

The Impact of ChatGPT Integration on Arabic Language Learning Among Generation Z Students

Nurfaizah¹, Muhamad Hasyim², Yanti Puspita³, Sumiati⁴, Yuli Azizah⁵, Ummatin Najla⁶, Syifaunopus⁷, Samiah⁸, Yanti Uspita⁹

^{1, 2, 3, 4, 5, 6, 7, 8, 9} Institut Agama Islam Hamzanwadi NW Lombok Timur, Indonesia

Email: nurfaizahbintualimun@gmail.com¹, hasiemmuhammad88@gmail.com², suntitula@gmail.com³, sumiatibomban@gmail.com⁴, yuliazizah1404@gmail.com⁵, ummatinnajla87@gmail.com⁶, nfssyifaun@gmail.com⁷, samiahmiha308@gmail.com⁸, yantiyupita@gmail.com⁹

ABSTRACT

This study addresses the limited evidence on the impact of ChatGPT integration on Arabic language learning among Generation Z students, particularly across linguistic, cognitive, motivational, and ethical dimensions. The study aims to synthesize recent evidence on ChatGPT-supported Arabic language learning and identify emerging patterns, challenges, and research gaps. A systematic literature review was employed using a PRISMA-informed approach to identify, screen, evaluate, and synthesize relevant studies published primarily between 2021 and 2026. The literature was analyzed thematically across linguistic development, motivation and engagement, learner autonomy, self-regulated learning, higher-order thinking, AI literacy, personalization, ethical use, and potential dependency. The findings indicate that ChatGPT can support Arabic reading, writing, speaking, grammar, vocabulary, translation, language practice, personalized feedback, motivation, and self-directed learning. Cognitive benefits were also identified when ChatGPT was integrated through structured guidance and active learner engagement. However, concerns regarding academic integrity, dependency, information accuracy, privacy, and inadequate human validation remain. The review concludes that ChatGPT is most effective as a guided human-AI learning support system rather than a replacement for teachers or learners. The study contributes theoretically by integrating technological, cognitive, and learner-centered perspectives and practically by highlighting the need for AI literacy, responsible prompting, teacher guidance, and human verification. Future research should examine Generation Z learners through longitudinal, experimental, and cross-skill studies in Arabic language education.

ARTICLE INFO

Article history:

Received: July 25, 2026

Revised: July 28, 2026

Accepted: July 30, 2026

Published: July 31, 2026

Keywords:

ChatGPT; Arabic Language Learning; Generation Z; Generative AI; AI Literacy

Correspondance Author:

Nurfaizah

Program Studi Pendidikan Bahasa Arab, Institut Agama Islam Hamzanwadi NW Lombok Timur, Jln Utama Mataram - Lb Lombok KM.49 Anjani - Suralaga Lombok Timur NTB, Indonesia.

INTRODUCTION

The rapid development of generative artificial intelligence (GenAI), particularly ChatGPT, has created significant changes in the landscape of higher education and language learning. ChatGPT is increasingly viewed not merely as a technological innovation but as a potential

learning partner capable of providing immediate feedback, generating learning materials, facilitating interaction, and supporting learners' independent study. In foreign language education, ChatGPT can provide flexible opportunities for learners to practice language outside the classroom and adjust learning activities according to their individual needs and learning pace (Dizon, 2024). Its potential is particularly significant for university students who routinely interact with digital technologies and expect learning experiences that are accessible, interactive, personalized, and responsive. Nevertheless, the growing integration of ChatGPT also raises concerns regarding information accuracy, privacy, overdependence, academic integrity, and the changing role of teachers and learners (Klimova et al., 2024; Mah et al., 2024).

At the international level, research demonstrates that students generally perceive ChatGPT positively because of its accessibility, flexibility, perceived usefulness, and capacity to provide personalized assistance. Students' intention to use ChatGPT in language learning is associated with perceived usefulness, ease of use, and enjoyment of interaction with the technology (Wang et al., 2025). ChatGPT can consequently support more self-directed learning by enabling students to seek explanations, practice language, receive feedback, and conduct learning activities beyond formal classroom settings (Dizon, 2024). However, the international literature also indicates that evidence concerning the effectiveness of ChatGPT across different language skills remains uneven. A systematic review found that empirical research has concentrated heavily on certain language skills and that evidence concerning reading, speaking, listening, and broader learning outcomes remains comparatively limited (Lo et al., 2024). Similarly, a recent systematic review of GenAI in language classrooms found that higher education and English language learning remain dominant contexts, indicating that research on other languages and educational settings is still underdeveloped (Lee et al., 2026). This international imbalance is particularly important for Arabic language education, which has linguistic, cultural, grammatical, and sociocultural characteristics that cannot necessarily be generalized from findings in English language learning.

The Indonesian context presents a related but more specific issue. Artificial intelligence has increasingly entered university students' academic activities, including Arabic language learning. Anwar and Mufidah (2024) found that students in Arabic language research courses were already using AI extensively, with ChatGPT becoming one of the most frequently used platforms. ChatGPT has also demonstrated potential in supporting Arabic-Indonesian translation and helping learners understand Arabic texts (Faris & Abdurrahman, 2023). Nevertheless, the quality of AI-generated Arabic translation cannot be assumed to be consistently accurate. Salma et al. (2025) demonstrated that ChatGPT can assist in translating classical Arabic texts, but human evaluation remains necessary to ensure accuracy, acceptability, and readability, particularly when religious and conceptually complex terminology is involved. These findings indicate that the increasing use of ChatGPT in Indonesian Arabic language education is accompanied by an important pedagogical challenge: students need to develop the capacity not only to access AI-generated information but also to evaluate, verify, and critically interpret it.

In responding to these developments, scholars have argued that AI should be integrated into language education as a pedagogical instrument rather than regarded as a substitute for teachers or learners' intellectual engagement. AI can transform language teaching, learning, and assessment by providing adaptive materials, interactive support, and linguistic assistance, but meaningful educational outcomes still depend on pedagogical design and human involvement (Pan, 2024). A multiliteracies perspective further suggests that GenAI can expand learners' capacity to interpret,

produce, and evaluate diverse forms of language and media while maintaining active participation and critical thinking (Zapata, 2025). Similarly, ChatGPT can facilitate personalized learning through interactive dialogue, feedback, learning support, and assistance tailored to learners' needs (Lahby, 2025). Such personalization is potentially valuable in Arabic language learning because students differ considerably in their levels of vocabulary, grammar, reading, writing, and communicative competence.

The effectiveness of ChatGPT, however, depends strongly on how learners interact with the technology. Generative AI becomes more pedagogically meaningful when learners are capable of constructing appropriate prompts that correspond to their learning objectives. Effective prompting can support reading comprehension, writing, problem solving, creativity, and personalized learning, while responsible use requires attention to safety, accuracy, and ethical considerations (Narciso, 2024). Therefore, technological familiarity alone is insufficient. Students also require critical-digital literacy that enables them to evaluate AI-generated outputs, recognize its limitations, and maintain academic integrity (Beckingham et al., 2024). This argument is particularly relevant to Generation Z students, whose familiarity with digital technology may facilitate technology adoption but does not necessarily guarantee critical, effective, or ethical AI use.

Previous empirical studies provide important evidence regarding the potential of ChatGPT in language education. In Arabic language learning, Leleparry et al. (2023) reported positive student responses to ChatGPT, particularly in relation to reading skills, motivation, enthusiasm, and assistance with academic tasks. Nasaruddin (2024) demonstrated that ChatGPT can assist teachers of Arabic as a foreign language in producing flexible and engaging learning materials through carefully constructed prompts. In foreign language education more broadly, Karataş et al. (2024) found positive effects of ChatGPT on writing, grammar, vocabulary acquisition, motivation, and engagement, although limitations were identified in speaking and concerns were raised regarding technological dependence. These findings are complemented by Zubaidi et al. (2025), who found improvements in Arabic writing through a human-AI collaboration model, while emphasizing that humans must remain conceptors, validators, and executors of the learning process. Mulloh et al. (2025) further conceptualized ChatGPT as a cognitive partner that can support scaffolding, reflective dialogue, the Zone of Proximal Development, and learner autonomy. Likewise, Kholid and Majeed (2025) reported improvements in Arabic speaking fluency, grammatical accuracy, and communicative confidence when ChatGPT was used as a conversational learning tool. More broadly, Chen et al. (2026) identified opportunities provided by GenAI for language practice, personalized feedback, interaction, authentic language input, and real-time linguistic assistance.

Despite these promising findings, the evidence also indicates that the impact of ChatGPT is not uniformly positive and should not be interpreted as an automatic consequence of AI integration. Song and Song (2023) found improvements in academic writing and motivation among EFL learners using ChatGPT, while also identifying concerns about contextual accuracy and dependency. Al-Abdullatif and Alsubaie (2024) further demonstrated that students' acceptance of ChatGPT is associated not only with perceived usefulness and ease of use but also with AI literacy. The cognitive dimension is equally important. Guided use of ChatGPT can improve self-regulated learning, higher-order thinking, and knowledge construction when students are encouraged to attempt problem solving before receiving AI assistance (Lee et al., 2024). ChatGPT has also been found to influence students' critical, creative, and reflective thinking, indicating that its educational value depends partly on the quality of learning activities surrounding its use (Essel et al., 2024). In addition, students' perceptions of usefulness, ease of use, risks, and psychosocial effects shape their

attitudes toward ChatGPT (Sallam et al., 2024). These findings suggest that ChatGPT integration should be understood as a complex pedagogical process involving technological, cognitive, motivational, and ethical dimensions.

An important gap emerges when these findings are considered from the perspective of Generation Z and Arabic language learning. Generation Z students are frequently characterized by their high exposure to digital environments; nevertheless, technological familiarity does not necessarily lead to effective or ethical use of AI. Acosta-Enriquez et al. (2024) found that knowledge, attitudes, ethical perceptions, and concerns influence how Generation Z university students use ChatGPT. At the same time, the international literature remains dominated by English language learning and higher education contexts (Lee et al., 2026; Lo et al., 2024), while empirical research specifically examining the intersection of ChatGPT, Arabic language learning, and Generation Z remains comparatively limited. Existing Arabic studies have tended to focus on particular skills such as reading, writing, translation, or speaking rather than examining the broader impact of ChatGPT integration across cognitive, affective, linguistic, and behavioral dimensions. Therefore, there remains a need for a comprehensive synthesis that connects the growing evidence on ChatGPT with the distinctive characteristics of Generation Z learners in Arabic language education.

Theoretically, this research is grounded in the understanding that effective AI-supported learning requires active learner participation, scaffolding, self-regulation, critical evaluation, and meaningful human-AI collaboration. The potential of ChatGPT as a cognitive and dialogic partner is consistent with the use of scaffolding and the Zone of Proximal Development identified in Arabic language learning research (Mulloh et al., 2025). At the same time, self-regulated learning and higher-order thinking provide an important theoretical basis for understanding why guided rather than purely answer-generating interaction can produce stronger educational outcomes (Lee et al., 2024). The multiliteracies perspective also strengthens the argument that learners should not merely consume AI-generated texts but should actively interpret, produce, evaluate, and transform language across multiple digital modes (Zapata, 2025). Consequently, the novelty of the present study lies in bringing these perspectives together to examine ChatGPT integration not simply as a technological tool, but as a multidimensional educational phenomenon situated within Arabic language learning and Generation Z digital learning practices.

Based on these considerations, this literature review focuses on the impact of ChatGPT integration on Arabic language learning among Generation Z students. It synthesizes recent international and national evidence concerning linguistic outcomes, motivation, learner autonomy, cognitive development, AI literacy, personalization, ethical use, and potential dependency. By integrating empirical findings and theoretical perspectives, this study seeks to clarify the educational opportunities and limitations of ChatGPT in Arabic language learning and to identify research directions that can support more effective, critical, ethical, and pedagogically grounded AI integration for Generation Z learners.

METHODS

1. Research Design

This study employed a systematic literature review (SLR) to examine the impact of ChatGPT integration on Arabic language learning among Generation Z students. The review was designed to systematically identify, screen, evaluate, and synthesize empirical evidence concerning the linguistic, cognitive, affective, behavioral, and ethical dimensions of ChatGPT use in Arabic language

learning. The methodological procedure was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020), which emphasizes transparent reporting of the identification, screening, eligibility, and inclusion processes in systematic reviews (Page et al., 2021). Recent systematic reviews of ChatGPT in education have similarly demonstrated that a structured protocol, predefined eligibility criteria, systematic searching, data extraction, quality evaluation, and evidence synthesis provide a rigorous basis for reviewing emerging AI research (Albadarin et al., 2024).

The review was selected because the research question addressed a rapidly developing educational phenomenon for which empirical findings remained diverse across contexts, language skills, learning outcomes, and student characteristics. Rather than collecting primary data from students, the study synthesized findings from previously published research to identify consistent patterns, contradictions, limitations, and research gaps concerning ChatGPT-supported Arabic language learning.

2. Literature Sources and Search Strategy

The literature search was structured around major multidisciplinary and education-oriented academic databases, namely Scopus, Web of Science, ERIC, ScienceDirect, SpringerLink, Taylor & Francis Online, and Google Scholar. These sources were selected to provide broad coverage of international educational technology, applied linguistics, language education, and Arabic language learning research. Similar systematic reviews of ChatGPT and language education have used combinations of Scopus, Web of Science, ERIC, and other scholarly databases to maximize the identification of relevant empirical studies (Albadarin et al., 2024). In language-learning reviews, systematic database searching has also been employed to identify empirical evidence across diverse digital learning environments (Inayati et al., 2025).

The search strategy was developed using combinations of keywords and Boolean operators. The principal search string was formulated as: (“ChatGPT” OR “generative AI” OR “generative artificial intelligence” OR “large language model*”) AND (“Arabic language learning” OR “Arabic language education” OR “Arabic learning” OR “Arabic as a foreign language”) AND (“Generation Z” OR “Gen Z” OR “university students” OR “higher education students”). Additional searches were conducted using related terms such as “Arabic writing,” “Arabic speaking,” “Arabic reading,” “Arabic translation,” “AI-assisted language learning,” “learner autonomy,” and “AI literacy.” The search covered studies published from 2021 to 2026 in order to capture recent developments in generative AI while retaining the most relevant research on digitally mediated language learning. The search strategy was designed to balance sensitivity and specificity so that studies directly focusing on Arabic could be identified while relevant evidence from foreign-language education could also be incorporated for comparative interpretation.

3. Eligibility Criteria and Study Selection

The unit of analysis was the published scholarly study rather than an individual student. Studies were considered eligible when they met the following substantive criteria: they addressed ChatGPT, generative AI, or closely related large language model applications; they examined language learning, language education, or a directly relevant educational outcome; and they provided findings applicable to university-level learners, Generation Z learners, or digital language-learning contexts relevant to Generation Z. Studies specifically addressing Arabic language learning were prioritized, while high-quality studies involving foreign-language learning were retained when they provided theoretically or empirically relevant evidence for interpreting ChatGPT integration in Arabic language education.

Peer-reviewed journal articles published between 2021 and 2026 were prioritized. Studies were excluded when they were editorials, opinion pieces, news articles, conference announcements, duplicate records, non-scholarly web materials, or studies that discussed artificial intelligence without a meaningful connection to language learning or educational use. Studies that focused exclusively on the technical architecture of ChatGPT without examining learning, teaching, learner experience, or educational outcomes were also excluded. This approach was consistent with recent systematic reviews that applied explicit inclusion and exclusion criteria to distinguish empirically relevant educational evidence from studies that merely examined technological capabilities (Albadarin et al., 2024; Lee et al., 2025).

The identification and screening process followed four stages: identification, duplicate removal, title and abstract screening, and full-text eligibility assessment. Records retrieved from the selected databases were compiled into a reference-management database, and duplicate publications were removed before screening. Titles and abstracts were then assessed against the eligibility criteria. Potentially relevant studies were subsequently subjected to full-text assessment, and only studies meeting all predefined criteria were included in the final synthesis. The complete screening process was documented in a PRISMA flow diagram so that the number of identified, screened, excluded, and included studies could be transparently reported (Page et al., 2021).

4. Data Extraction and Research Instrument

A structured data-extraction matrix was used as the primary instrument for organizing evidence from the included studies. The matrix was developed before the final synthesis and included bibliographic information, publication year, country or educational context, research objectives, participants, research design, data-collection methods, ChatGPT application, Arabic language skill or learning domain, theoretical framework, major findings, reported benefits, reported limitations, and implications for Generation Z learners.

Additional extraction categories were used to capture the multidimensional impact of ChatGPT. These categories included linguistic outcomes, such as reading, writing, speaking, grammar, vocabulary, and translation; cognitive outcomes, including critical thinking, higher-order thinking, and problem solving; affective outcomes, including motivation, confidence, engagement, and enjoyment; behavioral outcomes, such as learner autonomy and self-regulated learning; and ethical dimensions, including academic integrity, dependence, information accuracy, privacy, and AI literacy. The use of a structured extraction framework was consistent with previous systematic reviews of ChatGPT in education, in which predefined data-extraction procedures were employed to compare findings across heterogeneous empirical studies (Albadarin et al., 2024). This approach also supported the systematic comparison of language-learning studies that differed in research design, learning context, and outcome measures (Inayati et al., 2025).

5. Quality Appraisal of Included Studies

The methodological quality of the included studies was appraised before final synthesis. The appraisal considered the clarity of research objectives, appropriateness of research design, adequacy of participant or data selection, transparency of data-collection procedures, suitability of analytical methods, consistency between evidence and conclusions, and acknowledgement of methodological limitations. Because the review incorporated qualitative, quantitative, mixed-methods, descriptive, and case-based studies, methodological quality was assessed according to the characteristics of each research design rather than applying a single outcome criterion to all studies.

Particular attention was given to the credibility of evidence concerning ChatGPT's impact on language learning. Studies with transparent procedures, clearly defined samples, appropriate analytical techniques, and explicit limitations were considered stronger evidence than studies relying solely on perceptions or conceptual claims. This quality-oriented approach was important because recent reviews have identified substantial heterogeneity in empirical ChatGPT research and have emphasized the need to distinguish robust evidence from early exploratory findings (Albadarin et al., 2024). Recent GenAI language-learning reviews have similarly highlighted variations in research design, learner populations, outcomes, and methodological approaches across the emerging evidence base (Lee et al., 2025).

6. Data Analysis and Synthesis

The extracted data were analyzed using a qualitative thematic synthesis approach. First, the included studies were classified according to publication year, geographical context, research design, educational level, language skill, and type of ChatGPT integration. Second, findings were coded according to major impact domains, namely linguistic development, learner motivation and engagement, personalization and learner autonomy, cognitive development, self-regulated learning, AI literacy, and ethical concerns. Third, relationships and recurring patterns among the themes were compared across studies.

The synthesis focused on identifying convergent findings, divergent findings, methodological limitations, and under-researched areas. Findings concerning Arabic language learning were treated as the core evidence, whereas findings from broader foreign-language research were used to contextualize and interpret the Arabic-language evidence. Particular attention was paid to whether ChatGPT functioned as a tutor, conversational partner, feedback provider, translator, writing assistant, cognitive scaffold, or personalized learning tool. This categorization was important because the educational impact of ChatGPT may vary according to the pedagogical role assigned to the technology. Recent systematic evidence has shown that GenAI in language classrooms has been used for multiple functions and that research remains unevenly distributed across language skills and learning contexts (Lee et al., 2025). Similarly, systematic evidence on online language learning demonstrates the value of thematic synthesis for identifying patterns across heterogeneous studies and revealing directions for future research (Inayati et al., 2025).

Finally, the synthesized evidence was interpreted through the conceptual relationship between ChatGPT integration and Generation Z learning characteristics. The analysis therefore considered how digital familiarity, learner autonomy, motivation, AI literacy, critical evaluation, and ethical awareness may shape the educational impact of ChatGPT. The synthesis did not treat the presence of ChatGPT as an automatically beneficial intervention; instead, it examined the conditions under which ChatGPT appeared to support or constrain Arabic language learning. This analytical approach enabled the review to move beyond a simple description of ChatGPT adoption toward a more comprehensive explanation of its potential educational impact, limitations, and implications for Arabic language education.

7. Reporting and Reproducibility

The review procedures, search strategy, eligibility criteria, extraction framework, quality-appraisal principles, and thematic synthesis procedures were documented to support reproducibility. The PRISMA 2020 framework was used to guide the reporting of study identification and selection, while the final review reported the characteristics and methodological diversity of the included studies in a transparent manner (Page et al., 2021). The methodological structure was also informed by recent systematic reviews of ChatGPT and language education,

which have demonstrated the importance of transparent search procedures, predefined screening criteria, structured extraction, and critical synthesis when reviewing rapidly developing AI research (Albadarin et al., 2024; Lee et al., 2025).

RESULTS

1. ChatGPT and Arabic Language Learning Outcomes

a. Linguistic Development

The synthesis of the reviewed studies identified a predominantly positive pattern regarding the contribution of ChatGPT to Arabic language development, particularly in writing, speaking, reading, vocabulary, grammar, and translation. In Arabic language learning, Lelepary et al. (2023) reported positive student responses to ChatGPT, particularly in reading activities, task completion, and learning enthusiasm. More specifically, Zubaidi et al. (2025) reported an improvement in Arabic writing performance following the implementation of a human-AI collaboration model, indicating the potential of ChatGPT to support productive language development when its use was integrated into a structured learning process.

Evidence for speaking development was also identified. Kholid and Majeed (2025) found improvements in Arabic speaking fluency, grammatical accuracy, and communicative confidence following ChatGPT-supported conversational activities. In broader foreign-language learning contexts, Karataş et al. (2024) reported positive outcomes in writing, grammar, and vocabulary acquisition, although the effect on speaking was comparatively less evident. Across these studies, the reported outcomes therefore varied according to the language skill addressed and the manner in which ChatGPT was incorporated into instructional activities.

Translation was another recurring application of ChatGPT in Arabic language learning. Faris and Abdurrahman (2023) found that ChatGPT could assist Arabic-Indonesian translation and support learners' comprehension of Arabic texts. Similarly, Salma et al. (2025) identified substantial potential for ChatGPT-assisted translation of classical Arabic texts. However, their findings also indicated that AI-generated translations required human evaluation to maintain accuracy, acceptability, and readability, particularly when dealing with religious terminology and conceptually complex expressions. Thus, translation outcomes were generally favorable but remained dependent on human verification.

The reviewed literature further identified immediate feedback, individualized assistance, and opportunities for repeated language practice as recurring functions of ChatGPT. Nasaruddin (2024) reported that ChatGPT could assist in generating and adapting Arabic learning materials according to learners' topics and proficiency levels. At the broader language-learning level, Chen et al. (2026) identified language practice, personalized feedback, authentic language input, interaction, and real-time linguistic assistance as major affordances of generative AI. These findings are consistent with systematic evidence showing that personalized learning support and expanded opportunities for language practice constitute important applications of ChatGPT, although empirical evidence remains uneven across language skills and educational contexts (Lee et al., 2026; Lo et al., 2024).

Overall, the reviewed studies showed that the strongest evidence for linguistic development was concentrated in writing, speaking practice, translation, grammar, and vocabulary-related activities. Reading and broader receptive-language outcomes were also reported, but the evidence base was comparatively less extensive. The findings therefore demonstrate a differentiated pattern

of ChatGPT effectiveness across language domains rather than a uniform effect on all Arabic language skills.

b. Learning and Cognitive Outcomes

The synthesis also identified positive outcomes beyond linguistic proficiency, particularly in motivation, engagement, learner autonomy, self-regulated learning, and higher-order cognitive processes. In Arabic language education, Lelepary et al. (2023) reported increased motivation and enthusiasm among students using ChatGPT. Similar patterns were observed in foreign-language contexts, where ChatGPT-supported learning was associated with increased engagement and positive learning experiences (Karataş et al., 2024). Wang et al. (2025) further reported that perceived usefulness, ease of use, and enjoyment were associated with students' intention to use ChatGPT for language learning, indicating that motivational and perceptual factors were closely related to continued technology use.

Learner autonomy and self-directed learning represented another recurrent finding. Dizon (2024) reported that ChatGPT facilitated self-directed foreign-language learning by allowing learners to obtain assistance according to their individual learning needs and schedules. Within Arabic language education, Mulloh et al. (2025) characterized ChatGPT as a potential cognitive partner that could provide scaffolding and support dialogic learning. These findings were reinforced by Lee et al. (2024), who found that guided ChatGPT-assisted learning improved self-regulated learning and knowledge construction when students were supported through structured interaction rather than simply receiving direct answers.

The evidence concerning higher-order cognitive outcomes was also favorable. Lee et al. (2024) reported improvements in higher-order thinking, problem solving, creativity, and knowledge construction when ChatGPT was integrated through a guidance mechanism. Similarly, Essel et al. (2024) found that ChatGPT-supported learning was associated with critical, creative, and reflective thinking outcomes. These findings indicate that cognitive benefits were more consistently reported when ChatGPT was embedded within learning activities requiring active engagement, reflection, and knowledge construction.

Across the reviewed studies, therefore, ChatGPT was associated with two interconnected outcome dimensions: language development and learning development. The first included improvements in selected productive and receptive language skills, while the second encompassed motivation, engagement, autonomy, self-regulation, and higher-order thinking. However, the evidence also showed that these outcomes were influenced by the mode of integration. Studies employing structured guidance, human-AI collaboration, feedback, or interactive practice generally reported broader learning outcomes than studies in which ChatGPT was used primarily as an answer-generating tool.

2. ChatGPT Integration Among Generation Z Students

a. Digital Learning Characteristics

1). Technology Use and Learning Flexibility

The reviewed literature demonstrated that ChatGPT integration was closely associated with the digital learning characteristics of contemporary university students, particularly flexibility, accessibility, and continuous technology-mediated learning. Dizon (2024) found that ChatGPT supported self-directed language learning by allowing students to access assistance according to their individual schedules and learning needs. This flexibility was also reflected in broader language-learning research, where generative AI was reported to extend opportunities for language practice and interaction beyond conventional classroom settings (Chen et al., 2026). In the

Indonesian Arabic language-learning context, Anwar and Mufidah (2024) reported substantial student use of artificial intelligence, with ChatGPT emerging as one of the most frequently used platforms for academic activities. These findings indicate that ChatGPT has become increasingly embedded in students' existing digital learning practices rather than functioning solely as a supplementary classroom technology.

The evidence further showed that technology use among digitally oriented students was characterized by accessibility and continuity. ChatGPT enabled learners to obtain explanations, linguistic assistance, and learning support without being restricted by classroom time or location (Dizon, 2024). In Arabic language education, the capacity to generate or adapt learning materials according to topic and learner needs also increased the flexibility of instructional support (Nasaruddin, 2024). Thus, the reviewed studies identified learning flexibility as one of the most consistent characteristics associated with ChatGPT-supported learning.

2). Personalized and Interactive Learning

Personalization and interaction emerged as additional characteristics of ChatGPT-supported learning among digitally engaged students. Nasaruddin (2024) reported that ChatGPT could be used to produce and modify Arabic learning materials according to topics, proficiency levels, and specific instructional requirements. In broader language-learning contexts, Chen et al. (2026) identified personalized feedback, interaction, authentic language input, and real-time assistance as major affordances of generative AI. These findings indicate that ChatGPT can provide learners with more individualized forms of support than conventional one-directional digital materials.

The interactive dimension was also associated with students' learning experiences and engagement. Wang et al. (2025) found that enjoyment and perceived usefulness were related to students' intention to continue using ChatGPT for language learning. Similarly, Klimova et al. (2024) reported generally positive student perceptions of ChatGPT as a flexible and personalized language-learning tool. The evidence therefore indicates that Generation Z students' integration of ChatGPT was characterized not only by frequent technology use but also by an expectation of immediate interaction, individualized assistance, and responsive learning support.

b. AI Literacy and Learning Behavior

1). AI Literacy and Effective Prompting

AI literacy emerged as an important factor in determining how effectively students used ChatGPT. Al-Abdullatif and Alsubaie (2024) found that AI literacy contributed to students' perceptions of the value and usefulness of ChatGPT, indicating that technology adoption was not independent of students' understanding of AI capabilities and limitations. This finding is particularly relevant to Generation Z because high familiarity with digital technologies did not necessarily imply an equivalent ability to critically evaluate AI-generated content.

The reviewed literature also indicated that effective interaction with ChatGPT depended on students' ability to formulate appropriate prompts. Nasaruddin (2024) demonstrated that carefully designed prompts enabled the generation of Arabic learning materials that could be adjusted to topic, proficiency level, and learning requirements. The use of structured guidance was similarly associated with stronger learning outcomes, as Lee et al. (2024) found that guided ChatGPT interaction supported self-regulated learning and higher-order thinking more effectively than unrestricted answer generation. Therefore, the evidence identified prompting and AI literacy as important components of effective ChatGPT-supported learning behavior.

2). Perceived Usefulness, Ease of Use, and Acceptance

Students' acceptance of ChatGPT was consistently associated with perceived usefulness, ease of use, enjoyment, and perceived risks. Wang et al. (2025) reported that usefulness, ease of use, and enjoyment contributed to students' intention to use ChatGPT in language learning. Similarly, Sallam et al. (2024) found that perceived usefulness, ease of use, risk perceptions, and psychosocial factors significantly influenced students' attitudes toward ChatGPT. Klimova et al. (2024) likewise reported generally favorable perceptions among university students, particularly because ChatGPT could provide rapid and flexible language-learning assistance.

However, acceptance was not determined by usefulness alone. Al-Abdullatif and Alsubaie (2024) showed that AI literacy also shaped students' perceptions and intentions to use ChatGPT. Among Generation Z students, Acosta-Enriquez et al. (2024) further found that knowledge, attitudes, ethical perceptions, and concerns were relevant to ChatGPT use. These findings demonstrate that acceptance represented a multidimensional construct involving functional, experiential, cognitive, and ethical considerations.

3). Patterns of ChatGPT Use in Academic Learning

The reviewed studies identified several recurring patterns in students' academic use of ChatGPT. These included assistance with language practice, writing, translation, content generation, explanation of difficult concepts, feedback, and completion of academic tasks (Anwar & Mufidah, 2024; Dizon, 2024; Nasaruddin, 2024). In Arabic language learning specifically, students used AI-supported tools for research-related activities and language tasks, while ChatGPT was reported as one of the most frequently accessed AI platforms (Anwar & Mufidah, 2024). ChatGPT was also used to support translation and comprehension of Arabic texts (Faris & Abdurrahman, 2023).

The evidence further indicated that students did not use ChatGPT exclusively for passive information retrieval. Its functions included interactive practice, feedback, writing assistance, conversational activities, and personalized learning support (Chen et al., 2026; Dizon, 2024). Nevertheless, the pattern of use varied according to learners' technological competence, perceived usefulness, AI literacy, and learning objectives (Al-Abdullatif & Alsubaie, 2024; Sallam et al., 2024). Overall, the findings characterize Generation Z students' ChatGPT use as flexible, interactive, personalized, and task-oriented, with the extent and quality of use influenced by students' AI literacy and perceptions of the technology.

3. Challenges and Emerging Patterns of ChatGPT Use

a. Ethical and Pedagogical Challenges

1). Academic Integrity and Dependency

The reviewed literature identified academic integrity and excessive dependence on AI as recurring challenges in ChatGPT-supported learning. Mah et al. (2024) reported that the boundary between legitimate learning support and academic cheating remained ambiguous among students and teachers. This ambiguity was particularly relevant when ChatGPT was used to generate complete responses rather than to support students' own learning processes. Similar concerns regarding dependency were reported by Karataş et al. (2024), who identified overreliance on ChatGPT as a potential limitation of its use in foreign-language education. In Arabic language learning, Zubaidi et al. (2025) emphasized the importance of human-AI collaboration in which learners and teachers remained responsible for conceptualization, validation, and execution rather than delegating the entire learning process to AI.

The evidence also indicated that pedagogical guidance influenced whether ChatGPT functioned as a learning scaffold or as a substitute for student effort. Lee et al. (2024) found

stronger self-regulated learning and higher-order thinking outcomes when ChatGPT was incorporated through a structured guidance mechanism. This pattern suggests that unrestricted access to AI did not necessarily produce equivalent educational benefits. Instead, the reviewed studies showed that dependency and academic integrity concerns were closely associated with the way ChatGPT was positioned within learning activities.

2). Accuracy, Privacy, and Responsible Use

Accuracy and responsible use constituted another major group of challenges. Salma et al. (2025) found that ChatGPT-assisted translation of classical Arabic texts required human evaluation to ensure accuracy, acceptability, and readability, particularly when religious terminology and complex concepts were involved. Similar concerns regarding the reliability of AI-generated information have been identified in broader foreign-language education research, where students recognized information accuracy as one of the limitations associated with ChatGPT (Klimova et al., 2024). These findings indicate that AI-generated linguistic content cannot consistently be treated as an authoritative source.

Privacy and responsible use also emerged as important considerations. Klimova et al. (2024) identified privacy and ethical concerns among university students, while Sallam et al. (2024) reported that perceived risks and psychosocial factors influenced students' attitudes toward ChatGPT. Among Generation Z students, Acosta-Enriquez et al. (2024) further demonstrated that ethical perceptions and awareness of risks were associated with students' use of ChatGPT. Consequently, the reviewed studies identified responsible AI use as involving not only technical competence but also critical evaluation, awareness of limitations, ethical judgment, and appropriate human oversight.

b. Trends and Research Gaps

1). Increasing Integration of ChatGPT in Language Learning

The findings demonstrated a clear trend toward the increasing integration of ChatGPT into language learning activities. ChatGPT has been employed for writing, speaking practice, translation, reading support, feedback, content generation, conversational interaction, and self-directed learning (Chen et al., 2026; Dizon, 2024; Nasaruddin, 2024). In the Arabic language context, students have increasingly incorporated AI into academic activities, with ChatGPT reported as one of the most frequently used AI platforms (Anwar & Mufidah, 2024). Systematic reviews have likewise identified a rapidly expanding body of research on GenAI applications in language education (Lee et al., 2026; Lo et al., 2024).

The reviewed evidence therefore indicates a transition from experimental or exploratory use toward broader educational integration. However, the increasing adoption of ChatGPT has developed more rapidly than the accumulation of robust evidence concerning its long-term educational effects. This discrepancy was particularly apparent across studies employing different learning contexts, outcomes, and methodological designs (Lee et al., 2026; Lo et al., 2024).

2). Limited Evidence on Generation Z in Arabic Language Learning

Despite the growing body of literature on ChatGPT and language education, direct evidence concerning Generation Z students in Arabic language learning remained limited. Existing research on Generation Z has primarily examined students' knowledge, attitudes, ethical perceptions, and general use of ChatGPT in higher education rather than its specific effects on Arabic language acquisition (Acosta-Enriquez et al., 2024). In contrast, Arabic-language studies have primarily

concentrated on individual competencies, including reading, writing, speaking, translation, or general learning assistance (Kholid & Majeed, 2025; Lelepary et al., 2023; Zubaidi et al., 2025).

This separation between research on Generation Z and research on Arabic language learning represents an important gap in the current evidence base. Although Generation Z students constitute a highly digitally engaged learner population, their specific interaction with ChatGPT in Arabic language learning has not yet been sufficiently synthesized across linguistic, cognitive, motivational, and behavioral dimensions.

3). Need for Broader Evidence Across Arabic Language Skills

The final pattern identified in the literature was the uneven distribution of evidence across Arabic language skills. Existing Arabic-language studies have provided useful evidence concerning writing (Zubaidi et al., 2025), speaking (Kholid & Majeed, 2025), reading and general learning experiences (Lelepary et al., 2023), and translation (Faris & Abdurrahman, 2023; Salma et al., 2025). However, the evidence remains fragmented and does not provide an equally developed understanding of ChatGPT's contribution across reading, listening, speaking, writing, grammar, vocabulary, and translation.

This limitation is consistent with broader systematic evidence showing that ChatGPT research in language education remains unevenly distributed across language skills and learning outcomes (Lo et al., 2024). Although generative AI offers opportunities for language practice, personalized feedback, interaction, and real-time linguistic support, comprehensive evidence across different language domains remains insufficient (Chen et al., 2026). The reviewed findings therefore indicate a need for broader and more integrated evidence that examines Arabic language learning as a multidimensional process rather than focusing on isolated language skills.

Overall, the findings reveal a dual pattern in the development of ChatGPT-supported Arabic language learning. On the one hand, integration is expanding across language activities and is associated with increased flexibility, personalization, and learning support. On the other hand, concerns regarding academic integrity, dependency, accuracy, privacy, and responsible use remain evident. At the same time, the literature continues to show limited direct evidence concerning Generation Z students and an uneven distribution of research across Arabic language skills. These patterns define the principal areas requiring further investigation in the emerging field of ChatGPT-supported Arabic language education.

DISCUSSION

1. Results Analysis

The findings of this literature review indicate that the integration of ChatGPT has a multidimensional influence on Arabic language learning among Generation Z students. Rather than functioning as a single-purpose technological tool, ChatGPT appears to operate across linguistic, cognitive, motivational, and behavioral dimensions. The most consistent evidence concerns its contribution to writing, speaking practice, translation, language practice, motivation, personalized support, and learner autonomy, while its effects are less consistently established across all language skills. This pattern directly addresses the central objective of the present review, namely, to determine how ChatGPT integration influences Arabic language learning and to identify the conditions under which such integration produces meaningful educational outcomes.

The linguistic findings suggest that ChatGPT is particularly valuable because it expands opportunities for language production and immediate interaction. Improvements reported in Arabic writing and speaking indicate that ChatGPT can provide learners with repeated opportunities to practice productive skills, receive feedback, and revise their language output

(Kholid & Majeed, 2025; Zubaidi et al., 2025). This is particularly significant in Arabic learning, where opportunities for sustained interaction and individualized feedback may vary considerably across learning contexts. The findings from Leleparry et al. (2023) and Nasaruddin (2024) further indicate that ChatGPT can support reading comprehension and the development of adaptable learning materials. Consequently, its value appears to lie not simply in generating language content, but in increasing the frequency and accessibility of learner interaction with the target language.

The findings concerning translation provide a more nuanced interpretation. ChatGPT can facilitate Arabic-Indonesian translation and improve access to difficult texts, but the persistence of inaccuracies in complex or culturally specific content demonstrates that linguistic assistance does not necessarily equal linguistic reliability (Faris & Abdurrahman, 2023; Salma et al., 2025). This distinction is important for interpreting the overall impact of ChatGPT. Its effectiveness should therefore be understood in terms of supported learning, in which AI accelerates access to linguistic resources while human learners remain responsible for interpretation and verification. The same principle applies to grammar, vocabulary, and textual comprehension, where immediate AI assistance may increase practice opportunities but does not eliminate the need for human judgment.

The cognitive and affective findings further suggest that the impact of ChatGPT depends on the nature of learner interaction with the technology. The positive patterns in motivation, engagement, and enjoyment indicate that ChatGPT can make language-learning activities more responsive and less constrained by classroom time and resources (Karataş et al., 2024; Wang et al., 2025). However, the evidence on self-regulated learning and higher-order thinking suggests that positive cognitive outcomes are more likely when students are required to think, evaluate, and construct responses rather than simply accept AI-generated answers. Lee et al. (2024) demonstrated that guided ChatGPT use was associated with stronger self-regulation, higher-order thinking, and knowledge construction, while Essel et al. (2024) reported potential benefits for critical, creative, and reflective thinking. Thus, the findings indicate that the pedagogical mode of ChatGPT integration is more consequential than technological access alone.

This interpretation also explains why the reviewed literature simultaneously reports educational benefits and concerns about dependency. When ChatGPT is used as a scaffold, feedback mechanism, conversational partner, or source of guided assistance, it can support learner autonomy and cognitive engagement (Dizon, 2024; Mulloh et al., 2025). In contrast, when it is used primarily to generate complete answers, its convenience may reduce opportunities for independent reasoning and academic effort. The concern over academic integrity identified by Mah et al. (2024) and the dependency reported by Karataş et al. (2024) should therefore not be interpreted as evidence that ChatGPT is inherently harmful. Rather, these findings indicate that its educational effect is conditioned by the degree of learner agency, task design, and pedagogical supervision.

The Generation Z dimension provides an additional explanation for these patterns. The reviewed evidence suggests that Generation Z students are likely to value accessibility, speed, interaction, and personalization, characteristics that are strongly represented in ChatGPT-supported learning environments. Nevertheless, technological familiarity does not automatically translate into critical or ethical AI use. Acosta-Enriquez et al. (2024) showed that knowledge, attitudes, ethical perceptions, and concerns influence how Generation Z students engage with ChatGPT, while Al-Abdullatif and Alsubaie (2024) demonstrated the importance of AI literacy in

shaping students' perceptions and intentions. Therefore, the impact of ChatGPT among Generation Z should be understood as the interaction between digital readiness and AI literacy, rather than as a simple consequence of belonging to a digitally familiar generation.

The findings also provide an explanation for the uneven evidence across Arabic language skills. Existing studies have concentrated more strongly on writing, speaking, translation, and general learning support than on listening and other receptive skills (Kholid & Majeed, 2025; Lelepariy et al., 2023; Zubaidi et al., 2025). This distribution suggests that the affordances of ChatGPT are currently more visible in language domains in which text-based interaction and immediate AI feedback are directly applicable. Consequently, the apparent impact of ChatGPT on Arabic language learning may partly reflect the technological affordances of the tool itself. The limited evidence for some skills should therefore not be interpreted as evidence of limited potential, but as an indication that those domains remain underinvestigated.

Overall, the findings support a conditional interpretation of ChatGPT integration in Arabic language learning. ChatGPT can enhance access to language practice, personalized assistance, motivation, and selected cognitive outcomes, but these benefits are not automatic. They depend on structured integration, active learner participation, AI literacy, human validation, and pedagogical guidance. In relation to the objective of this review, ChatGPT should therefore be conceptualized not as a replacement for conventional Arabic language instruction, but as a human-AI learning support system whose educational effectiveness depends on how learners, teachers, and instructional tasks are organized around the technology.

2. Comparison with Previous Studies

The findings of this literature review are broadly consistent with previous studies that have identified ChatGPT as a potentially beneficial tool for language learning, while also revealing several differences in the scope and conditions of its impact. The positive evidence concerning Arabic writing, speaking, reading, translation, motivation, and learner support is consistent with earlier findings showing that ChatGPT can enhance language practice and provide immediate assistance. Lelepariy et al. (2023), for example, reported positive student perceptions of ChatGPT in Arabic language learning, particularly for reading, motivation, and academic task completion. Similarly, Zubaidi et al. (2025) found measurable improvement in Arabic writing through a human-AI collaboration model, while Kholid and Majeed (2025) reported improvements in Arabic speaking fluency, grammatical accuracy, and communicative confidence. These findings correspond with broader foreign-language research indicating positive effects on writing, grammar, vocabulary, motivation, and engagement (Karataş et al., 2024; Song & Song, 2023).

The present synthesis also supports the growing view that ChatGPT can facilitate learner autonomy and personalized learning. Dizon (2024) demonstrated that ChatGPT can support self-directed foreign-language learning, while Nasaruddin (2024) showed that the technology can assist in adapting Arabic learning materials to learners' topics and proficiency levels. These findings are consistent with the broader affordances identified by Chen et al. (2026), particularly personalized feedback, continuous language practice, real-time assistance, and interaction. The consistency across these studies indicates that flexibility and personalization are not isolated effects but recurring characteristics of ChatGPT-supported language learning.

However, the present findings also extend previous research by emphasizing that the quality of integration influences the magnitude and direction of learning outcomes. Earlier studies sometimes emphasized students' positive perceptions or specific skill improvements, whereas the current synthesis indicates that stronger cognitive outcomes were associated with guided

interaction and structured pedagogical use. Lee et al. (2024) found that a guidance-based ChatGPT approach improved self-regulated learning, higher-order thinking, creativity, and knowledge construction. Similarly, Essel et al. (2024) reported positive associations between ChatGPT use and critical, creative, and reflective thinking. These findings are consistent with the cognitive scaffolding perspective proposed by Mulloh et al. (2025), in which ChatGPT functions as a cognitive partner rather than merely an answer-generating system. Thus, the current review strengthens the argument that the pedagogical conditions surrounding AI use are central to explaining its educational effects.

The findings concerning the limitations of ChatGPT also correspond with earlier literature. The need for human verification in Arabic translation identified by Faris and Abdurrahman (2023) and Salma et al. (2025) is consistent with concerns regarding the accuracy and reliability of AI-generated information reported in foreign-language education (Klimova et al., 2024). Likewise, concerns about excessive dependence and academic integrity correspond with findings from Karataş et al. (2024) and Mah et al. (2024). These consistencies indicate that the advantages of ChatGPT are accompanied by persistent pedagogical risks, particularly when learners treat AI-generated content as authoritative or use it to replace rather than support their own cognitive effort.

The present findings also align with technology-acceptance research concerning the role of perceived usefulness, ease of use, enjoyment, and risk. Wang et al. (2025) found that these factors influenced students' intention to use ChatGPT for language learning, while Sallam et al. (2024) similarly reported that usefulness, ease of use, perceived risks, and psychosocial factors shaped students' attitudes toward the technology. Al-Abdullatif and Alsubaie (2024) extended this perspective by demonstrating that AI literacy also influences students' perceptions and intentions. The convergence of these findings suggests that ChatGPT adoption cannot be explained solely by technological availability. Instead, acceptance is shaped by a combination of functional value, user experience, perceived risk, and learners' capacity to understand AI.

An important difference emerges, however, when the findings are examined through the Generation Z perspective. Although previous studies have investigated ChatGPT use among university students, relatively few have directly connected Generation Z characteristics with Arabic language learning outcomes. Acosta-Enriquez et al. (2024) specifically demonstrated that Generation Z students' knowledge, attitudes, ethical perceptions, and concerns influence their use of ChatGPT. Nevertheless, much of the existing language-learning literature has focused either on general university populations or on English and other foreign-language contexts rather than on Generation Z learners of Arabic (Lo et al., 2024; Lee et al., 2026). The present review therefore connects two research strands that have largely developed separately: research on Generation Z and AI use, and research on ChatGPT-supported Arabic language learning.

Another difference concerns the distribution of evidence across language skills. Previous studies have demonstrated relatively strong findings for writing, speaking, translation, vocabulary, and grammar, but evidence regarding listening and other receptive competencies remains less developed. This pattern is consistent with systematic reviews showing that ChatGPT research in language education is unevenly distributed across skills and learning outcomes (Lo et al., 2024; Lee et al., 2026). The present review confirms this imbalance within the Arabic context and suggests that current evidence does not yet support an assumption of equivalent effects across all dimensions of Arabic language proficiency.

From a theoretical perspective, the findings are compatible with a view of AI-supported learning in which technology functions as a scaffold for active learner participation. The evidence supporting self-regulated learning, higher-order thinking, personalization, and dialogic interaction is consistent with the cognitive and scaffolding orientation identified by Mulloh et al. (2025) and Lee et al. (2024). At the same time, the emphasis on multiliteracy and critical engagement is consistent with the view that learners must not simply consume AI-generated content but should interpret, evaluate, and transform information through active participation (Zapata, 2025). Thus, the findings generally converge with contemporary theoretical perspectives that position GenAI as a support for learning processes rather than a replacement for human cognition.

Overall, the comparison with previous studies demonstrates substantial convergence regarding the potential benefits of ChatGPT for language learning, particularly in personalization, language practice, motivation, and selected cognitive outcomes. The principal contribution of the present synthesis is the integration of these findings within the specific intersection of ChatGPT, Arabic language learning, and Generation Z. At the same time, the comparison reveals important differences in the depth of evidence across language skills and highlights the continuing importance of pedagogical guidance, AI literacy, human validation, and ethical regulation.

3. Implications of Findings

The findings have important theoretical and practical implications for Arabic language education. Theoretically, ChatGPT should be conceptualized as a human-AI learning support system rather than a substitute for teachers or learners. Its potential to support language development, motivation, self-regulated learning, and higher-order thinking is strongest when learners remain actively involved in evaluating and constructing knowledge (Lee et al., 2024; Mulloh et al., 2025). The findings also indicate that AI literacy should be incorporated into models of Generation Z learning because technological familiarity alone does not guarantee effective or ethical AI use (Acosta-Enriquez et al., 2024; Al-Abdullatif & Alsubaie, 2024).

Practically, ChatGPT can support Arabic writing, speaking, translation, reading, vocabulary, grammar, and personalized practice. However, its integration should be guided by clear learning objectives, structured prompting, teacher supervision, and human validation, particularly for linguistically or culturally sensitive content (Nasaruddin, 2024; Salma et al., 2025). Arabic language curricula should therefore combine AI-supported learning with critical evaluation, academic integrity, and independent learner engagement. The findings further suggest that teacher education should include AI literacy so that educators can design meaningful human-AI learning activities rather than merely permit unrestricted technology use.

4. Research Limitations

This review has several limitations. First, the included studies differed substantially in research design, sample characteristics, learning contexts, and outcome measures, limiting direct comparison across findings. Second, research on ChatGPT in language education remains strongly concentrated on English and higher education, while direct evidence concerning Generation Z in Arabic language learning is still limited (Lee et al., 2026; Lo et al., 2024). Third, several studies relied on small samples, short intervention periods, or self-reported perceptions, which may limit conclusions about long-term language-learning effects.

In addition, ChatGPT continues to evolve rapidly, and differences in model versions, prompting strategies, and technological access may influence reported outcomes. Publication and database-selection biases may also have affected the evidence available for synthesis. Finally, the heterogeneity of the studies prevented the calculation of a common quantitative effect size. Future

research should therefore employ larger and more diverse samples, longitudinal and experimental designs, validated Arabic language measures, and explicit identification of Generation Z characteristics to establish more robust evidence regarding the long-term effects of ChatGPT integration.

CONCLUSION

This literature review concludes that ChatGPT integration has substantial potential to support Arabic language learning among Generation Z students, particularly through enhanced language practice, writing, speaking, translation, personalized feedback, motivation, learner autonomy, and selected higher-order cognitive outcomes. The findings indicate that ChatGPT is most effective when integrated as a guided learning support system rather than as a replacement for teachers or students' own cognitive efforts. Its educational value is therefore shaped not only by technological accessibility but also by pedagogical design, AI literacy, learner engagement, human validation, and responsible use. At the same time, challenges related to academic integrity, dependency, information accuracy, privacy, and unequal evidence across Arabic language skills remain significant.

This review also identifies an important research gap: empirical evidence specifically connecting ChatGPT integration, Generation Z learners, and Arabic language learning remains limited compared with the rapidly expanding literature on English and broader foreign-language education. The evidence base is further constrained by methodological heterogeneity, relatively small samples in some studies, short intervention periods, and differences in ChatGPT applications and learning outcomes. Future research should therefore employ larger and more diverse Generation Z samples, longitudinal and experimental designs, validated measures of Arabic language proficiency, and comparative studies across reading, listening, speaking, writing, grammar, vocabulary, and translation. Further studies should also examine how AI literacy, prompting competence, learner autonomy, and ethical awareness influence the effectiveness of ChatGPT-supported Arabic language learning. Such research would contribute to the development of more evidence-based, critical, and pedagogically responsible models of human-AI collaboration in Arabic language education.

REFERENCES

- Acosta-Enriquez, B. G., Arbulú Ballesteros, M. A., Arbulu Perez Vargas, C. G., Orellana Ulloa, M. N., Gutiérrez Ulloa, C. R., Pizarro Romero, J. M., Gutiérrez Jaramillo, N. D., Cuenca Orellana, H. U., Ayala Anzoátegui, D. X., & López Roca, C. (2024). Knowledge, attitudes, and perceived ethics regarding the use of ChatGPT among Generation Z university students. *International Journal for Educational Integrity*, 20, 10. <https://doi.org/10.1007/s40979-024-00157-4>
- Al-Abdullatif, A. M., & Alsubaie, M. A. (2024). ChatGPT in learning: Assessing students' use intentions through the lens of perceived value and the influence of AI literacy. *Behavioral Sciences*, 14(9), 845. <https://doi.org/10.3390/bs14090845>
- Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education*, 3, 60. <https://doi.org/10.1007/s44217-024-00138-2>
- Anwar, M. R., & Mufidah, N. (2024). Istikhdām al-dhakhā' al-iṣṭinā'ī (AI) fī ta'lim ṭullāb ṣaff al-takhaṣṣuṣ fī baḥṭh al-lughah al-'arabiyyah: Penggunaan kecerdasan buatan (AI) dalam

- pembelajaran mahasiswa kelas peminatan penelitian bahasa Arab. *Al-Kafaah: Journal of Arabic Language and Linguistics Education (ALLE)*, 1–17. <https://doi.org/10.52491/alle.vi.186>
- Beckingham, S., Lawrence, J., Powell, S., & Hartley, P. (Eds.). (2024). *Using generative AI effectively in higher education: Sustainable and ethical practices for learning, teaching and assessment*. Routledge.
- Chen, Z., Wei, W., & Zou, D. (2026). Generative AI technology and language learning: Global language learners' responses to ChatGPT videos in social media. *Interactive Learning Environments*, 34(2), 907–920. <https://doi.org/10.1080/10494820.2025.2511248>
- Dizon, G. (2024). ChatGPT as a tool for self-directed foreign language learning. *Innovation in Language Learning and Teaching*, 20(2), 334–350. <https://doi.org/10.1080/17501229.2024.2413406>
- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, 6, 100198. <https://doi.org/10.1016/j.caeai.2023.100198>
- Faris, H., & Abdurrahman, M. (2023). Analisis penerjemahan bahasa Arab ke bahasa Indonesia melalui artificial intelligence ChatGPT. *Al-Af'idah: Jurnal Pendidikan Bahasa Arab dan Pengajarannya*, 7(2), 222–233. <https://doi.org/10.52266/al-afidah.v7i2.1967>
- Inayati, N., Sanjani, M. I., Jayanti, F. G., Gao, X., & Nguyen, H. T. M. (2025). Online informal language learning (OILL): A systematic review of studies (2014–2023). *Innovation in Language Learning and Teaching*, 19(4), 335–351. <https://doi.org/10.1080/17501229.2024.2355296>
- Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 29, 19343–19366. <https://doi.org/10.1007/s10639-024-12574-6>
- Kholid, A. A., & Majeed, M. Y. (2025). Enhancing Arabic speaking skills via ChatGPT: An exploratory study at Nahdlatul Ulama Sunan Giri University. *Al-Ittijah: Jurnal Keilmuan dan Kependidikan Bahasa Arab*, 17(2), 148–162. <https://doi.org/10.32678/alittijah.v17i2.12206>
- Klimova, B., Pikhart, M., & Al-Obaydi, L. H. (2024). Exploring the potential of ChatGPT for foreign language education at the university level. *Frontiers in Psychology*, 15, 1269319. <https://doi.org/10.3389/fpsyg.2024.1269319>
- Lahby, M. (Ed.). (2025). *Empowering digital education with ChatGPT: From theoretical to practical applications*. CRC Press.
- Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21, 16. <https://doi.org/10.1186/s41239-024-00447-4>
- Lee, S., Choe, H., Zou, D., & Jeon, J. (2026). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 34(1), 335–359. <https://doi.org/10.1080/10494820.2025.2498537>
- Leleparry, H. L., Rachmawati, R., Zani, B. N., & Maharjan, K. (2023). ChatGPT: Opportunities and challenges in the learning process of Arabic language in higher education. *Journal International of Lingua and Technology*, 2(1), 11–23. <https://doi.org/10.55849/jiltech.v2i1.439>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of

- empirical studies. *Smart Learning Environments*, 11, 50. <https://doi.org/10.1186/s40561-024-00342-5>
- Mah, C., Walker, H., Phalen, L., Levine, S., Beck, S. W., & Pittman, J. (2024). Beyond CheatBots: Examining tensions in teachers' and students' perceptions of cheating and learning with ChatGPT. *Education Sciences*, 14(5), 500. <https://doi.org/10.3390/educsci14050500>
- Mulloh, T., Uloom, S., & Nurhadi. (2025). Istikhdamu minaṣah ChatGPT fi ta'allum al-lughah al-'Arabiyah 'alā ḍau'i an-naẓariyah al-ma'rifiyyah: Dirāsah taḥlīl al-muḥtawā [The use of ChatGPT platform in Arabic learning based on cognitive theory: A content analysis]. *Lisanudhad: Jurnal Bahasa, Pembelajaran, dan Sastra Arab*, 12(1), 185–205. <https://doi.org/10.21111/lisanudhad.v12i1.13052>
- Narciso, P. (2024). *Generative AI in education: A guide for parents and teachers*. Apress. <https://doi.org/10.1007/979-8-8688-0844-9>
- Nasaruddin. (2024). Using ChatGPT in teaching Arabic as a foreign language. *Arabiyatuna: Jurnal Bahasa Arab*, 8(1), 1–24. <https://doi.org/10.29240/jba.v8i1.9413>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pan, F. (Ed.). (2024). *AI in language teaching, learning, and assessment*. IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-0872-1>
- Salma, A., Ma'arif, M. S., & Afyuddin, M. S. (2025). The quality of ChatGPT's Arabic-Indonesian translation of *Matan al-Ghayah wa at-Taqrīb*: Implications for Arabic language learning in pesantren. *ATHLA: Journal of Arabic Teaching, Linguistic and Literature*, 6(2), 81–97. <https://doi.org/10.22515/athla.v6i2.12555>
- Sallam, M., Elsayed, W., Al-Shorbagy, M., Barakat, M., El Khatib, S., Ghach, W., Alwan, N., Hallit, S., & Malaeb, D. (2024). ChatGPT usage and attitudes are driven by perceptions of usefulness, ease of use, risks, and psycho-social impact: A study among university students in the UAE. *Frontiers in Education*, 9, 1414758. <https://doi.org/10.3389/feduc.2024.1414758>
- 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, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Wang, L., Zou, B., Wang, C., & Huang, F. (2025). Understanding learner motivation and intention to use ChatGPT for English language learning. *Discover Computing*, 28(1), 293. <https://doi.org/10.1007/s10791-025-09805-w>
- Zapata, G. C. (Ed.). (2025). *Generative AI technologies, multiliteracies, and language education*. Routledge.
- Zubaidi, A., Munip, A., Widodo, S. A., & Zerrouki, T. (2025). Enhancing Arabic writing skills using ChatGPT-based AI learning models: A tridimensional human-AI collaboration framework. *Indonesian Journal of Applied Linguistics*, 15(1), 87–101. <https://doi.org/10.17509/ijal.v15i1.75378>