

Examining the Effectiveness of ChatGPT as a Conversational Partner for Indonesian EFL Learners

Khaisya Sabina Nazarul Ummi¹

¹Faculty of Languages and Arts, English Language Education, Indraprasta PGRI University

ARTICLE INFO

Received: 27 May 2025
Revised: 18 July 2025
Accepted: 15 September 2025
Available online: 29 September 2025

Keywords:

ChatGPT
EFL Learners
Conversational AI

Corresponding Author:

Khaisya Sabina Nazarul Ummi

Email:

khaisy@gmail.com

Copyright © 2025, Language Inquiry & Exploration Review, Under the license [CC BY-SA 4.0](#)



ABSTRACT

Purpose: This study evaluates the efficiency of ChatGPT as a conversational partner in enhancing English proficiency among Indonesian learners within the English as a Foreign Language (EFL) context. It examines ChatGPT's potential as both a learning tool and a scalable alternative for institutions with limited instructional resources.

Subjects and Methods: A quantitative pre-experimental design was adopted, employing a one-group pretest-posttest method with 40 intermediate-level university students. Over a period of four weeks, participants engaged in structured conversations with ChatGPT. Speaking performance was videotaped and assessed using a standard rubric measuring fluency, vocabulary, and grammatical accuracy. In addition, learner perceptions were gathered through a structured questionnaire.

Results: Findings revealed statistically significant improvements across all assessed components, with the greatest gains in fluency and vocabulary. Questionnaire data indicated high levels of satisfaction, motivation, and perceived benefit, with a positive correlation between learner satisfaction and performance growth. These results suggest that ChatGPT enhances communicative competence while also increasing learner engagement.

Conclusions: ChatGPT demonstrates strong potential as an instructional aid in EFL learning and as a cost-effective solution for large programs. Effective integration, however, requires careful planning, ethical oversight, and teacher preparedness to ensure sustainable adoption.

INTRODUCTION

The educational sector has been invaded by artificial intelligence, which has significantly transformed the modern language acquisition in an environment where access to the real-life experience with vocabulary or conversation has been limited (Negrila, 2023; Bekou et al., 2024). Some of its most creative applications include those of large language models (LLM) like ChatGPT, which provide automated and interactive language communications in English available to all learners at various levels of proficiency.

With the current development of the digital revolution and the need to acquire EFL growing rapidly in globally linked education and laboratory systems of work-related frameworks, the English learners in Indonesia have started to massively use digital tools in complementing the traditional pedagogical games (Makhachashvili & Semenist, 2022). Particularly, ChatGPT works as an extremely promising conversational partner optimized to imitate the interaction experience with a human and promote language learning via continuous interaction.

English has always been seen as one of the main triggers of academic and professional mobility in Indonesia, rather than a second language, but a foreign language which has certain areas of use

in the social sphere (Taufik, 2024). The use of limited opportunities of real and meaningful oral communication in and out of the classroom often proves to be a challenge to attain communicative competence among the Indonesian EFL learners (Dewan & Sharma, 2025). Indonesia has teacher-centered conventional pedagogical conditions whereby, there is an increased focus on grammar, translation, and written evaluation when compared to speaking and hearing.

This discrepancy between what the curricular stresses and what is required to adequately use language in the real world continues to increase the gap on the oral proficiency of the learners. In consequence, the realization of the virtual spaces in the ability to foster interactive, communicative practice, including the one that ChatGPT offers, would make a convincing retort to such a pedagogical limitation. Vadivel et al. (2023) and Praveena & Anupama (2025) said that The recent technologies of AI-powered language software have led to the strikingly increased ability of learners to experience English usage in natural and meaningful ways. ChatGPT is an application developed based on the Generative Pre-trained Transformer architecture of OpenAI, which provides real-time, context-specific, and consistent responses making it possible to train learners in oral discourse without the pressure of speaking to rushing partners (Brown et al., 2020; Floridi & Chiriatti, 2020).

Unlike traditional computer-assisted language learning (CALL) programs that are limited either to structured forms or programmed dialogs, ChatGPT provides free and open-ended discussion, which is more similar to the interaction with a person. Such interactivity can have the potential to develop enhanced skills which include speaking fluency, lexical acquisition and pragmatic competence which are essential skills to English-as-a-foreign-language (EFL) learners when they are not exposed to English on a daily basis. The effectiveness of AI chatbots in the language learning process is more and more supported by the empirical evidence (Huang et al., 2022; Yang & Kyun, 2022). The evidence of studies carried out by Du (2025) suggests that using conversational agents can result in related gains in vocabulary retention and fluency in learners using them.

Additionally, the emerging AI-based spaces have proven capable of alleviating the anxiety of learners and empowering them, especially in an oral domain, which can be viewed as the most problematic area of EFL students. The affective value of ChatGPT such as an increase in motivation, engagement, and decrease in speaking anxiety is relevant to Indonesian learners in particular because of the cultural and educational limitations on the use of oral English. Despite the growing attention of scholars being focused on the idea of instructional settings mediated by AI, there is a lack of empirical studies that evaluate the effectiveness of ChatGPT as the means of increasing conversational efficacy of EFL learners in developing countries, like Indonesia. Most studies conducted so far have been conceptual, anecdotal, or have occurred in educational settings that are fairly well equipped in developed areas.

Due to the fact that learners in Indonesia are often exposed to a shortage of access to native speakers and quality pedagogy or full language immersion environments, it is imperative that ChatGPT English language support is scientifically and data-driven evaluated in a rigorous manner (Yawan, 2024). The topics of conversational abilities of ChatGPT have not been extensively scrutinized in terms of the possible effect on linguistic sub-skills of EFL learners, that is fluency, lexical richness, and grammatical correctness. All these dimensions are a fundamental measure of communicative competence yet most times remain untapped in the traditional model of EFL instructions (AlShareef, 2024).

According to Li (2024) and Prather et al. (2023) The ability of ChatGPT to provide the instant feedback and provide varied linguistic input as well as scaffold learner productions consequently provides a pertinent solution to redressing such shortcomings. Unlike traditional media content or the predecessors of AI-related applications, ChatGPT allows interactive negotiation of meaning, which is critical to the processes of second language acquisition. This research question seeks to research into the effectiveness of ChatGPT as a conversational interlocutor to improve second language spoken English formal level proficiency in Indonesia EFL learners (Werdiningsih et al., 2024; Salsabil & Rakhmawati, 2025). This was based on a quantitative process in which data was measured about the level of investigating performance of the subjects

before and after a longer interaction in the chatbot, specifically their fluency, lexical accuracy, and vocabulary, all of which tend to be limitable by the traditional classroom environments in Indonesia (Haugeland et al., 2022).

The outcomes demonstrate statistically insignificant improvements in all the three criteria, which consisted of fluency, accuracy, and vocabulary, implying that the greater the exposure and interaction with ChatGPT, the greater change in oral proficiency can be observed. The results also have a wider implication on the teaching practice, introduction of instructional technology and curriculum design in teaching and learning of EFL. With educators moving to hybrid and technology-centered teaching conditions, an advanced appreciation of the ways of how AI tools can be implemented reasonably and ethically becomes paramount. Its findings are an available piece of evidence-based scientific research that can be used to the knowledge of the EFL learners and instructors and the policy makers of Indonesia and contribute to academic discourse on the role of AI in language learning.

METHODOLOGY

Research Design

The current research was based on quantitative and pre-experimental one group experimental pretest design study that explores the effectiveness of ChatGPT as a conversational partner to improving English proficiency among Indonesian EFL students. The choice of this design was informed by the ability of capturing the intragroup change using multiple measures. Possible limitation is the lack of a control group but at this point no attempt was to establish causality and the aim was to find out initial effectiveness. The comparison of the pre-intervention and post-intervention results in terms of fluency, vocabulary, and grammatical correctness enables the findings to add to the literature research describing the role of ChatGPT in language learning.

Population and Sample

In the current research, the sample size was obtained through candidates who enrolled in a course of English as a Foreign Language (EFL) in various faculties across the city of Makassar in South Sulawesi and totaled 40 intermediate level of EFL learners. The participants would have been selected through a purposive sampling method and they had to show a score in placement that must have exceeded a decided value in standardized English-Language examination as offered by the institution and also showed that they were ready to continue to participate in the four weeks of the experimental period. About the inclusion criterion of intermediate-level proficiency, it was designed to allow finding the participants with adequate linguistic competence who would be able to interact with ChatGPT productively but still had an enumerable number of weaknesses in the field of spoken English. In addition, everywhere access to smartphone or laptop with stable connection line required as the condition of possibility to interact in the language model.

Research Instrumentation

The present study utilized two main tools to achieve its research aim, including (1) the speaking performance rubric the learners completed prior and following the intervention and (2) patterned questionnaire the learners took at the end of the intervention phase to identify their findings of the experience with ChatGPT. The speaking performance rubric measured three of the underlying dimensions of oral proficiency including fluency, lexical resource, and grammatical accuracy by converting IELTS band descriptors to speaking by ensuring standardized and reliable testing. The test was taken in a front of a camera, and participants were told to make a five-minute speech about a common topic (e.g., education, hobbies, culture) before and after intervention; each time both recordings were taken and evaluated later by two certified EFL instructors who have experience with oral language testing. Analyses on concordance were carried out using inter-rater reliability. The items in the post-intervention questionnaire were recalled based on 20 questions scored on a 5 points Likert scale of 1 (strongly disagree) to 5 (strongly agree). These items were related to the ease of use of ChatGPT, improvement in speaking skills perception, motivation, and engagement, less speaking anxiety, and satisfaction with the process. Clarity, reliability, and construct validity of the questionnaire were confirmed with the help of pilot study used with 10 non-participating students.

Procedure

This study took place in a duration of four weeks. A pre-test was run to assess the initial level of speaking skill as all the participants were analyzed during Session 1. The speech activities took place under an experimental setting with the audio being recorded later to be analyzed. Following the pre-test, participants had access to the introduction of the power of ChatGPT and were guided on how to practice conversations using this chatbot in a structure mode. They were asked to participate in three ChatGPT sessions a week spending about 30 minutes on each session, and the overall number of sessions 12. In the interactions, the participants were expected to engage and sustain a dialogue with ChatGPT on various topics in everyday life and academics. The list of suggested themes and guiding questions was presented but subjects were welcome to leave the conversation in free flow. In an effort to foster the sense of accountability and continuity, the participants provided weekly entries in their logs where they reported on the subjects of their interaction, the length of time, and notes on their reflections on the experience. Also, the quality of interaction was monitored through the sampling of some of the sessions (by agreement) with the use of screen recordings by the researcher. At the end of the fourth week, the same speaking performance protocol was used to complete the post-test by all the participants. The pre-test and post-test recordings being assessed by the original raters were graded with the same speaking rubric. During the last session, participants filled in the questions pertaining to their impressions and satisfaction with ChatGPT as the speaking partner.

Data Analysis

To understand whether statistically significant changes appeared in the areas of fluency, grammatical accuracy, and lexical diversity of the learners after interacting with ChatGPT, the speaking pretest and posttest data were submitted to paired-samples t-tests. Before determining the t-tests, there was consideration of the assumption of normal distribution through Shapiro-Wilk test. The effect sizes (Cohen d) were then estimated to determine the size of the noted achievements. The data obtained on the answers of the questionnaire were interpreted through descriptive statistical methods such as the mean score, standard deviations and frequencies. In addition, a Pearson correlation test was used to compare the perceived usefulness of ChatGPT and the actual performance improvements of learners, which was an attempt at achieving this balance between subjectivity and objectivity.

RESULTS AND DISCUSSION

As Indonesian EFL learners continue to face limitations in accessing authentic, consistent speaking practice, artificial intelligence offers a scalable and adaptive solution that traditional classroom settings often fail to provide. The integration of ChatGPT into the language learning process reflects a shift in how educational systems can manage and deliver communicative competence through non-human agents. This study measures the effectiveness of ChatGPT not only in terms of linguistic outcomes such as fluency, vocabulary, and grammatical accuracy but also in terms of learner perception and engagement, which are essential metrics in educational management and policy. The results that follow offer a data-driven perspective on how AI tools can enhance speaking proficiency and reshape the management of language education in resource-constrained settings.

Table 1. Descriptive Statistics of Speaking Test Scores (Pretest and Posttest)

Component	N	Pretest Mean	Posttest Mean	Mean Difference	Std. Deviation
Fluency	40	5.60	7.05	1.45	0.91
Vocabulary	40	5.45	6.80	1.35	1.02
Grammatical Accuracy	40	5.10	6.20	1.10	1.15
Total Speaking Score	40	16.15	20.05	3.90	1.20

The results of the study demonstrate clear improvements across the three dimensions of speaking performance fluency, vocabulary, and grammatical accuracy after students engaged in structured conversations with ChatGPT. Among these, fluency recorded the strongest effect, with an average

score increase of 1.45 points, indicating that learners became more confident, smoother, and less hesitant in their oral expression. Vocabulary followed with a 1.35-point improvement, reflecting a greater range of lexical items, more precise word choice, and an enhanced ability to use idiomatic expressions. Grammatical accuracy, though showing the smallest increase of 1.10 points, still represents a meaningful gain in the correct use of tense, agreement, and clause structure. Altogether, the combined speaking score improved by 3.90 points, signifying that ChatGPT contributed to measurable and well-rounded growth in oral proficiency. In proportional terms, fluency accounted for about 37 percent of the total gain, vocabulary 35 percent, and grammar 28 percent. These findings suggest that while conversational practice with ChatGPT is particularly effective in strengthening fluency and lexical development, grammatical accuracy requires additional support through explicit feedback and targeted instruction.

Table 2. Paired Samples t-Test Results

Component	t	df	p-value	Cohen's d (Effect Size)	Interpretation
Fluency	9.81	39	0.000	1.55	Large effect, significant
Vocabulary	8.76	39	0.000	1.39	Large effect, significant
Grammatical Accuracy	6.45	39	0.000	1.02	Large effect, significant
Total Score	11.22	39	0.000	1.77	Very large effect, highly significant

The statistical analysis using paired-samples t-tests confirmed that the ChatGPT intervention produced significant improvements across all measured aspects of speaking proficiency. In every dimension fluency, vocabulary, and grammatical accuracy the differences between pretest and posttest scores were statistically significant, with p-values below the 0.05 threshold. This indicates that the observed improvements were highly unlikely to be due to chance and can be confidently attributed to the intervention. Beyond statistical significance, the analysis also revealed strong practical significance, as reflected in the effect sizes. All effect sizes were greater than 0.8, which, according to conventional benchmarks, represents a large effect. This means that the magnitude of improvement was not only measurable but also meaningful in real-world terms. The most notable result was found in the total speaking score, which recorded the largest effect size (Cohen's d = 1.77). This exceptionally high value verifies that ChatGPT is particularly effective in enhancing overall oral proficiency, producing substantial progress within a relatively short intervention period. Taken together, these results confirm both the statistical reliability and the practical relevance of ChatGPT as a powerful tool for supporting EFL learners' speaking development.

Table 3. Learners' Perceived Usefulness of ChatGPT (Likert Scale Summary)

Item	Mean	Std. Deviation
ChatGPT helped me speak more fluently	4.50	0.60
I learned new vocabulary through ChatGPT conversations	4.45	0.67
ChatGPT improved my confidence in speaking English	4.60	0.55
ChatGPT was easy and convenient to use	4.70	0.48
I would continue using ChatGPT for English speaking practice	4.65	0.52
Overall satisfaction with ChatGPT as a speaking partner	4.58	0.51

Learner satisfaction with the use of ChatGPT as a conversational partner was reported to be very high, reflecting both its functional effectiveness and its broader psychological impact on the learning process. Among the most positively rated dimensions were the ease of use and the boost in learner confidence. The ease of use highlights ChatGPT's accessibility and user-friendly design, which allowed students to engage with the tool without requiring extensive technical knowledge or additional training. This simplicity made it possible for learners to focus fully on practicing the language rather than struggling with the medium of delivery. Equally significant was the increase in learner confidence, which suggests that ChatGPT not only provided a space for practicing fluency, vocabulary, and grammar, but also played a role in reducing anxiety and creating a supportive environment. For many learners, interacting with ChatGPT offered a low-pressure opportunity to rehearse language use repeatedly, free from the fear of judgment that might occur

in traditional classroom settings. This combination of practical usability and psychological support demonstrates that ChatGPT functions not just as a linguistic tool but also as a motivational and confidence-building aid, reinforcing its value as a holistic support system in EFL learning.

Table 4. Pearson Correlation between Perceived Usefulness and Speaking Score Gain

Variables	r	p-value	Interpretation
Perceived Usefulness (Questionnaire)	0.67	0.000	Moderate to strong positive correlation

The analysis also revealed a statistically significant positive correlation between learners' perceptions of the usefulness of ChatGPT and their actual improvements in speaking test scores. With a correlation coefficient of $r = 0.67$ at a significance level of $p < 0.01$, the relationship is both strong and reliable, indicating that students who reported the highest levels of perceived improvement were also the ones who achieved the greatest measurable gains in oral proficiency. In other words, positive user experiences with ChatGPT were not merely subjective impressions but were closely aligned with objective learning outcomes. This finding reinforces the idea that the effectiveness of educational technologies depends not only on their instructional design but also on how learners perceive and engage with them. When students view a tool as useful, motivating, and supportive, they are more likely to invest effort and practice consistently, which in turn leads to better performance. Thus, the results strongly support the connection between language-learning outcomes and user experience, underscoring the importance of designing and implementing AI-based learning tools in ways that maximize both pedagogical value and learner satisfaction.

Discussion

The current inquiry suggests more than a methodological realignment of language pedagogy; hence, it is establishing a cross-sectoral directive towards educational management to rethink its organizational operations to align with the emerging technological frameworks in favor of alternative practices that offer greater levels of autonomy, scalability, and flexibility to the deployment of resources towards the students. In management enquiries, which lie specifically within the context of education systems and the leadership within the teaching and learning process, the intra-systemically introduction of AI within the teaching and learning process cannot be understood by the criteria of a pedagogical innovation; it is more a strategic reconfiguration of the institutional arrangement that coordinates human and non-human teaching agent. Within the framework of the fact that the ability of Indonesian EFL learners to fully perform their linguistic functions has been an ongoing challenge (Gayatri et al., 2023), the results of the study confirm that ChatGPT, viewed as one of the recent members of the virtual assistants club, should be considered not as a technological sidekick but rather as a performance-enhancing aid that needs a systematic integration in the instructional practices to be administered.

The fact that ChatGPT reacts is not what makes it so different and valuable among the traditional CALL interventions but also the fact that it can be used in a limited and highly resource-restricted setting. In an institutional environment, where a significant number of Indonesian schools face the dilemma of immense classrooms, insufficient time allocation of teacher and students, and economic underinvestment in oral language training (Madhakomala et al., 2022), the evidence would imply that ChatGPT would take over a managerial role traditionally played by human corporates. It directly supports previous claims who pointed out that mobile and AI tools are regarded as the instructive amplifiers, especially in low-access contexts. The challenge that now faces the managers of schools and universities in turn is to structuralize these tools and not merely incorporate them into classroom practices, but also into things such as institutional performance measurement, faculty developmental frameworks and systems of tracking learner progress. In the larger context of strategic resource management in education, this study provides first-hand evidence that the reduced cost-performance gap in the process of teaching oral languages may be attained with the assistance of digital agents such as ChatGPT. Analysis of cost-benefit in language education technology was previously speculative (Camilleri, 2024), mainly because user outcome data have not previously been quantifiable.

The measurable increase in speaking proficiency of EFL learners in this study, as well as their satisfaction with the increased utility, productivity and stakeholder alignment is tangibly supported by statistical evidence of the improvements. It validates statements provided by Morgado et al. (2025) that by aligning AI-mediated instruments with clearly defined and measurable skills, learning outcomes are enforced but the economy of teaching becomes transformed as well. Tied to this debate as well is the managerial view of the engagement of learners and their behavioral change. In any study conducted on management, in particular, an organization based on learning, embracement of emerging technologies is hardly effective without the change of user behavior, perception, and appetite (Giannakos et al., 2022). It was discovered this study that learners considered ChatGPT to be a good and empowering partner in speaking. This is agreeable with previous studies who demonstrated that learners are willing to increase their engagement when granted authority over their learning pathways and adaptive paths of giving feedback. These results require a rethinking of the role of the learner agency in the eyes of academic managers it is not merely considered as one of the side-effects of curriculum development but it is an important indicator of the innovation at the system level.

Furthermore, close connection between the perceived utility of ChatGPT usage to the actual improvement of performance levels among the learners highlights the importance of subjective experience when dealing with digital pedagogy management. The convergence qualifies the technology acceptance Abu et al. (2022), that directly predicts the behavioral intention through perceived usefulness, and subsequently the performance. Managers who are creating the policies of language education and digital learning conditions will find this alignment as an opportunity to incorporate student feedback not as post-hoc evaluations but as immediate data in the decision making. It also reflects the study, which established that the learner satisfaction predicts digital tool sustainability better than the test-only results. Moreover, one should put these data into the context of international comparative paradigms of innovative approach to language teaching. The potential practical application of chatbot technologies in EFL resulted in comparable gains in those studies conducted in Taiwan (Hsu et al., 2023), though without assessing how this development may influence success in management. Those in charge of Indonesian education should be aware that as the effectiveness of ChatGPT is being globally validated, its local use necessitates its strategic approach to the infrastructural variability, the gaps in digital literacy and the rigidity of curriculum that has been put in writing in the context of Indonesia.

Meanwhile, we have to be cautious about the techno-optimism that is currently prevailing in the discourses of AI in education. Although the present study provides convincing evidence about the efficiency of ChatGPT, the legal and management issues associated with the use of non-human language partners should also be discussed. Reminding us that AI tools can also reinforce the patterns of dependency, genres of discourse, and hide the issues with teacher preparation or curriculum design, note that the practices based on AI tools could be used to shape the digital citizens who are dependent on the provided technology and therefore unable to question it or challenge it. Accordingly, it would be futile to expect those in charge of institutions to perceive ChatGPT as a replacement of ineffective teaching, rather than an adjunct to dialogically dense instruction. In order to be pedagogically relevant, AI should be an element of an ongoing system of human control, employee professional growth, and data-based curriculum assessment. This view brings into doubt the method of measurement of institutional performance. It is already possible to quantify the ways in which people improve as they speak through ChatGPT, and the study demonstrates that, as a result, why are language skills still being tested largely through written tests? There has been a long literature of management scholarship terming the lack of alignment between the assessment design and learning design in education (Maki, 2023). Implementing AI-based tools into the process of delivering instruction must come with a shift in the practice of performance measurement and assessment. Institutions should redesign assessment patterns to feature formative, dialogic, and process assessments of speaking ability, such assessments which can be directly supported and followed by tools such as ChatGPT.

CONCLUSION

The current study clearly shows that ChatGPT, as an AI-powered conversational agent, can be used indeed as a functional instrument of enhancing the English-speaking proficiency of

Indonesian EFL students, especially fluency, vocabulary, and grammatical correctness. On management level, ChatGPT can be incorporated into EFL instruction not only as a pedagogical tool, but also as a strategic one that helps manage the lack of resources, paid-instructional opportunities, and engagement issues of learners. Participants reacted favorably to the course, which fueled their growing motivation and confidence, supporting how the scalable low-cost digital interventions are viable in language learning. The correspondence between perceived usefulness and performance in practice manifests practical significance of this AI tool in aiding quantifiable learning performance. To the managers and policymakers in education, this paper provides evidence in the favor of the guided implementation of AI tools in the curriculum, professional development programmed, and the support system of learners. But to be responsible and sustain this adoption there must be ethical guidance, teacher preparedness, and curriculum integration. The encouraging outcomes attract the necessity to conduct additional studies that would include control groups, longer interventions, and work with more varied learner populations to get a better idea of the pedagogical implications of ChatGPT. However, the existing results seem to place ChatGPT as a feasible, productive, and tactically logical extension to EFL learning in Indonesia.

REFERENCES

- Abu-Taieh, E. M., Alhadid, I., Abu-Tayeh, S., Masa'deh, R. E., Alkhawaldeh, R. S., Khwaldeh, S., & Alrowwad, A. A. (2022). Continued intention to use of M-Banking in Jordan by integrating UTAUT, TPB, TAM and service quality with ML. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(3), 120. <https://doi.org/10.3390/joitmc8030120>
- AlShareef, S. N. A. (2024). Applying the Communicative Approach in Assessing EFL Young Learners. *Eurasian Journal of Applied Linguistics*, 10(1), 14-27.
- Bekou, A., Mhamed, M. B., & Assissou, K. (2024). Exploring opportunities and challenges of using ChatGPT in English language teaching (ELT) in Morocco. *Focus on ELT Journal*, 6(1), 87-106. <https://doi.org/10.14744/felt.6.1.7>
- Camilleri, M. A. (2024). Metaverse applications in education: a systematic review and a cost-benefit analysis. *Interactive Technology and Smart Education*, 21(2), 245-269. <https://doi.org/10.1108/ITSE-01-2023-0017>
- Dewan, M. S., & Sharma, B. D. (2025). Enhancing Oral Communication in English as a Second Language. *Journal of Tikapur Multiple Campus*, 8, 18-41. <https://doi.org/10.3126/jotmc.v8i8.75949>
- Du, Q. (2025). How artificially intelligent conversational agents influence EFL learners' self-regulated learning and retention. *Education and Information Technologies*, 1-67. <https://doi.org/10.1007/s10639-025-13602-9>
- Gayatri, P., Sit, H., Chen, S., & Li, H. (2023). Sustainable EFL blended education in Indonesia: Practical recommendations. *Sustainability*, 15(3), 2254. <https://doi.org/10.3390/su15032254>
- Giannakos, M. N., Mikalef, P., & Pappas, I. O. (2022). Systematic literature review of e-learning capabilities to enhance organizational learning. *Information Systems Frontiers*, 24(2), 619-635. <https://doi.org/10.1007/s10796-020-10097-2>
- Haugeland, I. K. F., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022). Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. *International Journal of Human-Computer Studies*, 161, 102788. <https://doi.org/10.1016/j.ijhcs.2022.102788>
- Hsu, M. H., Chen, P. S., & Yu, C. S. (2023). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments*, 31(7), 4297-4308. <https://doi.org/10.1080/10494820.2021.1960864>

- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of computer assisted learning*, 38(1), 237-257. <https://doi.org/10.1111/jcal.12610>
- Li, Y. (2024). Usability of ChatGPT in Second Language Acquisition: Capabilities, Effectiveness, Applications, Challenges, and Solutions. *Studies in Applied Linguistics & TESOL*, 24(1), 24-37.
- Madhakomala, R., Hakim, M. A., & Syifauzzuhrah, N. (2022). Problems of education in Indonesia and alternative solutions. *International Journal of Business, Law, and Education*, 3(2), 135-144. <https://doi.org/10.56442/ijble.v3i3.64>
- Makhachashvili, R., & Semenist, I. V. (2022). Dynamic assessment of e-learning in foreign languages programs in sustainable and emergency digitization formats. *Education, REsearch and Innovation*, 15, 4589-4594.
- Maki, P. L. (2023). *Assessing for learning: Building a sustainable commitment across the institution*. Routledge.
- Morgado, E., Leonido, L., Pereira, A., & Gouveia, L. B. (2025). Technology-Mediated Education: impact of AI on the main distance learning modalities. <https://doi.org/10.22521/edupij.2025.16.211>
- Negrila, A. M. C. (2023). The new revolution in language learning: The power of artificial intelligence and education 4.0. *Bulletin of" Carol I" National Defence University (EN)*, 12(02), 16-27.
- Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Craig, M., ... & Savelka, J. (2023). The robots are here: Navigating the generative ai revolution in computing education. In *Proceedings of the 2023 working group reports on innovation and technology in computer science education* (pp. 108-159). <https://doi.org/10.1145/3623762.3633499>
- Praveena, T., & Anupama, K. (2025). Machine Learning Meets Language Learning: The Transformative Potential of Artificial Intelligence in English Language Instruction. *Human Research in Rehabilitation*, 15(1). <https://doi.org/10.21554/hrr.042511>
- Salsabil, A. D., & Rakhmawati, L. A. (2025). From Silent Learners to Confident Speakers: The Effect of AI Voice Chat with ChatGPT on EFL Speaking Skills. *Journal of English Education*, 3(1), 38-50. <https://doi.org/10.61994/jee.v3i1.1137>
- Taufik, M. (2024). English Language Proficiency and Career Opportunities:: Perceptions of Indonesian University Graduates. *Language Value*, 17(1), 85-107.
- Vadivel, B., Shaban, A. A., Ahmed, Z. A., & Saravanan, B. (2023). Unlocking English proficiency: Assessing the influence of AI-powered language learning apps on young learners' language acquisition. *International Journal of English Language, Education and Literature Studies*, 5(2), 123-139. <https://dx.doi.org/10.22161/ijeel>
- Werdiningsih, I., Marzuki, Indrawati, I., Rusdin, D., Ivone, F. M., Basthomi, Y., & Zulfahreza. (2024). Revolutionizing EFL writing: unveiling the strategic use of ChatGPT by Indonesian master's students. *Cogent Education*, 11(1), 2399431. <https://doi.org/10.1080/2331186X.2024.2399431>
- Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, 38(5), 180-210. <https://doi.org/10.14742/ajet.7492>
- Yawan, H. (2024). Challenging Native Speakerism: TEFL in Indonesia Beyond Native Speaker Fallacy. *JLE: Journal of Literate of English Education Study Program*, 5(1), 62-72. <https://doi.org/10.47435/jle.v5i1.2692>