OPEN ACCESS
Lingeducа: Journal of Language and Education Studies
Vol. 4 No. 3. December 2025, pp. 169-182 DOI. 10.70177/lingeduca.v4i3.3244

Research Article

THE EFFICACY OF AI-POWERED TUTORS VS. HUMAN TUTORS IN DEVELOPING EFL ORAL FLUENCY: A COMPARATIVE EXPERIMENTAL STUDY

Miku Fujita[1], and Hindri Febri Ana Sari[2]
[1] University of Kyoto, Japan
[2] Politeknik Negeri Ambon, Indonesia

Corresponding Author:

Miku Fujita,
Faculty of Agriculture, University of Kyoto.
Yoshidahonmachi, Sakyo Ward, Kyoto, 606-8501, Japan
Email: mikufujita@gmail.com

Article Info

Received: June 13, 2025
Revised: September 27, 2025
Accepted: November 30, 2025
Online Version: December 31, 2025

Abstract

The rapid advancement of artificial intelligence has led to the increasing use of AI-powered tutors in English as a Foreign Language (EFL) instruction, particularly for speaking skills development. Despite growing adoption, empirical evidence comparing the effectiveness of AI-powered tutors and human tutors in developing oral fluency remains limited and inconclusive. This study aims to comparatively examine the efficacy of AI-powered tutors and human tutors in enhancing EFL learners' oral fluency across multiple fluency dimensions. This research employed a comparative experimental design involving two groups of EFL learners who received equivalent instructional content over a fixed intervention period, differentiated only by tutoring modality. One group engaged in AI-powered tutoring, while the other received instruction from human tutors. Pretest and posttest measures of oral fluency were administered, focusing on temporal, interactional, and discourse-related indicators. Quantitative data were analyzed using inferential statistical techniques to determine within-group and between-group differences. The findings reveal that both AI-powered tutors and human tutors significantly improved learners' oral fluency. AI-powered tutors were particularly effective in enhancing temporal fluency, including speech rate and reduced hesitation. Human tutors demonstrated superior effectiveness in developing interactional competence, discourse coherence, and pragmatic appropriateness. The study concludes that AI-powered tutors serve as effective supplementary tools for oral fluency practice, while human tutors remain essential for higher-level communicative development. A hybrid instructional approach is recommended to maximize EFL speaking outcomes..

Keywords: AI-Powered Tutors, Efl Oral Fluency, Experimental Study, Human Tutors, Speaking Skills

Journal Homepage https://research.adra.ac.id/index.php/lingeduca
How to cite: Fujita, M., & Sari, H. F. A. (2025). The Efficacy of Ai-Powered Tutors Vs. Human Tutors in Developing Efl Oral Fluency: A Comparative Experimental Study. *Lingeduca: Journal of Language and Education Studies*, *4*(3), 121–134. https://doi.org/10.70177/lingeduca. v4i3.3244
Published by: Yayasan Adra Karima Hubbi

INTRODUCTION

The rapid advancement of artificial intelligence has profoundly transformed educational practices, particularly in the field of language learning. AI-powered tutors, equipped with speech recognition, adaptive feedback, and personalized learning algorithms, are increasingly being integrated into English as a Foreign Language instruction (Heydarnejad, 2025; Mahmoudi-Dehaki & Nasr-Esfahani, 2025). These technologies promise scalable, flexible, and individualized learning experiences that can potentially address long-standing challenges in oral fluency development, such as limited speaking opportunities and inconsistent feedback. At the same time, human tutors remain central to communicative language teaching, offering nuanced interaction, emotional support, and contextualized feedback that are difficult to fully replicate through technology (Zhuang et al., 2025). The coexistence of these two tutoring modalities raises critical questions about their relative effectiveness in fostering oral fluency, a core yet challenging component of EFL competence (Al-Bogami & Alahmadi, 2025; Yang & Zhao, 2024).

Oral fluency in EFL contexts is widely recognized as a multidimensional construct involving speed, accuracy, coherence, and communicative confidence. Many EFL learners struggle to achieve fluency due to anxiety, lack of exposure to authentic interaction, and limited instructional time (Dwivedi et al., 2023; Zhang, 2025). Traditional classroom instruction often prioritizes grammar and reading skills, leaving speaking practice underdeveloped. AI-powered tutors have been proposed as a solution to this imbalance by providing learners with frequent, low-stakes speaking opportunities and immediate corrective feedback. Nevertheless, concerns persist regarding the depth of interaction, pragmatic appropriateness, and socio-emotional dimensions of AI-mediated communication when compared to human tutors (Alsswey et al., 2025; Xie et al., 2025).

The growing reliance on AI-based tutoring systems in educational institutions has outpaced rigorous empirical evaluation of their pedagogical effectiveness, particularly in comparison with human tutors. While proponents highlight efficiency, accessibility, and personalization, skeptics question whether AI can truly support complex communicative skills such as spontaneous speech production and interactive fluency (Brezovec et al., 2025; Ng et al., 2025). This tension between technological optimism and pedagogical caution forms the broader context of the present study, which seeks to empirically examine how AI-powered tutors and human tutors differentially influence the development of EFL oral fluency (Mim et al., 2025).

The central problem addressed in this study concerns the lack of clear empirical evidence regarding the comparative effectiveness of AI-powered tutors and human tutors in developing EFL oral fluency. Existing instructional practices increasingly adopt AI tools without sufficient understanding of their pedagogical strengths and limitations (Buciuman & Potra, 2025). Educators and institutions face uncertainty when deciding whether AI-based tutoring can serve as an effective alternative or complement to human tutoring, particularly for speaking skills that rely heavily on interaction and feedback quality (von Garrel & Mayer, 2024).

Another aspect of the problem lies in the inconsistent operationalization and measurement of oral fluency across studies examining technology-enhanced language learning. Many investigations rely on self-reported perceptions or limited performance indicators, making it difficult to draw robust conclusions about actual fluency development (Aladini et al., 2025). Furthermore, studies that examine AI-assisted learning often focus on vocabulary acquisition or grammar accuracy, leaving oral fluency comparatively underexplored. This methodological inconsistency complicates efforts to evaluate the true instructional value of AI-powered tutors (Avci et al., 2025; Kohnke & Moorhouse, 2025).

The problem is further intensified by contextual differences in EFL learning environments, where access to qualified human tutors may be limited due to financial, geographical, or institutional constraints. In such contexts, AI-powered tutors are frequently

promoted as cost-effective substitutes rather than pedagogically grounded tools (Chen & Cheung, 2025; Görer & Aydemir, 2024). Without comparative experimental evidence, the adoption of AI tutoring risks being driven by practicality rather than effectiveness. This study responds to these issues by systematically comparing AI-powered and human tutoring under controlled experimental condition (Namaziandost, 2025)s.

The primary objective of this study is to examine the relative efficacy of AI-powered tutors and human tutors in improving EFL learners' oral fluency. The research seeks to determine whether statistically significant differences exist in fluency gains between learners who receive AI-mediated tutoring and those who engage in human-led tutoring sessions. By focusing on oral fluency as the central outcome, the study aims to provide empirical clarity on how different tutoring modalities influence spoken language development (Brunton et al., 2025; Shahini, 2025).

A secondary objective involves analyzing specific dimensions of oral fluency, including speech rate, pause phenomena, lexical diversity, and communicative confidence. Rather than treating fluency as a single construct, the study seeks to capture its multifaceted nature and explore how each tutoring modality supports different aspects of spoken performance. This nuanced analysis is intended to inform pedagogical decision-making and instructional design in EFL contexts (Y. Liu et al., 2024; Nandagopal, 2025).

The study also aims to contribute methodological rigor by employing a comparative experimental design with standardized fluency measures and controlled instructional variables (G. L. Liu et al., 2025). Through this approach, the research aspires to generate evidence that is both pedagogically meaningful and practically applicable. The findings are expected to inform educators, curriculum designers, and policymakers about the strategic integration of AI tools alongside human instruction in language education (Law, 2024).

Despite the expanding body of research on technology-enhanced language learning, significant gaps remain in the literature concerning direct comparisons between AI-powered tutors and human tutors (Lai & Lee, 2024; Park & Kim, 2025). Many existing studies examine AI tools in isolation, without benchmarking their effectiveness against traditional human-mediated instruction. As a result, claims about the superiority or equivalence of AI tutoring often rest on indirect or incomplete evidence (Lissack & Meagher, 2024).

Another notable gap lies in the limited use of experimental designs that control for instructional content, duration, and learner proficiency (Menon & Shilpa, 2023). Studies frequently rely on quasi-experimental or observational methods, making it difficult to attribute learning outcomes specifically to the tutoring modality. Moreover, few studies employ objective fluency metrics derived from speech analysis, further limiting the robustness of their conclusions (Tian & Zheng, 2025).

The literature also shows a tendency to emphasize learner attitudes and perceptions toward AI tutors rather than measurable performance outcomes. While affective factors are important, they do not substitute for empirical evidence of skill development (Şimşek et al., 2025; Tram et al., 2024). This study addresses these gaps by implementing a controlled experimental design that directly compares AI-powered and human tutors using standardized oral fluency measures. By doing so, it seeks to provide a clearer and more balanced understanding of how AI and human tutoring function in EFL speaking instruction (Huwer et al., 2025; Korzynski et al., 2025).

The novelty of this research lies in its direct, experimental comparison of AI-powered tutors and human tutors with oral fluency as the primary outcome variable. Unlike prior studies that focus on technological feasibility or learner satisfaction, this study foregrounds communicative performance and fluency development (Adorni et al., 2024). The comparative design allows for a more precise evaluation of the pedagogical affordances and limitations of each tutoring modality.

Another innovative aspect of the study is its multidimensional approach to oral fluency assessment. By incorporating both temporal and qualitative indicators of spoken performance, the research moves beyond simplistic measures and captures the complexity of fluency development. This approach provides deeper insight into how AI and human tutors differentially support various components of speaking proficiency (Zhai & Li, 2025).

The justification for this study is grounded in the growing institutional reliance on AI-driven educational technologies and the need for evidence-based integration. As educational systems increasingly invest in AI-powered tutoring solutions, understanding their pedagogical impact becomes imperative (Adorni & Piatti, 2025; Sun et al., 2024). This research contributes to the field by offering empirically grounded guidance on when and how AI tutors can be effectively employed, and where human tutors remain indispensable. Through its comparative and experimental focus, the study advances scholarly discourse on artificial intelligence in language education while addressing pressing practical concerns in EFL instruction.

RESEARCH METHOD

Research Design

This study employed a comparative experimental research design to examine the differential effects of AI-powered tutors and human tutors on the development of EFL oral fluency. A pretest-posttest control group design was adopted to ensure systematic comparison between the two instructional conditions. Participants were randomly assigned to one of two groups: an experimental group receiving AI-powered tutoring and a comparison group receiving human-led tutoring (Lachheb et al., 2025). Both groups were exposed to identical instructional content, learning objectives, and duration of treatment, with the tutoring modality serving as the sole independent variable. This design enabled a rigorous assessment of causal relationships between the type of tutor and gains in oral fluency.

Research Target/Subject

The population of this study consisted of undergraduate EFL learners enrolled in a non-English-speaking higher education institution where English is taught as a foreign language. A total of 80 students with intermediate-level English proficiency were selected using stratified random sampling to ensure balanced representation in terms of gender and prior speaking ability. The sample was divided equally into two groups, with 40 students assigned to the AI-powered tutor group and 40 students assigned to the human tutor group. All participants had similar instructional backgrounds and had not previously received intensive speaking-focused tutoring outside the formal curriculum. These criteria were applied to minimize confounding variables related to prior exposure or proficiency differences (Ebadi et al., 2025).

Research Procedure

The research procedures were conducted in four stages. The first stage involved administering a pretest oral fluency assessment to all participants to establish baseline equivalence between groups. The second stage consisted of an eight-week intervention period, during which the AI-powered tutor group engaged in structured speaking practice using an AI system equipped with speech recognition, automated feedback, and adaptive prompts, while the human tutor group participated in interactive speaking sessions facilitated by trained EFL instructors. The third stage involved administering a posttest oral fluency assessment using tasks parallel to those in the pretest (Guo et al., 2025). The final stage focused on data analysis, where recorded speech samples were transcribed and analyzed quantitatively using fluency metrics and statistically examined using inferential techniques to determine differences in fluency gains between the two groups. This systematic procedure ensured consistency, comparability, and methodological rigor throughout the study.

Instruments, and Data Collection Techniques

The instruments used in this study included an Oral Fluency Assessment Task, an analytic speaking rubric, and a background questionnaire. The Oral Fluency Assessment Task required participants to perform structured and semi-spontaneous speaking activities, such as picture-based narration and opinion-based monologues, which were audio-recorded for analysis. Oral fluency was evaluated using an analytic rubric measuring speech rate, mean length of runs, frequency of pauses, lexical diversity, and overall communicative smoothness. Inter-rater reliability was established by involving two trained raters, yielding a high level of agreement. The background questionnaire was administered to collect demographic information and control for variables such as prior experience with AI-based learning tools (Schmidt et al., 2025).

Data Analysis Technique

Data analysis was conducted using a combination of descriptive and inferential statistical techniques to examine differences in oral fluency development between the two groups. Descriptive statistics, including means and standard deviations, were first calculated to summarize pretest and posttest fluency scores. To test the effectiveness of the instructional treatments, an independent samples t-test was employed to compare posttest gains between the AI-powered tutor group and the human tutor group, while a paired samples t-test was used to examine within-group improvements from pretest to posttest. Prior to hypothesis testing, assumptions of normality and homogeneity of variance were assessed to ensure the appropriateness of parametric analysis. Statistical significance was determined at the 0.05 level, and effect size measures were reported to indicate the magnitude of observed differences, thereby providing a comprehensive interpretation of the quantitative findings.

RESULTS AND DISCUSSION

The descriptive analysis of the pretest and posttest oral fluency scores indicates measurable improvement in both the AI-powered tutor group and the human tutor group. At baseline, the mean oral fluency score for the AI-powered tutor group was 62.45 (SD = 6.18), while the human tutor group recorded a comparable mean score of 63.10 (SD = 6.02), indicating equivalent initial proficiency levels. After the eight-week intervention, the AI-powered tutor group achieved a posttest mean score of 71.38 (SD = 5.47), whereas the human tutor group reached a higher mean score of 75.92 (SD = 5.21). These results suggest overall progress in oral fluency across both instructional modalities.

Table 1 presents a detailed comparison of pretest and posttest oral fluency scores for both groups. The table demonstrates that although both groups experienced improvement, the magnitude of gain differed, with the human tutor group showing a larger mean increase. The consistency of standard deviation values across testing phases indicates stable score dispersion and supports the reliability of the observed improvement patterns.

Table 1. Descriptive Statistics of EFL Oral Fluency Scores by Tutoring Modality

Group	Test Phase	Mean	SD	Gain
AI-Powered Tutor	Pretest	62.45	6.18	—
AI-Powered Tutor	Posttest	71.38	5.47	+8.93
Human Tutor	Pretest	63.10	6.02	—
Human Tutor	Posttest	75.92	5.21	+12.82

The descriptive data show that the AI-powered tutor group demonstrated notable gains in temporal fluency indicators such as speech rate and reduced silent pauses. Learners in this group produced longer stretches of uninterrupted speech and displayed improved automaticity

during posttest performances. These improvements were particularly evident in structured monologue tasks, where learners benefited from repeated practice and immediate automated feedback.

The human tutor group showed stronger gains across both temporal and interactional fluency dimensions. In addition to faster speech rates, learners exhibited greater discourse coherence, pragmatic appropriateness, and confidence during interactive speaking tasks. These descriptive patterns suggest that human tutoring may provide advantages in managing spontaneous interaction and contextualized feedback, which are critical for advanced oral fluency development.

Inferential statistical analysis was conducted using paired-sample and independent-sample t-tests to examine within-group and between-group differences. Paired-sample t-tests revealed statistically significant improvements in oral fluency for both the AI-powered tutor group ($t(39) = -9.84$, $p < .001$) and the human tutor group ($t(39) = -14.27$, $p < .001$). These results confirm that both instructional modalities were effective in enhancing EFL oral fluency over the intervention period.

Independent-sample t-test analysis of posttest scores revealed a statistically significant difference favoring the human tutor group ($t(78) = -3.62$, $p < .001$). The effect size, calculated using Cohen's d, was 0.81, indicating a large practical effect. These findings demonstrate that while AI-powered tutors support fluency development, human tutors yielded significantly greater gains in overall oral fluency performance.

Correlation analysis was conducted to explore relationships between specific fluency subcomponents and overall oral fluency gains. In the AI-powered tutor group, speech rate showed a strong positive correlation with overall fluency improvement ($r = .72$, $p < .001$), indicating that increased automatization was a key contributor to performance gains. Pause frequency demonstrated a moderate negative correlation with fluency scores ($r = -.58$, $p < .01$), suggesting that reduced hesitation played an important role.

In the human tutor group, lexical diversity and discourse coherence were more strongly correlated with overall fluency improvement ($r = .75$, $p < .001$ and $r = .78$, $p < .001$, respectively). These relationships indicate that qualitative aspects of speech production were central to fluency development in human-mediated instruction. The relational data suggest that different tutoring modalities may influence distinct dimensions of oral fluency.

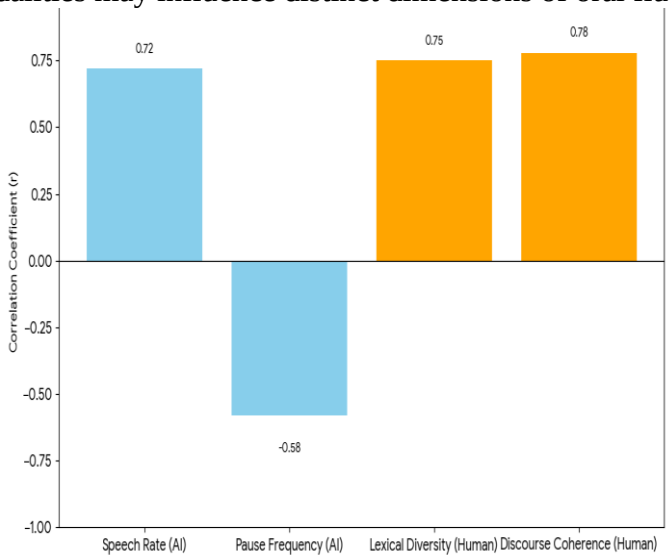

Figure 1. Correltaion Between Fluency Subcomponent and Overall Gains

A focused case analysis further illustrates these quantitative findings. One learner from the AI-powered tutor group demonstrated substantial improvement in speech rate and reduced filler use during monologic tasks, progressing from fragmented speech to more continuous

production. Audio analysis showed increased mean length of runs and faster articulation rates, reflecting gains in automatic speech processing (Heydarnejad, 2025).

A contrasting case from the human tutor group revealed improvement not only in speech smoothness but also in interactional competence. The learner demonstrated improved turn-taking, appropriate use of discourse markers, and increased responsiveness to interlocutor cues during dialogic tasks. Tutor-mediated scaffolding and immediate corrective feedback appeared to support these multidimensional gains, highlighting the pedagogical value of human interaction.

The explanatory analysis suggests that AI-powered tutors are particularly effective in supporting repetitive practice, immediate feedback, and anxiety-free speaking environments. These features appear to facilitate temporal fluency development by reducing cognitive load and encouraging frequent oral production (Mahmoudi-Dehaki & Nasr-Esfahani, 2025; Zhuang et al., 2025). Learners benefited from the non-judgmental nature of AI systems, which promoted risk-taking and sustained practice.

Human tutors, by contrast, appear to offer advantages in adaptive feedback, pragmatic modeling, and socio-interactional support. The presence of human tutors allowed for real-time negotiation of meaning and personalized scaffolding, which may explain the stronger gains observed in qualitative fluency dimensions. These explanatory patterns align with interactionist theories of second language acquisition.

Overall interpretation of the results indicates that both AI-powered tutors and human tutors contribute positively to EFL oral fluency development, but they do so through different mechanisms. AI-powered tutors effectively enhance automatization and temporal fluency, while human tutors exert a stronger influence on interactional and discourse-level fluency. The statistically significant advantage of human tutoring suggests that human interaction remains critical for advanced oral fluency development.

The findings imply that AI-powered tutors should be viewed not as replacements but as complementary tools within EFL speaking instruction. Strategic integration of AI systems for practice and automatization, alongside human-led interaction for higher-order communicative skills, may represent the most pedagogically effective approach. These results provide empirical guidance for evidence-based adoption of AI technologies in language education.

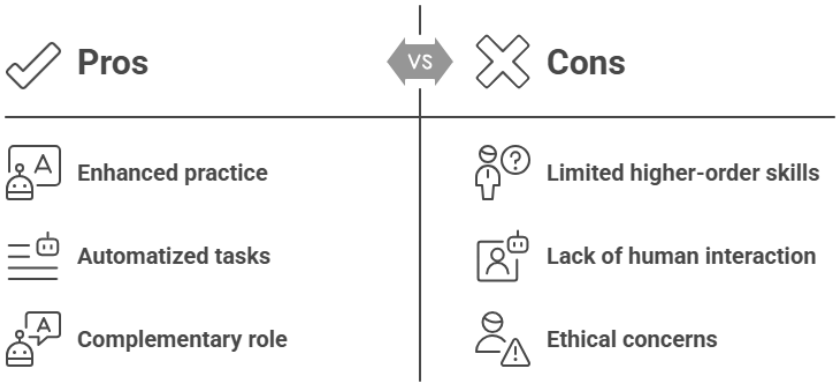

Figure 2. AI Tutors in EFL

The findings of this study demonstrate that both AI-powered tutors and human tutors contribute positively to the development of EFL oral fluency, albeit with differing degrees and mechanisms of impact. Learners in both instructional conditions showed statistically significant improvements from pretest to posttest, confirming that structured, speaking-focused tutoring regardless of modality enhances oral performance. The comparative analysis revealed that the human tutor group achieved significantly higher overall fluency gains than the AI-powered tutor group, particularly in discourse coherence, interactional competence, and pragmatic

appropriateness. These results suggest that while AI-powered tutors effectively support fluency development, human tutors remain more effective in fostering higher-level communicative abilities.

The results further indicate that AI-powered tutors were especially effective in improving temporal aspects of fluency, such as speech rate, reduced pausing, and length of fluent runs. Learners interacting with AI systems benefited from frequent, low-pressure speaking practice and immediate automated feedback, which facilitated automatisation of speech production. In contrast, learners guided by human tutors demonstrated more balanced development across both temporal and qualitative dimensions of fluency. Improvements in lexical variety, discourse organization, and conversational responsiveness were more pronounced in this group, highlighting the pedagogical value of human-mediated interaction (Al-Bogami & Alahmadi, 2025; Yang & Zhao, 2024).

The experimental design strengthens confidence in these findings by controlling instructional content, duration, and learner proficiency. The observed differences therefore reflect genuine modality effects rather than instructional inconsistencies. Overall, the results confirm that AI-powered tutoring is a viable tool for supporting EFL oral fluency development, but it does not yet match the comprehensive communicative impact of human tutoring.

The results of this study align with previous research suggesting that technology-enhanced language learning can significantly improve speaking performance, particularly in terms of practice frequency and learner autonomy. Prior studies have shown that AI-based systems are effective in promoting pronunciation accuracy, speech rate, and confidence through repetitive and individualized practice. The present findings support these conclusions by demonstrating meaningful gains in temporal fluency among learners using AI-powered tutors. The consistency of these results reinforces the view that AI technologies are well-suited to supporting automatization processes in second language acquisition.

At the same time, the findings diverge from studies that claim parity between AI tutors and human tutors in speaking development. While some research suggests that advanced AI systems can approximate human feedback, the current study provides empirical evidence that human tutors still outperform AI in fostering interactional and discourse-level fluency. This difference may stem from the human tutor's ability to interpret learner intent, negotiate meaning dynamically, and provide context-sensitive scaffolding that AI systems currently lack (Dwivedi et al., 2023; Zhang, 2025).

The results also resonate with interactionist and sociocultural theories of second language acquisition, which emphasize the role of meaningful interaction and social mediation in language development. Human tutoring environments inherently provide opportunities for negotiation of meaning, corrective feedback embedded in discourse, and affective support, all of which are less developed in AI-mediated contexts. The present study thus contributes to the literature by empirically confirming theoretical claims about the enduring importance of human interaction in oral language development.

The findings signal an important developmental distinction between different dimensions of oral fluency. Improvements driven by AI-powered tutors appear to reflect gains in proceduralization and automaticity, indicating progress in lower-level speech processing. This pattern suggests that AI-supported practice effectively reduces cognitive load during speech production, enabling learners to speak more smoothly and with fewer hesitations. Such outcomes are indicative of fluency growth at the performance level.

The stronger gains observed in the human tutor group signal the importance of social and communicative engagement in developing higher-order fluency skills. These results indicate that oral fluency is not merely a function of speed or continuity, but also of interactional competence, pragmatic awareness, and discourse management. The findings therefore highlight oral fluency as a multidimensional construct that benefits from both automated practice and socially mediated learning.

The results also signal a transitional phase in educational technology adoption. AI-powered tutors are no longer experimental novelties but functional instructional tools capable of producing measurable learning gains. However, the findings simultaneously indicate that AI systems have not yet reached a level of pedagogical sophistication sufficient to replace human tutors in communicative skill development. This dual signal underscores the need for balanced and evidence-based integration of AI into language education.

The implications of these findings are significant for EFL pedagogy and instructional design. The results suggest that AI-powered tutors can be effectively integrated into speaking curricula as supplementary tools that provide learners with increased practice opportunities outside the classroom. Their ability to offer immediate feedback and unlimited practice makes them particularly valuable in contexts where instructional time and access to qualified tutors are limited. Educators can leverage AI systems to reinforce fluency practice and reduce learner anxiety associated with speaking (Alsswey et al., 2025; Xie et al., 2025).

The findings also imply that human tutors should remain central to speaking instruction, especially for developing interactive and pragmatic aspects of oral fluency. Human-mediated instruction is particularly important for tasks involving negotiation of meaning, spontaneous interaction, and discourse construction. These results caution against replacing human tutors entirely with AI systems, especially in advanced stages of oral proficiency development.

At an institutional level, the study provides evidence to inform policy decisions regarding investment in educational technology. Rather than adopting AI-powered tutors as cost-saving substitutes, institutions should consider hybrid instructional models that combine AI-based practice with human-led interaction. Such models can maximize learning outcomes while addressing practical constraints related to scalability and resource allocation.

The observed results can be explained by fundamental differences in how AI-powered tutors and human tutors mediate learning. AI systems excel at providing consistent, immediate, and non-judgmental feedback, which supports repetitive practice and reduces performance anxiety. These features create favorable conditions for developing automaticity and temporal fluency, particularly for learners who are hesitant to speak in front of others.

Human tutors, by contrast, offer adaptive feedback that responds to learner intent, emotional cues, and communicative context. Their ability to scaffold interaction, model pragmatic language use, and adjust instruction in real time likely explains the stronger gains in discourse-level fluency. Human tutors also facilitate meaningful social interaction, which is a critical driver of language development according to sociocultural perspectives.

The differential outcomes observed in this study therefore reflect complementary rather than competing strengths. AI-powered tutors support the mechanical and procedural aspects of fluency, while human tutors facilitate the social and cognitive dimensions of communication. The results suggest that the observed superiority of human tutors arises not from greater practice quantity, but from richer interaction quality.

Future research should explore instructional models that strategically combine AI-powered tutors and human tutors to optimize oral fluency development. Longitudinal studies examining sustained hybrid instruction would provide insight into how different tutoring modalities interact over time. Such research could determine whether early AI-supported automatization enhances the effectiveness of later human-mediated instruction.

Further research should also investigate learner variables such as motivation, anxiety, and self-regulation in relation to AI and human tutoring. Understanding how different learners respond to each modality would allow for more personalized instructional designs. Additionally, qualitative studies examining learner perceptions and interactional dynamics could deepen understanding of how AI systems are experienced in real learning contexts.

Finally, future studies should explore advancements in AI conversational agents, particularly those incorporating pragmatic modeling and adaptive discourse strategies. As AI technology evolves, ongoing empirical evaluation will be essential to ensure pedagogical

decisions remain grounded in evidence rather than technological optimism. The present study provides a foundational benchmark against which future developments in AI-assisted language learning can be evaluated.

CONCLUSION

The most significant finding of this study is the differentiated impact of AI-powered tutors and human tutors on EFL oral fluency development. Both instructional modalities produced statistically significant gains, confirming their effectiveness in enhancing learners' speaking performance. Distinct patterns emerged in the dimensions of fluency development, where AI-powered tutors demonstrated stronger effects on temporal fluency indicators such as speech rate and reduced hesitation, while human tutors yielded superior outcomes in interactional fluency, discourse coherence, and pragmatic appropriateness. These findings underscore that oral fluency development is multidimensional and that instructional modality shapes the nature of learners' communicative gains.

The primary contribution of this research lies in its conceptual and methodological advancement of comparative EFL instruction. Conceptually, the study provides empirical clarification on how AI-mediated and human-mediated tutoring differentially support distinct components of oral fluency, moving beyond simplistic effectiveness claims. Methodologically, the use of a controlled experimental design with equivalent instructional content allows for a robust comparison that isolates modality effects. This dual contribution strengthens the theoretical understanding of technology-assisted language learning while offering evidence-based guidance for hybrid instructional models in EFL contexts.

The study is subject to several limitations that suggest directions for future research. The relatively short intervention period limits insight into long-term fluency development and retention effects. The participant sample, drawn from a single educational context, constrains generalizability across proficiency levels and cultural settings. Future studies should employ longitudinal designs, larger and more diverse samples, and mixed-method approaches to examine affective factors, learner engagement, and interactional dynamics. Further investigation into advanced AI conversational capabilities is also warranted to assess whether emerging technologies can more effectively approximate human-mediated communicative scaffolding.

AUTHOR CONTRIBUTIONS

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

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Adorni, G., Artico, I., Piatti, A., Lutz, E., Gambardella, L. M., Negrini, L., Mondada, F., & Assaf, D. (2024). Development of algorithmic thinking skills in K-12 education: A comparative study of unplugged and digital assessment instruments. *Computers in Human Behavior Reports*, *15*, 100466. https://doi.org/https://doi.org/10.1016/j.chbr.2024.100466

Adorni, G., & Piatti, A. (2025). Designing the virtual CAT: A digital tool for algorithmic thinking assessment in compulsory education. *International Journal of Child-Computer Interaction*, *45*, 100760. https://doi.org/https://doi.org/10.1016/j.ijcci.2025.100760

Al-Bogami, R. M., & Alahmadi, N. A. (2025). Effects of an AI-based reading progress tool on third-grade EFL learners' oral reading fluency. *Computers and Education Open*, *9*, 100283. https://doi.org/https://doi.org/10.1016/j.caeo.2025.100283

Aladini, A., Ismail, S. M., Ahmad Saleem Khasawneh, M., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2 writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports*, *17*, 100566. https://doi.org/https://doi.org/10.1016/j.chbr.2024.100566

Alsswey, A., El-Qirem, F. A., & Omar, F. (2025). 3D holograms and emotional intelligence: Transforming interactive learning in higher education. *Acta Psychologica*, *261*, 105758. https://doi.org/https://doi.org/10.1016/j.actpsy.2025.105758

Avci, H., Lunn, S. J., & Hazari, Z. (2025). Exploring STEM educators' perspectives on the integration of AI-enabled technologies in teaching and learning. *Computers and Education Open*, *9*, 100304. https://doi.org/https://doi.org/10.1016/j.caeo.2025.100304

Brezovec, E., Zelić, M., & Zagode, A. M. (2025). Stabilizing truth in educational sciences: a systematic review of generative AI in education. *Kybernetes*, *55*(13), 1–17. https://doi.org/https://doi.org/10.1108/K-09-2025-2339

Brunton, R. J., Rhazzafe, S., Moodley, R., Kuhn, S., Caraffini, F., Wilford, S., Higginbottom, R., Colreavy-Donnelly, S., & Gongora, M. (2025). Using generative artificial intelligence to enhance the performance of disadvantaged students in secondary education. *Social Sciences & Humanities Open*, *12*, 102110. https://doi.org/https://doi.org/10.1016/j.ssaho.2025.102110

Buciuman, C.-F., & Potra, S. (2025). Revolutionizing Education in Industry 4.0: Eye-Tracking and AI for Personalized Learning. *Procedia Computer Science*, *253*, 1658–1667. https://doi.org/https://doi.org/10.1016/j.procs.2025.01.228

Chen, S., & Cheung, A. C. K. (2025). Effect of generative artificial intelligence on university students learning outcomes: A systematic review and meta-analysis. *Educational Research Review*, *49*, 100737. https://doi.org/https://doi.org/10.1016/j.edurev.2025.100737

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2023.102642

Ebadi, S., Velayati, S., Ramezanzadeh, A., & Rawdhan Salman, A. (2025). Exploring the impact of AI-powered speaking tasks on EFL learners' speaking performance and anxiety: An activity theory study. *Acta Psychologica*, *259*, 105391. https://doi.org/https://doi.org/10.1016/j.actpsy.2025.105391

Görer, B., & Aydemir, F. B. (2024). Exploring the REIT architecture for requirements elicitation interview training with robotic and virtual tutors. *Journal of Systems and Software*, *212*, 112018. https://doi.org/https://doi.org/10.1016/j.jss.2024.112018

Guo, K., Zhang, J., & Ansari, H. W. A. (2025). Teacher care and mental wellbeing: Exploring the role of grit, resilience, and AI-interaction in education management. *Acta*

Psychologica, *261*, 105977. https://doi.org/https://doi.org/10.1016/j.actpsy.2025.105977

Heydarnejad, T. (2025). Unmasking the Impacts of Self-Evaluation in AI-Supported Writing Instruction on EFL Learners' Emotion Regulation, Self-Competence, Motivation, and Writing Achievement. *Computers and Education: Artificial Intelligence*, 100494. https://doi.org/https://doi.org/10.1016/j.caeai.2025.100494

Huwer, J., Thyssen, C., Becker-Genschow, S., von Kotzebue, L., Finger, A., Kremser, E., Berber, S., Brückner, M., Maurer, N., Bruckermann, T., Meier, M., & Thoms, L.-J. (2025). Competencies for teaching with and about artificial intelligence in the natural sciences — DiKoLAN AI. *Computers and Education Open*, *9*, 100303. https://doi.org/https://doi.org/10.1016/j.caeo.2025.100303

Kohnke, L., & Moorhouse, B. L. (2025). Enhancing the emotional aspects of language education through generative artificial intelligence (GenAI): A qualitative investigation. *Computers in Human Behavior*, *167*, 108600. https://doi.org/https://doi.org/10.1016/j.chb.2025.108600

Korzynski, P., Edwards, A., Gupta, M. C., Mazurek, G., & Wirtz, J. (2025). Humanoid robotics and agentic AI: reframing management theories and future research directions. *European Management Journal*, *43*(4), 548–560. https://doi.org/https://doi.org/10.1016/j.emj.2025.06.002

Lachheb, A., Leung, J., Abramenka-Lachheb, V., & Sankaranarayanan, R. (2025). AI in higher education: A bibliometric analysis, synthesis, and a critique of research. *The Internet and Higher Education*, *67*, 101021. https://doi.org/https://doi.org/10.1016/j.iheduc.2025.101021

Lai, W. Y. W., & Lee, J. S. (2024). A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents. *Computers and Education: Artificial Intelligence*, *7*, 100291. https://doi.org/https://doi.org/10.1016/j.caeai.2024.100291

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, *6*, 100174. https://doi.org/https://doi.org/10.1016/j.caeo.2024.100174

Lissack, M., & Meagher, B. (2024). Responsible Use of Large Language Models: An Analogy with the Oxford Tutorial System. *She Ji: The Journal of Design, Economics, and Innovation*, *10*(4), 389–413. https://doi.org/https://doi.org/10.1016/j.sheji.2024.11.001

Liu, G. L., Lee, J. S., & Zhao, X. (2025). Critical digital literacies, agentic practices, and AI-mediated informal digital learning of English. *System*, *134*, 103797. https://doi.org/https://doi.org/10.1016/j.system.2025.103797

Liu, Y., Zhang, H., Jiang, M., Chen, J., & Wang, M. (2024). A systematic review of research on emotional artificial intelligence in English language education. *System*, *126*, 103478. https://doi.org/https://doi.org/10.1016/j.system.2024.103478

Mahmoudi-Dehaki, M., & Nasr-Esfahani, N. (2025). Empowering stuttering female English learners: AI vs. human-AI hybrid tutoring for alleviating social anxiety. *Journal of Responsible Technology*, *24*, 100141. https://doi.org/https://doi.org/10.1016/j.jrt.2025.100141

Menon, D., & Shilpa, K. (2023). "Hey, Alexa" "Hey, Siri", "OK Google" …." exploring teenagers' interaction with artificial intelligence (AI)-enabled voice assistants during the COVID-19 pandemic. *International Journal of Child-Computer Interaction*, *38*, 100622.

https://doi.org/https://doi.org/10.1016/j.ijcci.2023.100622

Mim, J. F., Islam, M. M., & Raza, A. H. (2025). A hybrid MCDM approach for unveiling ChatGPT's effect on students' learning. *Computers and Education: Artificial Intelligence*, *9*, 100457. https://doi.org/https://doi.org/10.1016/j.caeai.2025.100457

Namaziandost, E. (2025). Integrating flipped learning in AI-enhanced language learning: Mapping the effects on metacognitive awareness, writing development, and foreign language learning boredom. *Computers and Education: Artificial Intelligence*, *9*, 100446. https://doi.org/https://doi.org/10.1016/j.caeai.2025.100446

Nandagopal, S. (2025). Transforming the self: Individual-level changes arising from collaboration with generative AI. *Computers in Human Behavior: Artificial Humans*, *6*, 100232. https://doi.org/https://doi.org/10.1016/j.chbah.2025.100232

Ng, M. L., Behforouz, B., & Ghaithi, A. Al. (2025). Grammar and engagement in focus: Evaluating Gemini AI's impact on an educational environment. *Computers and Education Open*, *9*, 100302. https://doi.org/https://doi.org/10.1016/j.caeo.2025.100302

Park, A., & Kim, T. (2025). Code suggestions and explanations in programming learning: Use of ChatGPT and performance. *The International Journal of Management Education*, *23*(2), 101119. https://doi.org/https://doi.org/10.1016/j.ijme.2024.101119

Schmidt, D. A., Alboloushi, B., Thomas, A., & Magalhaes, R. (2025). Integrating artificial intelligence in higher education: perceptions, challenges, and strategies for academic innovation. *Computers and Education Open*, *9*, 100274. https://doi.org/https://doi.org/10.1016/j.caeo.2025.100274

Shahini, A. (2025). Emotional dimensions of feedback: How AI and human responses shape ESL learning outcomes. *Ampersand*, *15*, 100235. https://doi.org/https://doi.org/10.1016/j.amper.2025.100235

Şimşek, A. C., Anders, G., Göth, J., Specht, L., & Huff, M. (2025). Is ChatGPT a good study companion? The role of AI-generated summaries and reflective prompts in learning from educational videos. *Computers and Education: Artificial Intelligence*, *9*, 100512. https://doi.org/https://doi.org/10.1016/j.caeai.2025.100512

Sun, H., Zhang, T., Han, J., & Chu, H. (2024). A fast transfer reinforcement learning model for transferring force-based human speed adjustment skills to robots for collaborative assembly posture alignment. *Advanced Engineering Informatics*, *62*, 102836. https://doi.org/https://doi.org/10.1016/j.aei.2024.102836

Tian, Q., & Zheng, X. (2025). The impact of artificial intelligence on students' 4C skills: A meta-analysis. *Educational Research Review*, *49*, 100728. https://doi.org/https://doi.org/10.1016/j.edurev.2025.100728

Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. *System*, *127*, 103528. https://doi.org/https://doi.org/10.1016/j.system.2024.103528

von Garrel, J., & Mayer, J. (2024). Which features of AI-based tools are important for students? A choice-based conjoint analysis. *Computers and Education: Artificial Intelligence*, *7*, 100311. https://doi.org/https://doi.org/10.1016/j.caeai.2024.100311

Xie, L., Jiang, Y., Chang, C.-N., Zeng, X.-Y., Hong, J., & Mo, F. (2025). How are faculty and college students embracing AI? — A multi-informant mixed method study. *Computers and Education: Artificial Intelligence*, *9*, 100506. https://doi.org/https://doi.org/10.1016/j.caeai.2025.100506

Yang, L., & Zhao, S. (2024). AI-induced emotions in L2 education: Exploring EFL students' perceived emotions and regulation strategies. *Computers in Human Behavior*, *159*, 108337. https://doi.org/https://doi.org/10.1016/j.chb.2024.108337

Zhai, X., & Li, S. (2025). The roles of growth mindset, resilience, and self-efficacy in student Engagement with AI-enhanced Chinese learning: A self-determination theory perspective. *Learning and Motivation*, *92*, 102183. https://doi.org/https://doi.org/10.1016/j.lmot.2025.102183

Zhang, Z. (2025). Enhancing English listening comprehension via AI - based adaptive learning platforms incorporating speech - to - text and predictive analytics. *Systems and Soft Computing*, *7*, 200418. https://doi.org/https://doi.org/10.1016/j.sasc.2025.200418

Zhuang, M., Long, S., Martin, F., & Castellanos-Reyes, D. (2025). The affordances of Artificial Intelligence (AI) and ethical considerations across the instruction cycle: A systematic review of AI in online higher education. *The Internet and Higher Education*, *67*, 101039. https://doi.org/https://doi.org/10.1016/j.iheduc.2025.101039