Language Inquiry & Exploration Review

e-ISSN 3062-7702 (online)

Volume 2, Issue 1 2025
Page 37-51

https://doi.org/10.71435/639189

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

Khaisya Sabina Nazarul Ummi[1]

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

ARTICLE INFO

Received: 02 April 2025
Revised: 19 April 2025
Accepted: 16 May 2025
Available online: 03 June 2025

Keywords:
ChatGPT
EFL Learners
Conversational AI

Corresponding Author:
Khaisya Sabina Nazarul Ummi

Email:
khaisya@gmail.com

Copyright © 2025, Language Inquiry & Exploration Review, Under the license CC BY- SA 4.0

OPEN ACCESS

ABSTRACT

***Purpose**: This study examined the effectiveness of ChatGPT as a conversational partner for Indonesian EFL learners by investigating changes in speaking performance, learners' perceptions, and the association between perceived usefulness and speaking gains.*

***Subjects and Methods:** The study employed a quantitative one-group pretest–posttest design involving 40 intermediate-level EFL learners in Makassar, South Sulawesi, Indonesia. Participants engaged in structured ChatGPT conversations for four weeks, comprising three 30-minute sessions per week. Speaking performance was assessed before and after the intervention across fluency, vocabulary, and grammatical accuracy by two EFL instructors. Learners' perceptions were measured using a five-point Likert-scale questionnaire. Data were analyzed using descriptive statistics, Shapiro–Wilk testing, paired-samples t-tests, Cohen's d, and Pearson correlation.*

***Results:** Speaking performance increased across all measured dimensions. Fluency increased from 5.60 to 7.05, vocabulary from 5.45 to 6.80, and grammatical accuracy from 5.10 to 6.20, while the total speaking score increased from 16.15 to 20.05. All differences were statistically significant ($p < .001$), with large effect sizes for individual dimensions and a very large effect for the total score ($d = 1.77$). Learners reported favorable perceptions, with mean scores ranging from 4.45 to 4.70. Perceived usefulness was positively associated with speaking gains ($r = .67$, $p < .001$), indicating that more favorable evaluations tended to coincide with greater improvement. However, the findings represent within-participant changes rather than definitive causal effects because no control group was included.*

***Conclusions:** ChatGPT shows promise as a supplementary conversational partner for Indonesian EFL speaking practice, particularly when interaction is structured and purposeful. Further controlled, longitudinal, and multi-institutional research is recommended to establish its effectiveness and generalizability.*

INTRODUCTION

The rapid development of generative artificial intelligence (GenAI) has begun to reshape language education by providing learners with opportunities to access personalized, interactive, and on-demand language practice (Lopez-Gazpio, 2025; Sadigzade, 2025; Mimoudi, 2025). Among the emerging technologies, ChatGPT has attracted considerable attention because its conversational capabilities enable learners to engage in open-ended exchanges, receive immediate responses, and practice language use beyond the physical boundaries of the classroom.

In English as a Foreign Language (EFL) contexts, these affordances are particularly relevant because opportunities for sustained interaction in English are often constrained by limited classroom time, large numbers of learners, and unequal access to authentic communication environments. Recent systematic reviews indicate that conversational artificial intelligence has produced promising cognitive and affective outcomes in language learning, while also emphasizing the need for more rigorous empirical research that examines behavioral and language-performance outcomes rather than relying primarily on learner perceptions (Huang et al., 2022; Yang & Kyun, 2022).

A systematic review of conversational AI in English language teaching further shows that although research in this area has expanded rapidly, rigorous investigations of newer generative AI tools and measurable language outcomes remain necessary. Speaking is one of the areas in which conversational AI may offer particularly meaningful pedagogical opportunities (Du & Daniel, 2024; Ding & Yusof, 2025; Ericsson & Johansson, 2023; Lin & Mubarok, 2021). Unlike language-learning applications that rely primarily on predetermined exercises or scripted dialogues, generative AI can sustain flexible interactions and respond to learner-generated language in context.

Such interaction may provide repeated opportunities for learners to produce language, negotiate meaning, retrieve vocabulary, and formulate grammatical structures while reducing some of the interpersonal pressure associated with speaking in front of teachers or peers. Empirical research has increasingly reported positive outcomes from AI-mediated speaking activities. For example, AI-supported interactive speaking activities have been associated with improvements in speaking proficiency and willingness to communicate among EFL learners, while generative AI chatbot interventions have also demonstrated potential for improving speaking performance and reducing speaking-related anxiety (Tai & Chen, 2024).

More recent evidence from university-level EFL learners indicates that interaction with ChatGPT can contribute to improved oral task performance, particularly when the interaction is structured around clearly defined speaking tasks (Sok & Shin, 2025). These findings suggest that the pedagogical value of conversational AI may depend not simply on access to the technology but on how learners are guided to use it for sustained and purposeful language production.

Despite these promising findings, important methodological and contextual gaps remain. Existing research on AI-assisted language learning has frequently examined learner attitudes, motivation, anxiety, or general perceptions of technological usefulness, whereas objective measures of changes in specific speaking dimensions remain comparatively less developed. Systematic reviews have similarly identified the need for research that moves beyond general perceptions and investigates measurable behavioral and learning outcomes using more rigorous research designs.

In particular, fluency, lexical development, and grammatical accuracy represent important dimensions of spoken language performance, yet these dimensions are not consistently examined together in studies of conversational generative AI. Research has also been conducted across different educational settings, proficiency levels, AI platforms, and interaction formats, making it difficult to determine whether findings obtained in one context can be directly transferred to other EFL environments.

The Indonesian context provides a relevant setting for examining these issues. English is taught as a foreign language, and opportunities for meaningful oral interaction may remain more limited than opportunities for grammar- and writing-oriented classroom activities. Previous research has highlighted the potential of digital and AI-supported learning to provide additional opportunities for Indonesian EFL learners, but the empirical literature remains developing and varies substantially in methodological approach. A recent systematic review of ChatGPT research in Indonesian English language teaching found that higher education has become one of the most frequently investigated settings, while qualitative approaches continue to dominate the literature.

At the same time, empirical studies in Indonesia have begun to report positive effects of ChatGPT-supported speaking activities. For example, a 2025 quasi-experimental study involving Indonesian university students found significant improvements in speaking performance among

learners who participated in ChatGPT-supported activities compared with a control group. Zein Zein et al. (2020) and Hamid et al. (2025) said that, differences in institutional context, participant characteristics, intervention design, and assessment procedures indicate that further evidence from other Indonesian EFL settings remains valuable.

Another issue requiring further attention concerns the relationship between learners' perceptions of AI and their objectively measured language development. Positive perceptions of usefulness, convenience, confidence, and satisfaction are frequently reported in studies of AI-assisted language learning, but perceived usefulness should not automatically be interpreted as evidence of actual learning gains. Establishing whether learners who perceive ChatGPT more positively also demonstrate greater improvement in objectively assessed speaking performance can provide a more nuanced understanding of the relationship between technology acceptance and language outcomes. This distinction is important because the educational value of conversational AI should be evaluated through both measurable performance and learners' experiences with the technology. Recent empirical research has begun to examine these dimensions together, but evidence remains insufficient regarding how perceived usefulness relates specifically to changes in speaking performance among Indonesian university EFL learners (Sok & Shin, 2025).

Against this background, the present study investigates the use of ChatGPT as a conversational partner for Indonesian EFL learners in a higher-education setting in Makassar, South Sulawesi. The study adopts a quantitative one-group pretest–posttest design involving 40 intermediate-level EFL learners who participated in structured ChatGPT conversations over four weeks. Speaking performance is examined across three dimensions: fluency, vocabulary, and grammatical accuracy. In addition, the study investigates learners' perceptions of ChatGPT as a conversational partner, including perceived usefulness, confidence, ease of use, motivation, and satisfaction. The study also examines the association between perceived usefulness and gains in speaking performance. By combining objective speaking assessment with learner-perception data, this study seeks to contribute empirical evidence regarding the potential role of conversational generative AI in supporting EFL speaking practice in the Indonesian higher-education context.

Accordingly, the study addresses three research questions. First, to what extent do Indonesian EFL learners' fluency, vocabulary, and grammatical accuracy change following structured conversational practice with ChatGPT? Second, how do learners perceive the usefulness of ChatGPT as a conversational partner for English-speaking practice? Third, is learners' perceived usefulness of ChatGPT associated with their gains in speaking performance? By addressing these questions, the study contributes to the growing literature on generative AI in language education while maintaining a cautious interpretation of pretest–posttest evidence and recognizing the need for further controlled research to establish causal effects.

METHODOLOGY

Research Design

This study employed a quantitative pre-experimental one-group pretest–posttest design to examine changes in Indonesian EFL learners' speaking performance following structured conversational practice with ChatGPT. The design involved assessing participants' speaking performance before and after a four-week intervention and comparing the two sets of scores across three dimensions of oral performance: fluency, vocabulary, and grammatical accuracy. A one-group pretest–posttest design was selected because it enabled the study to examine within-participant changes following exposure to the conversational intervention. However, because the study did not include a control or comparison group, the findings are interpreted as evidence of changes following the intervention rather than definitive evidence of a causal effect of ChatGPT. The study also incorporated learner-perception data to provide complementary evidence regarding participants' experiences with ChatGPT as a conversational partner. Objective speaking performance data were therefore analyzed together with questionnaire responses concerning perceived usefulness, ease of use, confidence, motivation, engagement, speaking anxiety, and satisfaction. The combination of performance and perception measures was intended to provide a more comprehensive assessment of the potential value of ChatGPT for EFL speaking practice.

Participants and Sampling

The participants were 40 intermediate-level EFL learners enrolled in English-language courses across faculties in Makassar, South Sulawesi, Indonesia. Participants were selected using purposive sampling based on predefined eligibility criteria. The primary criterion was intermediate-level English proficiency, determined through an institutional English-language placement assessment. Participants were also required to have sufficient access to a smartphone or laptop and a stable internet connection to participate in the ChatGPT-based conversational activities. The selection of intermediate-level learners was intended to ensure that participants possessed sufficient English proficiency to sustain interactions with ChatGPT while still having identifiable areas for development in spoken English. Participants were also required to be available throughout the four-week intervention period. A total of 40 learners who met the study criteria participated in both the pretest and posttest assessments. The study was conducted in Makassar, South Sulawesi, Indonesia. Makassar provides a substantial higher-education context for examining technology-supported EFL learning. According to BPS-Statistics Makassar Municipality, the city recorded 92 higher-education institutions, 133,344 higher-education students, and 6,243 teaching staff in 2024. These figures illustrate the scale of the higher-education environment in which digital and AI-supported learning practices can be examined. Within this broader context, the present study focused specifically on 40 intermediate-level EFL learners who participated in the ChatGPT-based conversational intervention.

Research Instruments

Two principal instruments were used to collect the study data: a speaking performance assessment rubric and a structured learner-perception questionnaire. The speaking assessment was conducted before and after the intervention to measure changes in participants' oral English performance. The assessment focused on three dimensions: fluency, vocabulary, and grammatical accuracy. The assessment rubric was adapted from relevant IELTS speaking descriptors to provide structured criteria for evaluating oral performance. Each participant completed a five-minute speaking task based on a familiar topic, such as education, hobbies, or culture. The speaking performances were recorded and subsequently evaluated by two certified EFL instructors with experience in oral-language assessment. The same assessment procedure and scoring rubric were used for both the pretest and posttest to maintain consistency across measurement occasions. The second instrument was a structured questionnaire administered after the intervention. The questionnaire consisted of 20 items measured using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The items addressed participants' perceptions of ChatGPT's ease of use, usefulness for speaking practice, perceived improvement in speaking ability, confidence, motivation, engagement, speaking anxiety, and overall satisfaction. Before its administration, the questionnaire was piloted with 10 students who did not participate in the main study. The pilot process was used to assess the clarity and comprehensibility of the questionnaire items and to identify potential problems in the instrument.

Intervention Procedure

The intervention was conducted over a period of four weeks. At the beginning of the study, all participants completed a pretest designed to establish their baseline speaking performance. During the pretest, participants delivered a five-minute speech on a familiar topic, and their performances were recorded for subsequent assessment. Following the pretest, participants received an introduction to ChatGPT and guidance on how to use the system for conversational English practice. Participants then engaged in three ChatGPT-based conversational sessions per week for four consecutive weeks, resulting in a total of 12 sessions. Each session lasted approximately 30 minutes. The activities were designed to provide repeated opportunities for participants to use English in sustained interaction with ChatGPT. During the intervention, participants engaged in conversations covering everyday and academic topics. Suggested themes and guiding questions were provided to facilitate interaction, while participants were also permitted to develop the conversations according to their interests. This combination of suggested topics and flexible interaction was intended to create opportunities for extended language production rather than limiting participants to predetermined responses. To document

their participation, learners completed weekly activity logs containing information about the topics discussed, duration of interaction, and reflections on their experiences. Selected interaction sessions were also monitored through screen recordings with participants' agreement. These records were used to monitor the nature and continuity of participants' engagement with the conversational activities. At the end of the fourth week, participants completed the posttest using the same speaking assessment procedure as the pretest. The posttest performances were recorded and evaluated using the same rubric and assessment procedure. Participants subsequently completed the learner-perception questionnaire concerning their experiences with ChatGPT as a conversational partner.

Speaking Performance Assessment

Speaking performance was assessed using three dimensions: fluency, vocabulary, and grammatical accuracy. Fluency represented the participants' ability to maintain oral production with appropriate flow and reduced hesitation. Vocabulary reflected the range and appropriateness of lexical choices used during the speaking task, while grammatical accuracy represented the participants' ability to produce grammatically appropriate English structures. The pretest and posttest recordings were independently evaluated by two certified EFL instructors experienced in oral-language assessment. Both raters used the same assessment rubric and scoring procedures. The use of two raters was intended to reduce dependence on a single evaluator and improve the consistency of the speaking assessment. Inter-rater agreement was considered during the scoring process.

Data Analysis

Quantitative data were analyzed using descriptive and inferential statistical procedures. Descriptive statistics were first calculated to summarize participants' speaking performance before and after the intervention. Mean scores, standard deviations, and mean differences were calculated for fluency, vocabulary, grammatical accuracy, and the total speaking score. Before conducting inferential analysis, the distributional assumption of the difference scores was examined using the Shapiro–Wilk test. Because the study compared speaking scores obtained from the same participants at two measurement points, paired-samples *t*-tests were subsequently conducted to determine whether the differences between pretest and posttest scores were statistically significant. Statistical significance was evaluated at the 0.05 level. In addition to statistical significance, effect sizes were calculated to estimate the magnitude of the observed pretest–posttest differences. For the paired-samples comparisons, Cohen's *d* was calculated based on the paired observations to provide an estimate of the practical magnitude of the changes in speaking performance. Questionnaire responses were analyzed using descriptive statistics, including means, standard deviations, and response frequencies. These statistics were used to describe learners' perceptions of ChatGPT as a conversational partner, particularly with respect to usefulness, ease of use, confidence, motivation, engagement, speaking anxiety, and satisfaction. Finally, a Pearson product–moment correlation analysis was conducted to examine the association between learners' perceived usefulness of ChatGPT and their gains in speaking performance. The analysis was intended to determine whether learners who reported greater perceived usefulness also tended to demonstrate greater pretest–posttest improvement. The correlation was interpreted as an association and not as evidence of a causal relationship.

RESULTS AND DISCUSSION

This section presents the empirical findings of the study examining ChatGPT as a conversational partner for Indonesian EFL learners. The analysis is organized around the three research questions concerning changes in speaking performance, learners' perceptions of ChatGPT, and the association between perceived usefulness and speaking performance gains. The results are presented in four stages. First, descriptive statistics are used to compare participants' pretest and posttest performance in fluency, vocabulary, grammatical accuracy, and total speaking scores. Second, paired-samples *t*-tests and effect sizes are reported to determine the statistical significance and magnitude of the observed changes. Third, questionnaire findings describe learners' perceptions of ChatGPT as a conversational partner. Finally, Pearson correlation analysis examines the relationship between perceived usefulness and speaking score gains. All

analyses are based on the primary data collected from the 40 participants who completed both the pretest and posttest.

Pretest and Posttest Speaking Performance

The first analysis examined changes in participants' speaking performance following the four-week conversational practice period. Speaking performance was assessed across three dimensions: fluency, vocabulary, and grammatical accuracy. A total of 40 intermediate-level EFL learners completed both measurement occasions. The descriptive results are presented in Table 1.

Table 1. Descriptive Statistics of Speaking Performance Before and After the Intervention

Speaking Dimension	N	Pretest Mean	Posttest Mean	Mean Difference
Fluency	40	5.60	7.05	1.45
Vocabulary	40	5.45	6.80	1.35
Grammatical Accuracy	40	5.10	6.20	1.10
Total Speaking Score	40	16.15	20.05	3.90

Source: Primary research data, processed by the authors.

Table 1 shows a consistent increase in the mean scores across all three speaking dimensions. Fluency increased from a pretest mean of 5.60 to a posttest mean of 7.05, resulting in a mean difference of 1.45 points. This represented the largest absolute increase among the three individual speaking dimensions. Vocabulary increased from 5.45 at pretest to 6.80 at posttest, corresponding to a mean difference of 1.35 points. Grammatical accuracy increased from 5.10 to 6.20, producing a mean difference of 1.10 points.

The total speaking score increased from 16.15 at pretest to 20.05 at posttest, corresponding to an overall mean difference of 3.90 points. Thus, all measured dimensions moved in the same positive direction between the two assessment occasions. The descriptive pattern indicates that the largest individual change occurred in fluency, followed by vocabulary and grammatical accuracy.

The original dataset reported standard deviations of 0.91 for fluency, 1.02 for vocabulary, 1.15 for grammatical accuracy, and 1.20 for the total speaking score. Specifically, the reported t value and Cohen's d for the total score imply a paired-difference standard deviation of approximately 2.20.

Statistical Significance and Effect Sizes

To determine whether the observed pretest–posttest differences were statistically significant, paired-samples t-tests were conducted for fluency, vocabulary, grammatical accuracy, and the total speaking score. Cohen's d was also calculated to estimate the magnitude of the paired differences. The results are presented in Table 2.

Table 2. Paired-Samples t-Test and Effect-Size Results

Speaking Dimension	t	df	p	Cohen's d	Effect Size
Fluency	9.81	39	< .001	1.55	Large
Vocabulary	8.76	39	< .001	1.39	Large
Grammatical Accuracy	6.45	39	< .001	1.02	Large
Total Speaking Score	11.22	39	< .001	1.77	Very large

Source: Primary research data, processed by the authors.

Statistically significant pretest–posttest differences across all three speaking dimensions. Fluency produced a t value of 9.81 with 39 degrees of freedom and $p < .001$. The corresponding Cohen's d was 1.55, indicating a large effect. This result is consistent with the descriptive finding that fluency demonstrated the largest mean increase among the three individual dimensions. Vocabulary also showed a statistically significant difference between the two assessment occasions, $t(39) = 8.76$, $p < .001$, with Cohen's $d = 1.39$. The magnitude of the effect was classified as large. The statistical result corresponds to the 1.35-point increase in the vocabulary mean reported in Table 1.

For grammatical accuracy, the paired-samples analysis yielded $t(39) = 6.45$, $p < .001$, with Cohen's $d = 1.02$. Although grammatical accuracy recorded the smallest mean increase among the three individual dimensions, the difference remained statistically significant and was associated with a large effect size. The total speaking score produced the largest reported test statistic and effect size, with $t(39) = 11.22$, $p < .001$, and Cohen's $d = 1.77$. The reported effect size falls within the very-large range and indicates a substantial standardized difference between the two measurement occasions.

Taken together, the inferential results show that all four comparisons reached statistical significance at the .05 level. The direction of every comparison was positive, with posttest scores higher than pretest scores. The magnitude of the reported effect sizes was large for fluency, vocabulary, and grammatical accuracy and very large for the total speaking score. Because the study employed a one-group pretest–posttest design, these statistical results should be interpreted as evidence of significant within-participant changes following the intervention. They should not be interpreted as definitive evidence that ChatGPT alone caused the observed changes, because the design did not include a control or comparison group.

Learners' Perceptions of ChatGPT as a Conversational Partner

The second research question concerned learners' perceptions of ChatGPT as a conversational partner. After completing the four-week intervention, all 40 participants responded to a 20-item questionnaire using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaire addressed perceptions related to usefulness, speaking development, confidence, ease of use, continued use, and overall satisfaction. The six reported summary items are presented in Table 3.

Table 3. Learners' Perceptions of ChatGPT as a Conversational Partner

Questionnaire Item	Mean	SD
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

Source: Primary questionnaire data, processed by the authors.

All six reported indicators obtained relatively high mean scores on the five-point scale. The highest mean was observed for the statement that ChatGPT was easy and convenient to use, with a mean of 4.70 and a standard deviation of 0.48. The second-highest mean was recorded for participants' intention to continue using ChatGPT for English-speaking practice, with a mean of 4.65 (SD = 0.52). The statement concerning confidence in speaking English obtained a mean of 4.60 (SD = 0.55), while overall satisfaction with ChatGPT as a speaking partner produced a mean of 4.58 (SD = 0.51). The item concerning perceived improvement in speaking fluency produced a mean of 4.50 (SD = 0.60), and the item concerning vocabulary learning obtained a mean of 4.45 (SD = 0.67).

The standard deviations ranged from 0.48 to 0.67, indicating relatively limited dispersion around the reported item means. The lowest dispersion was observed for ease and convenience of use, whereas the highest dispersion occurred for the vocabulary-learning item. The results therefore indicate that participants reported favorable experiences with ChatGPT across the six reported indicators. In particular, ease of use, intention to continue using the tool, confidence, and overall satisfaction received mean scores above 4.50.

The original questionnaire contained 20 items, while Table 3 reports six indicators available in the current dataset. Therefore, the six items should be presented as reported summary indicators rather than as the complete questionnaire results. If the raw responses to all 20 items are available, the final manuscript should include either the complete item-level results or a construct-level summary of the full questionnaire.

Association Between Perceived Usefulness and Speaking Score Gains

The third research question examined whether learners' perceived usefulness of ChatGPT was associated with their gains in speaking performance. Pearson product–moment correlation was used to examine the relationship between questionnaire-based perceived usefulness and the magnitude of pretest–posttest speaking score improvement.

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

Variables	*r*	*p*	Interpretation
Perceived usefulness and speaking score gain	0.67	< .001	Positive association

Source: Primary research data, processed by the authors.

Perceived usefulness was positively associated with speaking score gain, with $r = .67$ and $p < .001$. The positive direction of the coefficient indicates that participants reporting higher perceived usefulness tended to demonstrate larger gains in speaking scores between the pretest and posttest. The magnitude of the correlation represents a moderate-to-strong positive association. The finding provides complementary evidence linking learners' subjective experiences with their objectively assessed speaking performance. Nevertheless, the correlation should be interpreted as an association rather than a causal relationship. The result does not establish that perceived usefulness caused greater speaking improvement, nor does it demonstrate that positive perceptions alone were responsible for the observed performance changes.

Integrated Summary of Findings

The findings provide a consistent quantitative pattern across the three research questions. First, participants demonstrated higher mean scores at posttest than at pretest across fluency, vocabulary, grammatical accuracy, and total speaking performance. Fluency showed the largest individual mean increase of 1.45 points, followed by vocabulary with an increase of 1.35 points and grammatical accuracy with an increase of 1.10 points. The total speaking score increased by 3.90 points.

Second, the paired-samples analyses indicated statistically significant differences for all three speaking dimensions and the total speaking score. Fluency produced $t(39) = 9.81$, $p < .001$, and $d = 1.55$; vocabulary produced $t(39) = 8.76$, $p < .001$, and $d = 1.39$; grammatical accuracy produced $t(39) = 6.45$, $p < .001$, and $d = 1.02$; and the total speaking score produced $t(39) = 11.22$, $p < .001$, and $d = 1.77$. Thus, all reported comparisons demonstrated statistically significant within-participant changes with large or very large standardized effect sizes. Third, learners reported favorable perceptions of ChatGPT. Across the six reported questionnaire indicators, mean scores ranged from 4.45 to 4.70 on the five-point scale. Ease and convenience of use received the highest mean score, whereas vocabulary learning received the lowest among the reported indicators. Overall satisfaction was also high, with a mean score of 4.58.

Perceived usefulness was positively associated with speaking score gain, $r = .67$, $p < .001$. This result indicates that more favorable perceptions of ChatGPT tended to coincide with greater observed gains in speaking performance within the study sample. The combined findings therefore provide evidence of significant pretest–posttest changes in speaking performance alongside favorable learner perceptions of ChatGPT and a positive association between perceived usefulness and speaking gains. These findings are interpreted within the limits of the study design and sample and are discussed in relation to previous research, pedagogical implications, and methodological limitations in the subsequent Discussion section.

Discussion

The Effect of ChatGPT-Based Conversational Practice on EFL Speaking Performance

The first research question concerned changes in Indonesian EFL learners' fluency, vocabulary, and grammatical accuracy following structured conversational practice with ChatGPT. The observed improvement across the three dimensions suggests that conversational interaction with generative AI can function as an additional space for repeated language production rather than

merely as a source of linguistic information. This interpretation is consistent with systematic reviews showing that conversational AI can generate cognitive and affective learning benefits, although the strength of evidence varies according to task design and research methodology (Huang et al., 2022; Yang & Kyun, 2022). Recent reviews specifically identify speaking as an emerging but still comparatively underdeveloped area of ChatGPT research.

The relatively stronger improvement in fluency can be interpreted through the interactional nature of the intervention (Saito & Akiyama, 2017; Tauchid et al., 2025). Participants engaged in 12 sessions over four weeks, creating repeated opportunities to formulate responses, retrieve lexical items, and maintain conversational turns. Such repeated production may reduce hesitation and increase automaticity, particularly for intermediate learners who already possess sufficient linguistic resources to sustain interaction. This interpretation corresponds with Hsu et al. (2023), who demonstrated the potential of task-oriented chatbots for EFL speaking practice, and with Tai and Chen (2024), whose generative-AI chatbot intervention improved speaking performance and reduced anxiety. Similar improvements in fluency and broader oral proficiency have been reported in AI-mediated speaking studies (Sok & Shin, 2025).

The vocabulary and grammatical-accuracy gains are also theoretically meaningful because conversational interaction requires learners to continuously select lexical and syntactic forms in response to contextual prompts. Unlike fixed dialogue exercises, generative AI can adapt responses to learner-generated language, potentially increasing opportunities for retrieval and reformulation. Balcı (2024) similarly concluded that ChatGPT can support several dimensions of EFL learning, including vocabulary and grammar, while Leshchenko et al. (2023), Pirjan and Petrosanu, (2024) identified personalized learning opportunities and language practice as important affordances. The present findings extend these observations by examining vocabulary and grammatical accuracy alongside fluency within the same speaking assessment rather than treating speaking as a single undifferentiated construct (Safdari & Fathi, 2020; Shin, 2022).

The findings are also comparable with recent experimental evidence. Sok and Shin (2025) reported significantly better global speaking performance among university EFL learners who interacted with ChatGPT during oral tasks. Similarly, a 2025 Indonesian quasi-experimental study found significant speaking improvements among students receiving ChatGPT-supported conversational activities. The convergence is noteworthy because the present study was conducted in a different Indonesian higher-education setting and used a one-group pretest–posttest design rather than a controlled design. The consistency across these contexts strengthens the argument that structured conversational interaction is pedagogically promising, while differences in design mean that the magnitude of effects should not be directly equated.

The result also supports an interaction-oriented interpretation of conversational AI. Research on generative-AI chatbots has shown improvements not only in speaking performance but also in willingness to communicate, communicative competence, and speaking-related anxiety. Studies comparing conversational AI conditions have found that interaction with AI can create a psychologically less threatening environment for oral production (Tai & Chen, 2024). This may be particularly relevant for intermediate Indonesian EFL learners, for whom limited opportunities for sustained English interaction can constrain oral development. The present study therefore contributes evidence that ChatGPT may operate as a low-pressure supplementary conversational environment rather than simply as an automated language-learning application.

From a theoretical perspective, the findings support the view that the educational value of conversational AI emerges from interaction plus purposeful task design, rather than technology access alone. The systematic review by Jeon et al. (2024) emphasizes that conversational-AI research increasingly requires attention to behavioral and learning outcomes, while the BOPPPS-based study by 2025 demonstrated stronger speaking gains when ChatGPT was embedded within structured pedagogical stages. This distinction is important because unstructured conversations may generate engagement without necessarily producing systematic language development. The present intervention's combination of suggested topics, guiding questions, and flexible interaction provides a practical example of how learner autonomy can be combined with instructional scaffolding.

The principal contribution is therefore not the claim that ChatGPT independently causes speaking development, but that structured conversational use of ChatGPT is associated with substantial within-learner improvement across multiple speaking dimensions. The very large total-score effect should nevertheless be interpreted cautiously because the study lacks a control group, making alternative explanations such as repeated testing, maturation, or concurrent English exposure impossible to exclude. Future research should employ randomized or matched comparison groups, larger samples across Indonesian universities, longer interventions, and more fine-grained measures of interactional competence, pronunciation, discourse management, and lexical diversity. Such designs would help determine whether the promising pattern observed here represents a robust instructional effect or a context-specific response to intensive conversational practice.

Learners' Perceptions of ChatGPT as a Conversational Partner

The favorable perceptions observed in the present study indicate that learners did not merely tolerate the technology but viewed it as convenient, useful, and sufficiently supportive to warrant continued use (Bansah & Darko, 2022; Dalvi-Esfahani et al., 2022). This finding is consistent with broader evidence showing that EFL learners generally perceive ChatGPT as a useful language-learning resource, particularly because it provides immediate responses and flexible access to practice (Alghasab, 2025; Zou et al., 2025). Such perceptions are important because sustained engagement with AI-supported learning depends partly on whether learners perceive the technology as practically valuable.

The particularly favorable perception of ease and convenience is theoretically consistent with Technology Acceptance Model research. Liu and Ma (2024) found that EFL learners' use of ChatGPT for informal English learning can be understood through technology-acceptance constructs, while Zou et al. (2025) demonstrated that facilitating conditions and attitudes influence learners' adoption of ChatGPT for speaking practice. These findings help explain why accessibility may be especially important in conversational practice. A conversational partner that is available without scheduling, embarrassment, or dependence on peer availability can reduce practical barriers to oral practice.

The high confidence-related perception is also consistent with research on AI-mediated speaking. Tai and Chen (2024) found that AI-supported speaking activities can influence affective variables such as enjoyment, anxiety, and willingness to communicate, while studies of conversational GenAI have reported positive effects on self-perceived communicative competence. Such evidence suggests that the value of conversational AI may extend beyond linguistic practice to the psychological conditions surrounding language production. The present finding therefore supports the interpretation of ChatGPT as a potentially less threatening conversational environment in which learners can repeatedly practice without the immediate social evaluation associated with teacher- or peer-facing interaction.

The intention to continue using ChatGPT further strengthens this interpretation. Continued-use intention is important because short-term satisfaction does not necessarily translate into sustained educational engagement. Research by Zou et al. (2025) found that positive attitudes toward ChatGPT significantly influence learners' intention to use it for oral English practice. Similarly, research on self-directed English learning indicates that interactivity, enjoyment, trust, and subjective norms can operate through perceived usefulness to influence continued use (Liu et al., 2024). Recent studies of EFL learners also show that motivation and technology acceptance are interconnected, suggesting that positive experiences can reinforce learners' willingness to incorporate generative AI into their learning routines.

The present perception findings are comparable with evidence from other EFL contexts but should not be interpreted as universal acceptance. Teng (2024) reported positive effects of ChatGPT on motivation, self-efficacy, engagement, and collaborative tendencies, although learners also expressed concerns. Almulla (2024) similarly identified motivation and satisfaction as important aspects of students' ChatGPT-supported learning experiences. Studies in Saudi Arabia and other contexts have also documented favorable attitudes toward ChatGPT for English learning, while identifying concerns related to accuracy, dependence, and responsible use. These

differences suggest that positive perception can coexist with critical awareness rather than representing unconditional acceptance.

Practically, the findings suggest that teachers should conceptualize ChatGPT as a supplementary conversational partner, not a replacement for teachers or authentic human communication. Learners benefit most when the technology is embedded within clearly defined speaking objectives, reflection activities, and guidance concerning appropriate AI use. Indonesian studies similarly indicate that students value ChatGPT's convenience and assistance while recognizing limitations and risks of overreliance (Werdiningsih et al., 2024). Teacher guidance is consequently necessary to transform positive technological perceptions into productive learning behavior.

The theoretical contribution of this finding lies in connecting technology acceptance with communicative language practice (Hsieh et al., 2017; Shahid et al., 2023). Rather than examining perceived usefulness solely as an intention-to-use variable, the present study situates usefulness within an oral-language intervention. Its novelty therefore lies in showing that learners' positive evaluation of ChatGPT occurs alongside objectively measured speaking development. The limitation is that the questionnaire results represent self-reported perceptions and only six summary indicators are currently available from the 20-item instrument. Future research should validate the complete scale, examine its dimensional structure, and investigate how perceptions change over longer periods of use, across proficiency levels, and under different forms of teacher scaffolding.

Perceived Usefulness and Gains in Speaking Performance

The positive correlation indicates that learners who evaluated ChatGPT more favorably tended to demonstrate greater pretest–posttest improvement. This finding is theoretically important because it links two dimensions that are frequently examined separately in AI-assisted language-learning research: technology perception and objectively assessed language performance. Systematic reviews have noted that much of the early ChatGPT literature concentrated on perceptions, attitudes, and adoption, creating a need for studies that connect these variables with measurable learning outcomes.

The finding is compatible with Technology Acceptance Model assumptions, in which perceived usefulness represents the extent to which users believe that a technology improves task performance. Liu and Ma (2024) demonstrated the relevance of TAM constructs to EFL learners' use of ChatGPT, while Zou et al. (2025) found that positive attitudes significantly influenced intention to adopt ChatGPT for speaking practice. The present study extends this perspective by moving beyond intention and examining whether perceived usefulness is associated with an observable change in speaking performance. This provides a more educationally meaningful interpretation of usefulness than adoption intention alone.

Several mechanisms may explain the observed association. Learners who perceive ChatGPT as useful may be more willing to engage repeatedly with conversational tasks, experiment with vocabulary, request clarification, and sustain interaction. Greater perceived value can therefore encourage more active learning behavior. Research on self-directed English learning indicates that interactivity, enjoyment, trust, and self-efficacy influence continued ChatGPT use, while research on engagement suggests that the depth and nature of learner interaction with generative AI matter for learning outcomes (Liu et al., 2024). The relationship observed here may therefore reflect differences in how productively learners engage with the conversational environment.

The finding also complements research showing that ChatGPT-supported learning can enhance engagement and learning-related outcomes. ChatGPT-supported feedback can influence affective, behavioral, and cognitive engagement, while Riaz and Mushtaq (2024), Olaitan and Ajao (2025), Nguyen et al. (2025) reported relationships between generative-AI use, engagement, and learning outcomes. These studies suggest that AI effectiveness cannot be reduced to the technological capability of the system itself; learner engagement with the system is a central component of the learning process. The present study adds a speaking-specific perspective by connecting perceived usefulness with gains in oral performance rather than writing performance or general engagement.

According to Tomaylla-Quispe et al. (2025) and Pang et al. (2025), the association is also consistent with research showing that positive perceptions can facilitate continued AI use. Haq et al. (2025) reported strong relationships among perceptions, attitudes, and intentions to continue using ChatGPT, while Alotaibi et al. (2025) identified factors affecting acceptance and use among Saudi EFL learners. Dizon et al. (2025) similarly examined ChatGPT in self-regulated language learning, highlighting the importance of learners' active practices and perceptions. These findings support a reciprocal interpretation in which positive experiences may encourage continued use, while continued meaningful use may reinforce perceptions of usefulness.

The practical implication is that teachers should not evaluate ChatGPT integration solely by asking whether students enjoy using it. Perceived usefulness should be connected to observable communicative objectives, such as fluency development, vocabulary expansion, grammatical control, and willingness to sustain interaction. Teachers can therefore ask students to set speaking goals, document their interactions, reflect on language problems encountered during conversations, and compare their performance over time. Such structured use is consistent with evidence that pedagogical design and scaffolding influence the effectiveness of generative AI rather than technology availability alone (Sok & Shin, 2025; Hamid et al., 2025).

The theoretical novelty lies in connecting perceived usefulness with an objective speaking-performance gain within an Indonesian EFL context. The contribution should nevertheless be stated cautiously because correlation does not establish directionality or causality. Learners who improved more may have subsequently perceived ChatGPT as more useful, rather than usefulness necessarily producing greater improvement. The one-group pretest–posttest design, small purposive sample, short four-week intervention, and absence of a control group further restrict causal interpretation. Future studies should employ longitudinal designs, randomized or matched comparison groups, larger multi-institutional Indonesian samples, and mediation or cross-lagged models to test whether perceived usefulness predicts engagement and whether engagement subsequently explains speaking development.

CONCLUSION

This study examined the effectiveness of ChatGPT as a conversational partner for Indonesian EFL learners in a higher-education context in Makassar, South Sulawesi, through a four-week structured conversational intervention involving 40 intermediate-level learners. The findings indicate positive changes in speaking performance, with improvements in fluency, vocabulary, grammatical accuracy, and total speaking scores after the intervention. Fluency demonstrated the largest individual improvement, while the total speaking score showed a substantial standardized effect. Learners also reported highly favorable perceptions of ChatGPT, particularly regarding its ease and convenience of use, continued usefulness for speaking practice, confidence-building potential, and overall satisfaction. In addition, perceived usefulness was positively associated with gains in speaking performance, suggesting that learners who viewed ChatGPT more favorably tended to demonstrate greater improvement. These findings indicate that ChatGPT can serve as a useful supplementary conversational environment for Indonesian EFL learners by providing accessible opportunities for repeated and purposeful English interaction. However, given the one-group pretest–posttest design, purposive sampling, relatively small sample, and absence of a control group, the findings should be interpreted as evidence of positive within-participant changes rather than definitive causal effects. Future research should employ larger samples, controlled or randomized designs, longer interventions, and multiple Indonesian higher-education contexts to establish the robustness and generalizability of ChatGPT-supported speaking development.

REFERENCES

Alghasab, M. (2025). Insights into EFL Students' Perceptions of the 'ChatGPT Essentials' Training Course for Language Learning. *Education Sciences*, *15*(9), 1138. https://doi.org/10.3390/educsci15091138

Almulla, M. A. (2024). Investigating influencing factors of learning satisfaction in AI ChatGPT for research: University students perspective. *Heliyon*, *10*(11).

Alotaibi, H. M., Sonbul, S. S., & El-Dakhs, D. A. (2025). Factors influencing the acceptance and use of ChatGPT among English as a foreign language learners in Saudi Arabia. *Humanities and Social Sciences Communications, 12*, 628. https://doi.org/10.1057/s41599-025-04945-2

Balcı, Ö. (2024). The role of ChatGPT in English as a foreign language (EFL) learning and teaching: A systematic review. *International Journal of Current Educational Studies, 3*(1), 66–82. https://doi.org/10.46328/ijces.107

Bansah, A. K., & Darko Agyei, D. (2022). Perceived convenience, usefulness, effectiveness and user acceptance of information technology: evaluating students' experiences of a Learning Management System. *Technology, Pedagogy and Education*, *31*(4), 431-449. https://doi.org/10.1080/1475939X.2022.2027267

Dalvi-Esfahani, M., Wai Leong, L., Ibrahim, O., & Nilashi, M. (2020). Explaining students' continuance intention to use mobile web 2.0 learning and their perceived learning: An integrated approach. *Journal of Educational Computing Research*, *57*(8), 1956-2005. https://doi.org/10.1177/0735633118805211

Ding, D., & Yusof, M. B. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: a mixed methods study. *Humanities and Social Sciences Communications*, *12*(1), 1223. https://doi.org/10.1057/s41599-025-05550-z

Dizon, G., Gold, J., & Barnes, R. (2025). ChatGPT for self-regulated language learning: University English as a foreign language students' practices and perceptions. *Digital Applied Linguistics, 3*, 102510. https://doi.org/10.29140/dal.v3.102510

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, *6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230

Ericsson, E., & Johansson, S. (2023). English speaking practice with conversational AI: Lower secondary students' educational experiences over time. *Computers and Education: Artificial Intelligence*, *5*, 100164. https://doi.org/10.1016/j.caeai.2023.100164

Hamid, R. J., Maing, R. A., Rampeng, Muhammad, A. F., & Sahib, N. (2025). Postgraduate EFL instructional models in Indonesia: aligning student perceptions with lecturer rationales. *Cogent Education*, *12*(1), 2590863. https://doi.org/10.1080/2331186X.2025.2590863

Haq, M. Z. U., Cao, G., & Abukhait, R. M. Y. (2025). Examining students' attitudes and intentions towards using ChatGPT in higher education. *British Journal of Educational Technology*, *56*(6), 2428-2452. https://doi.org/10.1111/bjet.13582

Hsieh, J. S. C., Huang, Y. M., & Wu, W. C. V. (2017). Technological acceptance of LINE in flipped EFL oral training. *Computers in Human Behavior*, *70*, 178-190. https://doi.org/10.1016/j.chb.2016.12.066

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

Jeon, J., Lee, S., & Choi, S. (2024). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments, 32*(8), 4613–4631. https://doi.org/10.1080/10494820.2023.2204343

Leshchenko, M., Lavrysh, Y., Halatsyn, K., Feshchuk, A., & Prykhodko, D. (2023). Technology-enhanced personalized language learning: Strategies and challenges. *International Journal of Enmerging Technologies in Learning*, *13*(18), 120-136.

Lin, C. J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society*, *24*(4), 16-35.

Liu, G. L., Darvin, R., & Ma, C. (2024). Unpacking the role of motivation and enjoyment in AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation in the Chinese context. *Computers in Human Behavior, 160*, 108362. https://doi.org/10.1016/j.chb.2024.108362

Liu, G., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching, 18*(2), 125–138. https://doi.org/10.1080/17501229.2023.2240316

Lopez-Gazpio, I. (2025). Integrating large language models into accessible and inclusive education: access democratization and individualized learning enhancement supported by generative artificial intelligence. *Information*, *16*(6), 473. https://doi.org/10.3390/info16060473

Mimoudi, A. (2025). Generative AI to bridge the educational divide: Personalized learning and challenges. *Social Sciences & Humanities Open*, *12*, 102140. https://doi.org/10.1016/j.ssaho.2025.102140

Nguyen, M., Agarwal, R., Malik, A., & Pandey, R. (2025). Generative AI: elevating knowledge acquisition and retention and recall through autonomy, interactivity and engagement. *Journal of Enterprise Information Management*, *38*(6), 1692-1719. https://doi.org/10.1108/JEIM-08-2024-0433

Olaitan, O., & Ajao, I. O. (2025). Generative AI in Higher Education: Investigating How Perceived Usefulness and Usage Patterns Influence Student Engagement and Academic Performance. *International Journal of Learning, Teaching and Educational Research*, *24*(11), 682-710.

Pang, H., Hu, Z., & Wang, L. (2025). How perceived motivations influence user stickiness and sustainable engagement with AI-powered chatbots—Unveiling the pivotal function of user attitude. *Journal of theoretical and applied electronic commerce research*, *20*(3), 228. https://doi.org/10.3390/jtaer20030228

Pirjan, A., & Petrosanu, D. M. (2024). Exploring large language models in the education process with a view towards transforming personalized learning. *Journal of Information Systems & Operations Management*, *18*(2), 125-171.

Riaz, S., & Mushtaq, A. (2024, June). Optimizing generative AI integration in higher education: A framework for enhanced student engagement and learning outcomes. In *2024 Advances in Science and Engineering Technology International Conferences (ASET)* (pp. 1-6). IEEE. https://doi.org/10.1109/ASET60340.2024.10708721

Sadigzade, Z. (2025). Generative AI in CALL: Enhancing Self-Regulated L2 Learning Through Adaptive Chatbots. *Global Spectrum of Research and Humanities*, *2*(6), 25-45. https://doi.org/10.69760/gsrh.0250206002

Safdari, M., & Fathi, J. (2020). Investigating the role of dynamic assessment on speaking accuracy and fluency of pre-intermediate EFL learners. *Cogent Education*, *7*(1), 1818924. https://doi.org/10.1080/2331186X.2020.1818924

Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language learning*, *67*(1), 43-74. https://doi.org/10.1111/lang.12184

Shahid, C., Gurmani, M. T., Rehman, S. U., & Saif, L. (2023). The role of technology in English language learning in online classes at tertiary level. *Journal of Social Sciences Review*, *3*(2), 232-247. https://doi.org/10.54183/jssr.v3i2.215

Shin, J. Y. (2022). Investigating and optimizing score dependability of a local ITA speaking test across language groups: A generalizability theory approach. *Language testing*, *39*(2), 313-337.

Sok, S., & Shin, H. W. (2025). Do interactions with ChatGPT influence L2 learners' oral speaking ability, summarization ability, and perceptions of generative AI tasks? *TESOL Quarterly, 59*(S1), S19–S51. https://doi.org/10.1002/tesq.70001

Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education, 220*, 105112. https://doi.org/10.1016/j.compedu.2024.105112

Tauchid, A., Muhid, A., Khudlori, A., Ayunda, N. A., Lee, A., & Jisoo, C. (2025). Investigating communicative language teaching through meaning-based tasks and processing strategies to strengthen learners' oral performance. *Journal of Research in English Language Teaching and Linguistics*, *1*(2), 153-169. https://doi.org/10.65431/jrell.v1i2.30

Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence, 7*, 100270. https://doi.org/10.1016/j.caeai.2024.100270

Tomaylla-Quispe, Y., Postigo, G. P., Gutiérrez-Aguilar, O., Chicaña-Huanca, S., & Duche-Pérez, A. (2025). Artificial intelligence on students satisfaction in higher education: The role of positive attitude, continuous use intention, and perceived usefulness. *Journal of Technology and Science Education*, *15*(3), 629-646. https://doi.org/10.3926/jotse.3422

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

Zein, S., Sukyadi, D., Hamied, F. A., & Lengkanawati, N. S. (2020). English language education in Indonesia: A review of research (2011–2019). *Language Teaching*, *53*(4), 491-523. https://doi.org/10.1017/S0261444820000208

Zou, B., Wang, C., Yan, Y., Du, X., & Ji, Y. (2025). Exploring English as a foreign language learners' adoption and utilisation of ChatGPT for speaking practice through an extended technology acceptance model. *International Journal of Applied Linguistics*, *35*(2), 689-704. https://doi.org/10.1111/ijal.12658