




Automated Feedback vs. Human Feedback: A Study on AI-Driven Language Assessment

Mohammad Aliakbari¹, Pooria Barzan^{1*}, Mehran Sayyadi²

¹ Department of English Language and Literature, Faculty of Humanities, Ilam University, Ilam, Iran

² Department of English Language and Literature, Faculty of Humanities, Ilam University, Iran

* Corresponding author email address: pooriabarzan@yahoo.com

Article Info

Article type:

Original Research

How to cite this article:

Aliakbari, M., Barzan, P., & Sayyadi, M. (2025). Automated Feedback vs. Human Feedback: A Study on AI-Driven Language Assessment. *AI and Tech in Behavioral and Social Sciences*, 3(2), 113-126.

<https://doi.org/10.61838/kman.aitech.3.2.9>



© 2025 the authors. Published by KMAN Publication Inc. (KMANPUB), Ontario, Canada. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

ABSTRACT

This study examines the comparative effectiveness of automated feedback from AI-driven language assessment tools and traditional human feedback in improving language learning outcomes. As artificial intelligence (AI) becomes increasingly integrated into educational technology, understanding its role in supporting language learners is critical. To address this, we conducted a 12-week controlled experiment with 80 EFL (English as a Foreign Language) learners, dividing participants into two groups: one receiving feedback from an AI-driven tool and the other from human instructors. Language proficiency improvements were assessed in grammar, vocabulary, and pronunciation. Findings reveal that both feedback methods significantly enhanced language skills. Human feedback proved superior in fostering gains in pronunciation and contextual understanding, while AI-driven feedback excelled in detecting and correcting grammatical errors. These results suggest that a hybrid model, blending AI and human feedback, could provide the most effective support for language learners. This research enriches the literature on AI in education and offers practical insights for incorporating AI tools into language learning curricula.

Keywords: Assessment, Feedback, Artificial Intelligence (AI), Human Feedback.

1. Introduction

In an increasingly globalized world, the demand for communicative competence in second languages, particularly English as a Foreign Language (EFL), has surged, driving educators and researchers to explore innovative strategies for language acquisition (Wu et al., 2024; Yunina, 2023). The rapid digitalization of teaching methodologies has positioned artificial intelligence (AI) as a transformative force, revolutionizing traditional paradigms of language education and assessment (Mahmud et al., 2023; Trang & Thu, 2024). This shift has catalyzed a critical examination of feedback mechanisms—both AI-driven and human—as pivotal tools for fostering learner autonomy, enhancing engagement, and improving linguistic proficiency across grammar, vocabulary, and pronunciation (Dewi et al., 2023; M. Liu, 2023; Nurmala et al., 2023). While AI offers scalable, immediate, and data-driven solutions, human feedback provides nuanced, context-sensitive insights, sparking debates about their comparative efficacy and potential synergies (Long & Kim, 2024; Zou et al., 2023). This study synthesizes these perspectives, aiming to elucidate the differential impacts of automated and human feedback and propose a hybrid framework that optimizes language learning outcomes (Tiwari, 2024; Wei et al., 2023).

1.1. *The Emergence and Capabilities of AI in Language Assessment*

AI-driven language assessment tools represent a significant milestone in educational technology, leveraging sophisticated machine learning algorithms and natural language processing (NLP) to evaluate and enhance learners' performance (Qin, 2024; Wei, 2023). These systems deliver rapid, personalized feedback, excelling in error detection and correction for grammar and vocabulary with remarkable precision (Bulut & Wongvorachan, 2022; Cheung & Luk, 2024). For instance, AI-powered automatic scoring applications utilize deep learning to provide immediate corrective insights, reducing error fossilization and reinforcing structural accuracy (Huang, 2023; Wei, 2023). Beyond technical proficiency, these tools foster immersive, interactive learning experiences that enhance engagement and promote self-regulated learning, addressing cognitive load and metacognitive awareness (Vadivel et al., 2023; Yang & Kyun, 2022). The scalability of AI systems further democratizes access to high-quality instruction, overcoming resource constraints in EFL

contexts and supporting large cohorts with consistent feedback (Bulut & Wongvorachan, 2022; Y. Liu, 2023). However, limitations persist: AI struggles to capture socio-pragmatic nuances such as tone, contextual inference, and cultural subtleties, necessitating ongoing refinement to align with holistic language development goals (Huang, 2023; Long & Kim, 2024).

1.2. *The Persistent Relevance of Human Feedback*

Despite AI's advancements, traditional human feedback remains a cornerstone of effective language instruction, offering irreplaceable depth and adaptability (Morales et al., 2024; Zhou & Hou, 2024). Human instructors excel in providing individualized guidance on complex domains like pronunciation, intonation, and pragmatic appropriateness, fostering deeper interpersonal connections that enhance learner motivation, confidence, and self-efficacy (Byeon, 2023; Liu et al., 2023; Moybeka et al., 2023). This contextual richness—rooted in experiential knowledge and cultural awareness—addresses affective factors critical for sustained language development (Dewi et al., 2023; Long & Kim, 2024). For example, human feedback proves superior in refining prosodic elements and idiomatic expressions, areas where AI's algorithmic approach often falls short (Zou et al., 2023). Yet, challenges such as labor intensity, inconsistency, and instructor bias highlight scalability and standardization issues, underscoring the need for complementary technological integration (Dewi et al., 2023; Zhou & Hou, 2024).

1.3. *Bridging the Divide: Toward a Hybrid Feedback Model*

The complementary strengths of AI-driven and human feedback have spurred calls for hybrid models that synergize technological efficiency with pedagogical depth (Fitria, 2023; Zulkarnain & Yunus, 2023). AI excels in rapid, quantifiable tasks—such as grammatical precision and vocabulary retention—reducing cognitive load and enabling self-directed learning (Chen et al., 2024; Wei et al., 2023). In contrast, human feedback provides emotional support and contextual insight, particularly for pronunciation and communicative competence, aligning with constructivist and sociocultural theories that emphasize scaffolding and social interaction (Şişianu & Puşcaşu, 2024; Wu et al., 2024). Empirical evidence suggests AI outperforms in structural accuracy, while

human instructors dominate in nuanced, higher-order skills, supporting the viability of blended approaches (M. Liu, 2023; Zou et al., 2023).

1.4. Research Significance

This research addresses a critical gap by systematically comparing automated and human feedback in EFL contexts, focusing on measurable improvements in grammar, vocabulary, and pronunciation, alongside qualitative dimensions like learner engagement and motivation (Cohen et al., 2024; Yin & Hanif, 2024). Employing a mixed-methods approach grounded in constructivist and sociocultural frameworks, the study assesses proficiency via pre-/post-tests, error analysis, and learner surveys (Wu et al., 2024; Zhou & Hou, 2024). The convergence of AI and education marks a paradigmatic shift, challenging traditional models and inviting interdisciplinary dialogue among educators, developers, and policymakers (Moybeka et al., 2023; Yunina, 2023). Beyond performance metrics, the research explores practical implications—such as digital literacy, accessibility, and equity—offering evidence-based insights for curriculum design in a digital age (Alzahrani & Alotaibi, 2024; Hashim et al., 2023). By proposing a hybrid feedback model, it contributes to the evolving literature on educational technology, balancing technological innovation with pedagogical integrity (Chen et al., 2024; Ma et al., 2024).

1.5. Implications for EFL Pedagogy

The evolution of language assessment—from static drills to dynamic AI applications—underscores AI's potential to reshape EFL instruction (Wu et al., 2024; Yunina, 2023). While AI's immediacy and scalability address resource constraints and diverse learner needs (Belda-Medina & Goddard, 2024; Trang & Thur, 2024), human feedback's empathetic and adaptive nature remains vital for holistic development (Dewi et al., 2023; Long & Kim, 2024). Meta-analyses affirm AI's efficacy in structured tasks, yet emphasize human interaction's irreplaceable role in fostering communicative effectiveness (Chen et al., 2024; Wu et al., 2024).

In conclusion, this investigation establishes a foundation for understanding the interplay of AI-driven and human feedback in language assessment. As AI technologies advance, harmonizing their precision with human insight promises adaptive, learner-centered environments that meet

the diverse needs of EFL learners worldwide (Ma et al., 2024; Tiwari, 2024). The findings are poised to inform best practices, bridging empirical gaps and guiding the future of language pedagogy in an era of digital transformation.

2. Literature Review

The integration of AI into educational technology has emerged as a transformative force, particularly in the domain of language learning and assessment (Abuhassna, 2024; Ifraheem et al., 2024; Xu & Ouyang, 2022; Yao & Chung, 2024). This literature review synthesizes current research to examine the comparative effectiveness of automated feedback from AI-driven language assessment tools and traditional human feedback in enhancing language learning outcomes, with a specific focus on English as a Foreign Language (EFL) learners. It explores the evolution of AI applications in language education, evaluates the strengths and limitations of both feedback modalities, and considers broader implications, including ethical challenges and the necessity of AI literacy among educators. The review advocates for a hybrid approach that leverages the complementary strengths of AI and human feedback to optimize language instruction.

2.1. AI in Language Learning and Assessment

The advent of AI-driven technologies has significantly reshaped language education by introducing innovative tools for assessment and feedback (Chen, 2024; Yunina, 2023). These tools, underpinned by advanced NLP and machine learning algorithms, enable rapid, personalized, and data-driven feedback (Chen et al., 2024; Chen, 2024). For instance, AI systems employ deep learning architectures to analyze linguistic data with high precision, facilitating consistent error detection in areas such as grammar and vocabulary (Chen et al., 2024; Yunina, 2023). Studies demonstrate that AI-powered platforms, including chatbots and automated grading systems, enhance learner engagement and motivation by adapting to individual needs and providing immediate responses (Moybeka et al., 2023; Nurjanah et al., 2024; Zhou & Hou, 2024). This adaptability fosters a dynamic and interactive learning environment, positioning AI as a valuable supplement to traditional pedagogy (Vadivel et al., 2023; Wiyaka et al., 2024).

The scalability of AI tools further amplifies their utility, allowing for real-time analysis of large datasets and consistent feedback delivery across diverse learner

populations (Wei, 2023; Yunina, 2023). This capability is particularly advantageous in large classroom settings, where human instructors may struggle to provide timely, individualized support (Abernathy, 2024). Moreover, AI's capacity for personalization—tailoring feedback to learners' unique error patterns—enhances self-regulated learning and promotes autonomy (Alhusaiyan, 2024; Li, 2024). However, despite these advancements, AI-driven feedback is not without limitations, particularly in capturing the nuanced, contextual aspects of language use that remain a stronghold of human expertise.

2.2. Comparative Effectiveness of Automated and Human Feedback

Empirical research consistently highlights the distinct yet complementary roles of automated and human feedback in language education. Automated feedback excels in delivering immediate, consistent corrections, particularly in grammatical and syntactic domains (Chen et al., 2024; Chen, 2024; M. Liu, 2023). Studies show that AI systems can detect and correct errors with algorithmic precision, reducing biases inherent in human evaluations and supporting standardized assessment processes (Pack & Maloney, 2024; Yunina, 2023). For example, AI-driven tools have proven highly effective in improving grammatical accuracy and vocabulary acquisition by offering real-time suggestions and contextual examples (Alhusaiyan, 2024; Chen, 2024). This immediacy prevents error fossilization and encourages learners to engage actively with their mistakes (Chen et al., 2024; Ojeda-Ramirez et al., 2023).

In contrast, human feedback retains a critical edge in areas requiring contextual judgment and socio-pragmatic understanding. Research indicates that human instructors outperform AI in providing nuanced feedback on pronunciation, intonation, and pragmatic language use—skills essential for communicative competence (Amin, 2023; Chen, 2024; Mohammadkarimi, 2024). Human feedback also incorporates affective elements, such as encouragement and cultural insights, which foster emotional intelligence and interpersonal skills (Koka et al., 2023; Zhou & Hou, 2024). Comparative studies in EFL contexts reveal that while automated feedback drives improvements in technical accuracy, human feedback significantly enhances pronunciation and contextual comprehension (Chen et al., 2024; Davis, 2024; Kumari, 2023; Wei, 2023). These findings underscore the domain-specific strengths of each approach, suggesting that neither

modality alone fully addresses the complexities of language acquisition.

2.3. Challenges and Ethical Considerations

The integration of AI into language education introduces several challenges and ethical concerns that warrant careful consideration. Despite their technical prowess, AI-driven tools often struggle to interpret socio-pragmatic nuances and emotional dimensions of language learning (Chen et al., 2024; Davoodifard & Eskin, 2024; Dhanapal et al., 2024). This limitation can hinder their effectiveness in fostering holistic language proficiency, particularly in culturally sensitive contexts. Additionally, ethical issues such as data privacy, algorithmic bias, and the potential displacement of human roles pose significant risks (Gupta et al., 2024; Igbokwe, 2024; Stracke et al., 2023; Yu & Yu, 2023). For instance, reliance on AI may exacerbate inequities if access to technology or digital literacy varies across learner populations (Alhusaiyan, 2024; Trang & Thur, 2024). Addressing these challenges requires robust institutional support, educator training, and policies that ensure equitable and responsible AI implementation (Gupta et al., 2024; Yu & Yu, 2023).

2.4. The Need for Hybrid Models and AI Literacy

Given the complementary strengths of automated and human feedback, the literature increasingly advocates for hybrid models that integrate AI's scalability and precision with human instructors' contextual and affective insights (Aghaziarati et al., 2023; AlAfnan, 2024; Davis, 2024; Kim et al., 2024; Kumari, 2023; Muslimin et al., 2024). Such models leverage AI to handle routine error correction while allowing educators to focus on higher-order skills like creativity and pragmatic competence (Chen, 2024; Konyrova, 2024). Experimental evidence suggests that this blended approach yields superior outcomes in grammar, vocabulary, and pronunciation, enhancing both linguistic accuracy and learner engagement (Amin, 2023; Chen et al., 2024). Furthermore, hybrid systems support collaborative learning environments, where AI provides auxiliary corrections and human-led discussions deepen cognitive engagement (Trang & Thur, 2024; Wei, 2023).

Central to the success of AI integration is the development of AI literacy among educators. As AI tools become ubiquitous, teachers must acquire the skills to navigate their applications and ethical implications effectively (Ibrahim, 2024; Rütli-Joy et al., 2023; Yim &

Wegerif, 2024; Yin & Goh, 2024). AI literacy empowers educators to harness technology responsibly, ensuring it enhances rather than supplants traditional instruction (Ibrahim, 2024; Rütli-Joy et al., 2023). This competency is vital for designing curricula that balance technological innovation with pedagogical integrity, fostering a sustainable and inclusive educational ecosystem.

The literature underscores the transformative potential of AI in language education while affirming the enduring value of human feedback. Automated systems offer unparalleled speed, consistency, and scalability, excelling in technical domains like grammar and vocabulary correction (Chen et al., 2024; Yunina, 2023). Conversely, human feedback remains indispensable for addressing pronunciation, contextual nuances, and emotional aspects of learning (Chen, 2024; Mohammadkarimi, 2024). By synthesizing empirical findings and theoretical insights, this review advocates for a strategic hybrid model that optimizes language learning outcomes through the synergy of AI and human expertise. Continued research and thoughtful implementation are essential to address ethical challenges, enhance AI literacy, and maximize the positive impact of these technologies on EFL instruction. As AI continues to evolve, its integration into language education promises to redefine pedagogical practices, enriching the experiences of learners worldwide.

3. Methods and Materials

This study employed a mixed-methods approach to compare the effectiveness of automated feedback from AI-driven language assessment tools and traditional human feedback in enhancing language learning outcomes among EFL learners. The methodology was designed to assess improvements in grammar, vocabulary, and pronunciation, while also exploring qualitative dimensions such as learner engagement and motivation. A 12-week controlled experiment was conducted with 80 EFL learners, who were randomly assigned to either an AI feedback group or a human feedback group. Data were collected through pre-tests, post-tests, error analysis, and learner surveys, and analyzed using both quantitative and qualitative techniques.

3.1. Research Design

The study utilized a pre-test/post-test controlled experimental design to evaluate the differential impacts of AI-driven and human feedback on language proficiency. Participants were divided into two groups:

- **AI Feedback Group:** Received automated feedback from an AI-driven language assessment tool.
- **Human Feedback Group:** Received feedback from experienced human instructors.

Both groups followed the same curriculum and used identical learning materials, ensuring that the source of feedback was the only variable manipulated. This design enabled a direct comparison of the two feedback modalities over the 12-week period, aligning with the study's aim to assess their comparative effectiveness.

3.2. Participants

The study involved 80 EFL learners enrolled in an intensive English language program at Ilam university in Ilam, Iran. Participants were selected through convenience sampling based on their enrollment in the program and were adults aged 18 to 25 with intermediate English proficiency, as determined by standardized Oxford Placement Test. To ensure group comparability and minimize bias, participants were randomly assigned to either the AI feedback group (n=40) or the human feedback group (n=40) using a computer-generated randomization sequence.

3.3. Setting

The experiment took place in a blended learning environment, integrating face-to-face classroom instruction with online components. Participants attended weekly in-person classes and completed additional tasks via an online platform. Both groups used the same textbooks, multimedia resources, and practice exercises to maintain consistency in instructional content, with the only difference being the feedback delivery method.

3.4. AI-Driven Feedback Tool

The AI feedback group received feedback from ChatGPT. Leveraging NLP and machine learning, the tool provided automated feedback on written and spoken tasks, focusing on:

- **Grammar and Vocabulary:** Error detection, correction suggestions, and explanations for written submissions.
- **Pronunciation:** Analysis of audio recordings, offering feedback on accuracy, fluency, and intonation.

Feedback was delivered in real-time for written tasks and within 24 hours for spoken tasks, ensuring timely support tailored to intermediate proficiency levels.

3.5. Human Feedback

The human feedback group received personalized feedback from certified EFL instructors with at least five years of teaching experience. Feedback was provided on the same tasks as the AI group, using a standardized rubric to ensure consistency. For written tasks, instructors offered detailed comments and corrections. For spoken tasks, feedback was delivered through individual consultations or recorded audio comments, emphasizing pronunciation, fluency, and communicative effectiveness.

3.6. Data Collection

Data were gathered using multiple methods to capture both quantitative proficiency improvements and qualitative learner perceptions:

1. Pre-test and Post-test:

- Administered at the start and end of the 12-week period, these tests assessed grammar, vocabulary, and pronunciation.
- Grammar: Multiple-choice questions, cloze tests, and sentence correction tasks.
- Vocabulary: Word matching, synonym identification, and contextual usage exercises.
- Pronunciation: Oral test involving reading a passage and a short conversation, recorded and evaluated by two independent raters using a standardized rubric.

2. Error Analysis:

- Bi-weekly collections of written assignments and spoken recordings were analyzed to categorize and quantify errors in grammar, vocabulary, and pronunciation, tracking progress over time.

3. Learner Surveys:

- Conducted at the experiment's conclusion, surveys included Likert-scale items and open-ended questions to gather feedback on the perceived effectiveness of the feedback, engagement, and motivation.

3.7. Data Analysis

Data analysis combined quantitative and qualitative approaches:

▪ Quantitative Analysis:

- Pre-test and post-test scores were summarized using descriptive statistics (means, standard deviations).
- Independent samples t-tests compared improvements between groups for grammar, vocabulary, and pronunciation.
- Error rates from bi-weekly tasks were analyzed similarly to assess differences in error reduction.

▪ Qualitative Analysis:

Survey responses underwent thematic analysis, with open-ended answers coded to identify recurring themes related to engagement, motivation, and feedback perceptions.

4. Findings and Results

This section presents the findings from a 12-week controlled experiment comparing the effectiveness of automated feedback from an AI-driven language assessment tool (ChatGPT) and traditional human feedback in enhancing language learning outcomes among 80 EFL learners at Ilam University, Iran. Participants were randomly assigned to two groups: the AI feedback group (n=40) and the human feedback group (n=40). Results are organized into three subsections: proficiency improvements based on pre-test and post-test scores, error analysis from bi-weekly tasks, and learner perceptions from surveys.

4.1. Proficiency Improvements: Pre-test and Post-test Scores

Pre-test and post-test scores were collected to assess improvements in grammar, vocabulary, and pronunciation for both groups. Table 1 summarizes the mean scores and standard deviations at both time points, along with the mean differences.

Table 1

Mean Pre-test and Post-test Scores for Grammar, Vocabulary, and Pronunciation

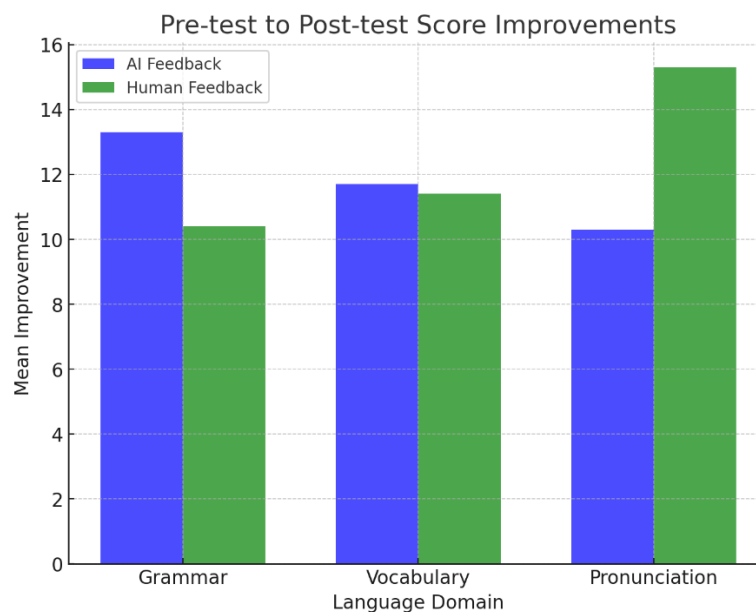
Language Domain	Group	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Difference
Grammar	AI Feedback	65.2 (8.3)	78.5 (7.1)	13.3
Grammar	Human Feedback	64.8 (7.9)	75.2 (8.2)	10.4
Vocabulary	AI Feedback	68.4 (9.1)	80.1 (6.8)	11.7
Vocabulary	Human Feedback	67.9 (8.5)	79.3 (7.4)	11.4
Pronunciation	AI Feedback	60.5 (10.2)	70.8 (9.3)	10.3
Pronunciation	Human Feedback	61.2 (9.8)	76.5 (8.1)	15.3

Independent samples t-tests were conducted to compare the mean differences between the AI feedback and human feedback groups for each language domain:

- **Grammar:** The AI feedback group exhibited a significantly greater improvement ($t(78) = 2.45$, $p = 0.016$).
- **Vocabulary:** No significant difference was observed ($t(78) = 0.32$, $p = 0.751$).
- **Pronunciation:** The human feedback group showed a significantly greater improvement ($t(78) = -3.12$, $p = 0.003$).

Figure 1

Pre-test to Post-test score improvement



4.2. Error Analysis

Error rates in grammar, vocabulary, and pronunciation were analyzed from bi-weekly written assignments and

spoken recordings collected over the 12-week period. Table 2 presents the average error rates at Week 1 and Week 12, along with the percentage reduction.

Table 2

Average Error Rates at Week 1 and Week 12

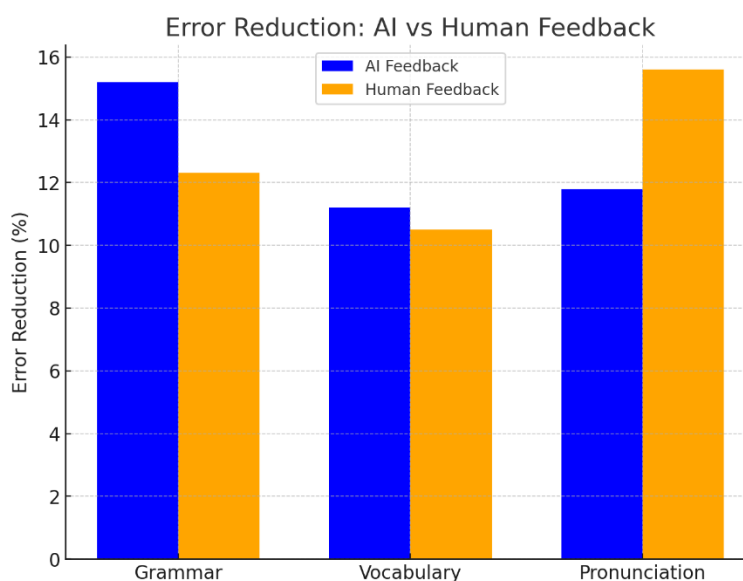
Language Domain	Group	Week 1 Error Rate (%)	Week 12 Error Rate (%)	Reduction (%)
Grammar	AI Feedback	25.3	10.1	15.2
Grammar	Human Feedback	24.8	12.5	12.3
Vocabulary	AI Feedback	20.1	8.9	11.2
Vocabulary	Human Feedback	19.7	9.2	10.5
Pronunciation	AI Feedback	30.5	18.7	11.8
Pronunciation	Human Feedback	29.8	14.2	15.6

Independent samples t-tests compared the reductions in error rates between the groups:

- **Grammar:** The AI feedback group demonstrated a significantly greater reduction ($t(78) = 2.78$, $p = 0.007$).
- **Vocabulary:** No significant difference was found ($t(78) = 0.56$, $p = 0.577$).
- **Pronunciation:** The human feedback group exhibited a significantly greater reduction ($t(78) = -2.95$, $p = 0.004$).

Figure 2

Error Reduction: AI vs Human Feedback



4.3. Learner Perceptions: Survey Results

At the end of the 12-week experiment, participants completed surveys with Likert-scale items (1 =strongly

Table 3 presents the mean ratings for key survey items.

Table 3

Mean Ratings for Survey Items

Survey Item	AI Feedback (Mean)	Human Feedback (Mean)
-------------	--------------------	-----------------------

disagree, 5 = strongly agree) and open-ended questions to assess perceptions of feedback effectiveness, engagement, and motivation.

4.3.1. Quantitative Survey Results

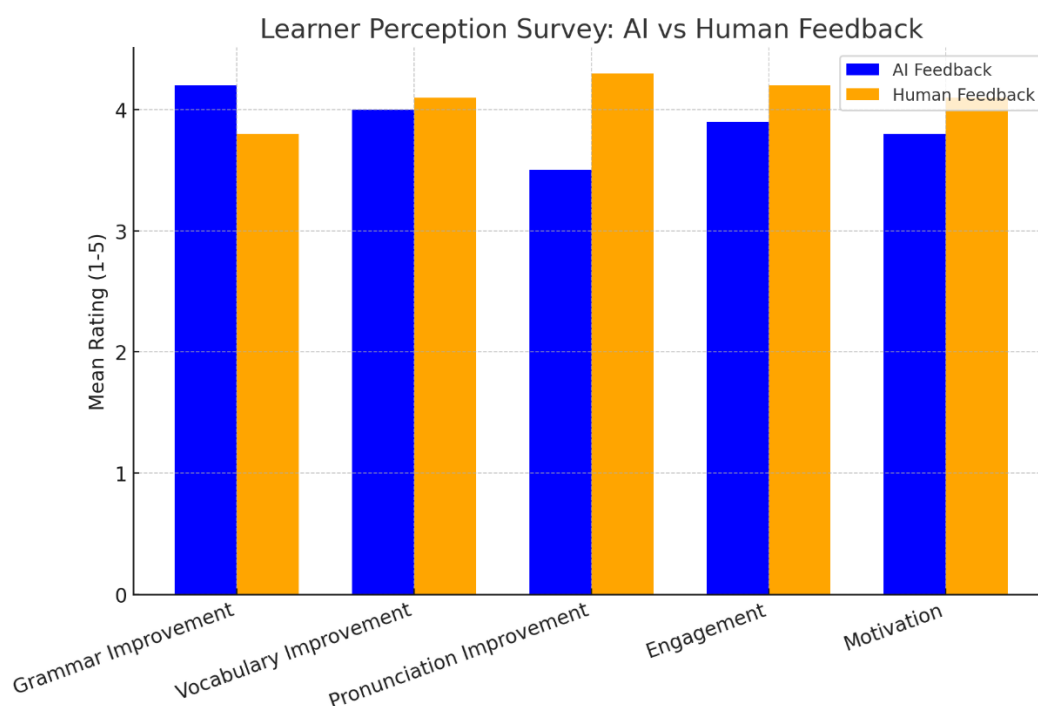
Survey Item	AI Feedback (Mean)	Human Feedback (Mean)
The feedback helped me improve my grammar.	4.2	3.8
The feedback helped me improve my vocabulary.	4.0	4.1
The feedback helped me improve my pronunciation.	3.5	4.3
I felt engaged with the feedback process.	3.9	4.2
The feedback motivated me to learn.	3.8	4.1

Independent samples t-tests revealed:

- The AI feedback group rated feedback as more helpful for grammar improvement ($t(78) = 2.10$, $p = 0.039$).
- The human feedback group rated feedback as more helpful for pronunciation improvement ($t(78) = -3.45$, $p = 0.001$).
- No significant differences were found for vocabulary improvement, engagement, or motivation.

Figure 3

Learner Perception Survey Results



4.3.2. Qualitative Survey Results

Thematic analysis of open-ended responses identified key themes for each group:

AI Feedback Group:

- Frequent mentions of immediate and detailed grammar corrections as a strength.
- Reports of frustration with pronunciation feedback lacking contextual depth.

- Positive remarks about the interactive and gamified features of ChatGPT.

Human Feedback Group:

- Emphasis on personalized and empathetic feedback, particularly for pronunciation.
- Appreciation for the ability to ask follow-up questions during instructor interactions.
- Some dissatisfaction with delays in receiving feedback due to instructor availability.

These results highlight distinct outcomes for the two feedback methods across grammar, vocabulary, and

pronunciation, with AI feedback showing greater efficacy in grammar improvement and human feedback excelling in pronunciation gains. Learner perceptions align with these findings, reflecting preferences for the strengths of each modality.

5. Discussion and Conclusion

The results revealed domain-specific advantages for each feedback type. The AI feedback group showed significantly greater improvements in grammar proficiency (mean difference = 13.3 vs. 10.4, $p = 0.016$) and a larger reduction in grammar errors (15.2% vs. 12.3%, $p = 0.007$) compared to the human feedback group. This aligns with AI's capacity for immediate, precise, and consistent error detection and correction, as evidenced by ChatGPT's real-time feedback on written tasks. Learners in the AI group also rated the feedback as more helpful for grammar improvement (mean = 4.2 vs. 3.8, $p = 0.039$), reflecting its effectiveness in addressing structural accuracy.

Conversely, the human feedback group outperformed the AI group in pronunciation, with greater gains in proficiency scores (mean difference = 15.3 vs. 10.3, $p = 0.003$) and a larger reduction in pronunciation errors (15.6% vs. 11.8%, $p = 0.004$). Human instructors' ability to provide nuanced, context-sensitive feedback—delivered through personalized consultations—likely drove these improvements. Survey responses supported this, with the human feedback group rating the feedback as more helpful for pronunciation (mean = 4.3 vs. 3.5, $p = 0.001$) and highlighting its empathetic and individualized nature.

For vocabulary, both groups exhibited comparable gains (AI: 11.7 vs. Human: 11.4, $p = 0.751$) and error reductions (11.2% vs. 10.5%, $p = 0.577$), suggesting that AI and human feedback are equally effective in this domain. This may reflect vocabulary learning's reliance on both immediate correction (via AI) and contextual reinforcement (via humans), as perceived by learners (AI: 4.0 vs. Human: 4.1, $p > 0.05$).

5.1. Relation to Existing Literature

These findings resonate with prior research on AI and human feedback in language education. The superiority of AI-driven feedback in grammar aligns with studies emphasizing its precision and immediacy in error correction (Chen, 2024; Yunina, 2023). The literature review noted AI's strengths in scalable, data-driven solutions for grammar and vocabulary (Bulut &

Wongvorachan, 2022; Wei, 2023), which is consistent with the AI group's performance. Qualitative survey responses praising ChatGPT's real-time grammar corrections further echo claims that AI reduces error fossilization and promotes self-regulated learning (Vadivel et al., 2023; Yang & Kyun, 2022).

Human feedback's advantage in pronunciation corroborates research highlighting instructors' ability to address complex, context-dependent skills (Mohammadkarimi, 2024; Zhou & Hou, 2024). The introduction and literature review underscored human feedback's role in refining prosodic elements and fostering interpersonal connections (Byeon, 2023; Liu et al., 2023), mirrored by learners' appreciation for personalized guidance. However, AI's limitations in capturing socio-pragmatic nuances—evident in frustrations with its pronunciation feedback—align with critiques of its contextual shortcomings (Huang, 2023; Long & Kim, 2024).

The lack of significant difference in vocabulary improvement supports the notion that both feedback types can effectively support lexical acquisition, as noted in studies comparing AI's contextual examples with human reinforcement (Chen et al., 2024). This convergence suggests that vocabulary learning benefits from a blend of technological and human inputs, reinforcing the complementary nature of these approaches.

5.2. Implications for Language Learning and Teaching

The findings advocate for a hybrid model that integrates AI-driven and human feedback to optimize language learning outcomes. AI's efficiency in handling technical aspects like grammar and vocabulary can complement human instructors' focus on pronunciation and communicative competence. This approach aligns with constructivist and sociocultural theories, which emphasize scaffolding and social interaction (Șișianu & Pușcașu, 2024; Wu et al., 2024). By leveraging AI for routine corrections, educators can dedicate more time to nuanced instruction, creating adaptive, learner-centered environments.

This hybrid model also has practical implications for resource allocation. AI tools can address scalability challenges in EFL contexts, while human expertise ensures holistic development. Additionally, the study underscores the importance of AI literacy among educators, enabling them to integrate technology effectively while maintaining

pedagogical integrity (Ibrahim, 2024; Rütli-Joy et al., 2023). Such literacy is crucial for balancing AI's capabilities with human insight, ensuring equitable and effective instruction.

5.3. Limitations and Directions for Future Research

Despite its contributions, this study has limitations. The sample size (n=80) and single-context setting (intermediate EFL learners in Iran) may limit generalizability. Future research could explore larger, more diverse populations across different proficiency levels and cultural contexts. The absence of significant differences in vocabulary improvement, engagement, and motivation also merits further investigation. Studies could examine why vocabulary outcomes were similar and identify strategies to enhance engagement with both feedback types.

Moreover, the 12-week duration provides a snapshot of short-term effects; longitudinal research could assess long-term retention and proficiency. The study did not explore learner autonomy or self-regulated learning, which are critical in AI-integrated settings. Future work could investigate how learners interact with AI tools independently and their impact on feedback efficacy.

In conclusion, this research highlights the complementary strengths of AI-driven and human feedback in EFL learning, with AI excelling in grammar and human feedback in pronunciation. A hybrid model offers a promising framework for enhancing language education, blending technological precision with human depth. As AI continues to shape educational landscapes, its thoughtful integration—supported by educator training and further research—can foster effective, inclusive, and innovative language learning environments.

This study investigated the comparative effectiveness of AI-driven and human feedback in enhancing language learning outcomes among EFL learners. The findings underscore the unique strengths of each feedback modality: AI-driven feedback demonstrated superior performance in grammar correction, offering precise, immediate, and scalable error detection, while human feedback proved more effective in improving pronunciation and fostering deeper contextual understanding. Vocabulary learning exhibited no significant differences between the two feedback approaches, suggesting that both AI and human input can effectively support lexical acquisition.

The results of this study align with existing literature on AI-assisted language learning, reinforcing the role of AI in

providing structured, data-driven feedback while emphasizing the irreplaceable value of human instruction in developing nuanced language skills. The mixed-methods analysis, integrating quantitative performance metrics and qualitative learner perceptions, provides a comprehensive understanding of how AI and human feedback function within an EFL context. Learners reported higher engagement with AI's immediate feedback on grammar, while human feedback was preferred for pronunciation due to its interactive and personalized nature.

These findings have significant implications for language pedagogy. A hybrid feedback model, combining AI's efficiency with human instructors' contextual expertise, presents an optimal strategy for enhancing language learning outcomes. AI can be leveraged for routine grammatical and lexical corrections, freeing educators to focus on complex, communicative aspects of language instruction. This balanced approach aligns with contemporary constructivist and sociocultural learning theories, advocating for adaptive, learner-centered environments.

However, the study is not without limitations. The sample size and contextual scope may restrict generalizability, necessitating further research across diverse learner populations and proficiency levels. Additionally, while the study captured short-term learning gains, longitudinal research is required to assess the sustained impact of AI and human feedback on language proficiency. Future investigations should also explore the role of learner autonomy in AI-integrated environments, examining how self-regulated learning strategies influence feedback effectiveness.

In conclusion, this research contributes to the ongoing discourse on AI in language education by highlighting the complementary nature of AI-driven and human feedback. As AI technologies continue to evolve, their thoughtful integration into EFL curricula can enhance learning efficiency while preserving the pedagogical depth of human instruction. By adopting a hybrid feedback model, educators can harness the benefits of both AI and human expertise, fostering a more effective, inclusive, and technologically enriched language learning experience.

Authors' Contributions

Authors equally contribute to this study.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethics Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants. The study received approval from Ilam university's Institutional Review Board. Participants provided informed consent, were assured of data confidentiality, and could withdraw at any time without penalty. Data were anonymized during analysis to protect privacy.

References

- Abernathy, P. (2024). Ai tools as supplementary support in language acquisition. <https://doi.org/10.20944/preprints202407.2213.v1>
- Abuhassna, H. (2024). The information age for education via artificial intelligence and machine learning: a bibliometric and systematic literature analysis. *International Journal of Information and Education Technology*, 14(5), 700-711. <https://doi.org/10.18178/ijiet.2024.14.5.2095>
- Aghaziarati, A., Nejatifar, S., & Abedi, A. (2023). Artificial Intelligence in Education: Investigating Teacher Attitudes. *AI and Tech in Behavioral and Social Sciences*, 1(1), 35-42. <https://doi.org/10.61838/kman.aitech.1.1.6>
- AlAfnan, M. (2024). Artificial intelligence and language: bridging Arabic and English with technology. *Journal of Ecohumanism*, 3(8). <https://doi.org/10.62754/joe.v3i8.4961>
- Alhusaiyan, E. (2024). A systematic review of current trends in artificial intelligence in foreign language learning. *Saudi Journal of Language Studies*, 5(1), 1-16. <https://doi.org/10.1108/sjls-07-2024-0039>
- Alzahrani, F., & Alotaibi, H. (2024). The impact of artificial intelligence on enhancing EFL writing skills among high school students. *Journal of Educational and Human Sciences*, 34. <https://doi.org/10.33193/jeahs.34.2024.482>
- Amin, M. (2023). Ai and ChatGPT in language teaching: enhancing EFL classroom support and transforming assessment techniques. *International Journal of Higher Education Pedagogies*, 4(4), 1-15. <https://doi.org/10.33422/ijhep.v4i4.554>
- Belda-Medina, J., & Goddard, M. (2024). Ai-driven digital storytelling: a strategy for creating English as a foreign language (EFL) materials. *International Journal of Linguistics Studies*, 4(1), 40-49. <https://doi.org/10.32996/ijls.2024.4.1.4>
- Bulut, O., & Wongvorachan, T. (2022). Feedback generation through artificial intelligence.
- Byeon, H. (2023). Advances in value-based, policy-based, and deep learning-based reinforcement learning. *International Journal of Advanced Computer Science and Applications*, 14(8). <https://doi.org/10.14569/ijcsa.2023.0140838>
- Chen, A., Zhang, Y., Jia, J., Liang, M., Cha, Y., & Lim, C. (2024). A systematic review and meta-analysis of ai-enabled assessment in language learning: design, implementation, and effectiveness. *Journal of Computer Assisted Learning*, 41(1). <https://doi.org/10.1111/jcal.13064>
- Chen, Y. (2024). Enhancing language acquisition: the role of ai in facilitating effective language learning. In (pp. 593-600). https://doi.org/10.2991/978-2-38476-253-8_71
- Cheung, W., & Luk, C. (2024). Implementing automated error correction and feedback loops in Kimi, a Chinese large language model.
- Cohen, S., Mompelat, L., Mann, A., & Connors, L. (2024). The linguistic leap: understanding, evaluating, and integrating ai in language education. *Journal of Language Teaching*, 4(2), 23-31. <https://doi.org/10.54475/jlt.2024.012>
- Davis, R. (2024). Korean in-service teachers' perceptions of implementing artificial intelligence (ai) education for teaching in schools and their ai teacher training programs. *International Journal of Information and Education Technology*, 14(2), 214-219. <https://doi.org/10.18178/ijiet.2024.14.2.2042>
- Davoodifard, M., & Eskin, D. (2024). Exploring the role of generative ai in second language education: insights for instruction, learning, and assessment. *Studies in Applied Linguistics and Tesol*, 24(1). <https://doi.org/10.52214/salt.v24i1.12863>
- Dewi, A., Putri, E., & Widiawati, U. (2023). Conventional written corrective feedback for EFL learners' writing skill enhancement. *International Journal of Educational Best Practices*, 7(2), 172. <https://doi.org/10.31258/ijebp.v7n2.p172-188>
- Dhanapal, C., Asharudeen, N., & Alfuruque, S. (2024). Impact of artificial intelligence versus traditional instruction for language learning: a survey. *World Journal of English Language*, 14(2), 182. <https://doi.org/10.5430/wjel.v14n2p182>
- Fitria, T. (2023). The use of artificial intelligence in education (AIED): can ai replace the teacher's role? *Epigram*, 20(2), 165-187. <https://doi.org/10.32722/epi.v20i2.5711>
- Gupta, P., Sreelatha, C., Latha, A., Raj, S., & Singh, A. (2024). Navigating the future of education: the impact of artificial intelligence on teacher-student dynamics. *EATP*, 6006-6013. <https://doi.org/10.53555/kuey.v30i4.2332>

- Hashim, H., Jamal, M., Esa, I., & Rafiq, K. (2023). Teall: technology-enhanced active language learning. *Creative Education*, 14(07), 1453-1462. <https://doi.org/10.4236/ce.2023.147092>
- Huang, J. (2023). Engineering ChatGPT prompts for EFL writing classes. *International Journal of Tesol Studies*. <https://doi.org/10.58304/ijts.20230405>
- Ibrahim, A. (2024). Assessing the knowledge and perception of artificial intelligence for teaching and research among lecturers in the faculties of arts in Nigeria. *Journal of Global Research in Education and Social Science*, 18(2), 25-33. <https://doi.org/10.56557/jogress/2024/v18i28671>
- Ifraheem, S., Rasheed, M., & Siddiqui, A. (2024). Transforming education through artificial intelligence: personalization, engagement and predictive analytics. *Journal of Asian Development Studies*, 13(2), 250-266. <https://doi.org/10.62345/jads.2024.13.2.22>
- Igbokwe, I. (2024). Artificial intelligence in educational leadership: risks and responsibilities. *EJAHSS*, 1(6), 3-10. [https://doi.org/10.59324/ejahss.2024.1\(6\).01](https://doi.org/10.59324/ejahss.2024.1(6).01)
- Kim, J., Ham, Y., & Lee, S. (2024). Differences in student-ai interaction process on a drawing task: focusing on students' attitude towards ai and the level of drawing skills. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.8859>
- Koka, N., Khan, A., & Jan, N. (2023). Exploring the attitudes and perceptions of foreign language lecturers on artificial intelligence-driven writing evaluation and feedback tools for improving undergraduate writing skills. *Journal of Southwest Jiaotong University*, 58(5). <https://doi.org/10.35741/issn.0258-2724.58.5.6>
- Konyrova, L. (2024). The evolution of language learning: exploring ai's impact on teaching English as a second language. *Eurasian Science Review*, 2(2), 133-138. <https://doi.org/10.63034/esr-42>
- Kumari, M. (2023). Perception of dental students in incorporating artificial intelligence into dental education. *Journal of Advanced Sciences*, 2(1), 41-45. <https://doi.org/10.58935/joas.v2i1.27>
- Li, Y. (2024). Usability of ChatGPT in second language acquisition: capabilities, effectiveness, applications, challenges, and solutions. *Studies in Applied Linguistics and Tesol*, 24(1). <https://doi.org/10.52214/salt.v24i1.12864>
- Liu, M. (2023). Exploring the application of artificial intelligence in foreign language teaching: challenges and future development. *SHS Web of Conferences*, 168, 03025. <https://doi.org/10.1051/shsconf/202316803025>
- Liu, S., Wright, A., Patterson, B., Wanderer, J., Turer, R., & Nelson, S. (2023). Assessing the value of ChatGPT for clinical decision support optimization. <https://doi.org/10.1101/2023.02.21.23286254>
- Liu, Y. (2023). A comparison of automated corrective feedback and traditional corrective feedback: a review study. *The Educational Review USA*, 7(9), 1365-1368. <https://doi.org/10.26855/er.2023.09.024>
- Long, L., & Kim, M. (2024). It is like a friend to me: critical usage of automated feedback systems by self-regulating English learners in higher education. <https://doi.org/10.14742/ajet.8821>
- Ma, W., Mat, A., Rahman, G., Hamzah, F., & Halim, R. (2024). Digital technologies in language education: a comprehensive review and analysis. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 41(1), 90-102. <https://doi.org/10.37934/araset.41.1.90102>
- Mahmud, M., Mat, N., Manap, M., Zakaria, S., Tazijan, F., & Wong, S. (2023). Sustainability of educational technologies: an approach to ratio and feedback. *Asian Journal of Contemporary Education*, 7(2), 56-65. <https://doi.org/10.55493/5052.v7i2.4895>
- Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Teaching*, 27(1), 98-115. <https://doi.org/10.24071/llt.v27i1.8151>
- Morales, M., Amado-Salvatierra, H., Medina, J., Barchino, R., Hernández-Rizzardini, R., & Teixeira, A. (2024). Personalized feedback in massive open online courses: harnessing the power of Lang chain and OpenAI API. *Electronics*, 13(10), 1960. <https://doi.org/10.3390/electronics13101960>
- Moybeka, A., Syariatn, N., Tatipang, D., Mushthoza, D., Dewi, N., & Tineh, S. (2023). Artificial intelligence and English classroom: the implications of ai toward EFL students' motivation. *Edumaspul - Jurnal Pendidikan*, 7(2), 2444-2454. <https://doi.org/10.33487/edumaspul.v7i2.6669>
- Muslimin, A., Mukminatien, N., & Maria, I. (2024). Evaluating Cami AI across smart stages: students' achievement and perceptions in EFL writing instruction. *Online Learning*, 28(2). <https://doi.org/10.24059/olj.v28i2.4246>
- Nurjanah, A., Salsabila, I., Azzahra, A., Rahayu, R., & Marlina, N. (2024). Artificial intelligence (AI) usage in today's teaching and learning process: a review. *Syntax Idea*, 6(3), 1517-1523. <https://doi.org/10.46799/syntax-idea.v6i3.3126>
- Nurmala, I., Irianto, S., Franchisca, S., Amsa, H., & Susanti, R. (2023). Technology-enhanced language learning: a meta-analysis study on English language teaching tools. *Journal on Education*, 6(1), 2188-2195. <https://doi.org/10.31004/joe.v6i1.3221>
- Ojeda-Ramirez, S., Rismanchian, S., & Doroudi, S. (2023). Learning about AI to learn about learning: artificial intelligence as a tool for metacognitive reflection.
- Pack, A., & Maloney, J. (2024). Using artificial intelligence in TESOL: some ethical and pedagogical considerations. *TESOL Quarterly*, 58(2), 1007-1018. <https://doi.org/10.1002/tesq.3320>
- Qin, L. (2024). Adaptive system of English-speaking learning based on artificial intelligence. *Journal of Electrical Systems*, 20(6s), 267-275. <https://doi.org/10.52783/jes.2637>
- Rütti-Joy, O., Winder, G., & Biedermann, H. (2023). Building AI literacy for sustainable teacher education. *Zeitschrift für Hochschulentwicklung*, 18(4), 175-189. <https://doi.org/10.21240/zfhe/18-04/10>
- Şişianu, A., & Puşcaşu, A. (2024). New technologies and foreign language learning: transforming language education. *Journal of Social Sciences*, 7(1), 91-99. [https://doi.org/10.52326/jss.utm.2024.7\(1\).07](https://doi.org/10.52326/jss.utm.2024.7(1).07)
- Stracke, C., Chounta, I., Holmes, W., Tlili, A., & Bozkurt, A. (2023). A standardised Prisma-based protocol for systematic reviews of the scientific literature on artificial intelligence and education. *Journal of Applied Learning & Teaching*, 6(2). <https://doi.org/10.37074/jalt.2023.6.2.38>
- Tiwari, H. (2024). Artificial intelligence in the classroom: revolutionizing English language teaching. *Journal of English Teaching and Linguistics Studies*, 6(1), 42-59. <https://doi.org/10.55215/jetli.v6i1.9757>
- Trang, N., & Thu, P. (2024). The role of AI in improving student learning outcomes: evidence in Vietnam. *International Journal of Multidisciplinary Research and Analysis*, 07(06). <https://doi.org/10.47191/ijmra.v7-i06-48>
- Vadivel, B., Shaban, A., Ahmed, Z., & Saravanan, B. (2023). Unlocking English proficiency: assessing the influence of ai-powered language learning apps on young learners' language

- p>acquisition.
- International Journal of English Language Education and Literature Studies (IJELLS)*
- , 2(6), 55-62.
- <https://doi.org/10.22161/ijeel.2.6.7>
- Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: a randomized controlled trial. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1249991>
- Wiyaka, W., Silitonga, L., Sunardi, S., & Pramudi, Y. (2024). From nervous to fluent: the impact of ai chatbot-assisted assessment on English reading anxiety and performance in Indonesia. *Theory and Practice in Language Studies*, 14(12), 3851-3860. <https://doi.org/10.17507/tpls.1412.20>
- Wu, L., Li, K., Yu, M., & Lin, Y. (2024). Application of artificial intelligence in teaching English as a foreign language: progress, challenges, and trends. *English Language Teaching and Linguistics Studies*, 6(4), 215. <https://doi.org/10.22158/eltls.v6n4p215>
- Xu, W., & Ouyang, F. (2022). The application of ai technologies in stem education: a systematic review from 2011 to 2021. *International Journal of Stem Education*, 9(1). <https://doi.org/10.1186/s40594-022-00377-5>
- 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*, 180-210. <https://doi.org/10.14742/ajet.7492>
- Yao, Y., & Chung, K. (2024). Application of neural network and structural model in ai educational performance analysis. *Sensors and Materials*, 36(3), 891. <https://doi.org/10.18494/sam4429>
- Yim, I., & Wegerif, R. (2024). Teachers' perceptions, attitudes, and acceptance of artificial intelligence (ai) educational learning tools: an exploratory study on ai literacy for young students. *Future in Educational Research*, 2(4), 318-345. <https://doi.org/10.1002/fer3.65>
- Yin, S., & Goh, D. (2024). Factors influencing university students' ai use and knowledge acquisition. *Proceedings of the Association for Information Science and Technology*, 61(1), 1162-1164. <https://doi.org/10.1002/pr2.1218>
- Yin, T., & Hanif, H. (2024). ELTC framework: revolutionizing MUET learning by crafting a dynamic chatbot learning experience. *International Journal of Academic Research in Progressive Education and Development*, 13(1). <https://doi.org/10.6007/ijarped/v13-i1/20208>
- Yu, L., & Yu, Z. (2023). Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSViewer and CitNetExplorer. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1061778>
- Yunina, O. (2023). Artificial intelligence tools in foreign language teaching in higher education institutions. *The Modern Higher Education Review*(8). <https://doi.org/10.28925/2617-5266.2023.85>
- Zhou, C., & Hou, F. (2024). Can ai empower l2 education? exploring its influence on the behavioural, cognitive and emotional engagement of EFL teachers and language learners. *European Journal of Education*, 59(4). <https://doi.org/10.1111/ejed.12750>
- Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *Sage Open*, 13(3). <https://doi.org/10.1177/21582440231193818>
- Zulkarnain, N., & Yunus, M. (2023). Primary teachers' perspectives on using artificial intelligence technology in English as a second language teaching and learning: a systematic review. *International Journal of Academic Research in Progressive Education and Development*, 12(2). <https://doi.org/10.6007/ijarped/v12-i2/17119>