Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 10, Number 1, pp: 306-319, June 2026
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v10i1.13433

EXPLORING THE IMPACT OF AI TOOLS ON EFL LEARNERS' COMMUNICATION AND AUTONOMY

Adrian, Pupung Purnawarman
Universitas Pendidikan Indonesia (UPI), Bandung, Indonesia
E-mail: adrianasgali@upi.edu

Received: 2026-04-27 Accepted: 2026-05-22 Published: 2026-06-10

Abstract

This study examines the impact of AI tools on EFL learners' communication skills and learner autonomy across varying proficiency levels. Employing an exploratory mixed-methods design that combined a 15-item Likert-scale survey with open-ended questions, data were collected from 16 purposively selected EFL learners enrolled in an undergraduate English program at an Indonesian university. The small sample was deemed appropriate for qualitative exploration rather than statistical generalization. Grammarly was used primarily for writing correction and grammatical feedback; ChatGPT supported conversational practice and essay drafting; and Duolingo provided gamified vocabulary and pronunciation exercises. Participants reported improvements in grammar accuracy, vocabulary range, and writing confidence, with gains most pronounced among beginner and intermediate learners. Advanced learners, however, expressed frustration with idiomatic expressions and cultural nuances, identifying a ceiling effect in the utility of AI tools at higher proficiency levels. Regarding autonomy, participants described using the tools for self-directed error review and goal setting, yet several acknowledged a tendency toward passive dependency, accepting AI corrections without critical reflection. These findings suggest that AI tools can complement traditional instruction but require human oversight, blended feedback strategies, and teacher scaffolding to ensure cultural and linguistic appropriateness. This study contributes to emerging research on AI in EFL by foregrounding learner autonomy and contextual challenges as proficiency-dependent variables that prior single-level studies have not fully addressed.

Keywords: *Artificial intelligence in education; communication skills; EFL learners; learner autonomy.*

1. Introduction

The integration of Artificial Intelligence (AI) in education has significantly transformed how learners acquire language skills, with AI tools, including chatbots, speech recognition systems, and language learning platforms, becoming essential components of English as Foreign Language (EFL) environments. Research consistently highlights three thematic benefits of these tools. In terms of learning outcomes, AI technologies provide immediate

feedback that enables real-time error correction and accuracy improvement (Jiang, 2022), while speech recognition features enhance pronunciation accuracy and listening comprehension (Huang, Hew, & Fryer, 2022). With respect to personalization and engagement, AI-driven platforms offer adaptive content and personalized recommendations that sustain learner motivation (Song & Song, 2023; Ma & Chen, 2024), and gamification elements such as streaks and rewards further reinforce engagement and a sense of progression (Luo, 2023). Regarding interactivity, AI-driven chatbots simulate real-life conversations in secure, non-judgmental environments (Han et al., 2023a), and tools such as ChatGPT and Grammarly provide structured feedback on grammar, vocabulary, and style (Barrot, 2023). These benefits align with Deci and Ryan's (1985) Self-Determination Theory, which situates autonomy as central to intrinsic motivation, and reflect broader advances in natural language processing that have enhanced the usability of AI platforms.

Despite these benefits, AI tools face significant limitations that constrain their effectiveness in complex language learning contexts, particularly regarding cultural nuance, idiomatic language, and emotional responsiveness. Tafazoli (2024) cautions that excessive reliance on AI may undermine critical thinking and reduce opportunities for authentic language practice. Sohnui, Mukhlis, and Pendri (2025) and Jiang (2022) both note that AI systems fail to incorporate cultural idioms and context-specific expressions vital to global communication, while Huang et al. (2023) advocate for culturally diverse datasets and scenario-based simulations to address these gaps. Moreover, Kohnke and Moorhouse (2025) observe that AI lacks the emotional intelligence and deeper comprehension of learner needs that human instructors provide. These interrelated shortcomings highlight a critical need for balanced integration of AI and traditional teaching methods, a need that remains insufficiently addressed, particularly across varying learner proficiency levels, and that situates this study within broader debates on the pedagogical role of AI in education.

Few studies have examined how AI tools affect communication skills and learner autonomy across proficiency levels, particularly in ways that balance their documented benefits against cultural and contextual limitations. This study addresses that gap by analyzing how EFL learners at beginner and advanced levels engage with AI tools, providing insights for optimizing these technologies in personalized language education. By identifying opportunities for developers to enhance AI sensitivity to cultural and contextual nuances (Lewis, 2025; Sohnui, Mukhlis, & Pendri, 2025) and by foregrounding the importance of integrating AI tools with conventional teaching methods to foster critical thinking and authentic language practice, this study contributes both theoretical insights and practical strategies to the broader discourse on AI in language education.

Based on the above, this study addresses the following research questions:

1. How do EFL learners perceive AI tools' impact on their communication skills?
2. How do AI tools influence learner autonomy in language learning?

2. Literature Review

2.1 AI Tools in EFL Learning: An Overview

AI-powered platforms, principally Grammarly, ChatGPT, and Duolingo, have reshaped EFL pedagogy by delivering immediate feedback, personalized learning pathways, and interactive practice environments. Adaptive learning systems adjust content difficulty based on individual learners' progress, enabling tailored experiences that promote engagement

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 10, Number 1, pp: 306-319, June 2026
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v10i1.13433

and retention (Ma & Chen, 2024), while automating routine tasks such as error correction and progress tracking, thereby allowing instructors to focus on higher-order pedagogical guidance (Jiang, 2022). These tools align with learners' intrinsic motivation through gamified and conversational features; Duolingo's leaderboards, streaks, and reward mechanics, for instance, make consistent practice more accessible for beginners and advanced learners alike (Luo, 2023; Deterding, Dixon, Khaled, & Nacke, 2011).

Despite these strengths, AI tools share common limitations that constrain their pedagogical value, particularly their inability to handle cultural nuance, idiomatic language, and higher-order critical thinking. Grammarly identifies grammatical errors reliably yet cannot evaluate tonal appropriateness for specific audiences or rhetorical contexts (Song & Song, 2023). Similarly, AI systems lack the emotional intelligence and contextual sensitivity that characterize effective human instruction (Kohnke & Moorhouse, 2025) and frequently fail to incorporate the cultural idioms and situation-specific expressions essential to global communication (Sohnui, Mukhlis, & Pendri, 2025). These shortcomings reinforce the case for hybrid models that combine AI efficiency with the contextual depth of human educators (Huang, Hew, & Fryer, 2022).

2.2 AI Tools and Communication Skills

Research converges on three dimensions through which AI tools strengthen EFL communication. In academic writing, consistent use of Grammarly has been shown to improve coherence, structure, and vocabulary choice, which is especially critical for non-native speakers in formal contexts (Song & Song, 2023). Detailed, context-specific feedback builds learners' writing confidence and deepens understanding of language usage beyond rote memorization (Jiang, 2022; Luo, 2023). While Grammarly demonstrably supports accuracy and surface-level correctness, it may not foster creativity or rhetorical awareness; its corrective logic privileges convention over voice, a limitation educators should account for when integrating it into writing curricula.

In oral communication, ChatGPT's simulated dialogues have produced significant gains in fluency, vocabulary use, and confidence by providing a judgment-free environment for language experimentation (Han et al., 2023b). Duolingo's speech recognition further advances pronunciation accuracy by detecting subtle phonetic errors and enabling learners to replicate complex phonetic distinctions (Huang, Hew, & Fryer, 2022; Song & Song, 2023). However, simulated interaction differs qualitatively from authentic discourse; AI interlocutors may not fully replicate the sociolinguistic unpredictability, register variation, or pragmatic negotiation that define real communicative encounters, and learners may develop fluency without the adaptive competence required in genuine cross-cultural exchanges.

2.3 AI Tools and Learner Autonomy

The autonomy-fostering potential of AI tools extends well beyond gamification. Adaptive platforms such as Duolingo enable learners to set objectives, monitor progress, and self-direct their learning pace, behaviors that operationalize the autonomy construct central to Deci and Ryan's (1985) Self-Determination Theory. More recent work within motivational self-systems frameworks similarly emphasizes the role of self-regulation in sustaining long-term language engagement (Huang, Hew, & Fryer, 2022). Crucially, however, genuine autonomy requires metacognitive engagement, where learners must evaluate their own

understanding, identify gaps, and adjust strategies accordingly (Mohebbi, 2025). When AI tools provide only corrective output without prompting reflection, they risk producing surface compliance rather than deep self-regulation (Achuthan, 2025). Educators can scaffold deeper autonomy by embedding reflective journaling, error-analysis tasks, and learner-generated goal setting alongside AI feedback routines, thereby ensuring that technology amplifies rather than substitutes for independent thinking.

Gamification reinforces motivation through rewards and progression mechanics (Deterding, Dixon, Khaled, & Nacke, 2011; Luo, 2023), yet its benefits are bounded. Tafazoli (2024) warns that overreliance on AI rewards may cultivate passive learning, in which students await algorithmic correction rather than actively consolidating knowledge. Sohnui, Mukhlis, and Pendri (2025) further caution that gamified mechanics can overshadow substantive learning objectives. A balanced integration of AI-driven support and guided human instruction remains essential, as AI tools are structurally unable to provide the cultural, contextual, and critical dimensions that human educators can (Huang et al., 2023).

2.4 Research Gap and Novelty of the Present Study

Prior research has established the effectiveness of individual AI tools at specific proficiency levels, yet three gaps persist. Most studies examine a single tool or learner population, limiting cross-tool and cross-proficiency comparability. The simultaneous relationship between AI-driven communication development and learner autonomy remains underexplored within a single empirical frame. Crucially, no study has directly compared the effectiveness of AI tools across the full EFL proficiency spectrum, from beginner (A1–A2) to advanced (C1–C2), while accounting for both communication and autonomy outcomes. The present study addresses all three gaps through a mixed-methods design that examines Grammarly, ChatGPT, and Duolingo across varied proficiency levels, offering educators, developers, and policymakers a comparative account of AI's dual pedagogical impact.

3. Research Method

3.1 Research Design

This study adopts a primarily qualitative design supplemented by quantitative data, reflecting the exploratory nature of the inquiry and the small purposive sample of 16 participants. Although the study incorporates Likert-scale survey items to identify preliminary perceptual trends, the sample size precludes generalizable statistical inference; quantitative data therefore serve an orienting function, providing a structured basis for the qualitative elaboration that follows. This approach is consistent with Creswell and Plano Clark's (2011) Explanatory Sequential Design, in which an initial phase of quantitative data collection and analysis is followed by a qualitative phase that explains and contextualizes the numerical patterns. Such a design is well-suited to educational research, where understanding the depth of learners' experiences is prioritized over statistical generalization (Merriam & Tisdell, 2016).

The mixed-methods framework is justified here as exploratory: the quantitative strand maps broad perceptual tendencies across proficiency levels, while the qualitative strand, through open-ended survey responses, illuminates the mechanisms and meanings behind those tendencies. Johnson, Onwuegbuzie, and Turner (2007) argue that integrating both strands within a single study enhances construct validity and produces richer, more credible findings than either method alone.

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 10, Number 1, pp: 306-319, June 2026
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v10i1.13433

3.2 Participants

Sixteen EFL learners enrolled in an undergraduate English program at an Indonesian university participated in this study. Participants ranged in age from 19 to 26 years. Proficiency was distributed as follows: five beginners (A1–A2), five intermediate learners (B1–B2), and six advanced learners (C1–C2), as verified through self-reported CEFR placement aligned with institutional diagnostic scores. All participants had used at least one of the three focal AI tools, Grammarly, ChatGPT, or Duolingo, for at least 3 months prior to data collection. Purposive sampling was employed to ensure representation across the full proficiency spectrum and to enable cross-level comparison (Creswell, 2014). Informed consent was obtained from all participants, who were briefed on the study's purpose and their right to withdraw at any time without consequence.

3.3 Research Instruments

Two instruments were used: a Likert-scale survey and a set of open-ended questions. Both were embedded within a single online form and administered sequentially, generating complementary quantitative and qualitative data.

Likert-scale survey. The survey comprised 15 statements rated on a five-point scale (1 = Strongly Disagree to 5 = Strongly Agree), organized into three categories aligned with the research questions: (1) Communication Skills (e.g., “AI tools have improved my ability to express myself in English”); (2) Learner Autonomy (e.g., “AI tools motivate me to learn English independently”); and (3) Engagement and Motivation (e.g., “Using AI tools makes learning English more enjoyable”). The survey was selected for its capacity to generate standardized, comparable data across participants (Dörnyei & Csizér, 2012).

Open-ended questions. Four open-ended questions were designed to elicit rich accounts from participants with experience using AI tools. Two questions targeted communication skills, and two addressed autonomy, directly mirroring the study’s research questions. This question-to-objective alignment follows best practice in qualitative instrument design (Patton, 2015) and ensures that responses yield data amenable to thematic analysis. Representative questions included: “Can you describe specific ways in which AI tools have influenced your communication skills?” and “In what ways, if any, have AI tools changed how independently you approach your language learning?” Open-ended formats allow participants to construct meaning without the constraints of predefined categories, producing narratives that surface contextual nuance unavailable through closed items (Merriam & Tisdell, 2016).

3.4 Data Collection Procedure

Data were collected via Google Forms over a two-week period. The online format was chosen for its accessibility, scheduling flexibility, and ability to reach participants across geographic locations, advantages that are particularly relevant in post-pandemic Indonesian higher education contexts (Mailizar, Almanthari, Maulina, & Bruce, 2020). The combined form required approximately 20 minutes to complete. Participants were prompted to reflect on their most recent interactions with the focal AI tools before responding, grounding answers in recent, concrete experience. A limitation of this approach is the absence of face-to-face probing; follow-up clarification of ambiguous responses was not possible, potentially

constraining the depth of the qualitative data. Future studies should consider supplementing online instruments with semi-structured interviews to mitigate this constraint.

3.5 Data Analysis

This research analyses the Likert-scale responses as research data by using descriptive statistics, means, standard deviations, and frequency distributions to identify perceptual patterns across the three survey categories and proficiency levels. Given the exploratory nature of the study and the sample size of 16, inferential statistics were not employed; the descriptive findings serve as an orienting map for the qualitative analysis rather than as generalizable conclusions (Field, 2013). Results are presented in summary tables.

Furthermore, the open-ended responses were analyzed using Braun and Clarke's (2006) six-phase thematic analysis framework: familiarization, initial coding, theme generation, theme review, theme definition, and reporting. Coding was conducted manually by the primary researcher, with a second coder independently analyzing 30% of the data to establish inter-coder reliability; Cohen's kappa was calculated to assess agreement (target $\kappa \geq .70$). Discrepancies were resolved through discussion until consensus was reached. This process ensured that emergent themes such as enhanced fluency, metacognitive awareness, and technology dependency were grounded in participant data rather than researcher interpretation.

3.6 Ethical Considerations

All participants provided written informed consent prior to participation and were advised of the study's purpose, data usage, and their unconditional right to withdraw. No personally identifiable information was retained; all data were anonymized using participant codes at the point of collection. Data are stored securely on password-protected institutional servers and will be deleted five years after publication, in accordance with institutional data governance policy.

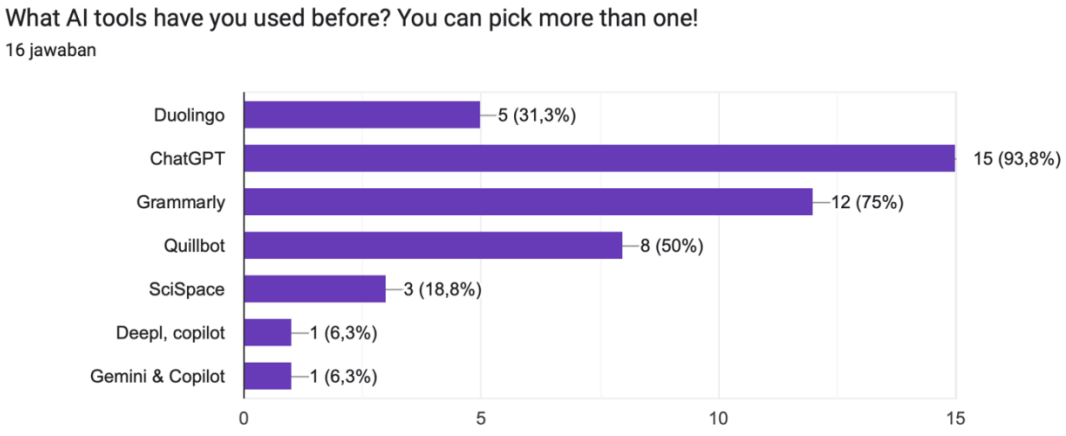
4. Results and Discussion

4.1 Results

4.1.1 AI Tool Preferences and Usage Frequency

Figure 1.

AI tool preferences and usage frequency among EFL learners (N = 16).

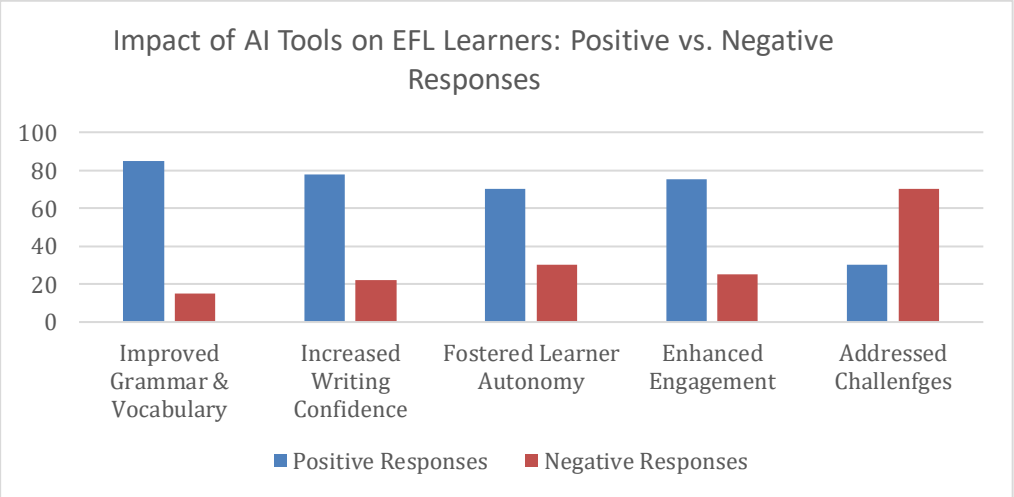


Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 10, Number 1, pp: 306-319, June 2026
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v10i1.13433

Figure 1 reveals a clear hierarchy of tool adoption. ChatGPT and Grammarly were the most commonly used, reported by 93.8% and 75% of participants, respectively, reflecting their broad applicability across writing, vocabulary, and conversational practice. QuillBot (50%) served a complementary paraphrasing function, while Duolingo (31.3%) appealed primarily to beginner learners through its gamified structure. Specialized tools such as SciSpace (18.8%), DeepL, Copilot, and Gemini (each 6.3%) attracted a narrower audience. The pattern potentially reflects learners' preference for tools offering immediate feedback rather than those structured around discrete exercises.

4.1.2 Learners' Perceptions of AI Impact

Figure 2.
Likert-scale responses on AI tools' perceived impact on communication skills, autonomy, and engagement (N = 16).



Three major themes emerged from the survey and open-ended responses: perceived improvements in communication skills, the development of learner autonomy, and proficiency-related differences in engagement and motivation. Together, these themes illuminate the multifaceted role of AI-powered language learning tools in shaping learners' experiences and perceptions of language development.

Communication Skills

Participants across all proficiency levels reported that AI-powered tools, particularly Grammarly and ChatGPT, contributed positively to their writing development by enhancing grammatical accuracy and expanding vocabulary repertoire. These perceived benefits were especially evident among beginner and intermediate learners, for whom real-time corrective feedback appeared to address foundational linguistic challenges. One beginner participant remarked:

> "AI makes it easier for me to gain a lot of new vocabulary in writing. This helps me better in my writing tasks."

This perception reflects a broader trend identified in the survey findings, where fourteen of the sixteen participants (87.5%) agreed or strongly agreed that AI tools improved their written expression. The convergence between quantitative and qualitative data

suggests that learners generally perceive AI-assisted feedback as a valuable mechanism for strengthening linguistic accuracy and lexical development. Nevertheless, these findings should be interpreted with caution due to the limited sample size and the reliance on self-reported perceptions. While participants reported noticeable improvements, the extent to which such gains are sustained over time or transferred to independent, unaided writing remains unclear.

Notably, advanced learners offered a more critical evaluation of AI-generated support. Although they acknowledged the usefulness of AI tools for correcting surface-level errors, they highlighted limitations in the tools' ability to account for contextual appropriateness, discourse conventions, and idiomatic language use. As one advanced learner explained:

> "AI isn't good at constructing sentences contextually, especially in discourse or idiomatic expressions."

This observation points to an important distinction between linguistic accuracy and communicative competence. Whereas AI tools appear effective in supporting lower-level language skills such as grammar and vocabulary, they may be less capable of fostering the rhetorical, pragmatic, and contextual awareness required for advanced language use. The absence of similar concerns among beginner learners suggests that the perceived effectiveness of AI tools may vary according to proficiency level. As learners develop greater linguistic sophistication, they become more sensitive to the limitations of AI-generated language, revealing a potential ceiling effect in the pedagogical value of such tools.

Learner Autonomy

A second prominent theme concerned the influence of AI tools on learner autonomy. Participants frequently described these technologies as facilitating independent learning by enabling self-correction, personalized feedback, goal setting, and ongoing monitoring of learning progress. Rather than relying exclusively on teachers for feedback, learners reported using AI tools to identify weaknesses and revise their work autonomously. One intermediate learner characterized AI as a "guide or partner," illustrating the perception of technology as a supportive learning companion rather than a substitute for human instruction.

However, the findings also revealed a tension between autonomy enhancement and technological dependency. While AI tools encouraged self-directed learning, some participants expressed concern that frequent reliance on automated feedback reduced opportunities for critical reflection. One participant commented:

> "Sometimes I find myself depending too much on AI instead of thinking critically about my mistakes."

This concern was echoed by four of the sixteen participants, indicating that the relationship between AI use and learner autonomy is not entirely straightforward. Rather than uniformly promoting independent learning, AI tools may simultaneously foster and constrain autonomy, depending on the manner in which learners engage with them. Learners who critically evaluate AI-generated feedback may develop greater metacognitive awareness and self-regulation, whereas those who accept suggestions uncritically may

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 10, Number 1, pp: 306-319, June 2026
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v10i1.13433

become increasingly dependent on technological assistance. These findings suggest that the educational value of AI tools lies not merely in their availability but in learners' ability to engage with them reflectively and strategically.

Engagement and Motivation

The third theme concerned learner engagement and motivation. Participants identified immediate feedback and gamified learning features as key factors sustaining their interest and participation in language-learning activities. Applications such as Grammarly and ChatGPT were valued for providing instant responses to learner input, while Duolingo's reward systems, streaks, and progress indicators were frequently mentioned as sources of motivation.

These motivational benefits were particularly salient among beginner learners, who appreciated the low-pressure learning environment and the sense of accomplishment generated by incremental achievements. The combination of immediate feedback and gamification appeared to reduce anxiety while promoting sustained engagement with language-learning tasks. Such findings suggest that AI-powered tools can enhance motivational outcomes by making language learning more interactive, accessible, and rewarding.

In contrast, advanced learners tended to view gamified elements as less relevant to their learning objectives. Their motivation appeared to be driven more by opportunities for nuanced language practice, contextualized communication, and higher-order skill development than by reward-based mechanisms. This divergence indicates that learner engagement is influenced not only by technological features but also by proficiency level and learning goals. Consequently, the design and implementation of AI-assisted language-learning environments should account for the differing motivational needs of learners at various stages of language development.

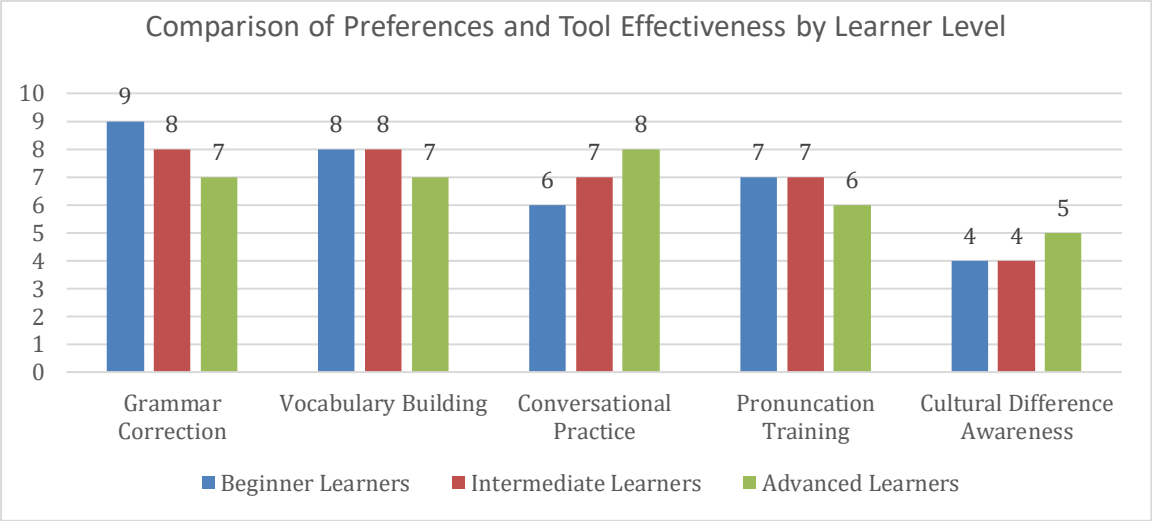
Overall, the findings demonstrate that AI-powered language learning tools are perceived as valuable resources for improving communication skills, supporting learner autonomy, and enhancing motivation. However, their effectiveness is not uniform across proficiency levels, and their pedagogical benefits depend significantly on how learners engage with the feedback and learning opportunities they provide.

4.1.3 Challenges and Limitations of AI Tools

Participants identified three recurring limitations. First, occasional inaccuracies in AI-generated output undermined confidence: "It's not always relevant or correct; sometimes the answers are false." Second, advanced learners noted the tools' inability to handle cultural nuance or idiomatic meaning: "AI tools lack the human-like understanding needed for translating nuanced concepts." Third, some learners struggled with prompt formulation, indicating that effective AI use requires its own metalinguistic skill set: "Sometimes it's difficult to make specific prompts so that AI can help effectively." These challenges are not incidental; they reflect structural constraints of current natural language processing systems that pedagogical integration must acknowledge.

4.1.4 Variations by Proficiency Level

Figure 3.
Mean effectiveness ratings (0–10) across language-learning categories by proficiency group (A1–A2, B1–B2, C1–C2).



Beginners rated AI tools most favorably for grammar correction and vocabulary building, with Grammarly and Duolingo receiving the highest endorsements. Their engagement was primarily receptive: they relied on AI for correction rather than generation. Intermediate learners extended this to essay outlining and conversational scaffolding, though they began to register tool limitations in complex compositional tasks. Advanced learners reported the lowest ratings, particularly in Cultural Difference Awareness and Idiomatic Expression, two categories where AI tools were consistently described as insufficient. This convergence of quantitative ratings and qualitative critique across proficiency levels lends coherence to the findings despite the small sample size.

4.2 Discussion

4.2.1 Communication, Autonomy, and Motivation: An Integrated Analysis

The present findings partially align with, but also complicate, established accounts of AI-enhanced language learning. Song and Song (2023) reported significant gains in coherence and structure among Grammarly users; the current participants echoed these gains in accuracy, though advanced learners' critiques suggest the benefit is uneven across proficiency levels, a nuance that Song and Song's (2023) single-level design could not capture. Similarly, while Luo (2023) and Deterding, Dixon, Khaled, and Nacke (2011) emphasize gamification as a generalized motivational scaffold, the present data indicate that its appeal diminishes substantially at advanced levels, where learners seek contextually rich feedback rather than reward-based progression.

The autonomy findings extend Deci and Ryan's (1985) self-determination framework in a productive direction. Although the theory posits autonomy as intrinsically motivating, the dependency accounts reported here suggest that perceived autonomy in AI-assisted contexts may be illusory: learners who accept corrections without critical reflection may be engaging in tool-directed rather than self-directed learning. More recent motivational frameworks, such as the Ideal L2 Self model (Dörnyei, 2009) and self-regulation theory

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 10, Number 1, pp: 306-319, June 2026
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v10i1.13433

(Zimmerman, 2002), are better equipped to account for this distinction, foregrounding the metacognitive and volitional dimensions that AI interaction alone cannot develop. The implication is that using autonomy-supportive AI requires deliberate pedagogical scaffolding; access to technology is necessary but not sufficient.

The persistent limitations identified by advanced learners, idiomatic inaccuracy, cultural opacity, and the absence of rhetorical feedback, corroborate Sohnui, Mukhlis, and Pendri (2025) and Jiang (2022), but the present study adds proficiency as a moderating variable. Crucially, these limitations were not merely incidental frustrations; they constrained the type of communicative practice available at higher levels, effectively creating a proficiency ceiling for unaided AI use. This finding has direct implications for how AI tools are positioned in curricula: presenting them as universally applicable risks overstating their pedagogical value for advanced learners.

4.2.2 Implications for Practice

For educators, the findings suggest that AI tools should be integrated selectively and critically rather than broadly. At beginner and intermediate levels, tools such as Grammarly and ChatGPT can productively scaffold accuracy and confidence when paired with reflection tasks that prompt learners to evaluate, rather than simply accept, AI-generated feedback. At advanced levels, AI is most appropriately positioned as a first-draft resource, with human feedback retained for rhetorical, cultural, and pragmatic dimensions. Across all levels, educators should design activities that make the limitations of AI visible, for instance, by comparing AI output with expert human feedback to counter passive dependency.

For developers, the proficiency-differentiated ratings in Figure 3 point to clear design priorities: culturally diverse training datasets, scenario-based simulations for professional and intercultural communication, and advanced-level features that move beyond surface correction toward discourse-level feedback (Huang, Hew, & Fryer, 2022; Huang et al., 2023). The gap between beginner satisfaction and advanced dissatisfaction represents both a pedagogical problem and a product opportunity.

5. Conclusion

This exploratory study examined how 16 EFL learners with varying proficiency levels perceived the impact of Grammarly, ChatGPT, and Duolingo on their communication skills and learner autonomy. Given the small purposive sample, the following observations are offered as preliminary rather than generalizable. Participants reported improvements in grammar accuracy, vocabulary range, and writing confidence, with gains most pronounced among beginner and intermediate learners who benefited from real-time corrective feedback and structured practice opportunities. Advanced learners, by contrast, reported a ceiling effect: while AI tools supported surface-level accuracy, they were perceived as inadequate for contextually sensitive tasks involving cultural nuance, idiomatic expression, and discourse-level coherence. Regarding autonomy, participants described the tools as enabling self-directed review and goal setting, yet several also acknowledged a tendency toward passive dependency, accepting AI corrections without critical reflection, suggesting that autonomy-supporting and autonomy-undermining effects may coexist within the same learner experience.

A recurring theme across proficiency levels was the bounded utility of AI tools: effective as technical scaffolds, insufficient as substitutes for human instruction. A hybrid approach in which AI handles preliminary correction and human educators provide contextual, cultural, and rhetorical guidance may offer a more productive balance than either resource alone (Huang, Hew, & Fryer, 2022; Kohnke & Moorhouse, 2025). This conclusion extends prior work by identifying proficiency level as a moderating variable that prior single-level studies could not capture.

For educators. AI tools should be embedded within critically reflective pedagogical tasks rather than positioned as independent correction utilities. Activities that require learners to evaluate, question, and selectively apply AI feedback rather than passively accept it are more likely to build the metacognitive and rhetorical competencies that current tools cannot develop on their own.

For developers. The consistently low ratings from advanced learners in the cultural and idiomatic categories indicate a clear design gap. Incorporating localized language datasets, culturally diverse scenario-based simulations, and discourse-level feedback mechanisms would extend the tool's utility beyond beginner and intermediate proficiency (Lewis, 2025; Sohnui, Mukhlis, & Pendri, 2025).

For policymakers. Institutional frameworks for AI integration in language education should address equitable access, data privacy, and academic integrity. Guidelines should avoid one-size-fits-all mandates and instead support differentiated, context-sensitive adoption that accounts for learner proficiency and instructional context.

The primary limitation of this study is its small purposive sample drawn from a single institutional context, which limits the generalizability of the findings. The cross-sectional survey design also captures perceptions at one point in time, precluding conclusions about sustained learning outcomes. Four directions for future research follow from these constraints: (1) longitudinal designs to track whether AI-reported gains translate into durable, unaided language performance; (2) investigation of culturally adaptive AI systems trained on diverse, localized datasets to determine whether enhanced sensitivity narrows the advanced-learner gap; (3) exploration of multimodal and voice-interactive platforms for pronunciation and real-time conversation training; and (4) replication across varied EFL contexts and cultural settings to build a more generalizable evidence base.

References

Achuthan, K. (2025). Artificial intelligence and learner autonomy: A meta-analysis of self-regulated and self-directed learning. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1738751

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, Article 100745. https://doi.org/10.1016/j.asw.2023.100745

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer. https://doi.org/10.1007/978-1-4899-2271-7

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning future media environments* (pp. 9–15). Association for Computing Machinery. https://doi.org/10.1145/2181037.2181040

Dörnyei, Z. (2009). The L2 motivational self-system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.

Dörnyei, Z., & Csizér, K. (2012). How to design and analyze surveys in second language acquisition research. In A. Mackey & S. M. Gass (Eds.). *Research methods in second language acquisition: A practical guide* (pp. 74–94). Wiley-Blackwell.

Field, A. (2013). *Discovering statistics using IBM SPSS statistic*s (4th ed.). Sage.

Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., Kim, J., Lee, T. Y., Hong, H., Ahn, S. Y., & Oh, A. (2023a). RECIPE: How to integrate ChatGPT into EFL writing education. In *Proceedings of the 10th ACM conference on learning @ scale*, (pp. 416–420). https://doi.org/10.1145/3573051.3596200

Han, J., Yoo, H., Myung, J., Kim, M., Lee, T. Y., Ahn, S.-Y., & Oh, A. (2023b). ChEDDAR: Student-ChatGPT dialogue in EFL writing education [Preprint]. *arXiv*. https://arxiv.org/abs/2309.13243v2

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

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues, and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112–131. https://doi.org/10.30191/ETS.202301_26(1).00019

Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review of artificial intelligence in the EFL context. *Frontiers in Psychology, 13*, 1049401. https://doi.org/10.3389/fpsyg.2022.1049401

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112–133. https://doi.org/10.1177/1558689806298224

Kohnke, L., & Moorhouse, B. L. (2025). Enhancing the emotional aspects of language education through generative artificial intelligence (GenAI): A qualitative investigation. *Computers in Human Behavior*, *158,* Article 108600. https://doi.org/10.1016/j.chb.2025.108600

Lewis, A. A. (2025). Unpacking cultural bias in AI language learning tools: An analysis of impacts and strategies for inclusion in diverse educational settings. *International Journal of Research and Innovation in Social Science, 9*(1), 1878–1892. https://doi.org/10.47772/IJRISS.2025.9010151

Luo, Z. (2023). The effectiveness of gamified tools for foreign language learning (FLL): A systematic review. *Behavioral Sciences, 13*(4), 331. https://doi.org/10.3390/bs13040331

Ma, Y., & Chen, M. (2024). AI-empowered applications' effects on EFL learners' engagement in the classroom and academic procrastination. *BMC Psychol 12*, Article 739. https://doi.org/10.1186/s40359-024-02248-w

Mailizar, Almanthari, A., Maulina, S., & Bruce, S. (2020). Secondary school mathematics teachers' views on e-learning implementation barriers during the COVID-19 Pandemic: The case of Indonesia. *Eurasia Journal of Mathematics, Science and Technology Education, 16*(7), em1860. https://doi.org/10.29333/ejmste/8240

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation (4th ed.).* Jossey-Bass.

Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: A systematic review. *Cogent Education, 12*(1). Article 2433814. https://doi.org/10.1080/2331186X.2024.2433814

Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Sage.

Sohnui, S., Mukhlis, M., & Pendri, A. (2025). Artificial intelligence in the perspective of students in BIPA (Bahasa Indonesia for Foreign Speakers) learning. *J-LELC: Journal of Language Education, Linguistics, and Culture, 5*(1), 42–48. https://doi.org/10.25299/j-lelc.2025.22163

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Tafazoli, D. (2024). Critical appraisal of artificial intelligence-mediated communication in language education. In H. P. Bui, R. Kumar, & N. K. Kamila (Eds.), *Innovations and applications of technology in language education* (pp. 62–79). Taylor & Francis. https://doi.org/10.1201/9781003473916-5

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2