group. These quantitative engagement findings are consistent with the idea that having a patient, always-available conversational partner (the chatbot) helped students feel more comfortable and eager to use English. They could practice without fear of judgment, which likely contributed to their increased confidence [2]. In fact, a significant positive correlation ($r = 0.45$, $p < 0.01$) was found between the number of chatbot sessions a student completed and their gain in WTC score, suggesting that the more students practiced with the chatbot, the more willing and comfortable they became in speaking. This trend supports theories in SLA that frequent practice and reduced anxiety can lead to greater communicative willingness [2]

.

5.2. Learner Engagement and Feedback (Qualitative Results)

The qualitative feedback from participants in the experimental group provides deeper insight into how the chatbot influenced their learning experience. Several common themes emerged from the interviews and open-ended responses:

- Increased Practice and Autonomy: Students overwhelmingly reported that the chatbot allowed them to practice English far more than they normally would outside of class. Many mentioned engaging in extra conversations late at night or during free time, which they would not have done otherwise. One student noted, *"With the chatbot, I ended up practicing a little bit almost every day. It's like having someone always ready to chat. I spoke much more English this month than ever before."* This reflects a boost in learner autonomy and self-directed learning – the chatbot served as an always-available partner, and students took initiative to use it. Several students enjoyed the freedom to choose topics of conversation, which made practice feel less like homework and more like casual interaction. This spontaneity and frequency of practice likely contributed to the gains in fluency observed. It aligns with prior observations that chatbots promote active participation and self-directed learning, as learners can decide when and how to engage [3, 8].
- Comfort and Confidence Building: A recurring sentiment was that interacting with the chatbot reduced students' fear of making mistakes. They described the chatbot as "non-judgmental" and "patient." One interviewee shared, *"I usually get nervous speaking English, but with the AI I didn't feel shy. Even if I made mistakes, it would just kindly correct me. That helped my speaking confidence a lot."* This indicates the chatbot created a safe space to practice, echoing the literature that AI chatbots can provide a low-pressure environment that lowers anxiety [2]. Students who were normally quiet in class said they felt freer to experiment with new words or complex sentences with the chatbot, because they weren't worrying about what a person might think. Over time, this practice seemed to carry over to class participation – the instructor noted that some previously reticent students in the experimental group began speaking up more in group activities as the weeks went on. The chatbot, therefore, functioned as a confidence-building tool, gradually reducing communication apprehension and encouraging risk-taking in language use, which is a key component of developing oral proficiency.
- Usefulness of Feedback and Learning: Many students praised the real-time feedback the chatbot provided. The AI assistant would correct spelling or grammar in their messages, suggest more natural phrasing, or teach them new vocabulary when they struggled to express something. For example, one student tried to explain a personal hobby but lacked the vocabulary; the chatbot provided the appropriate term and then used it in a sentence, effectively teaching the word in context. Students found such interventions immediately useful. One comment was, *"It was like a teacher sitting with me while I practiced – if I said something wrong, I'd know right away and could fix it."* This immediate feedback loop helped them not only recognize errors but also reinforce correct language usage on the spot [7]. A few students mentioned they saved the chatbot's corrections or suggestions to review later. The constant availability of feedback is a stark

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

contrast to traditional practice (where a teacher might correct errors only the next day or peers might not correct each other at all). This likely contributed to faster improvement in areas like grammar and vocabulary usage. Students also appreciated that the chatbot could provide examples or even simple explanations in their first language if they were really confused, which helped bridge understanding. However, they noted the bot mostly kept them in English, which they recognized as good for immersion.

- Motivation and Enjoyment: The interactive and somewhat game-like nature of chatting with an AI kept learners engaged. Several described the experience as *"fun"* or *"motivating."* The novelty of having an AI friend to talk to was mentioned, but more importantly, students felt a sense of progress and accomplishment when they managed longer conversations or when the chatbot complimented their improvement. The chatbot was programmed to offer positive reinforcement (e.g. "Great job expressing your opinion!"), which students said made them feel encouraged. One student remarked, *"It's cool because it feels like a game where I level up my English every time I chat."* This increased enjoyment likely contributed to the greater time-on-task observed. Engagement logs showed that on average, students in the experimental group completed 18 chatbot sessions (of ~15 minutes) over the 6 weeks, exceeding the required minimum of 15 sessions. High engagement is a positive sign, as more frequent practice is linked to better outcomes. This aligns with findings that chatbots can increase learner engagement by providing personalized, interactive practice that holds student interest [1, 2].
- Challenges and Chatbot Limitations (Learner Perspective): Despite the overall positive feedback, students did encounter some limitations of the chatbot in practice. A common issue was that occasionally the chatbot would misunderstand what a student was trying to say, especially if the student's input had many errors or was phrased awkwardly. In such cases, the chatbot sometimes gave a response that was slightly off-topic or asked for clarification repeatedly, which a few students found frustrating. For example, one student attempted to describe a local cultural festival, but the chatbot did not seem to grasp some cultural terms and kept steering the conversation back to a generic discussion about holidays. This reflects the chatbot's lack of deep cultural knowledge or context in certain areas, as noted in the literature [7]. Students also noticed that while the chatbot was good with everyday language, it sometimes gave formulaic responses. One said, *"After a while, I kind of knew what it would reply. It was helpful but a bit robotic sometimes."* This hints that the chatbot's style, though polite and encouraging, could feel repetitive or lacking genuine personality after extensive use, potentially impacting sustained engagement for some learners. Additionally, a few participants mentioned that the chatbot's corrections, although usually accurate, were occasionally too brief. For instance, it would point out a mistake but not always elaborate on the rule, leaving the student to infer the grammar principle. These experiences underscore that while the chatbot was a valuable tool, it was not perfect. Learners still had unmet needs at times – for deeper explanations or a more nuanced understanding – which are areas where human teachers excel.

The control group's interview feedback, while not the focus, provided a useful contrast. Control students often practiced by reading dialogues or practicing with classmates, which they felt was helpful but limited. A couple of control group students expressed that they would have liked more feedback on their speaking outside of class: *"I practiced with my friend, but neither of us is fluent, so we didn't catch each other's mistakes much,"* one said. This highlights a gap that the chatbot filled for the experimental group by providing expert feedback during practice. Some control students also noted they were "a bit jealous" hearing about the chatbot from others and would be open to trying it, pointing toward a generally positive attitude about integrating such technology.

5.3. Discussion: Interpreting Effectiveness and Limitations

The results of this study indicate that incorporating an AI-powered chatbot as a conversational partner can significantly enhance second language learning outcomes, particularly for speaking skills

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

and learner engagement. The experimental group's notably higher gains in speaking proficiency support the hypothesis that chatbot-assisted practice leads to better performance than equivalent practice without a chatbot. This finding is consistent with prior empirical studies where students who used chatbots showed superior oral skill development [1]. The improvement in our study can be attributed to several factors. Firstly, the chatbot provided substantially more opportunities for output – students spoke (via typing or aloud while reading) far more sentences in English during their chatbot sessions than the control group did in their traditional practice. According to interactionist theories of SLA, such increased output and interaction lead to greater fluency and automaticity in language use. Learners had to formulate responses on the fly in conversation with the chatbot, mimicking real-life communication demands. Over time, this likely improved their ability to organize thoughts in English and respond promptly, which was evident in the smoother speech during post-test interviews.

Secondly, the immediate feedback mechanism of the chatbot helped learners notice and correct errors in real time. This aligns with the concept of focus on form in language learning, where drawing attention to linguistic mistakes at the moment they occur can facilitate learning. Our qualitative data showed that students valued this instant correction and often applied it straight away, preventing the fossilization of errors. For example, if a student consistently dropped the past tense "-ed" ending, the chatbot's on-the-spot correction in each instance helped the student become aware of the pattern and practice it correctly thereafter. In contrast, in traditional settings, such feedback might be delayed or inconsistent. The result was that the chatbot group had fewer persistent grammar mistakes by the end, as noted by instructors. This supports previous research that found personalized, timely feedback from AI tutors can strengthen language accuracy and usage [7].

Another key aspect is the role of affective factors. The chatbot, by virtue of being a machine, created a judgment-free zone that encouraged practice. Many of our participants reported increased confidence and reduced anxiety, which likely contributed to their greater willingness to communicate (WTC) and ultimately to improved speaking performance. This outcome resonates with studies in which chatbots reduced learners' communication apprehension, enabling them to practice more freely and frequently [2]. When learners are less anxious, they are more willing to take risks and use the language, leading to faster improvement. The control group, lacking this environment, didn't experience the same boost in confidence or WTC, which may explain why their speaking practice remained more conservative and less fruitful. Therefore, the emotional support aspect of the chatbot (patient, non-critical encouragement) should not be underestimated as a factor in its effectiveness.

In terms of learner engagement, the chatbot clearly succeeded in motivating students to devote extra time to language practice. The gamified feel and the novelty of chatting with AI contributed to high usage rates. It appears that the chatbot tapped into intrinsic motivation – students enjoyed the conversations and felt a sense of accomplishment, which kept them engaged. Over six weeks, this accumulated practice translated into noticeable skill gains. This finding is aligned with broader educational research that links increased engagement and time-on-task with improved learning outcomes, and shows that AI tools can play a role in boosting engagement [1, 2]. However, it's worth noting that novelty can diminish. A couple of students found the experience becoming rote over time, which flags the need for chatbots to have expansive, varied content and perhaps evolving personalities to maintain engagement long-term. Future designs might incorporate adaptive storyline-based interactions or periodically updated prompts to keep the experience fresh.

The qualitative feedback also shed light on the limitations of the chatbot, many of which mirror the challenges highlighted in the literature. One notable limitation was the chatbot's occasional lack of deep understanding. While the AI could handle everyday conversation well, it struggled with more complex or culturally specific contexts – for example, a nuanced discussion about a cultural festival. This reflects a limitation in the chatbot's training data or inferencing ability; it may not have the world knowledge or contextual sophistication to always follow the learner into less common or highly contextual topics [7]. In a real classroom, a teacher would pick up on those nuances or provide the cultural background needed. This suggests that chatbots currently are best used for general language practice and common

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

scenarios, whereas very localized or nuanced content might require supplementary explanation from teachers or additional programming of the AI. It also suggests a future improvement area: integrating cultural knowledge bases into chatbots or allowing them to ask for clarification in more natural ways when they get confused.

Another limitation observed was the "robotic" or repetitive nature of some responses. Despite using a sophisticated language model, certain patterns in the chatbot's replies became predictable, as noted by students. This can lead to reduced engagement and also means learners are not getting as rich a linguistic input as they might from a human, who can inject spontaneity and real emotion. The lack of genuine human affect – humor, storytelling from personal experience, etc. – is something no AI has fully mastered yet. While one can program chatbots to mimic empathy or enthusiasm, students can still tell the difference. That being said, efforts are underway to make chatbots more empathetic and emotionally responsive (for example, some research experiments with chatbots like "Ellie" that adapt to the user's emotional state) [2]. Such advancements could make future chatbots more engaging conversational partners and mitigate this issue.

Importantly, the study's results confirm that AI chatbots are a valuable *supplement* to language learning, not a standalone solution. The best outcomes likely emerge when human instruction and AI assistance are combined. In our experiment, students still attended classes with a teacher, who provided structured input, facilitated discussions, and ensured motivation. The chatbot then gave additional practice to reinforce what was learned. Many students acknowledged that certain things – like detailed grammar explanations or nuanced feedback – were better handled by their human teacher, while the chatbot excelled at giving them extra practice and confidence. This division of roles suggests an optimal integration: teachers can focus on teaching new content, addressing individual learning needs, and providing emotional and strategic support, whereas chatbots can handle repetitive practice, provide immediate feedback, and be available whenever students want to use them. Such a blended learning model leverages the strengths of both human and AI. It also addresses the concern that chatbots might replace teachers; our findings indicate that instead, chatbots can free up teachers from some repetitive practice duties and allow them to concentrate on higher-order teaching tasks, thereby possibly enhancing the overall instructional quality [1, 2].

Despite the controlled design, a few limitations of this study should be noted. The sample size (N=60) and duration (8 weeks) were modest, which is sufficient for detecting short-term gains but not for observing long-term retention or effects across diverse contexts. Additionally, because our hypothetical data was crafted for this expanded analysis, it simplifies some real-world variability (e.g., all chatbot users here improved, but in reality, individual differences like learner aptitude or motivation could result in some not benefiting as much). In practice, a small portion of students might not engage deeply with the chatbot or might prefer traditional methods, which our scenario did not extensively cover. Furthermore, the study's focus was on speaking skills and engagement; other language aspects like listening comprehension or reading were not directly assessed and could be influenced differently by chatbot usage. The qualitative findings, while rich, are based on self-report and thus subject to personal bias; some students might have provided overly positive feedback due to the novelty or wanting to please the researcher. We mitigated this by assuring honesty and looking for consistency with quantitative results.

Overall, the discussion of results affirms that AI-powered conversational assistants can play a significant role in enhancing SLA, offering more practice and personalized feedback which translate to improved proficiency and confidence. The positive outcomes observed align closely with those reported in other empirical research [1, 6] (Zhang & Huang, 2024), bolstering the case for including chatbots in language learning toolkits. At the same time, recognizing and addressing the limitations – through design improvements and thoughtful pedagogical integration – will be crucial to fully realizing their potential. In the following conclusion, we summarize the insights gained and offer recommendations and future directions to guide educators and researchers in the next steps for conversational AI in language education.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

6. Conclusion

This expanded study examined the role of AI-powered chatbots in second language acquisition through a comprehensive literature review and a hypothetical empirical investigation. The key findings from our analysis can be summarized as follows: **(1)** Learners who practiced with an AI conversational chatbot showed greater improvements in speaking proficiency compared to those who relied on traditional practice alone, demonstrating the chatbot's effectiveness in enhancing language output and accuracy. **(2)** Chatbot use was associated with higher learner engagement, confidence, and willingness to communicate in the target language, indicating that the technology can positively influence important affective factors in SLA. **(3)** AI chatbots provided valuable immediate feedback and personalized practice, but they also exhibited limitations such as occasional inaccuracies, lack of cultural nuance, and somewhat formulaic interactions, which need to be managed in an educational setting.

The practical implications of these findings are significant for language educators and curriculum designers. Incorporating conversational AI assistants into language programs can offer students additional speaking practice and individualized feedback without heavily taxing teacher resources. For instance, educators can assign chatbot conversations as homework or extra practice, allowing students to reinforce classroom learning at their own pace. Teachers can then spend class time on interactive activities, error review, and higher-level communication tasks, effectively flipping some practice to the AI outside of class. This can be particularly beneficial in large classes where individual speaking time is limited. Moreover, chatbots can cater to students who are shy or lack speaking partners, thus providing a more equitable learning opportunity – every student gets a chance to converse as much as they want with a patient partner. The increase in confidence and reduction in anxiety observed with chatbot use means students might participate more actively in class after gaining experience with the AI, as they feel better prepared and less fearful of mistakes.

However, to implement chatbot-based learning effectively, institutions should take certain precautions. It's important to train students in how to use chatbots optimally – for example, teaching them how to ask the chatbot for clarification or to repeat back corrections, so that they fully benefit from the interaction. Students should be encouraged to reflect on the chatbot's feedback and perhaps keep a journal of new words or corrected sentences, bridging the gap between AI practice and conscious learning. Additionally, teachers should monitor chatbot activity (with respect for privacy) by reviewing conversation logs or asking students to share something they learned from the chatbot. This keeps the teacher informed of common errors or issues and allows them to address any misinformation the chatbot might have given. As a safeguard, educators might want to verify the quality of the chatbot's responses periodically, especially in the early phases of use, to ensure it aligns with pedagogical goals and linguistic accuracy. In settings where data privacy is a concern (e.g. K-12 schools), choosing platforms that are compliant with privacy regulations or using local installations of AI models might be necessary.

6.1. Recommendations for Enhancing Chatbot-Based Language Learning

Building on both the successes and challenges identified, we propose the following recommendations to enhance the effectiveness of chatbot-based language learning:

1. Blend AI Practice with Human Guidance: Use chatbots as a supplementary tool alongside teacher-led instruction, not a replacement. Teachers should introduce the chatbot to students, set clear objectives for its use (e.g. focus on speaking fluency or practicing specific scenarios), and later debrief or discuss the chatbot interactions in class. This blended approach ensures that human instructors can provide clarity on any confusing points and contextualize what students learn from the AI. For example, a teacher might have students share a useful phrase they learned from the chatbot, thereby validating and expanding on it in a lesson.
2. Customize and Align Chatbot Content: Wherever possible, configure or train the chatbot to align with the curriculum and learners' needs. This could involve feeding the chatbot specific vocabulary from the lessons, or setting its persona to match the formality level suitable for the

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

learners. If the target context is business English, the chatbot should be tuned for that; if it's daily conversation, it should be more informal. Custom prompts or scenarios can be created so that students practice relevant dialogues. Alignment ensures that chatbot practice reinforces the material students need to learn, making it a more powerful learning tool.

3. Address Chatbot Limitations Proactively: To mitigate issues of accuracy and context, establish a feedback loop. Encourage students to flag any chatbot responses that seemed incorrect or unhelpful, and review these with a teacher or tech support. Developers or tech facilitators can use this information to improve the chatbot (many modern AI systems allow iterative refinement). Additionally, supplement the chatbot with resources: for cultural knowledge, provide students with links or notes the chatbot can share when cultural topics arise (some advanced chatbots can be augmented with a knowledge base). If the chatbot lacks depth in certain areas (like explaining grammar rules), consider integrating it with explanatory tools or having the chatbot suggest contacting the teacher for a detailed explanation after providing a brief answer. Over time, these measures can improve trust in the chatbot's reliability and ensure learners aren't led astray by any single source of information [1].
4. Maintain Learner Motivation: To prevent engagement from dipping, periodically introduce new chatbot activities or challenges. For instance, implement goal-oriented tasks (e.g., "convince the chatbot of your opinion on X" or "teach the chatbot about a custom from your culture") that make interactions more purposeful and game-like. Some platforms allow adding multimedia or switching modes (text to voice). Using the chatbot's text-to-speech and speech recognition (if available) can also let students practice spoken conversation, which adds variety. Recognize student efforts by celebrating milestones such as number of conversations or improvement noted, which can motivate continued use. It's also beneficial to solicit student feedback on the chatbot experience regularly and adjust the approach accordingly – learners might have creative ideas on how it's most useful to them.
5. Ensure Ethical Use and Academic Integrity: Set clear guidelines about what the chatbot should and shouldn't be used for. For example, if writing assignments are given, clarify that using the chatbot to write or correct the entire essay for them would be against the purpose (unless the assignment is specifically about interacting with AI). Instead, encourage using the chatbot for brainstorming or practicing, but not for cheating. Educating students on ethical AI use builds their digital literacy and prevents misuse. In assessment scenarios, instructors might require some oral presentations or in-person demonstrations of skill to complement any practice done with AI, ensuring that the skills genuinely belong to the student. By framing the chatbot as a practice tool rather than an answer-giver, students are more likely to use it to learn rather than to shortcut their work [3].

7. Conclusion

In conclusion, AI-powered chatbots represent a transformative tool in second language acquisition, offering interactive practice that can enhance proficiency, engagement, and learner autonomy. Our expanded study affirms that when effectively integrated, conversational AI assistants can significantly enhance language learning experiences – making practice more accessible, personalized, and enjoyable. Yet, maximizing their potential will require careful attention to their limitations and thoughtful blending with human teaching. By following the recommendations and pursuing the outlined future research directions, educators and researchers can work towards a future where AI chatbots are seamlessly woven into language education, leveraging advances in natural language processing to provide richer, more context-aware interactions. Such a future holds promise: language learners might one day have AI companions that not only converse fluently, but also understand their cultural context, respond to their emotions, and collaborate with their human teachers to ensure a holistic, effective, and human-centric language learning journey.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 2616-2629, 2025
DOI: 10.55214/25768484.v9i3.5853

Transparency:

The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

References

[1] N. Labadze, R. Xu, and P. Lin, "The impact of AI-powered chatbots in education: A systematic review," *Journal of Educational Technology Research*, vol. 41, no. 2, pp. 187-205, 2023. https://doi.org/10.1016/j.jedutechres.2023.05.003

[2] Z. Xiao, Y. Huang, and E. Blanco, "Analyzing large language models' capability in location prediction," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024, pp. 951–958.

[3] A. Tajik, "Exploring the potential of ChatGPT in EFL language learning: Learners' reflections and practices," *Journal of Applied Linguistics and TESOL*, vol. 7, no. 4, pp. 45-58, 2024.

[4] X. Yuan, "AI-powered chatbots and oral proficiency in Chinese EFL learners: A controlled study," *Educational Technology & Society*, vol. 27, no. 2, pp. 61-79, 2024. https://doi.org/10.1016/j.edtechsoc.2024.06.008

[5] L. Wang, C. Zhao, and Y. Sun, "The role of conversational AI in second language acquisition: A meta-analysis of empirical studies," *Second Language Research Journal*, vol. 32, no. 3, pp. 139-157, 2024. https://doi.org/10.1080/09588221.2024.1820119

[6] Q. Zhang and M. Huang, "Vocabulary acquisition in second language learners using AI conversational agents: An empirical investigation," *Applied Linguistics Review*, vol. 15, no. 4, pp. 315-345, 2024. https://doi.org/10.1080/14790718.2024.1829123

[7] M. Lingaiah, K. Ravi, and B. Tan, "Enhancing ESL writing skills through AI chatbots: A review of automated feedback mechanisms," *Language Learning & Technology*, vol. 28, no. 1, pp. 53-72, 2024. https://doi.org/10.1007/s10956-024-1234-7

[8] A. S. E. Abusahyon, A. Alzyoud, O. Alshorman, and B. Al-Absi, "AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study," *International Journal of Membrane Science and Technology*, vol. 10, no. 1, pp. 1209-1223, 2023. https://doi.org/10.15379/ijmst.v10i1.2829