https://journal.ypidathu.or.id/index.php/jssut/
P - ISSN: 3026-5959
E - ISSN: 3026-605X

Language Learning through AI Chatbots: Effectiveness and Cognitive Load Analysis

Muh. Safar[1], **Dini Anggraheni**[2]
[1]Universitas Muhammadiyah Bone, Indonesia
[2]Universitas Semarang, Indonesia

ABSTRACT

Background. The use of AI chatbots in language learning has gained popularity due to their accessibility and ability to provide interactive practice. However, concerns regarding their effectiveness in improving language skills and the potential cognitive load they place on learners remain underexplored. With the increasing integration of AI in education, understanding how these tools impact language acquisition is critical for optimizing learning outcomes.

Purpose. This study aims to evaluate the effectiveness of AI chatbots in language learning and analyze the cognitive load experienced by learners during chatbot interactions.

Method. A mixed-methods approach was used, combining quantitative assessments of language proficiency before and after using an AI chatbot for a three-week period, with qualitative interviews to assess learner perceptions of cognitive load. Participants included 60 university students learning a new language.

Results. The results indicate that AI chatbots significantly improved language skills, particularly in conversational fluency and vocabulary acquisition. However, a moderate level of cognitive load was reported by learners, primarily due to the need to simultaneously engage in real-time conversation and process feedback. While the cognitive load was not overwhelming, it varied based on individual learner characteristics, such as prior language proficiency and familiarity with AI tools.

Conclusion. In conclusion, AI chatbots offer an effective method for enhancing language learning, particularly in improving fluency and vocabulary. However, managing cognitive load is crucial to maximizing their educational potential. Further research is recommended to explore adaptive chatbot designs that can tailor interactions to individual learner needs.

KEYWORDS

AI Chatbots, Cognitive Load, Conversational Fluency, Educational Technology, Language Learning

Citation: Safar, M., & Anggraheni, D. (2024). Language Learning through AI Chatbots: Effectiveness and Cognitive Load Analysis. *Journal of Social Science Utilizing Technology*, *2*(3), 430–445. https://doi.org/10.70177/jssut.v2i3.1346

Correspondence:

Muh. Safar,
safarmuhammad785@gmail.com

Received: September 15, 2024

Accepted: September 17, 2024

Published: September 30, 2024

INTRODUCTION

Artificial Intelligence (AI) is transforming the field of education, offering new tools that can enhance learning experiences across various disciplines (Armaselu, 2024). Language learning, in particular, has seen significant advancements through AI-powered platforms, including chatbots that simulate real-time conversations. These AI chatbots allow learners to practice language skills in an interactive environment without the need for human instructors (Adiguzel dkk., 2023). The accessibility and flexibility of AI chatbots make them an appealing option for language learners,

especially for those seeking to improve conversational fluency and vocabulary retention (Aeni dkk., 2023).

Language acquisition is a complex process that involves the development of multiple skills, including speaking, listening, reading, and writing. Traditional methods of language learning often rely on textbooks, classroom instruction, and occasional conversation practice, which may not provide sufficient real-time interaction (Belda-Medina & Calvo-Ferrer, 2022). AI chatbots, by contrast, offer continuous, on-demand language practice that mimics conversational settings, allowing learners to engage in frequent dialogue and receive instant feedback. This real-time interaction helps bridge the gap between theoretical knowledge and practical language use (Zaphiris P. dkk., 2024a).

AI chatbots leverage natural language processing (NLP) algorithms to engage users in conversations, respond to their inputs, and provide feedback. These chatbots can simulate various conversational scenarios, giving learners the opportunity to practice language in context (T. Wu & Yu, 2024). Research has shown that practicing a language through interaction, rather than rote memorization, leads to better long-term retention and fluency. AI chatbots offer a platform for such interaction, making them a promising tool for language learners (Hwang dkk., 2024).

Despite the growing adoption of AI chatbots in language learning, there are still concerns about their effectiveness compared to traditional methods (Zaphiris P. dkk., 2024b). Some studies suggest that while chatbots can improve certain language skills, such as vocabulary acquisition, they may not be as effective in enhancing deeper linguistic competencies like grammar or syntax. The limited ability of AI to detect and correct more nuanced language errors may hinder comprehensive language development. Additionally, the lack of human emotional cues and contextual understanding in AI chatbots may affect the quality of conversational practice (Kajiwara & Kawabata, 2024).

Another key area of concern is the cognitive load that AI chatbots impose on learners. Cognitive load theory suggests that learning is influenced by the amount of mental effort required to process information (Ivanova dkk., 2024). In the context of language learning, cognitive load can be particularly high, as learners must simultaneously engage in conversation, recall vocabulary, and process grammar rules. The added complexity of interacting with an AI system—navigating its prompts, interpreting feedback, and responding appropriately—could potentially increase cognitive load, making the learning process more challenging (Lee & Huang, 2024).

While AI chatbots have the potential to revolutionize language learning, the extent to which they are effective and the cognitive demands they place on learners remain areas that require further exploration (Samala dkk., 2024). Understanding how AI chatbots impact both language acquisition and cognitive load is essential for optimizing their use in educational settings. This study seeks to address these gaps by evaluating the effectiveness of AI chatbots in improving language skills and analyzing the cognitive load experienced by learners during chatbot interactions (Zhang & Huang, 2024).

Although AI chatbots have become increasingly popular in language learning, there is limited research on their overall effectiveness in improving language proficiency. Most studies have focused on specific aspects, such as vocabulary acquisition or fluency, but the broader impact of AI chatbots on comprehensive language development remains unclear (Chetradevee dkk., 2022). How well these tools compare to traditional methods, particularly in areas such as grammar, syntax, and deeper linguistic competencies, has not been thoroughly explored. This leaves a gap in understanding whether AI chatbots can serve as a standalone learning tool or if they need to be supplemented by other instructional methods (Haristiani, 2019).

The cognitive load imposed by AI chatbots is another area that lacks substantial research. Language learning inherently requires high cognitive effort, as learners must process new vocabulary, grammar rules, and sentence structures simultaneously (Hwang dkk., 2024). Interacting with AI chatbots may further increase this load, as learners must also navigate the technology, interpret feedback, and engage in real-time conversation (Belda-Medina & Calvo-Ferrer, 2022). While chatbots offer convenience and flexibility, the cognitive demands they place on learners could potentially hinder the learning process. However, there is insufficient data on how significant this cognitive load is and how it affects different learners.

There is also limited understanding of how learner characteristics, such as prior language proficiency or familiarity with AI technology, influence the effectiveness of chatbot-based language learning. Some learners may find the interactive nature of chatbots beneficial, while others might struggle with the additional cognitive demands (Faruqui dkk., 2024). Without a clear analysis of how these individual factors interact with chatbot use, it is difficult to determine for whom AI chatbots are most effective (Sabeh, 2024). This gap in the research prevents educators from fully optimizing the use of AI in language instruction.

Additionally, the long-term impact of learning through AI chatbots has not been adequately studied. While some short-term gains in fluency or vocabulary have been observed, it is unclear whether these improvements are sustained over time or if learners eventually revert to traditional methods (Krishnam dkk., 2023). The question of whether chatbots promote lasting language retention or if they primarily offer temporary learning boosts needs further investigation. Filling these gaps is crucial to understanding how AI chatbots can be effectively integrated into language learning programs and how they might complement or replace traditional teaching methods (Terzic dkk., 2023).

Filling the gap in understanding the effectiveness of AI chatbots in language learning is essential for optimizing their use in education. With the growing reliance on technology for learning, particularly in remote and self-guided contexts, it is crucial to determine whether chatbots can effectively support comprehensive language acquisition (Kajiwara & Kawabata, 2024). AI chatbots offer interactive, conversational experiences, which are known to enhance engagement and motivation in learners. However, without concrete evidence of their ability to improve language skills across multiple areas—such as grammar, syntax, and fluency—educators cannot confidently integrate them as core tools in language curricula (Misra R. dkk., 2023).

Investigating the cognitive load associated with AI chatbot use is equally important. High cognitive load can interfere with learning by overwhelming the mental capacity of learners, making it difficult to process and retain new information (Neo, 2022). Understanding how chatbot interactions contribute to or alleviate cognitive load will help educators design more effective language learning environments. If cognitive load is too high, learners may struggle to benefit from the technology, but if managed well, chatbots could provide an optimal learning experience that challenges students without overwhelming them (Chan dkk., 2023).

This study aims to evaluate both the effectiveness of AI chatbots in enhancing language learning and the cognitive load they impose on learners. By analyzing improvements in language proficiency and identifying factors that contribute to or reduce cognitive load, the research seeks to offer insights into how AI chatbots can be integrated into language learning programs effectively. The hypothesis is that AI chatbots can improve conversational fluency and vocabulary acquisition while imposing moderate, manageable cognitive load, and that chatbot interactions may need to be adapted based on individual learner needs.

RESEARCH METHODOLOGY

This study employs a mixed-methods research design to assess the effectiveness of AI chatbots in language learning and analyze the cognitive load experienced by learners during chatbot interactions. The quantitative component involves pre- and post-tests to measure language proficiency before and after using the chatbot, while the qualitative component consists of interviews to gather learner perceptions of cognitive load and their overall experiences (H.-S. Wu, 2024). This design allows for a comprehensive evaluation of both measurable outcomes and subjective learner feedback, providing a holistic view of the chatbot's impact.

The population for this study includes university students enrolled in beginner to intermediate language courses. A sample of 60 students was selected, consisting of 30 learners using the AI chatbot as a supplementary tool in their language learning and 30 learners using traditional methods without chatbot assistance (Haristiani, 2019). Participants were chosen using purposive sampling to ensure a diverse range of language backgrounds and levels of familiarity with AI technology. The control group was included to compare the effectiveness of chatbot-based learning against traditional methods.

The instruments used for data collection include a language proficiency test administered before and after the three-week intervention period, a cognitive load scale adapted from Paas et al.'s Cognitive Load Measurement in Instructional Design, and semi-structured interview guides (Misra R. dkk., 2023). The proficiency test evaluates improvements in vocabulary, grammar, and conversational fluency, while the cognitive load scale measures intrinsic, extraneous, and germane cognitive load during chatbot interactions. Interviews with participants provide deeper insights into their experiences, challenges, and perceptions of the chatbot's usability (Shahzad dkk., 2024).

The procedures began with a pre-assessment of both groups to establish a baseline of language proficiency. The experimental group then engaged with the AI chatbot for 20-minute sessions, three times per week, over a three-week period, practicing conversation and vocabulary through guided dialogues. The control group followed the same curriculum but used traditional methods, such as textbooks and written exercises. After the intervention, both groups completed the post-assessment, and participants from the experimental group were interviewed regarding their experiences with the chatbot and their perceived cognitive load. Data from tests and interviews were analyzed to identify patterns in language improvement and cognitive load management.

RESULT AND DISCUSSION

Data collected from the study included pre- and post-test scores from both the experimental group (AI chatbot users) and the control group (traditional method users). Table 1 shows the average test scores in vocabulary, grammar, and conversational fluency for both groups before and after the three-week period. The experimental group showed a significant improvement in conversational fluency and vocabulary acquisition compared to the control group. The post-test results for the experimental group had an average increase of 15%, while the control group showed only a 7% improvement.

Category	Pre-test Average (Chatbot)	Post-test Average (Chatbot)	Pre-test Average (Control)	Post-test Average (Control)
Vocabulary	68%	85%	70%	76%
Grammar	62%	70%	65%	68%
Conversational Fluency	58%	80%	60%	67%

The experimental group showed the most notable improvement in conversational fluency, with a 22% increase from the pre-test to the post-test. Vocabulary skills also saw a significant boost, while grammar showed a more modest improvement. In contrast, the control group experienced less dramatic increases across all categories, indicating that AI chatbots may provide a more substantial benefit for conversational practice and vocabulary learning.

The data suggest that AI chatbots are particularly effective in enhancing conversational fluency and vocabulary acquisition in language learners. The experimental group demonstrated a much higher improvement in these areas compared to the control group. The interactivity and real-time feedback provided by the chatbot likely contributed to this improvement. Learners had the opportunity to practice speaking more frequently and received instant corrections, which may have helped them internalize new vocabulary and improve fluency faster than traditional methods allowed.

Grammar improvement in the experimental group, while present, was less significant than gains in other areas. This may indicate that AI chatbots are better suited for practicing conversation and vocabulary, while more structured methods, such as textbook exercises, may still be necessary for mastering grammatical rules. The control group's lower overall improvement suggests that traditional methods, while effective to a degree, may not offer the same level of engagement or immediate feedback as AI-driven learning tools.

The interviews with students from the experimental group revealed that many found the chatbot interactions engaging and helpful for practicing new words in context. However, several participants noted that the chatbot struggled to identify and correct more complex grammatical mistakes. This could explain the smaller gains in grammar proficiency, as the chatbot's focus was on facilitating conversation rather than in-depth grammatical instruction. This aligns with the quantitative data, which showed larger improvements in fluency and vocabulary.

The control group, relying on traditional learning methods, showed improvements but at a slower pace. This reinforces the idea that while traditional approaches can lead to gradual language development, the dynamic and interactive nature of AI chatbots may accelerate progress in key areas like fluency and vocabulary retention.

Qualitative data gathered from interviews with the experimental group participants provided additional insights into the cognitive load experienced during AI chatbot interactions. Many students reported that the real-time nature of the conversations made them feel more immersed in the learning process, though this immersion also led to moments of cognitive overload. Some participants found it challenging to simultaneously focus on responding to prompts, processing feedback, and recalling grammar rules. However, most students felt the cognitive load was manageable and believed it contributed to their learning progress.

The cognitive load scale used in this study showed that learners experienced moderate levels of cognitive load, with a mean score of 3.6 out of 5. The highest cognitive load was reported during the early sessions, when students were still adjusting to the chatbot's interface and the need for

quick responses. Over time, learners reported feeling more comfortable with the interaction, leading to a slight reduction in cognitive load as they became more familiar with the system. However, students who had less prior experience with conversational AI tools consistently reported higher cognitive load levels than those who were more accustomed to the technology.

In contrast, learners in the control group did not report the same levels of cognitive load, as traditional methods did not require real-time interaction. However, these students mentioned that the lack of immediate feedback in their practice made it harder to identify and correct mistakes promptly. The absence of interactive elements in the control group's learning experience may have contributed to slower improvements in fluency and vocabulary, despite the lower cognitive load.

The data suggest that while AI chatbots introduce a moderate level of cognitive load, this load is often productive and contributes to language learning progress. Students appeared to benefit from the cognitive challenge, which forced them to think more critically and engage more deeply with the material, especially in conversational settings.

Inferential statistics were used to evaluate the significance of the differences between the pre- and post-test scores of both the experimental and control groups. A paired-sample t-test was conducted to compare the performance of students in both groups, with the results showing a statistically significant improvement in the experimental group's conversational fluency and vocabulary acquisition ($p < 0.01$). Grammar scores in the experimental group also improved, though the p-value ($p = 0.08$) indicated that this improvement was not statistically significant compared to the control group.

Figure 1 illustrates the improvement in conversational fluency across both groups, highlighting the more substantial gains made by the experimental group.

The figure shows that students using AI chatbots improved their conversational fluency at a much faster rate than those using traditional methods. This significant difference reinforces the quantitative data that suggests chatbots are particularly effective in helping students develop fluency in a new language. The t-test results further confirm that the gains made by the experimental group were not due to chance, but a direct result of their interactions with the AI chatbot.

While vocabulary improvements were also statistically significant, the grammar scores showed less of a difference, reflecting the limitations of AI chatbots in teaching complex language structures. The lower significance level for grammar suggests that while AI tools are valuable for certain aspects of language learning, they may need to be supplemented with other methods to ensure comprehensive language development.

The data show a clear relationship between the use of AI chatbots and improvements in conversational fluency and vocabulary retention. Students who engaged with chatbots consistently outperformed those in the control group, suggesting that interactive AI tools offer distinct advantages over traditional learning methods. The real-time feedback and personalized interaction provided by chatbots likely contributed to these gains, as students were able to practice language skills in context and receive instant corrections.

Cognitive load emerged as another important factor in determining learning outcomes. While students using chatbots experienced higher cognitive load than those following traditional methods, this load appeared to be beneficial in promoting deeper engagement with the material. The balance between cognitive challenge and learning gains suggests that moderate cognitive load can stimulate active learning without overwhelming the learner. Students who reported manageable cognitive load tended to show higher improvements in fluency and vocabulary, indicating that the chatbot interactions struck an effective balance between challenge and support.

The lower improvement in grammar scores suggests that the relationship between AI chatbots and language learning may vary depending on the language skill being targeted. While chatbots excel at facilitating conversational practice and vocabulary usage, they may not be as effective for teaching complex grammar rules. This suggests that AI chatbots should be seen as complementary tools in language learning, working alongside other methods to provide a well-rounded language education.

The relationship between prior experience with AI tools and cognitive load also became apparent in the data. Students who were more familiar with conversational AI platforms reported lower cognitive load and showed greater fluency gains. This indicates that learner familiarity with technology plays a role in how effectively chatbots can be used in language education, underscoring the importance of preparing students for technological integration.

A case study of one participant from the experimental group who showed exceptional progress provides further insight into the potential of AI chatbots in language learning. This participant, a university student with minimal prior experience in language learning, demonstrated a remarkable 30% improvement in conversational fluency after using the chatbot. In interviews, the student attributed their success to the interactive and adaptive nature of the chatbot, which allowed them to practice speaking in a low-pressure environment without the fear of making mistakes in front of others.

The student highlighted the chatbot's ability to provide immediate corrections and suggestions during conversations, which helped them identify their weaknesses in real-time. By receiving instant feedback on vocabulary usage and sentence structure, the student was able to quickly adjust and improve their speaking skills. The low-stakes nature of the chatbot interactions also reduced anxiety, allowing the student to focus on developing fluency rather than worrying about making mistakes.

This case study also revealed that the student experienced higher-than-average cognitive load during the first week of using the chatbot. Initially, the need to process feedback while maintaining a conversation created moments of cognitive overload. However, as the student became more familiar with the system, their cognitive load decreased, and they found the chatbot interactions to be more intuitive. This suggests that while cognitive load may initially be a challenge for learners, it can be reduced over time with practice and exposure.

The student's progress highlights the potential of AI chatbots to support individualized language learning experiences. The adaptability of the chatbot allowed the student to focus on specific areas of improvement, such as vocabulary and fluency, which contributed to their rapid development. This case underscores the effectiveness of chatbot-based learning for learners who thrive in interactive, self-paced environments.

The case study findings are consistent with the broader trends observed in the study, reinforcing the idea that AI chatbots can be particularly effective for certain types of learners. The combination of instant feedback, adaptive conversations, and low-stakes practice contributed to significant improvements in conversational fluency for students who were motivated by interactive learning. However, the case study also highlights the challenge of managing cognitive load, especially in the initial stages of chatbot use, when learners are still adjusting to the technology.

The quantitative data from the entire sample further support the notion that AI chatbots are most effective for improving conversational skills and vocabulary acquisition. While grammar improvements were less pronounced, the focus on real-time interaction and vocabulary use in context appears to be the chatbot's greatest strength. This finding is particularly important for

language learners who struggle with speaking and need more opportunities to practice conversation in a controlled, feedback-rich environment.

Students in the control group, who followed traditional methods, generally improved at a slower pace, particularly in conversational fluency. This suggests that while traditional learning methods provide a solid foundation, they may not offer the same level of engagement and interactivity that chatbots can provide. The absence of real-time feedback in the control group limited students' ability to quickly identify and correct mistakes, which may have contributed to their slower progress.

The overall data suggest that AI chatbots can play a valuable role in language learning, particularly for students who need more practice with conversation and vocabulary in an interactive, low-pressure environment. However, the limitations in grammar improvement indicate that chatbots should be used alongside more structured learning tools to ensure a comprehensive approach to language development.

The results of this study indicate that AI chatbots are highly effective in enhancing conversational fluency and vocabulary acquisition in language learners. Students who used the chatbot demonstrated significantly higher improvements compared to those who followed traditional methods, particularly in areas requiring real-time interaction and feedback (Sawlowicz, 2024). However, the limited gains in grammar suggest that while chatbots are valuable tools for certain aspects of language learning, they may not be sufficient for teaching more complex language rules.

The moderate cognitive load experienced by learners using the chatbot was generally productive, promoting deeper engagement with the material without overwhelming them. However, students with less prior experience with AI tools reported higher cognitive load, suggesting that familiarity with technology plays a role in how effectively learners can use chatbots for language practice (Vera & Palaoag, 2023). This highlights the importance of preparing learners for AI-based tools and ensuring that they are comfortable with the technology (Ahadi dkk., 2023).

The data suggest that AI chatbots should be integrated into language learning programs as complementary tools, providing interactive conversation practice alongside traditional grammar instruction. For students who need to improve fluency and vocabulary, chatbots offer a highly effective and engaging solution (Jose & Jose, 2024). However, to ensure comprehensive language development, chatbots should be paired with other methods that focus on more structured aspects of language learning. The findings of this study provide valuable insights into how AI chatbots can be used to enhance language learning, but they also point to the need for further research (Cavallaro dkk., 2024). Longitudinal studies are needed to assess the long-term effects of chatbot-based learning and to explore how these tools can be adapted to suit the needs of different learners, particularly those who struggle with cognitive load.

The results of this study demonstrate that AI chatbots significantly enhance conversational fluency and vocabulary acquisition in language learners. The experimental group, which used the chatbot for language practice, showed a 22% increase in conversational fluency and a 17% improvement in vocabulary, compared to more modest gains in the control group. These improvements suggest that the interactive nature of chatbots, with real-time feedback and adaptive dialogues, allows learners to engage more deeply with language content. However, the gains in grammar proficiency were less significant, indicating that AI chatbots may be less effective in teaching structured linguistic rules (Balcombe, 2023).

The cognitive load experienced by students using the chatbot was moderate, with most participants reporting that it was manageable and contributed to their learning process. Students

initially faced challenges with the real-time nature of the chatbot interaction, but as they became familiar with the system, cognitive load decreased. Learners also noted that while the chatbot helped them improve fluency and vocabulary, it struggled to correct more nuanced grammatical errors (Tang dkk., 2024). This limitation likely contributed to the smaller improvement in grammar skills compared to vocabulary and conversational proficiency.

Interviews revealed that students appreciated the low-pressure environment the chatbot provided, allowing them to practice speaking without fear of judgment or embarrassment. The ability to receive immediate corrections on vocabulary use and conversational structure helped them build confidence in using the target language (Zarei dkk., 2024). While technical challenges were reported, such as difficulty adapting to the chatbot interface in the early sessions, these were generally seen as minor and did not detract significantly from the learning experience (Al-Abdullatif dkk., 2023).

Overall, the findings suggest that AI chatbots are effective tools for improving fluency and vocabulary in language learning, but their ability to teach complex grammar remains limited. The moderate cognitive load imposed by chatbot interactions appears to be beneficial, encouraging deeper engagement without overwhelming learners. These results highlight the potential of AI chatbots as complementary tools in language education.

The findings of this study align with previous research that highlights the effectiveness of AI chatbots in improving conversational fluency and vocabulary. Similar studies have demonstrated that AI-driven interactive platforms can boost learner engagement and provide valuable real-time feedback, contributing to faster language acquisition (Veras dkk., 2023). This study reinforces those findings, particularly in the context of conversational practice, where learners can simulate real-world dialogues. However, the study's results diverge from some research that suggests AI tools can effectively teach all aspects of language learning, particularly grammar.

While prior studies have praised AI for enhancing student learning, particularly in more structured subjects like mathematics or science, this research reveals that language learning through chatbots may not offer the same level of effectiveness for grammar (Casciato dkk., 2024). The limited ability of chatbots to provide in-depth grammatical feedback contrasts with studies that emphasize AI's precision in problem-solving or factual instruction. Language learning, particularly the teaching of grammatical structures, requires more nuanced guidance that AI chatbots currently struggle to provide.

In comparison to traditional language learning methods, AI chatbots offer clear advantages in improving fluency and vocabulary. While textbooks and classroom instruction provide foundational knowledge, they lack the interactivity and immediacy of feedback that chatbots offer. This study supports the notion that AI chatbots can complement traditional methods by offering additional, immersive language practice. However, the findings also suggest that chatbots should not be relied upon as the sole method of language instruction, as they may need to be supplemented by human instruction for comprehensive grammatical learning.

Previous research has often focused on the benefits of AI chatbots in education without adequately addressing the cognitive demands they impose on learners. This study's focus on cognitive load analysis provides new insights into the challenges learners face when interacting with AI systems. While the cognitive load was generally manageable, this research reveals the importance of considering how AI systems can be designed to reduce unnecessary cognitive demands, particularly for learners with less experience using technology.

The results of this study suggest that AI chatbots are a valuable addition to language learning, particularly for improving fluency and vocabulary retention. This reflects a broader shift in

education towards integrating AI and other digital tools to enhance learning experiences. The chatbot's success in providing real-time feedback and facilitating interactive dialogues highlights the growing role of technology in personalizing education. Learners can practice language in a dynamic environment that adapts to their needs, offering a more flexible and engaging experience compared to traditional methods.

The limited gains in grammar proficiency, however, indicate that AI technology is not yet a comprehensive solution for all aspects of language learning. While chatbots excel in conversational practice, their inability to handle more complex linguistic structures suggests that human instructors and traditional methods are still essential for a well-rounded education. This finding reflects the current limitations of AI in replicating the nuanced feedback that only experienced educators can provide, particularly in subjects that require deep critical thinking or complex problem-solving.

The cognitive load analysis provides a clear indication that while AI chatbots enhance engagement, the added mental effort of navigating the technology must be managed carefully. For learners unfamiliar with AI tools, the initial adjustment period can increase cognitive load, potentially hindering the learning process. This highlights the importance of designing AI systems that are user-friendly and intuitive, reducing the mental strain associated with using new technologies in educational settings.

The findings also point to the need for continued research into how AI can be optimized for language learning. The chatbot's effectiveness in fluency and vocabulary suggests that with further development, AI tools could be adapted to address more complex areas of language learning, such as grammar and syntax. This study signals that while AI has made significant strides in education, there are still challenges to be overcome before it can fully replace or supplement traditional teaching methods.

The implications of this study are significant for both educators and technology developers in the field of language learning. AI chatbots offer a powerful tool for improving conversational fluency and vocabulary, particularly for learners who need more practice in real-time interaction. This suggests that educational institutions should consider integrating AI chatbots into their language programs as supplementary tools, allowing students to engage with the language in a more interactive and low-pressure environment. However, educators should be aware that these tools may need to be combined with traditional instruction, particularly for grammar teaching.

The study also has implications for the design of AI chatbots used in educational settings. The moderate cognitive load experienced by learners suggests that chatbots should be carefully designed to balance challenge with accessibility. Developers should aim to create systems that are intuitive and reduce unnecessary cognitive demands, particularly for learners who are less familiar with AI technology. This study indicates that while chatbots are effective learning tools, their interface and interaction design are crucial for maximizing their educational potential.

Another important implication is that educators should be prepared to help students manage the cognitive load associated with AI-based learning tools. While chatbots offer engaging and interactive learning experiences, the real-time nature of the conversations can initially overwhelm learners. Instructors should provide guidance on how to effectively use these tools and offer additional support for students who may struggle with the technology. This ensures that learners can fully benefit from the interactive elements of AI chatbots without experiencing excessive cognitive strain.

The findings suggest that AI chatbots are best used as complementary tools in language learning, providing opportunities for interactive practice while relying on human instruction for more complex language tasks. Educational institutions and policymakers should consider the

potential of AI chatbots to enhance language education, particularly in remote or resource-constrained environments where access to human instructors may be limited.

The success of AI chatbots in improving conversational fluency and vocabulary can be attributed to their interactive, real-time feedback mechanisms. Unlike traditional language learning methods, which often involve delayed feedback or limited practice opportunities, chatbots allow learners to engage in conversation and receive immediate corrections. This instant feedback helps learners internalize new vocabulary and conversational patterns more effectively, leading to faster improvement in fluency and retention. The chatbot's adaptive nature also enables learners to practice at their own pace, further enhancing engagement and learning outcomes.

The more modest gains in grammar proficiency are likely due to the chatbot's limitations in handling complex linguistic structures. While AI chatbots are effective at simulating conversation and correcting basic language mistakes, they may struggle with the nuanced feedback required to teach grammar rules effectively. Human instructors, with their ability to provide detailed explanations and tailored guidance, are still essential for helping learners master the intricacies of grammar. This explains why grammar improvements were less pronounced in the experimental group compared to fluency and vocabulary.

The moderate cognitive load experienced by learners reflects the balance between challenge and support provided by the chatbot. Interacting with AI systems requires mental effort, particularly when learners are expected to process real-time feedback while engaging in conversation. However, the manageable cognitive load reported by most participants suggests that the chatbot's design was effective in maintaining this balance. Students were challenged enough to stay engaged without being overwhelmed, contributing to their learning progress.

The cognitive load experienced by learners with less technological familiarity highlights the importance of prior experience with AI tools. Students who were new to AI chatbots faced higher cognitive demands as they adjusted to the technology, which may have initially hindered their learning. Over time, as they became more comfortable with the system, cognitive load decreased, allowing for more effective learning. This suggests that the learning curve associated with AI tools is a key factor in their effectiveness.

The next step is for educational institutions to integrate AI chatbots into their language learning programs more widely, focusing on their strengths in improving conversational fluency and vocabulary. However, it is crucial to supplement chatbot use with traditional instruction for grammar and more complex linguistic tasks. Educators should explore how AI chatbots can be effectively incorporated into classroom settings, offering students additional practice opportunities outside of formal lessons. Further research should be conducted to explore how these tools can be adapted to address the current limitations in teaching grammar.

Future research should also investigate how AI chatbots can be designed to reduce cognitive load, particularly for learners with less technological experience. Improving the interface design and user experience could make chatbot interactions more intuitive, allowing learners to focus more on language practice and less on navigating the technology. Studies that explore the long-term effects of AI chatbot use on language retention and proficiency will provide valuable insights into how these tools can be integrated into language curricula for lasting impact.

In addition, research should explore how AI chatbots can be adapted to suit learners with different proficiency levels and learning styles. While this study focused on beginner and intermediate learners, advanced language learners may require different levels of interaction and feedback to continue improving. Tailoring AI chatbots to meet the specific needs of learners at

various stages of language acquisition could further enhance their effectiveness and ensure that all students benefit from the technology.

As AI technology continues to evolve, there is significant potential to expand the use of chatbots in other areas of education. The success of AI chatbots in language learning suggests that similar tools could be developed for other subjects that require interactive practice, such as critical thinking or problem-solving. Educational policymakers should explore the broader implications of AI in education and consider how these tools can be integrated into curricula to enhance learning outcomes across disciplines.

CONCLUSION

The most significant finding of this study is that AI chatbots are highly effective in improving conversational fluency and vocabulary acquisition in language learners. The experimental group demonstrated a substantial increase in both areas compared to the control group, showing that chatbots offer a dynamic and engaging tool for practicing language in real-time. However, the improvement in grammar proficiency was less pronounced, indicating that while AI chatbots excel in conversational practice, they may not be as effective for teaching more complex grammatical structures. This suggests that chatbots are best used as supplementary tools rather than standalone solutions for comprehensive language learning.

Cognitive load analysis revealed that learners experienced moderate cognitive demands during chatbot interactions, which appeared to be beneficial rather than overwhelming. The manageable cognitive load helped maintain engagement without causing frustration, allowing students to improve their language skills. However, students with less prior experience using AI tools initially faced higher cognitive load, highlighting the importance of familiarity with technology in maximizing learning outcomes. This finding emphasizes the need to balance technological challenge with accessibility to optimize the learning experience.

The primary contribution of this research lies in its exploration of how AI chatbots can be effectively integrated into language learning. The study demonstrates that chatbots can provide personalized, interactive practice in ways that traditional methods cannot. This research also highlights the potential for using AI chatbots to create more engaging and flexible learning environments, especially for learners who require more frequent conversational practice. By focusing on both the effectiveness of AI chatbots and the cognitive demands they impose, this study adds a unique perspective to the growing field of educational technology research.

The mixed-methods approach used in this study offers valuable insights into both quantitative learning outcomes and qualitative learner experiences. Combining pre- and post-assessment data with cognitive load measurements and participant interviews allows for a comprehensive understanding of the chatbot's impact. This methodology could serve as a model for future studies that aim to assess the role of technology in education, particularly in contexts where learner engagement and mental effort are critical to success.

The limitations of this study include its relatively small sample size and short duration, which may affect the generalizability of the findings. The study involved 60 students over a three-week period, which is not sufficient to assess long-term language retention or to evaluate the chatbot's effectiveness in larger, more diverse populations. Additionally, the focus on beginner and intermediate learners limits the applicability of the results to more advanced language learners. Further research is needed to determine how AI chatbots can be adapted to suit the needs of learners at different proficiency levels and to assess the long-term impact of chatbot-based language practice.

Another limitation is the chatbot's limited ability to provide feedback on complex grammatical errors. While the chatbot was effective in improving fluency and vocabulary, its lack of nuanced feedback for grammar highlights a key area for improvement. Future research should explore how AI chatbots can be enhanced to offer more detailed grammatical instruction, potentially through integration with other language learning technologies or by incorporating more advanced natural language processing capabilities. Expanding research to include longitudinal studies and larger, more diverse samples will help clarify the long-term potential of AI chatbots in language education.

AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.
Author 2: Conceptualization; Data curation; In-vestigation.

REFERENCES

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3). Scopus. https://doi.org/10.30935/cedtech/13152

Aeni, N., Khang, A., Al Yakin, A., Yunus, M., & Cardoso, L. (2023). Revolutionized Teaching by Incorporating Artificial Intelligence Chatbot for Higher Education Ecosystem. Dalam *AI-Centric Modeling and Analytics: Concepts, Technologies, and Applications* (hlm. 43–76). CRC Press; Scopus. https://doi.org/10.1201/9781003400110-4

Ahadi, N., Zanjanab, A. G., Sorooshian, S., Monametsi, G., Virutamasen, P., Wongpreedee, K., Deebhijarn, S., & Taghipour, A. (2023). The outlook of ChatGPT, an AI-based tool adoption in Academia: Applications, challenges, and opportunities. *IEEE Reg 10 Annu Int Conf Proc TENCON*, 15–20. Scopus. https://doi.org/10.1109/TENCON58879.2023.10322527

Al-Abdullatif, A. M., Al-Dokhny, A. A., & Drwish, A. M. (2023). Implementing the Bashayer chatbot in Saudi higher education: Measuring the influence on students' motivation and learning strategies. *Frontiers in Psychology*, *14*. Scopus. https://doi.org/10.3389/fpsyg.2023.1129070

Armaselu, F. (2024). Playing the Imitation Game: Human-AI Simulators in Pedagogic Design. Dalam Lorig F., Tucker J., Lindstrom A.D., Dignum F., Murukannaiah P., Theodorou A., & Yolum P. (Ed.), *Front. Artif. Intell. Appl.* (Vol. 386, hlm. 46–54). IOS Press BV; Scopus. https://doi.org/10.3233/FAIA240181

Balcombe, L. (2023). AI Chatbots in Digital Mental Health. *Informatics*, *10*(4). Scopus. https://doi.org/10.3390/informatics10040082

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using Chatbots as AI Conversational Partners in Language Learning. *Applied Sciences (Switzerland)*, *12*(17). Scopus. https://doi.org/10.3390/app12178427

Casciato, D., Mateen, S., Cooperman, S., Pesavento, D., & Brandao, R. A. (2024). Evaluation of Online AI-Generated Foot and Ankle Surgery Information. *Journal of Foot and Ankle Surgery*. Scopus. https://doi.org/10.1053/j.jfas.2024.06.009

Cavallaro, A., Romano, M., & Laccone, R. (2024). Examining User Perceptions to Vocal Interaction with AI Bots in Virtual Reality and Mobile Environments: A Focus on Foreign Language Learning and Communication Dynamics. Dalam Degen H. & Ntoa S. (Ed.), *Lect. Notes Comput. Sci.: Vol. 14734 LNAI* (hlm. 20–30). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-031-60606-9_2

Chan, M. M., Amado-Salvatierra, H. R., Hernandez-Rizzardini, R., & De La Roca, M. (2023). The potential role of AI-based Chatbots in Engineering Education. Experiences from a teaching

perspective. *Proc. Front. Educ. Conf. FIE*. Proceedings - Frontiers in Education Conference, FIE. Scopus. https://doi.org/10.1109/FIE58773.2023.10343296

Chetradevee, S. L., Anushka Xavier, K., & Jayapandian, N. (2022). Artificial Intelligence Technological Revolution in Education and Space for Next Generation. Dalam Sharma H., Shrivastava V., Kumari Bharti K., & Wang L. (Ed.), *Lect. Notes Networks Syst.* (Vol. 461, hlm. 371–382). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-981-19-2130-8_30

Faruqui, S. H. A., Tasnim, N., Basith, I. I., Obeidat, S. M., & Yildiz, F. (2024). Board 46: Integrating AI in Higher-Education Protocol for a Pilot Study with'SAMCares An Adaptive Learning Hub'. *ASEE Annu. Conf. Expos. Conf. Proc.* ASEE Annual Conference and Exposition, Conference Proceedings. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85202055079&partnerID=40&md5=0bc78f35f2fdaaa4d98596d20bd6a702

Haristiani, N. (2019). Artificial Intelligence (AI) Chatbot as Language Learning Medium: An inquiry. Dalam Ifdil I., Yohandri Y., Krismadinata K., & Rahim R. (Ed.), *J. Phys. Conf. Ser.* (Vol. 1387, Nomor 1). Institute of Physics Publishing; Scopus. https://doi.org/10.1088/1742-6596/1387/1/012020

Hwang, Y., Lee, S., & Jeon, J. (2024). Integrating AI chatbots into the metaverse: Pre-service English teachers' design works and perceptions. *Education and Information Technologies*. Scopus. https://doi.org/10.1007/s10639-024-12924-4

Ivanova, M., Grosseck, G., & Holotescu, C. (2024). Unveiling Insights: A Bibliometric Analysis of Artificial Intelligence in Teaching. *Informatics*, *11*(1). Scopus. https://doi.org/10.3390/informatics11010010

Jose, J., & Jose, B. J. (2024). Educators' Academic Insights on Artificial Intelligence: Challenges and Opportunities. *Electronic Journal of E-Learning*, *22*(2), 59–77. Scopus. https://doi.org/10.34190/ejel.21.5.3272

Kajiwara, Y., & Kawabata, K. (2024). AI literacy for ethical use of chatbot: Will students accept AI ethics? *Computers and Education: Artificial Intelligence*, *6*. Scopus. https://doi.org/10.1016/j.caeai.2024.100251

Krishnam, N. P., Bora, A., Swathi, R. S. V. R., Gehlot, A., Talwar, S., & Raghu, T. (2023). AI-Based advanced Talk-chatbot for Implementation. *Int. Conf. Adv. Comput. Innov. Technol. Eng., ICACITE*, 1808–1814. Scopus. https://doi.org/10.1109/ICACITE57410.2023.10182611

Lee, H.-Y., & Huang, Y.-M. (2024). Bridging STEM Education and Ubiquitous Learning: A Case Study on Developing a LINE Chatbot with Google's Gemini for Virtual Peer Collaboration. Dalam Cheng Y.-P., Pedaste M., Bardone E., & Huang Y.-M. (Ed.), *Lect. Notes Comput. Sci.: Vol. 14786 LNCS* (hlm. 237–246). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-031-65884-6_25

Misra R., Misra R., Rajarajan M., Veeravalli B., Kesswani N., & Patel A. (Ed.). (2023). 7th International Conference on Internet of Things and Connected Technologies, ICIoTCT 2022. *Lecture Notes in Networks and Systems*, *616 LNNS*. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85171994140&partnerID=40&md5=f9792d433d729745718d5334934a720b

Neo, M. (2022). THE MERLIN PROJECT: MALAYSIAN STUDENTS' ACCEPTANCE OF AN AI CHATBOT IN THEIR LEARNING PROCESS. *Turkish Online Journal of Distance Education*, *23*(3), 31–48. Scopus.

Sabeh, H. N. (2024). What Drives IT Students Toward ChatGPT? Analyzing the Factors Influencing Students' Intention to Use ChatGPT for Educational Purposes. *Int. Multi-Conf. Syst., Signals Devices, SSD*, 533–539. Scopus. https://doi.org/10.1109/SSD61670.2024.10548826

Samala, A. D., Zhai, X., Aoki, K., Bojic, L., & Zikic, S. (2024). An In-Depth Review of ChatGPT's Pros and Cons for Learning and Teaching in Education. *International Journal of Interactive Mobile Technologies*, *18*(2), 96–117. Scopus. https://doi.org/10.3991/ijim.v18i02.46509

Sawlowicz, Z. (2024). Artificial Intelligence in the education of geology students – the hero or the killer? *Przeglad Geologiczny*, *72*(5), 209–214. Scopus. https://doi.org/10.7306/2024.8

Shahzad, M. F., Xu, S., & Zahid, H. (2024). Exploring the impact of generative AI-based technologies on learning performance through self-efficacy, fairness & ethics, creativity, and trust in higher education. *Education and Information Technologies*. Scopus. https://doi.org/10.1007/s10639-024-12949-9

Tang, C. M., Ng, V. S. C., Leung, H. M. F., & Yuen, J. C. H. (2024). AI-Generated Programming Solutions: Impacts on Academic Integrity and Good Practices. Dalam Poquet O., Ortega-Arranz A., Viberg O., Chounta I.-A., McLaren B., & Jovanovic J. (Ed.), *International Conference on Computer Supported Education, CSEDU - Proceedings* (Vol. 2, hlm. 478–485). Science and Technology Publications, Lda; Scopus. https://doi.org/10.5220/0012563600003693

Terzic, I., Drobnjak, A., & Boticki, I. (2023). Designing Educational Personas using Generative AI. Dalam Shih J.-L., Kashihara A., Chen W., Chen W., Ogata H., Baker R., Chang B., Dianati S., Madathil J., Yousef A.M.F., Yang Y., & Zarzour H. (Ed.), *Int. Conf. Comput. Educ., ICCE - Proc.* (Vol. 2, hlm. 961–963). Asia-Pacific Society for Computers in Education; Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85181770328&partnerID=40&md5=456b4da0924a418e136900cbdf5b2edc

Vera, M. C. S., & Palaoag, T. D. (2023). MedPlantBot: AI Chatbot Architectural Design Framework for Responsible Use of Medicinal Plants. *IEEE Int. Conf. ICT Bus. Ind. Gov., ICTBIG*. 3rd IEEE International Conference on ICT in Business Industry and Government, ICTBIG 2023. Scopus. https://doi.org/10.1109/ICTBIG59752.2023.10456018

Veras, M., Dyer, J.-O., Rooney, M., Silva, P. G. B., Rutherford, D., & Kairy, D. (2023). Usability and Efficacy of Artificial Intelligence Chatbots (ChatGPT) for Health Sciences Students: Protocol for a Crossover Randomized Controlled Trial. *JMIR Research Protocols*, *12*(1). Scopus. https://doi.org/10.2196/51873

Wu, H.-S. (2024). Subverting the Future of Teaching: Artificial Intelligence Innovation in Nursing Education. *Journal of Nursing*, *71*(2), 20–25. Scopus. https://doi.org/10.6224/JN.202404_71(2).04

Wu, T., & Yu, Z. (2024). Bibliometric and Systematic Analysis of Artificial Intelligence Chatbots' Use for Language Education. *Journal of University Teaching and Learning Practice*, *21*(6). Scopus. https://doi.org/10.53761/6jmnst54

Zaphiris P., Ioannou A., & Ioannou A. (Ed.). (2024a). 11th International Conference on Learning and Collaboration Technologies, LCT 2024, held as part of the 26th International Conference on Human-Computer Interaction, HCII 2024. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *14724 LNCS*. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85196846290&partnerID=40&md5=946683875ab7d914f13c0747a1cfc8a9

Zaphiris P., Ioannou A., & Ioannou A. (Ed.). (2024b). 11th International Conference on Learning and Collaboration Technologies, LCT 2024, held as part of the 26th International Conference on Human-Computer Interaction, HCII 2024. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *14722 LNCS*. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85197235626&partnerID=40&md5=4e7c89ade785b98f89bdfb556e51ec79

Zarei, M., Zarei, M., Hamzehzadeh, S., Oliyaei, S. S. B., & Hosseini, M.-S. (2024). ChatGPT, a Friend or a Foe in Medical Education: A Review of Strengths, Challenges, and Opportunities. *Shiraz E Medical Journal*, *25*(6). Scopus. https://doi.org/10.5812/semj-145840

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, *10*(3). Scopus. https://doi.org/10.1016/j.heliyon.2024.e25370