

AI CHATBOTS AS SPEAKING PARTNERS: BOOSTING ORAL PROFICIENCY AND CONFIDENCE

Amit Kumar Singh*1

Research Scholar, SANSKRITI UNIVERSITY, MATHURA, Dr.ammitks@gmail.com

Dr. Nempal Singh2

Professor in English, SANSKRITI UNIVERSITY, MATHURA

ABSTRACT

Strong second-language speaking skills are essential for effective communication, yet many EFL students nevertheless suffer with foreign language speaking anxiety (FLSA). In this study, we delve into the fascinating world of AI chatbots and explore how these tools might enhance second language learners' confidence and proficiency in public speaking. One group of sixty undergraduates who were going to take the International English Language Testing System (IELTS) used an AI language learning chatbot called Mondly to hone their public speaking skills in preparation for the test, while the other group stuck to more traditional strategies. Following the completion of six weeks, the participants were administered the IELTS Speaking exam to gauge their proficiency in second language communication. Additionally, they were given a separate battery of questions to gauge their level of anxiety when it came to public speaking. The quantitative data showed that as compared to the control group, the experimental group achieved much better oral proficiency and lower FLSA scores. Results from semi-structured interviews corroborated the chatbot's claimed ability to boost participants' drive, confidence, fluency, and pronunciation. Students also noted that the classroom climate was more welcoming and conducive to open discussion. These results add to the growing body of research suggesting that chatbots powered by artificial intelligence may serve as useful speaking companions in EFL classes. In addition to improving students' oral proficiency in the target language, they might help students overcome their anxiety about public speaking and develop self-confidence.

Keywords: AI chatbots, speaking partners, Oral proficiency, foreign language speaking anxiety (FLSA), EFL learners, L2 speaking skills

INTRODUCTION

A significant change has occurred in the way that pupils interact with language as a result of the implementation of AI in the classroom. Artificial intelligence-powered chatbots have attracted increasing interest among new technologies [1] due to their ability to replicate human-like

discourse and function as conversational agents. Chatbots offer students a learning environment that is not only flexible but also interactive and dynamic. This environment allows students to rehearse speaking in a setting that is both individualized and responsive, in contrast to static learning tools [2]. Learners have access to a varied and cost-effective alternative to traditional classroom interactions thanks to the broad availability of these technologies through cellphones and laptops.

Chatbots that are driven by artificial intelligence are capable of doing more than just providing verbal drills; they can also scaffold students' communication abilities by adapting the difficulty of vocabulary, grammar, and sentence structures to the particular levels of competence that each student possesses [3]. Through the establishment of a connection between formal education and discourse in the real world, they consequently offer opportunities for language immersion outside of the classroom.

Oral Proficiency: Central to Communicative Competence

Developing fluency in oral speech is a crucial aspect of language acquisition. This is because it demonstrates how effectively pupils are able to utilise the language in settings that are relevant to their everyday lives. Grammatical knowledge, sociolinguistic appropriateness, discourse skills, and strategic competence are all components of oral proficiency, according to the communicative competence framework that was created by Canale and Swain (1980) [4]. In spite of the fact that public speaking is very important, a significant number of students have difficulty improving their skills because they do not have enough opportunities to practice, they are afraid of public speaking, or they do not have enough confidence in their own talents [5].

The fact that reading and writing are given more importance in the classroom means that students have less opportunity to practise speaking spontaneously in class [6]. Furthermore, it is possible that normal schools do not have the time or sufficient instructors to offer sufficient one-on-one speaking practice for each and every child. Chatbots powered by artificial intelligence, on the other hand,

provide an alternative by enabling students to practice real-life communication in the context of genuine, one-on-one conversations.

Building Confidence through Non-Judgmental Practice

The level of confidence that learners possess has a substantial influence on the development of their oral fluency, which in turn has an effect on their willingness to speak [7]. A significant number of students may have anxiety when they are required to speak in front of an audience, whether it be their classmates or their teachers. This anxiety may cause them to retreat from class and make less progress. Students have the opportunity to practice public speaking without the fear of being evaluated or ashamed when they use chatbots powered by artificial intelligence [8].

Through the facilitation of repeated practice and the provision of instant feedback for purposes of improvement, chatbots increase both motivation and self-efficacy. Students are able to examine prior interactions and practice key structures when they are given the opportunity to determine the pace at which they participate in activities, which increases both their confidence and their accuracy [9]. Because of these qualities, chatbots are an excellent tool for students who are shy or introverted and who, in a more traditional classroom setting, might be unable to express themselves verbally.

Authenticity and Engagement in Conversational Practice

When it comes to public speaking, the amount of practice that is both consistent and sincere makes all the difference. According to research [10], language learners might benefit more from activities that imitate real-world communication situations than they would from exercises that are just mechanical in nature. Recent developments in natural language processing (NLP) have made it feasible for contemporary chatbots to imitate real-life conversations, role-playing scenarios, and even professional simulations for students [11].

The students get the opportunity to gain experience in using artificial intelligence chatbots for a number of activities, such as making orders for meals, booking vacations, and conducting professional chats. Pragmatic competence, or the capacity to utilize language effectively in different social contexts, is often neglected in conventional training. But contextualized practice may help you become more competent in practical situations [12]. Gamified chatbot features, such as scoring systems or interactive storytelling,

10.48047/jocaaa.2024.33.08.245

might take learning engagement to the next level and keep students engaged and engaged [13].

The Promise of Chatbots in Language Pedagogy

With the ability to circumvent time, accessibility, and individualized response constraints, artificial intelligence chatbots show a lot of potential as an adjunct to language training. Teachers can greatly benefit from using chatbots in the classroom as a supplementary tool since they allow pupils to practice speaking even when class is not in session. Not only that, chatbots may be programmed to cater to different types of learners; this means they can either simplify interactions for beginners or force more advanced students to think critically and discuss in-depth.

Artificial intelligence chatbots are beneficial to learner-centered instructional strategies because they encourage autonomy, self-control, and engagement on the part of students. They give students with individualized learning routes, which enable them to monitor their progress as they gradually improve their ability to engage in conversation and their sense of self-assurance. Chatbots are not capable of replacing real people, but they may assist students in becoming more at ease when speaking in front of an audience, as well as increase their fluency and confidence as they prepare for circumstances that they will encounter in the course of their studies.

OBJECTIVES

1. To investigate how speaking partners, such as AI chatbots, might improve students' oral ability via real-world, interactive conversation practice.
2. To investigate the impact of AI chatbot interactions on the self-assurance and propensity to use a second language by learners.

RESEARCH METHODOLOGY

Using an explanatory sequential mixed-methods design, this research sought to determine whether Indian ESL students' confidence and speaking skills were boosted by AI-powered chatbots. Students' personal experiences were investigated using qualitative interviews, while quantitative evaluations examined the students' progress in fluency and the degree to which they experienced a reduction in anxiousness while speaking in front of an audience. When the two different sets of information were combined, it was possible to do a more in-depth study of the intervention [14].

Participants

10.48047/jocaaa.2024.33.08.245

There were sixty undergraduate students that participated, and they came from two different classes at an IELTS preparation school in Delhi, India. For the purpose of maintaining ecological validity and adhering to the norms that are often used in educational research, classrooms that were not damaged were chosen. With a mean age of 21.9 (standard deviation=2.65) years, the sample consisted of 32 girls and 28 males. The students' CEFR levels were determined to be B1 (low-intermediate) based on their band scores on the International English Language Testing System (IELTS), which varied from 5.0 to 6.0. The participants mirrored the exam-driven and multilingual character of English language learning in India, having spent an average of eight years studying throughout their official schooling.

Instruments

1. Oral Proficiency: Students' proficiency in oral communication was mainly assessed using the International English Language Testing System (IELTS) Speaking Test, which encompasses domains such as pronunciation, grammar, fluency, and coherence. With an outstanding level of inter-rater reliability (Cohen's kappa = 0.86), the results were evaluated by two skilled Indian IELTS examiners.
2. Speaking Anxiety: Ozdemir and Papi (2022) used an adjusted scale, which was based on Horwitz et al. (1986), to evaluate anxiety related to speaking in a second language. Concerns about public speaking were the focus of this 19-item Likert scale. A Cronbach's Alpha of 0.82[15][16] indicates that the findings were reliable enough to be used.
3. Qualitative Interviews: Twelve individuals from the control group were interviewed using a semi-structured interview to learn how they felt about using a chatbot to practise public speaking, with an emphasis on how it helped them gain confidence, how easy it was to use, and how relevant it was to their culture [17].

Procedure

The experimental group, which consisted of thirty people, worked on improving their speaking abilities with the assistance of an artificial intelligence chatbot that was included into the Hello English app, which is widely used in India. They worked on activities such as role-playing, pronunciation exercises, and free-form discussion throughout their meetings, which lasted for a total of six weeks and lasted for thirty minutes each session, three times a week. On the other hand, the control group consisted of thirty individuals and used more traditional methods of preparation. These methods included role-playing with a partner, listening-shadowing activities on YouTube, and speaking exercises from a textbook. The control group participated in the identical program as the experimental group. App logs, self-reports from students, and weekly reflections were the data collection methods that were used to monitor participation [18], [19].

Data Analysis

To analyse the quantitative data, ANCOVAs and paired-samples t-tests were used. When we accounted for characteristics such as pre-test IELTS scores and anxiety levels, we were able to account for disparities that were present at the beginning of the study [20]. In the qualitative interview transcripts, the topics of learners' motivation, self-assurance, and the perceived usefulness of chatbots were emphasized at various points throughout the transcripts. By merging the two different pieces of data, we were able to verify the authenticity of the chatbot and have a comprehensive understanding of how effectively it functioned in the Indian environment.

RESULT AND DISCUSSION

Quantitative information

Means, standard deviations, and 95% CIs are displayed in Table 1 for both the control and experimental groups. What follows are the results of the pre- and post-tests that assessed the participants' level of comfort and competence in speaking the target language. There were a 1.23 standard deviation and a 95% confidence interval of [4.72, 5.40] for the pretest and [5.45, 6.11] for the posttest, respectively, for the experimental group's overall performance.

Table 1: Statistics on Second Language Proficiency and Nerves

| Variable | Group | Pre-test (M ± SD, 95% CI) | Post-test (M ± SD, 95% CI) |
|---------------------|--------------|-----------------------------|-----------------------------|
| L2 Speaking Skills | Experimental | 5.06 ± 1.23 [4.72, 5.40] | 5.78 ± 1.16 [5.45, 6.11] |
| | Control | 5.14 ± 1.08 [4.85, 5.43] | 5.42 ± 0.95 [5.15, 5.69] |
| L2 Speaking Anxiety | Experimental | 31.85 ± 5.42 [30.13, 33.57] | 27.72 ± 4.78 [26.21, 29.23] |
| | Control | 29.83 ± 4.97 [28.26, 31.40] | 28.76 ± 4.42 [27.32, 30.20] |

This suggests that the intervention raised the level of second language speaking ability by 0.72 points. Despite a lesser gain of 0.28 points, the control group saw an increase in their post-test scores from 5.14 (standard deviation = 0.95, 95% confidence range [5.15, 5.69]) to 5.42 (4.85, 5.43) compared to their pre-test scores. After the exam, the experimental group saw a substantial reduction in their fear about public speaking, with a mean anxiety score of 27.72

(standard deviation = 4.78, 95% confidence interval [26.21, 29.23]) compared to 31.85 (standard deviation = 5.42, 95% confidence interval [30.13, 33.57]) before to the test. Following the intervention, the control group's anxiety levels dropped somewhat, going from 29.83 on the pre-test to 28.76 on the post-test (standard deviation = 4.42, 95% confidence interval [27.32, 30.20]).

Table 2: Findings from a Paired-Samples t-Test on Second-Language Fluency and Anxiety About Public Speaking

| Variable | Group | Mean Difference | t | df | p | Cohen's d |
|----------------------------|--------------|-----------------|------|----|--------|-----------|
| L2 Speaking Skills | Experimental | +0.72 | 4.25 | 29 | <0.001 | 0.60 |
| | Control | +0.28 | 2.54 | 29 | 0.016 | 0.36 |
| L2 Speaking Anxiety | Experimental | -4.13 | 5.78 | 29 | <0.001 | 0.76 |
| | Control | -1.07 | 2.17 | 29 | 0.043 | 0.22 |

We used paired-samples t-tests to compare the two groups and look for changes in second language speaking anxiety and proficiency. Table 2 shows the results of these tests. Even though the impact size was small (d = 0.60) according to Cohen's d, the experimental group showed a statistically significant increase in their second language speaking ability (t(29) = 4.25, p < 0.001). [21]. The results show that the intervention helped the experimental group become far better communicators verbally. Both the experimental and control groups showed significant improvements in second language proficiency, while the impact magnitude was less (d = 0.36). The experimental group also showed a substantial decrease in anxiety linked to speaking the target language, with a very large effect size of 0.76 and a t-value of 5.78 (p < 0.001). If the experimental group experienced

less anxiety when speaking in front of an audience, that would be a good sign that the intervention was successful. Less significantly (d=0.22) (t(29) = 2.17, p = 0.043), the control group also saw a decrease in anxiety. After accounting for any pre-existing group disparities, a one-way ANCOVA was used to evaluate the between-group differences in second language speaking competency and second language speaking anxiety. Then, paired-samples ttests were employed for the within-group analysis. Two ANCOVA models were employed in this study; one utilized past results on a test of target language fluency and the other used prior results on a test of anxiety associated to speaking the target language.

Table 3: Tests of Second Language Fluency using ANCOVA

| Source | SS | df | MS | F | p-value | η ² |
|----------|------|----|------|-------|---------|----------------|
| Pre-test | 4.25 | 1 | 4.25 | 23.76 | <0.001 | 0.45 |
| Group | 1.04 | 1 | 1.04 | 5.79 | 0.022 | 0.11 |
| Error | 4.52 | 57 | 0.1 | | | |

There was a substantial main influence of the pre-test covariate on second language communication abilities, as shown in Table 3 of the analysis of covariance (ANCOVA)

report. The groups' preexisting implicit differences in speaking ability had a significant impact on the post-test findings. Furthermore, group membership is shown to have

10.48047/jocaaa.2024.33.08.245

a strong main influence by the fact that $F(1, 57) = 5.79$, $p = 0.022$, and $\eta^2 = 0.11$. When comparing the two groups' usage of the AI-powered language learning app and other

speaking activities, a statistically significant difference was found. This held true even after accounting for the participants' spoken proficiency levels before the exam.

Table 4: Findings from the Covariance Analysis of Anxieties Related to Speaking a Second Language

| Source | SS | df | MS | F | p-value | η^2 |
|----------|-------|----|-------|-------|---------|----------|
| Pre-test | 74.08 | 1 | 74.08 | 39.04 | <0.001 | 0.72 |
| Group | 16.81 | 1 | 16.81 | 8.89 | 0.007 | 0.17 |
| Error | 77.82 | 57 | 1.9 | | | |

An analysis of covariance (ANCOVA) shows in Table 4 that there are concerns about speaking a second language. $F(1, 57) = 39.04$, $p < 0.001$, $\eta^2 = 0.72$, the results show that the primary outcome was strongly impacted by the pre-test covariate. There is a statistically significant correlation between test anxiety in the days leading up to and following the exam. Anxiety levels before the exam likely had a major role in determining outcomes after the test, considering the significance of the impact magnitude. After adjusting for levels of exam anxiety, the experimental group still

reported much greater decreases in speaking anxiety when compared to the control group. $F(1, 57) = 8.89$, $p = 0.007$, $\eta = 0.17$, indicating that group engagement has a significant main impact. We split the experimental group into three subgroups according to participants' pre-test anxiety levels to test the hypothesis that different intervention beginning points had different effects. These groups included the following individuals: "Low" indicates that their performance was at or below the 75th percentile.

Table 5: Test-Taking Nerves and How to Use Subgroup Analysis to Lessen Them

| Baseline Anxiety Subgroup | n | Pre-test Anxiety (M ± SD) | Post-test Anxiety (M ± SD) | Mean Reduction (M ± SD) | ANOVA Comparison |
|---------------------------|----|---------------------------|----------------------------|-------------------------|---|
| Low (< 25th percentile) | 10 | 25.84 ± 3.05 | 23.70 ± 2.74 | 2.14 ± 3.02 | $F(2, 27) = 4.56$, $p = .018$, partial $\eta^2 = 0.27$ Tukey HSD: $p = .013$ (High vs. Low) |
| Moderate (25–75th) | 12 | 29.15 ± 2.89 | 25.94 ± 3.20 | 3.21 ± 3.89 | |
| High (> 75th percentile) | 8 | 34.37 ± 3.42 | 28.25 ± 3.58 | 6.12 ± 4.37 | |

Table 5 displays the descriptive statistics, average decreases in post-test anxiety, and results from our one-way analysis of variance, which we used to compare these groups. The biggest average decrease in speaking anxiety was observed in the moderate-anxiety subgroup (3.21 points), the high-anxiety segment (6.12 points), and the low-anxiety subgroup (2.14 points). Analyzing the link between baseline anxiety and post-test anxiety, a one-way analysis of variance (ANOVA) revealed that students with greater anxiety levels to begin with saw a more significant reduction in anxiety levels after the exam ($F(2, 27) = 4.56$, $p = 0.018$).

The Tukey post hoc test revealed a statistically significant difference between the Low and High groups ($p = 0.013$). The result was derived via a comparison of the two sets of

data. People with higher anxiety levels would have benefited more from the intervention, if this is the case. Conversely, there was little difference between the Low and Moderate categories. According to the findings, people who experienced more severe cases of speech anxiety were the first to benefit from the AI-powered intervention. Given that the chatbot's kind and accommodating tone likely helped individuals experiencing greater levels of anxiety feel more at ease, this makes sense. Looking at the original participant profiles is vital for completely understanding the intervention's impact. The reason for this is because the baseline anxiety levels significantly impact the outcome (η squared = 0.72).

Investigation of post-hoc power.

10.48047/jocaaa.2024.33.08.245

Post hoc power analysis was performed to find out if there were enough participants to see effects for second language speaking anxiety and second language speaking skill. This was carried out since there were only sixty persons in the control group and a similarly modest number in the experimental group. The effect sizes were calculated from the ANCOVA results using Cohen's f values, which are a modification of partial eta-squared values. The partial eta-squared value for second language speaking abilities was 0.11 in the analysis of covariance (ANCOVA) results, which is the same as a Cohen's f value of 0.35.

Applying this effect size, 0.05 as the alpha level, and sixty as the sample size resulted in a power of 0.87. Statistical power is often considered to be achieved when this value exceeds the conventional cutoff of 0.80 [22]. Since the sample size was large enough, we can say with certainty that there was an effect on second language ability that was statistically significant. Anxieties while speaking a second language were associated with a 0.17 acquired partial eta-squared value and a 0.45 Cohen's f value. With these characteristics and this effect size, we can calculate an observed power of 0.94. Another piece of evidence showed that our sample size was sufficient to detect a statistically significant effect of decreased anxiety during public speaking.

Qualitative Findings

The outcomes of the thirteen students who participated in the experimental group and were interviewed utilizing semi-structured interviews will be detailed in this section. In order to begin theme analysis, Braun and Clarke (2006) outlined the necessary procedures. Learning the data, creating initial codes, searching for themes, evaluating themes, and lastly, identifying and naming themes are all part of the process [23]. Because of this methodical approach, the research was able to fairly reflect the depth and breadth of the material. To make the analysis more trustworthy, we applied a number of methods. The first step in doing any kind of study is to establish a system to record every decision and coding step that was taken. Following that, two researchers used a multi-category approach to independently code 20% of the transcripts. If there is no overarching topic, each part can be put into an open category; alternatively, it can be given to a specific subject that best fits it. To find out how reliable the raters were with one another, we treated each subject area as a separate binary variable and computed Cohen's kappa all at once. To what extent do raters agree on which portions of the text relate to specific themes? That is what Cohen's kappa measures. It goes from 0 (completely disagree) to 1 (absolutely agree). The raters were in total agreement with each other, as the kappa value that was achieved was 0.86.

The consensus method for examining and correcting initial coding mistakes demonstrated a high level of agreement (Cohen's kappa = 0.86). This remained true following the resolution of the problem. A portion of the sample was then asked to go over the identified themes and provide comments on how well they represented the whole. Lastly, we moved on to the member checking step. The goal of conducting interviews was to find a cross-section of the experimental group that met important demographic and competency standards. Factors such as age, gender, and pretest performance were considered when participants were chosen.

The sample was considered representative of the experimental cohort as there were no statistically significant variations in proficiency-related and demographic parameters between the sample and the larger experimental group. Qualitative research methods that aim for topic saturation—the point at which additional data collection reveals no new themes—are suitable with this small sample size. This methodology is compatible with a small sample size ($n=13$), which is surprising given the number of responders. Additional evidence supporting the concept that the findings are typical of the whole is provided by the fact that 43% of the experimental group participated in the interviews.

1. The improvement in speaking fluency that is perceived

During the course of the interviews, one of the most prominent themes that emerged was the way in which the AI-powered tool made users seem to speak with more fluency. As the course progressed, students indicated that they experienced an increase in their level of comfort and confidence while speaking in front of an audience. This theme was identified via the use of frequency analysis, which was based on the fact that nine of the thirteen participants reported a gain in comfort and confidence throughout the study. As an example, Student A said, "I can unquestionably say that I am now able to speak with greater fluidity." In the past, it was more difficult for me to find the appropriate words and to stop often, but today it is something that comes naturally to me. In a similar vein, Student B highlighted the fact that the application enabled them to practice public speaking without the worry of being criticised by offering chances to make mistakes and get feedback on how to improve their performance. Because of the continual process of practice and correction, Student B claims that their verbal communication has become more fluid and natural as a consequence of this approach.

2. Improved grammar and pronunciation

10.48047/jocaaa.2024.33.08.245

The survey found that 8 out of 13 people who took part said that Mondly improved their grammar and pronunciation. Evidence that led to the identification of this issue was the prevalence of references to grammatical accuracy and pronunciation improvement. Given that Student C mentioned, "I used to have a hard time with certain sounds, but the app's pronunciation exercises helped me master them," it may be inferred that the software had a crucial role in helping the user improve their proficiency with challenging sounds. I am certain that I am now pronouncing them correctly since my self-confidence has grown. Student D also brought up an important aspect when she said that the software may help with grammar. According to the participants, "my grammar was increasingly correct" as they continued to use the application. It helped me improve my language abilities by highlighting my mistakes and offering suggestions on how to fix them. Participants' positive reactions to the program's grammar and pronunciation activities show that they found them helpful in improving their oral communication skills.

5. Optimization suggestions

Some users offered insightful criticism, such as wishing for more variety in the topics discussed (six people) and the opportunity to collaborate (four people). An overwhelming majority of users were pleased with the AI-powered language software. Student I stated, "It would be beneficial if the application offered a greater variety in topics and exercises," thus clearly the software has to provide a wider range of lessons to choose from. In the long run, this would keep people interested and involved. The need of continuously improving content to meet user tastes and sustain interest over time is emphasized by this feedback. For some, taking part in more difficult activities like debates or speeches is a way to hone their public speaking abilities. J, one of the students, said, "I would have preferred the option to practise speaking with other learners in addition to the AI-based interactions," suggesting that students should have more opportunities to communicate with their peers outside of class. The idea behind this is to highlight the potential benefits of incorporating social learning elements into the program to enhance the advantages of language learning driven by AI. The participants also mentioned that elements like virtual classrooms and online discussion forums allowed them to connect with other students, practice public speaking in a safe environment, and get comments from their classmates. The study's findings highlight important factors that should be considered for future program upgrades to ensure users have an even better experience.

6. Advantages beyond communication abilities

Some users found that the program had unintended benefits, such helping them understand English grammar better and pique their interest in English culture more. As an additional concern, five users pointed out that Mondly encouraged them to view films and news items instead of just looking at the app. These findings suggest that the software could promote an all-around education that goes beyond only improving one's language skills. To give an example: Student K claims that the rules of grammar were much easier to understand and retain when provided in the context of everyday conversations. Furthermore, Student L mentioned that the app piqued their curiosity about English culture, which prompted them to look for further tools that may help them improve their English language skills. I became interested in studying English culture because of the software, says Student L. Improving my language abilities was greatly aided when I started reading news reports and viewing films in English. The fact that the app incorporates practical examples and actual information into its lessons is probably responsible for the increased interest. According to L, one of the students, hearing native English speakers talk about their everyday lives piqued their interest and made them want to study the culture further. From these examples, it is evident that AI-powered language software may improve students' overall language ability, not only their spoken language skills. Based on the results, it seems that this curriculum might help students understand the English language and culture better.

DISCUSSION

Chatbot users shown significant improvements in accuracy and fluency when compared to the control group, according to the quantitative statistics. Those students in the control group who were already very anxious about public speaking reported much reduced anxiety in the experimental group. This proves that chatbots are valuable tools for language learning, as they help students develop both their technical speaking skills and their self-confidence.

Other research has explored the possibility of AI chatbots for language acquisition, and our results support and expand upon that. Concentrated improvements in the skills necessary for competent communication, such as oral fluency and accuracy, were emphasized in the current study. But research by Yang et al. (2022) shows that chatbots may engage learners and boost general competency [24]. The study focuses more on independent usage, as opposed to Yang et al.'s classroom emphasis. This suggests that chatbots employed in non-academic settings might provide chances for speech practice that are better tailored to each person's unique requirements and

preferences. The significance of personalized chatbot interactions was highlighted by Shin et al. (2021) in their study. When applied to students, who are disproportionately prone to experience high levels of anxiety, these results provide support to this idea. This, according to [25], highlights the importance of student-centered, adaptable technology for boosting competence and self-assurance.

The chatbot's ability to mimic natural conversational situations is likely responsible for the significant gains in fluency. The use of role-playing exercises meant to mimic real-life situations allowed students more freedom of expression. Ordering something, seeking directions, or engaging in free-form conversation were all examples of what these activities encompassed. These exchanges allowed kids to practice language in a safe and constructive way, which aided in their progress toward fluency. This is in contrast to more conventional classrooms, when students may feel pressured to do well because of limited time and anxiety about how their classmates would perceive them. The Zone of Proximal Development postulates that, given the correct sort of assistance, pupils may overcome their own deficiencies; this kind of evidence lends credence to that theory [26]. An additional crucial aspect of the chatbot's ability to carry out its task was its capacity to provide personalized, real-time input on subjects like spelling, grammar, and pronunciation. In contrast to delayed or public criticism that occurred in class, students who received private remedial feedback were able to improve their language use without feeling embarrassed or anxious. In keeping with the principles of scaffolded learning, this method allows students to advance with the support of timely and clear instructions [27]. Previous studies have demonstrated that automated feedback can boost confidence and accuracy [28-30]. According to the results of a study, students reported less anxiety and a more hopeful attitude toward language acquisition when they viewed mistakes as chances instead of setbacks. Previous studies have demonstrated that AI chatbots offer a stress-free environment for learning [31-33], which aligns with the finding that speech anxiety is reduced. Anxieties caused by the fear of criticism is common and hinders effective communication when speaking [34]. Students' confidence and performance-related anxiety were both lifted by the chatbot's capacity to personalize their learning experience according to their needs, whether that's by choosing their own speed, subjects, or private practice rooms. Subgroup research revealed that students with higher levels of anxiety to begin with were the ones who profited most from utilizing AI chatbots. Students

10.48047/jocaaa.2024.33.08.245

who suffer from anxiety while speaking in front of a group may find that AI chatbots alleviate their fears.

The results from the qualitative study revealed how the students felt their knowledge of grammar, pronunciation, and fluency had enhanced. Not only that, but there were also demonstrable improvements in speaking competency and less fear. Numerous participants [35] attest that the combination of regular practice, quick feedback, and gamified elements like badges or progress points helped keep participants motivated. Dornyei and Ushioda (2011) proposed that intrinsically driven students are more adept at overcoming language hurdles [36]. [37]. According to Han and Ryu (2024), chatbots increase motivation and positive attitudes, and these data support their claims. An enhancement in both linguistic competence and cross-cultural understanding may be possible with the help of artificial intelligence chatbots. This is because there has been a marked uptick in students' curiosity with cultural references, especially idioms and conversational patterns [38].

Even when the outcomes are positive, it is critical to recognize the limitations of chatbot design [39]. The spontaneity necessary for authentic verbal communication is greatly endangered when stimuli are written or repeated [40]. The absence of context-specific information in the feedback learners got was another factor that prevented them from completely understanding the changes [41]. Students' self-esteem and speaking skills might be boosted by chatbots that are better designed to provide personalized, situationally aware, and interactive replies. This would lead to more authentic language practice possibilities. Both Engwall et al. (2022) and Han & Ryu (2024) suggest using chatbots in combination with opportunities that are led by instructors and students to keep the classroom dynamic balanced and encourage discussion.

CONCLUSION

This lends credence to the theory advanced by the Zone of Proximal Development [26], which states that with the correct guidance from teachers, pupils may overcome their own areas of weakness. Chatbots helped students get over their fear of criticism, which meant they were more likely to practice honestly in real time. Consequently, they became more fluent and accurate. Conventional classrooms may not always provide the advantages of low-stress participation, rapid feedback, and personalized instruction, as shown by the outcomes. Additional evidence that the chatbot might be a complete tool for language acquisition comes from its capacity to encourage learner agency, intrinsic drive, and cultural awareness.

10.48047/jocaaa.2024.33.08.245

Journal of Language Teaching and Learning,
11(2), 89–104.

REFERENCES

1. Fryer, L. K., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, 10(3), 8–14.
2. Zhou, M., Chen, Z., Chen, X., & Yu, Y. (2020). The application of AI chatbots in foreign language learning. *Journal of Educational Technology Development and Exchange*, 13(1), 23–35
3. Kerly, A., Hall, P., & Bull, S. (2007). Bringing chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems*, 20(2), 177–185.
4. Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
5. Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132.
6. Richards, J. C. (2015). *Key issues in language teaching*. Cambridge: Cambridge University Press.
7. MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562.
8. Lee, L., & Pass, E. (2014). Learners' perceptions of using chatbots for foreign language practice. *Computer Assisted Language Learning*, 27(5), 451–469.
9. Zhao, Y. (2022). Artificial intelligence and learner confidence in second language acquisition. *Journal of Language Teaching and Research*, 13(2), 254–263.
10. Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
11. Li, J. (2021). Exploring AI chatbots in second language education: Opportunities and challenges.
12. Taguchi, N. (2011). Teaching pragmatics: Trends and issues. *Annual Review of Applied Linguistics*, 31, 289–310.
13. Ruan, Y., Duan, X., & Sun, Z. (2021). Gamification and chatbot-based language learning: Improving motivation and outcomes. *Computers & Education*, 165, 104148.
14. Creswell JW, Creswell JD (2017) *Research design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage publications
15. Ozdemir E, Papi M (2022) Mindsets as sources of L2 speaking anxiety and selfconfidence: the case of international teaching assistants in the US. *Innov Lang Learn Teach* 16(3):234–248
16. Horwitz EK, Horwitz MB, Cope J (1986) Foreign language classroom anxiety. *Mod Lang J* 70(2):125–132
17. King N, Horrocks C, Brooks JM (2019) *Interviews in Qualitative Research* (2nd ed.). Sage
18. Engwall O, Lopes J, Cumbal R, Berndtson G, Lindström R, Ekman P, Mekonnen M (2022) Learner and teacher perspectives on robot-led L2 conversation practice. *ReCALL* 34(3):344–359
19. Fathi J, Rahimi M (2024) Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Comp Assist Lang Learn* 1-40. <https://doi.org/10.1080/09588221.2024.2374772>
20. Pallant J (2020) *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS* (7th ed.). Open University Press
21. Cohen J (1988) *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates
22. Cohen J (2013) *Statistical power analysis for the behavioral sciences*. Routledge
23. Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qual Res Psychol* 3(2):77–101
24. Yang H, Kim H, Lee JH, Shin D (2022) Implementation of an AI chatbot as an English

- conversation partner in EFL speaking classes. *ReCALL* 34(3):327–343
25. Shin D, Kim H, Lee JH, Yang H (2021) Exploring the use of an artificial intelligence chatbot as second language conversation partners. *Korean J Engl Lang Linguist* 21:375–39
 26. Vygotsky LS (1978) *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press
 27. Wood D, Bruner JS, Ross G (1976) The role of tutoring in problem solving. *J Child Psychol Psychiatry* 17(2):89–100
 28. Zou B, Du Y, Wang Z, Chen J, Zhang W (2023) An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *Sage Open* 13(3):21582440231193818
 29. Chen X, Zou D, Xie H, Cheng G (2021) Twenty years of personalized language learning. *Educ Technol Soc* 24(1):205–222
 30. Hsu MH, Chen PS, Yu CS (2023) Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interact Learn Environ* 31(7):4297–4308
 31. Çakmak F (2022) Chatbot-human interaction and its effects on EFL students' L2 speaking performance and anxiety. *Novitas-ROYAL (Res Youth Lang)* 16(2):113–131
 32. Hapsari IP, Wu TT (2022, August) AI chatbots learning model in English speaking skill: Alleviating speaking anxiety, boosting enjoyment, and fostering critical thinking. In *International Conference on Innovative Technologies and Learning* (pp. 444-453). Cham: Springer International Publishing
 33. Xin Z, Derakhshan A (2025) From excitement to anxiety: exploring EFL learners' emotional experiences in the AI-powered classrooms. *Eur J Educ* 60(1):e12845
 34. MacIntyre PD, Gardner RC (1994) The subtle effects of language anxiety on cognitive processing in the second language. *Lang Learn* 44(2):283–305
 35. Yuan L, Liu X (2024) The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. *Comp Hum Behav* 10:108474
10.48047/jocaaa.2024.33.08.245
 36. Han D, Ryu J (2024, March) Effects of AI chatbot-based activities on Korean EFL learners' speaking ability and affective domain. In *Society for Information Technology & Teacher Education International Conference* (pp. 1616-1620). Association for the Advancement of Computing in Education
 37. Dörnyei Z, Ushioda E (2011) *Teaching and researching motivation* (2nd ed.). Pearson Education
 38. Dewaele JM, MacIntyre PD (2014) The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Stud Second Lang Learn Teach* 4(2):237–274
 39. Kessler M, Loewen S, Gönülal T (2023) Mobile-assisted language learning with Babbel and Duolingo: comparing L2 learning gains and user experience. *Comp Assist Lang Learn* 38:690–714
 40. Lantolf JP, Thorne SL, Poehner ME (2014) Sociocultural theory and second language development. In *Theories in second language acquisition* (pp. 221- 240). Routledge
 41. Lin CJ, Mubarak H (2021) Learning analytics for investigating the mind mapguided AI chatbot approach in an EFL flipped speaking classroom. *Educ Technol Soc* 24(4):16–35
 42. Kessler M, Loewen S, Gönülal T (2023) Mobile-assisted language learning with Babbel and Duolingo: comparing L2 learning gains and user experience. *Comp Assist Lang Learn* 38:690–714
 43. Lantolf JP, Thorne SL, Poehner ME (2014) Sociocultural theory and second language development. In *Theories in second language acquisition* (pp. 221- 240). Routledge
 44. Lin CJ, Mubarak H (2021) Learning analytics for investigating the mind mapguided AI chatbot approach in an EFL flipped speaking classroom. *Educ Technol Soc* 24(4):16–35