Asalibuna
مجلة اللغة العربية وتعليمها
E-ISSN: 2656-3533
P-ISSN: 2615-5885
Volume 09, Issue 02, 2025
DOI: 10.30762/asalibuna.v9i02.6810

The Influence of Generative AI on Student Engagement and Academic Outcomes in Arabic Language Learning

Azkia Muharom Albantani[1]; Abd. Rozak[2]; Muhammad Sabri Sahrir[3]; Bilqis Khoiriyyah[4];
[1, 2, 4] UIN Syarif Hidayatullah Jakarta, Indonesia;
[3]International Islamic University, Malaysia (IIUM)

Abstract:

This study examines the influence of generative artificial intelligence (AI) on student engagement and academic outcomes in Arabic language learning through a quantitative descriptive survey design that captures learners' perceptions of AI-assisted instruction. Participants were students engaged in Arabic language learning contexts who had used generative AI tools to support their studies. The findings show that learners perceived AI as most effective in facilitating access to learning materials and enhancing vocabulary acquisition, while its role in sustaining focus and supporting writing skill development received the lowest ratings. Qualitative findings further indicate that students valued AI for its immediacy, flexibility, and personalized feedback, alongside concerns about overreliance and reduced depth of engagement. Overall, the results suggest that while generative artificial intelligence provides meaningful pedagogical support in Arabic language learning—particularly for vocabulary development and resource accessibility—intentional instructional guidance remains necessary to address its limitations, contributing novel mixed-methods evidence to the underexplored field of AI-assisted Arabic as a second language learning.

Keywords: Arabic language learning; AI technology; student perceptions; technology-assisted learning; language acquisition;

Submitted : September 10th 2025
Revised : January 23rd 2026
Accepted : January 31st 2026

Correspondence Author:
Azkia Muharrom Albantani
azki@uinjkt.ac.id

Open Access Journal
Availabel at: https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/asalibuna

Asalibuna
مجلة اللغة العربية وتعليمها
E-ISSN: 2656-3533
P-ISSN: 2615-5885
Volume 09, Issue 02, 2025
DOI: 10.30762/asalibuna.v9i02.6810

Pengajaran Bahasa Arab untuk Anak-Anak: Studi Kasus di Sekolah Dasar Terpadu Islam

Azkia Muharom Albantani[1]; Abd. Rozak[2]; Muhammad Sabri Sahrir[3]; Bilqis Khoiriyyah[4];
[1, 2, 4] UIN Syarif Hidayatullah Jakarta, Indonesia;
[3]International Islamic University, Malaysia (IIUM)

Abstrak:
Penelitian ini mengkaji pengaruh kecerdasan buatan generatif (AI) terhadap keterlibatan siswa dan hasil akademik dalam pembelajaran bahasa Arab melalui desain survei deskriptif kuantitatif yang menangkap persepsi siswa terhadap instruksi yang didukung AI. Peserta penelitian adalah siswa yang terlibat dalam konteks pembelajaran bahasa Arab dan telah menggunakan alat kecerdasan buatan generatif untuk mendukung studi mereka. Hasil penelitian menunjukkan bahwa peserta didik menganggap AI paling efektif dalam memfasilitasi akses ke bahan pembelajaran dan meningkatkan penguasaan kosakata, sementara perannya dalam mempertahankan fokus dan mendukung pengembangan keterampilan menulis mendapatkan penilaian terendah. Temuan kualitatif lebih lanjut menunjukkan bahwa siswa menghargai kecerdasan buatan (AI) karena kecepatan, fleksibilitas, dan umpan balik yang dipersonalisasi, meskipun ada kekhawatiran tentang ketergantungan berlebihan dan berkurangnya kedalaman keterlibatan. Secara keseluruhan, hasil penelitian menunjukkan bahwa meskipun kecerdasan buatan generatif memberikan dukungan pedagogis yang berarti dalam pembelajaran bahasa Arab—terutama dalam pengembangan kosakata dan aksesibilitas sumber daya—panduan instruksional yang terencana tetap diperlukan untuk mengatasi keterbatasannya, sehingga memberikan bukti baru berbasis metode campuran dalam bidang yang belum banyak dieksplorasi mengenai pembelajaran bahasa Arab sebagai bahasa kedua dengan bantuan kecerdasan buatan.

Kata Kunci: Pembelajaran bahasa Arab; teknologi kecerdasan buatan (AI); persepsi siswa; pembelajaran yang didukung teknologi; penguasaan Bahasa;

Submitted : September 10th 2025
Revised : January 23rd 2026
Accepted : January 31st 2026

Correspondence Author:
Azkia Muharrom Albantani
azki@uinjkt.ac.id

INTRODUCTION

Generative Artificial Intelligence (AI) has emerged as a transformative force in language education, offering unprecedented tools for personalized and adaptive learning. Unlike Google translate, which has many weaknesses (Rahmat Satria, dkk, 2023) AI can be a stronger tool to learn new languages. (In Arabic language pedagogy, recent advancements leverage generative AI for tasks ranging from automated feedback on grammar to immersive conversational simulations (Khan et al., 2024). These technologies align with Sahrir (2025) theory of Learning Effectiveness, which posits that optimized instructional tools enhance core competencies, such as vocabulary acquisition, grammatical accuracy, Generative AI–enabled instructional technologies align with Sahrir's (2025) Learning Effectiveness framework, which conceptualizes effective learning as the optimization of instructional tools to enhance measurable linguistic competencies. In the context of Arabic as a second language (L2), this study operationalizes learning effectiveness through Arabic-specific competencies, including vocabulary acquisition, morphological awareness (ṣarf), orthographic accuracy (harakāt), phonological precision (makhārij al-ḥurūf), and the comprehension of idiomatic expressions. These competencies reflect the distinctive structural and orthographic features of Arabic that require sustained exposure, feedback, and adaptive practice—capacities increasingly supported by generative AI systems.

Despite this potential, empirical research on generative AI in Arabic higher education remains limited, particularly regarding its integrated influence on student engagement and academic rigor. Existing literature reveals several critical gaps. First, while prior studies demonstrate AI's effectiveness in supporting foundational drills and lexical retrieval, its role in fostering higher-order Arabic skills, such as idiomaticity, extended writing proficiency, and culturally embedded language use, remains underexplored (Zapata, 2025). Second, although Sahrir's (2025) engagement-related constructs—motivation, interactivity, and perseverance—are theoretically central to learning effectiveness, few studies have examined whether AI-supported Arabic instruction sustains these dimensions over time, especially in cognitively demanding tasks such as morphological analysis and written production. Third, current research largely emphasizes technical affordances of AI tools, overlooking learner-centered outcomes, including emotional engagement, self-regulation, and persistence in Arabic L2 learning contexts (Pillai et al., 2024).

To address these gaps, this study adopts an integrated theoretical framework that bridges Learning Effectiveness, Student Engagement, and Computer-Assisted Language Learning (CALL) principles, further grounded in Universal Design for Learning (UDL). Specifically, generative AI is examined through UDL's multiple means of representation (e.g., adaptive explanations of morphology and harakāt), multiple means of action and expression (e.g., AI-supported writing and pronunciation practice), and multiple means of engagement (e.g., personalized feedback and autonomous learning pathways). This integration provides a principled justification for the study's instruments and analytic approach.

Methodologically, the study introduces a unified analytical lens that synthesizes Sahrir's (2025) Learning Effectiveness framework with Student Engagement constructs tailored to Arabic L2 learning. Pedagogically, it advances the field by evaluating culturally responsive generative AI applications that address the linguistic and orthographic complexity of Arabic rather than treating it as a generic

foreign language (Sahrir et al., 2025) and reading comprehension. Yet, despite its potential, empirical research on generative AI's role in Arabic higher education remains nascent, particularly regarding its holistic impact on engagement and academic rigor.

Current literature reveals several critical gaps in understanding generative AI's influence on advanced language competencies. Although systematic reviews identify broad applications of generative AI in language teaching and learning, much of the research remains at a general level without disaggregating effects on specific higher-order skills (e.g., idiomatic competence, writing processes, cultural nuance) or nuanced pedagogical outcomes (e.g., learner autonomy and self-regulation) in second language contexts (Center for Applied English Studies, 2024). Empirical studies on AI-assisted writing indicate that while AI feedback can support grammar, organization, and vocabulary development, detailed process-oriented research examining revision behaviors, rhetorical decision-making, and sustained writing proficiency is limited, especially outside dominant English contexts (Iin Almeina,dkk, 2025). Similarly, although large-scale analyses indicate that AI-enhanced language tools broadly improve learner engagement and motivation, these studies rarely distinguish whether such engagement translates into sustained development of cultural competence, idiomaticity, or complex discourse skills, leaving uncertainty regarding AI's role in deep L2 linguistic and cultural learning (Fatih Karataş et al., 2024). This limitation becomes particularly salient when viewed through Sahrir's (2025) engagement framework, which positions motivation, interactivity, and perseverance as essential conditions for effective learning. Despite this emphasis, few empirical studies have examined whether AI-driven Arabic language learning environments maintain these engagement dimensions longitudinally, especially in relation to cognitively demanding skills such as writing and morpho-syntactic processing. Moreover, existing research tends to prioritize technical functionality and tool performance over learner-centered outcomes, thereby overlooking how generative AI shapes students' emotional, behavioral, and self-regulatory engagement in Arabic L2 contexts (Pillai et al., 2024). Addressing these interconnected gaps, the present study investigates both the cognitive and affective dimensions of AI-integrated Arabic language learning.

This study advances three interrelated contributions. Methodologically, it operationalizes a dual-theory paradigm by systematically integrating Sahrir's (2025) Learning Effectiveness framework with the Student Engagement model in both instrument design and analysis. Survey items were explicitly mapped to Arabic-specific learning effectiveness indicators (e.g., vocabulary development, morphological accuracy, orthographic precision, and pronunciation) and to engagement dimensions (motivation, interactivity, and perseverance), enabling theory-driven measurement. Analytically, the study tests predicted relationships between these constructs, including the mediating role of student engagement in the relationship between generative AI use and learning effectiveness. Pedagogically, the research evaluates culturally responsive generative AI tools by examining how features such as adaptive feedback, idiomatic input, and pronunciation modeling align with Arabic linguistic and cultural norms (Sahrir et al., 2025). Technologically, it assesses generative AI's capacity to simulate immersive Arabic learning environments through real-time pronunciation correction and dynamic listening practice. Collectively, this empirically grounded dual-theory approach provides a

robust and replicable framework for evaluating the educational impact of generative AI in Arabic language learning.

Aligned with the above frameworks, this study pursues three objectives. 1) Assess Learning Effectiveness: Quantify generative AI's impact on Arabic grammar mastery, writing skills, and complex material comprehension. 2) Evaluate Student Engagement: Measure how AI tools influence motivation, interactive participation, and perseverance in Arabic courses. 3) Identify Synergies: Analyze correlations between engagement metrics and academic outcomes.

By bridging theoretical rigor with empirical analysis, this work provides educators and AI developers actionable insights to design culturally adaptive Arabic learning tools. It advances the discourse on AI's role in preserving linguistic heritage while accelerating proficiency—addressing a critical need in global higher education. Ultimately, it positions generative AI not as a replacement for human instruction, but as a catalyst for revolutionizing Arabic pedagogy through sustained engagement and demonstrable academic success.

METHOD

This study employed a quantitative descriptive survey design to examine university students' perceptions of the use of Generative AI in Arabic language learning. A descriptive survey approach is appropriate for systematically capturing respondents' views, experiences, and attitudes toward an educational phenomenon at a particular point in time (Fraenkel, Wallen, & Hyun, 2019). The study was conducted at a public Islamic university in Indonesia, within first-year undergraduate Arabic language courses offered as part of a formal language program. A total of 123 first-year university students participated in the study, as summarized in Table 1.

Participants were enrolled in compulsory Arabic language courses and represented novice to lower-intermediate proficiency levels, as typically expected of first-year learners in the program. No standardized Arabic proficiency pre-test was administered; however, course placement followed the institution's curriculum structure for beginner-level Arabic instruction. Inclusion criteria required that students (a) were officially registered in the Arabic course during the data collection period and (b) had prior experience using Generative AI tools (e.g., ChatGPT, Google Translate, Gemini) to support their Arabic learning. Students who reported no AI use or did not complete the questionnaire were excluded from the analysis. This contextual framing is essential for interpreting students' perceptions and for understanding the scope of generalizability of the findings.

Data were collected using a structured questionnaire consisting of closed-ended items measured on a six-point Likert scale. The instrument was designed to capture three key domains: demographic characteristics, frequency of Generative AI use, and students' perceptions of Generative AI in relation to learning effectiveness and student engagement. The construction of questionnaire items was theoretically informed by the Learning Effectiveness framework (Baker et al., 2010), which emphasizes knowledge acquisition and skill development, and the Student Engagement model (Fredricks, Blumenfeld, & Paris, 2004), encompassing emotional, behavioral, and cognitive dimensions of learner involvement. These frameworks were operationalized with reference to Arabic language learning

outcomes commonly targeted at the beginner level, including vocabulary development, basic grammar comprehension, pronunciation practice, and engagement with learning materials.

Survey data were analyzed using descriptive statistical techniques, including frequencies, percentages, and mean scores, to summarize response patterns and identify overall trends in students' perceptions. Descriptive statistics are widely used in educational research to provide clear representations of central tendencies and distributional characteristics of survey data (Field, 2018). This analytical approach enabled a focused examination of how first-year Arabic learners in an Indonesian higher education context perceive the role of Generative AI in supporting language learning outcomes and engagement-related processes.

Data analysis was conducted in two stages. First, descriptive statistics were calculated for all quantitative variables. Second, findings were interpreted through the lens of the theoretical frameworks, allowing for a nuanced discussion of the benefits, limitations, and usage patterns of Generative AI in Arabic language learning contexts.

This study of 123 university students in Arabic courses found a significant gender imbalance, with 84.55% female and 15.45% male students, aligning with trends in language education where female learners often show higher motivation. Regarding AI usage, most students (60.16%) used it weekly, indicating it serves as a supplementary tool for specific tasks rather than a daily necessity. This pattern supports Blended Learning Theory, where technology complements traditional methods. The findings also reflect the Technology Acceptance Model (TAM), as students likely adopt AI due to its perceived usefulness and ease of use, reinforcing classroom instruction and supporting independent learning.

Instruments

The main instrument implemented in this study was a structured questionnaire designed to capture students' demographic information, their patterns of AI use in Arabic learning, and their perceptions of learning effectiveness and student engagement. The first part of the questionnaire included demographic items (such as gender and frequency of AI use) that enabled researchers to map respondent characteristics. This is a common practice in educational research, as demographic data provide context and allow for meaningful interpretation of learning behaviors (Creswell & Creswell, 2018).

The second part of the instrument adopted a Likert-scale format (1 = strongly disagree to 6 = strongly agree) to measure students' perceptions of AI in Arabic learning. Two theoretical lenses guided the questionnaire design: Learning Effectiveness and Student Engagement. The Learning Effectiveness items measured students' views on how AI supported vocabulary, grammar, writing, pronunciation, and comprehension. Meanwhile, the Student Engagement items focused on the motivational and interactive aspects of AI, including persistence, focus, and access to materials. Likert scales are widely used in educational technology research because they capture the degree of agreement with specific statements and provide reliable quantitative data for analysis (Joshi, Kale, Chandel, & Pal, 2015).

Furthermore, the instrument incorporated items related to students' preferences for AI platforms (e.g., ChatGPT, Google Translate, Gemini). This allowed

the study to compare not only the frequency of AI use but also the tools students found most useful. Using platform-specific questions is supported in CALL (Computer-Assisted Language Learning) research, where technology adoption depends on perceived reliability and task alignment (Chapelle, 2001). By combining demographic data, Likert-scale items, and platform preferences, the questionnaire provided a comprehensive overview of students' perspectives on AI in Arabic learning.

Procedures

This study was conducted using a mixed-methods approach, combining quantitative survey data with descriptive qualitative interpretation. The procedure began with the administration of a structured questionnaire to 123 first-year university students enrolled in Arabic language courses. The questionnaire consisted of three main sections: demographic data (gender, AI usage, and frequency of use), platform preferences (such as ChatGPT, Google Translate, and Gemini), and Likert-scale items measuring perceptions of Learning Effectiveness and Student Engagement. Such a multi-layered design reflects Creswell and Creswell's (2018) recommendation that mixed-methods research should collect both contextual and evaluative data to capture the complexity of educational experiences.

After the data were collected, the responses were organized into descriptive statistics. Demographic data were presented in frequency tables, showing gender distribution, the proportion of students using AI, and the frequency of AI usage. This step provided an overview of the respondents' characteristics and contextualized their experiences. Subsequently, data regarding AI platform preferences were analyzed using bar charts, which allowed clearer visualization of students' choices across different generative AI tools. This visual analysis was essential to highlight the popularity of ChatGPT, Google Translate, and Gemini, as well as the underlying reasons for their selection. Such descriptive analysis aligns with Cohen, Manion, and Morrison's (2018) guidelines on using frequency and visual representation to summarize survey data in educational research.

The third stage involved analyzing students' perceptions of Learning Effectiveness and Student Engagement, measured on a six-point Likert scale. Mean scores for each item were calculated and ranked to identify the strongest and weakest dimensions of AI-assisted Arabic learning. For Learning Effectiveness, scores ranged from 4.52 to 5.02, with vocabulary acquisition rated the highest and writing skills the lowest. Similarly, for Student Engagement, scores ranged from 4.37 to 4.96, with accessibility of learning materials ranked the highest and maintaining focus the lowest. This approach follows Boone and Boone's (2012) recommendation that Likert-scale data should be summarized through mean and frequency analysis to identify trends in attitudes and perceptions.

Finally, the results were interpreted through the dual theoretical lenses of Learning Effectiveness and Student Engagement. Quantitative findings were contextualized with qualitative reasoning, explaining why certain tools were more frequently used and why specific learning dimensions were rated higher than others. This integration of numeric data with theoretical interpretation reflects the essence of the mixed-methods design, where results are not only reported but also explained in light of broader pedagogical theories. By following these procedures,

the study ensured both systematic data collection and theoretically informed interpretation, thereby enhancing the credibility and depth of its findings.

RESULTS AND DISCUSSION

The data collected from the 123 students were analyzed using a combination of descriptive statistics and theoretical interpretation. The demographic information (gender, AI usage, and frequency of use) was first summarized in percentage tables to provide an overview of the respondent profile. Frequency distributions were employed to determine patterns of participation, which is a standard approach for analyzing survey data (Cohen, Manion, & Morrison, 2018). The data indicated that female students dominated the sample, and a significant majority (85.37%) reported using AI in Arabic learning, with most students (60.16%) relying on AI on a weekly basis. These findings were interpreted through the Technology Acceptance Model (TAM) (Davis, 1989), which suggests that the widespread adoption of AI tools can be explained by students' perception of their usefulness and ease of use in supporting learning tasks.

The second stage of analysis involved examining students' preferences for specific AI platforms, represented through bar charts for clarity of comparison. This visual analysis highlighted the predominance of ChatGPT, followed by Google Translate and Gemini. The findings were further contextualized using theories from Computer-Assisted Language Learning (CALL), which emphasize that learners choose tools based on task alignment and the reliability of feedback provided (Chapelle, 2001). ChatGPT was preferred for its interactive and generative capacities, Google Translate for its accessibility and quick translation functions, and Gemini for its integration with broader Google-based learning resources. This comparative analysis allowed the study to move beyond raw numbers and explain why students gravitated toward particular AI tools.

The third stage focused on students' perceptions of Learning Effectiveness and Student Engagement, measured using a six-point Likert scale. Mean scores were calculated for each statement, with results ranging from 4.52 to 5.02 for Learning Effectiveness and 4.37 to 4.96 for Student Engagement. This type of Likert-scale analysis is widely recognized as an effective means of identifying attitudinal trends in educational research (Boone & Boone, 2012). The highest score for Learning Effectiveness (M = 5.02) was for vocabulary acquisition, highlighting the role of AI in supporting lexical development, while the lowest (M = 4.52) was for writing skills, reflecting the limitations of AI in guiding complex, productive tasks. For Student Engagement, the highest score (M = 4.96) was given to access to learning materials, suggesting that AI improves convenience and availability, while the lowest (M = 4.37) was for maintaining focus, indicating that students may still be prone to distraction despite AI integration. These results were interpreted through Blended Learning Theory (Graham, 2006), which argues that digital tools serve best as supplementary resources that enhance but do not replace classroom interaction.

Overall, the data analysis integrated both statistical summaries and theoretical frameworks to draw meaningful conclusions. Descriptive statistics provided clear insights into student behaviors and perceptions, while established theories such as TAM, CALL, and Blended Learning gave explanatory depth to the patterns observed. By combining these approaches, the study ensured methodological rigor and

strengthened the interpretation of how generative AI tools shape Arabic language learning experiences.

The Profile of the First-Year University Students

This study adopts a mixed-methods approach to investigate student perspectives on using Generative AI tools in Arabic courses. Data from 123 university students will be analyzed through the dual theoretical lenses of Learning Effectiveness and Student Engagement. There were 123 students who participated in this study. Further details about students are presented in Table 1.

Table 1. The Profile of the First-Year University Students

Gender	Students	Percentage
Male	19	15.45%
Female	104	84.55%
Total	123	100%
Using AI in Arabic Learning	**Students**	**Percentage**
Yes	105	85.37%
No	18	14.63%
Total	123	100%
Frequency of Using AI in Arabic Learning	**Students**	**Percentage**
Everyday	11	8.94%
Every three days	10	8.13%
Everyweek	74	60.16%
Every two weeks	9	7.32%
Everymonth	19	15.45%
Total	123	100%

The table 1 presents the gender distribution among 123 first-year undergraduates. Of this cohort, 19 students—or approximately 15.45 %—are male, whereas the remaining 104 students (84.55 %) are female. This indicates that female students constitute the overwhelming majority of the sample. Such demographic insights are critical for understanding the composition of the cohort and may suggest the need for differentiated educational approaches in terms of pedagogical methods, communication styles, or learning preferences.

According to the table titled "Using AI in Arabic Learning," a substantial majority of students—robustly 105 individuals (85.37 %)—incorporate artificial intelligence into their Arabic language study, while only 18 students (14.63 %) do not. This demonstrates a markedly high adoption rate of AI in this educational context.

Further examination of the table "Frequency of Using AI in Arabic Learning" reveals that the most prevalent usage pattern is weekly, reported by 74 students (60.16 %). The remaining distribution of usage frequency is as follows: 11 students (8.94 %) use AI daily, 10 students (8.13 %) every three days, 9 students (7.32 %) biweekly, and 19 students (15.45 %) on a monthly basis. Thus, although daily usage

rates are relatively modest, it is evident that the majority of students engage with AI on a consistent, routine schedule—especially on a weekly basis.

This regularity in AI usage suggests stable and sustained engagement, a pattern that aligns with research demonstrating a positive correlation between frequency of AI tool use and enhanced learning outcomes

The high adoption of artificial intelligence (AI)—utilized by 85 % of students—indicates that AI has become a critical tool in Arabic language instruction. Several factors may underlie this widespread uptake:

Personalized and Effective Learning

AI-powered adaptive learning systems dynamically customize instruction to each student's needs. For instance, Akbulut and Cardak's review of 70 web-based adaptive educational hypermedia studies reported positive impacts on students' academic performance, learning process, and satisfaction (Akbulut & Çardak, 2012). More broadly, a scoping review of 69 studies found that 59% showed increased academic performance, while the remainder exhibited no change, underscoring adaptive learning's generally beneficial—but not universal—effectiveness (Personalized adaptive learning in higher education, scoping review, 2020).. Similarly, a systematic review of 45 studies showed a medium to large positive effect size (g = 0.70) for adaptive systems compared to non-adaptive interventions (Colab, 2024). Other reviews confirm that personalized adaptive learning enhances academic performance, engagement, and outcomes, although certain technological and design challenges remain (Gao, 2023; Al-Tameemi & Hadi, 2024).

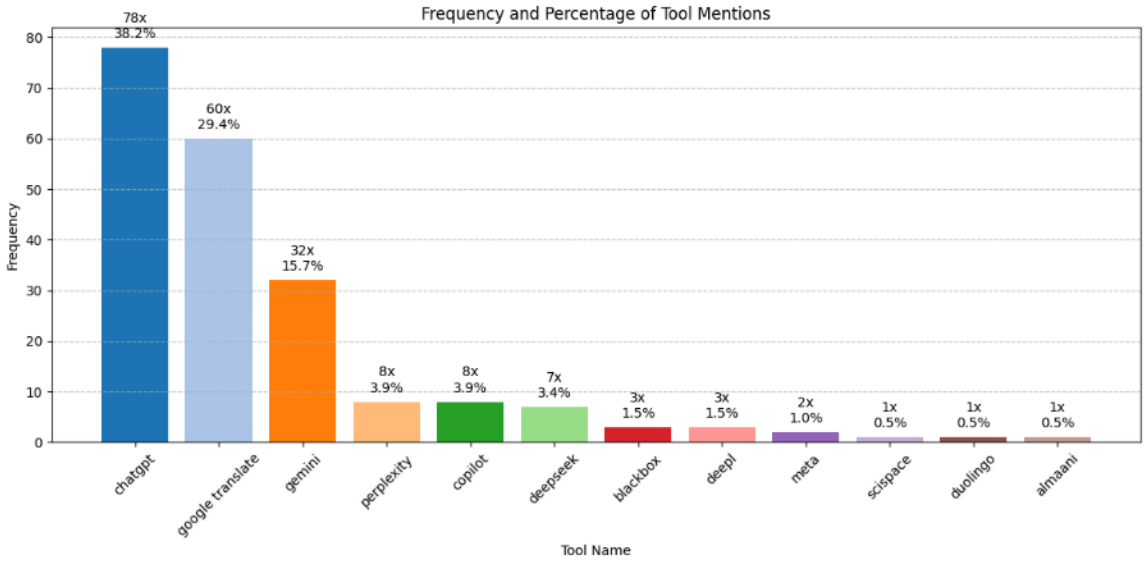
Efficiency and Learning Motivation

AI facilitates rapid, contextualized feedback while gamification elements such as points, levels, and challenges enhance engagement and intrinsic motivation. A recent framework emphasizes that AI-driven personalization combined with gamification fosters adaptability, motivation, and improved learning results (Chang & Chen, 2023). In science education, integrating ChatGPT into game-based learning significantly improved students' perceived competence, reduced cognitive load, and enhanced learning behavior compared to conventional methods (Zhou et al., 2024). Moreover, reviews of gamification in higher education show that when designed thoughtfully, these strategies increase motivation, performance, and enjoyment in the learning process (Deterding et al., 2011; Wikipedia, 2024).

Flexible Access and Support

Unlike traditional classroom learning, AI platforms such as ChatGPT and Duolingo offer flexible, anytime-anywhere access to learning resources. These tools provide adaptive assessments, virtual assistance, and gamified experiences that support self-paced study (Times of India, 2024). Duolingo in particular has introduced AI-powered features such as interactive chats and immersive "adventures," further enhancing personalized, on-demand language learning (Von Ahn, 2024).

Figure 1. Frequency and Percentage of Tool Mentions



The figure 1 illustrates the frequency and percentage of various AI and language-related tools mentioned by respondents. ChatGPT leads with the highest frequency at 78 mentions (38.2%), followed by Google Translate with 60 mentions (29.4%), and Gemini with 32 mentions (15.7%). Other tools such as Perplexity, Copilot, and Deepseek were mentioned far less frequently, each accounting for under 4%. Meanwhile, tools like Blackbox, DeepL, Meta, Scispace, Duolingo, and Almaani were rarely mentioned, each representing less than 2% of total responses. This distribution clearly highlights the dominance of a few key tools over others.

ChatGPT emerges as the most frequently mentioned tool, reflecting its popularity among users due to its versatility and advanced conversational abilities. Respondents likely prefer ChatGPT because it can generate coherent text, assist with academic writing, provide explanations, and support multiple disciplines. Its ease of access, adaptability across tasks, and interactive style make it an effective tool not only for translation or paraphrasing but also for critical thinking, summarization, and problem-solving (Apriani et al., 2025). The higher percentage suggests that users perceive ChatGPT as more reliable and capable compared to other AI tools.

Google Translate ranks second, with 29.4% of mentions, highlighting its continued relevance as a fast and practical translation tool. Its widespread accessibility on both web and mobile devices makes it highly convenient for users. Additionally, Google Translate supports more than 100 languages and is integrated into other applications, which explains its frequent use (Bin Dahmash, 2020). Despite limitations in accuracy for complex or nuanced texts, its immediate results and ease of use make it a go-to tool, particularly for users seeking quick translations rather than deep explanations.

Gemini, with 15.7% of mentions, stands as the third most frequently used tool. Its popularity may stem from its integration with Google's ecosystem and its potential for reliable responses supported by search-based features. Users might favor Gemini because it combines generative AI capabilities with access to real-time information, making it especially useful for academic queries and fact-checking (Thurzo & Varga, 2025). While not as dominant as ChatGPT or as widely accessible

as Google Translate, Gemini appeals to those who value AI tools that balance creativity with updated knowledge.

Learning Effectiveness

Table 2. Learning effectiveness after using generative AI in Arabic learning

No	Responses	Mean Score (1-6)
	Generative AI enhances learning new Arabic vocabulary.	5.02
	Generative AI significantly improves understanding of Arabic grammar.	4.65
	Generative AI contributes positively to reading comprehension in Arabic.	4.70
	Generative AI supports the development of Arabic writing skills.	4.52
	Generative AI provides clear and constructive feedback.	4.57
	Generative AI facilitates regular listening exercises in Arabic.	4.70
	Generative AI promotes faster learning compared to traditional methods.	4.54
	Generative AI aids in mastering Arabic idioms and expressions.	4.57
	Generative AI enhances pronunciation practice for Arabic.	4.57
0	Generative AI accelerates the understanding of complex Arabic materials.	4.61

The table shows that the highest mean score is **5.02**, indicating that respondents strongly agreed that generative AI enhances learning new Arabic vocabulary. On the other hand, the lowest mean score is **4.52**, which corresponds to the statement that generative AI supports the development of Arabic writing skills. Although both values are relatively high on the scale, the difference highlights that students perceive vocabulary acquisition as the greatest benefit of AI, while writing skills are seen as less directly supported.

The highest score, 5.02, suggests that learners find generative AI especially effective in helping them acquire new Arabic vocabulary. This result reflects how AI tools such as ChatGPT, Gemini, or Duolingo provide instant translations, contextual examples, and interactive practice, which are crucial for vocabulary expansion (Neupane et al., 2025). Since vocabulary is a fundamental building block of language learning, respondents appear to value AI's ability to introduce new words and reinforce them through adaptive exercises and repeated exposure.

The lowest score, 4.52, indicates that while generative AI is still beneficial for developing Arabic writing skills, students view its impact as relatively weaker compared to other areas. Writing requires not only grammar and vocabulary knowledge but also coherence, organization, and creativity—skills that may not be fully cultivated through AI tools alone (Deep & Chen, 2025). Respondents may feel that although AI provides corrections and suggestions, authentic writing practice and feedback from human instructors are still essential for mastering this complex skill.

Student Engagement

Table 3. Student engagement after using generative AI in Arabic learning

No	Responses	Mean Score (1-6)
	Generative AI usage in Arabic learning occurs regularly each week.	4.78
	Generative AI makes Arabic learning more interesting.	4.59
	Generative AI increases motivation to learn Arabic.	4.59
	Generative AI offers interactive learning experiences.	4.61
	Generative AI helps maintain focus during learning sessions.	4.37
	Generative AI encourages perseverance when facing learning challenges.	4.83
	Generative AI increases engagement in Arabic learning.	4.74
	Generative AI fosters active participation in Arabic lessons.	4.57
	Generative AI simplifies access to learning materials.	4.96
0	Generative AI provides useful progress-tracking features.	4.76

The table demonstrates that the highest mean score is 4.96, showing that respondents strongly agreed with the statement "Generative AI simplifies access to learning materials." In contrast, the lowest mean score is 4.37, corresponding to the statement "Generative AI helps maintain focus during learning sessions." Although both scores fall within the higher range, the difference highlights that learners value AI's role in providing easy access to resources more than its capacity to sustain concentration.

The largest value, 4.96, indicates that learners see simplified access to learning materials as the most significant benefit of generative AI. This result reflects AI's ability to provide quick, on-demand information and resources, which reduces

barriers to learning and supports independent study. By offering instant explanations, translations, and references, generative AI makes the learning process more flexible and efficient (Almelhes, 2024). For students of Arabic, this means they can access diverse texts, practice exercises, and supplementary materials without relying solely on traditional classroom settings.

The smallest value, 4.37, suggests that while generative AI contributes to maintaining focus during learning, students perceive this effect as relatively weaker compared to other areas. Focus and sustained attention are influenced by individual discipline, motivation, and learning environment, which AI tools may not directly control. Although interactive features and personalized feedback can help engagement, students might still experience distractions or reduced concentration when studying independently with AI (Bhatia et al., 2024). Therefore, maintaining focus remains an area where human guidance and structured learning environments may still play an essential role.

CONCLUSION

The study concludes that students value chatbots for their ability to make language learning tasks more efficient and manageable, particularly in completing routine exercises. Students also find the tool user-friendly and generally have a positive interaction experience, which supports its use as a supplementary resource in Arabic language education. However, the analysis also highlights significant shortcomings, particularly in areas like the chatbot's accuracy, design quality, and effectiveness in improving speaking skills. These limitations indicate that while chatbots can assist with certain aspects of language learning, they are not yet fully equipped to handle the complexities of Arabic language education. Given these findings, it is clear that chatbots should be integrated into a blended learning approach, where they complement rather than replace traditional teaching methods and human instructors.

ACKNOWLEDGMENT

The researchers expressed their gratitude to Pusat Penelitian dan Penerbitan (Puslitpen) The Institute for Research and Community Empowerment (LP2M) of Syarif Hidayatullah State Islamic University Jakarta, Indonesia, for sponsoring the research.

REFERENCES

Almelhes, S. (2024). Enhancing Arabic Language Acquisition: Effective Strategies for Addressing Non-Native Learners' Challenges. *Education Sciences*, *14*(10), 1116. https://doi.org/10.3390/educsci14101116

Al-Tameemi, H. M. & Hadi, H. F. 2024. A scoping review of adaptive learning systems: Effectiveness, challenges, and future directions. *Heliyon*, 10(2): e25538. https://doi.org/10.1016/j.heliyon.2024.e25538

Apriani, E., Hamidah Daulay, S., Aprilia, F., Marzuki, A. G., Warsah, I., Supardan, D., & Muthmainnah. (2025). A Mixed-Method Study on the Effectiveness of Using ChatGPT in Academic Writing and Students' Perceived Experiences. *Journal of*

Language and Education, *11*(1), 26–45. https://doi.org/10.17323/jle.2025.17913

Bacon, S. M., & Finnemann, M. D. (1992). Sex differences in self-reported beliefs about foreign-language learning and authentic oral and written input. Language Learning, 42(4), 471–495. https://doi.org/10.1111/j.1467-1770.1992.tb01041.x

Baker, D., Dube, C., & Gilbreath, B. (2010). Learning effectiveness: Definition and measurement. New Directions for Institutional Research, 2010(128), 5–13. https://doi.org/10.1002/ir.319

Bhatia, A., Bhatia, P., & Sood, D. (2024). *Leveraging AI to Transform Online Higher Education: Focusing on Personalized Learning, Assessment, and Student Engagement*. SSRN. https://doi.org/10.2139/ssrn.4959186

Bin Dahmash, N. (2020). I Can't Live Without Google Translate: A Close Look at the Use of Google Translate App by Second Language Learners in Saudi Arabia. *Arab World English Journal*, *11*(3), 226–240. https://doi.org/10.24093/awej/vol11no3.14

Boone, H. N., & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension, 50*(2), 1–5.

Chang, T. Y., & Chen, C. M. (2023). Designing adaptive gamified learning systems with AI to enhance motivation and performance. *Education Sciences, 15*(3), 301. https://doi.org/10.3390/educsci15030301

Chapelle, C. A. (2001). *Computer Applications in Second Language Acquisition: Foundations for*

Chou, C. M. (2023). ChatGPT in language learning: Opportunities and challenges. Journal of Language Teaching and Research, 14(5), 1012–1021.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th ed.). London: Routledge.

Colab. (2024). The Efficacy of Artificial Intelligence-Enabled Adaptive Learning Systems From 2010 to 2022 on Learner Outcomes: A Meta-Analysis, *Journal of Educational Technology*. https://doi.org/10.1177/07356331241240459

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed*

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Deep, P. D., & Chen, Y. (2025). The Role of AI in Academic Writing: Impacts on Writing Skills, Critical Thinking, and Integrity in Higher Education. *Societies*, *15*(9), 247. https://doi.org/10.3390/soc15090247

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 (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Dinata, Rahmat Satria dkk. (2023). Web-Based Google Translate Inconsistencies in Bahasa-Arabic Translation from the Arabic Thesis Writer Perspectives. *Asalibuna, 7*(2), 71. https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/asalibuna/article/view/2634/809

Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5th ed.). SAGE Publications.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). How to design and evaluate research in education (10th ed.). McGraw-Hill Education.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gao, Y. (2023). *The effectiveness of adaptive learning systems: A meta-analysis* (Doctoral dissertation, University of Alberta). University of Alberta Library. https://era.library.ualberta.ca/items/e8de2f25-be42-4c7b-967a-510fa4bf98a8

Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *The Handbook of Blended Learning: Global Perspectives, Local Designs* (pp. 3–21). San Francisco: Pfeiffer.

Hughes, A. *Teaching, Testing, and Research.* Cambridge University Press.

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology, 7*(4), 396–403. https://doi.org/10.9734/BJAST/2015/14975

Khan, T., Singh, M., & Raza, S. (Eds.). (2024). *Artificial Intelligence: A Multidisciplinary Approach towards Teaching and Learning*. Bentham Science Publishers. https://doi.org/10.2174/9789815305180124010l

Methods Approaches (5th ed.). Thousand Oaks, CA: Sage.

Neupane, B. P., Paudel, P., Dahal, N., Karki, S., Paudel, G. R., Ghimire, P., & Thapa, B. (2025). English Language Teaching in the Age of Artificial Intelligence: Tools, Techniques, and Methodologies. *Advances in Mobile Learning Educational Research*, *5*(1), 1356–1369. https://doi.org/10.25082/AMLER.2025.01.011

Pavlenko, A., & Piller, I. (2001). New directions in the study of gender and multilingualism. In A. Pavlenko, A. Blackledge, I. Piller, & M. Teutsch-Dwyer (Eds.), Multilingualism, Second Language Learning, and Gender (pp. 17–52). Berlin: Mouton de Gruyter.

Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. 2020. *Smart Learning Environments*, 7(1): 1–15. https://doi.org/10.1186/s40561-020-00132-6

Pillai, D., Doifode, A., Singh, A. S., & Bhosale, T. (2024). *Fostering learner-centric education through smart technology*. 080007. https://doi.org/10.1063/5.0245203

Sahrir, M. S., Albantani, A. M., Arifin, F., & Fathudin, F. (2025). The Use of Generative Artificial Intelligence (AI) in Arabic Language Education: Insights and Implications between Malaysia and Indonesia. *International Journal of Research and Innovation in Social Science*, *IX*(III), 3638–3646. https://doi.org/10.47772/IJRISS.2025.90300288

Thurzo, A., & Varga, I. (2025). Revisiting the Role of Review Articles in the Age of AI-Agents: Integrating AI-Reasoning and AI-Synthesis Reshaping the Future of Scientific Publishing. *Bratislava Medical Journal*, *126*(4), 381–393. https://doi.org/10.1007/s44411-025-00106-8

Times of India. (2024, May 18). Beyond screen time: 10 smart ways AI can revolutionise your studying. *The Times of India*. https://timesofindia.indiatimes.com/education/news/beyond-screen-time-10-smart-ways-ai-can-revolutionise-your-studying/articleshow/121756911.cms

Von Ahn, L. (2024, March 4). Duolingo's new AI-powered features. *The Verge*. https://www.theverge.com/24267841/luis-von-ahn-duolingo-owl-language-learning-gamification-generative-ai-android-decoder

Y., Akbulut, & Çardak, M. (2012). A content analysis of web-based adaptive educational hypermedia (AEH) systems: Effects on student performance, learning process, and satisfaction. *International Journal of E-Learning & Educational Technologies in the Digital Media*,

Yu, S., & Yu, L. (2022). Artificial intelligence in language learning: A systematic review. Computer Assisted Language Learning, 35(7), 1502–1524. https://doi.org/10.1080/09588221.2021.1975522

Zapata, G. C. (2025). *Generative AI Technologies, Multiliteracies, and Language Education* (1st ed.). Routledge. https://doi.org/10.4324/9781003531685

Zhou, J., Li, X., & Wang, Y. (2024). Exploring ChatGPT-assisted game-based learning in science education: Effects on motivation, cognitive load, and behavior. *Education and Information Technologies, 29*(5), 5431–5450. https://doi.org/10.1007/s10639-024-12553-x

Zou, B., & Xie, H. (2019). Personalized word-learning for academic purposes in mobile-assisted learning. Interactive Learning Environments, 27(4), 1–14. https://doi.org/10.1080/10494820.2019.1593196