



Challenges and Opportunities of Artificial Intelligence Use in Arabic Language Learning: A Systematic Literature Review

Muhammad Rachmatul Aziz¹, Durtam², Suci Ramadhini³

^{1,2,3} UIN Siber Syekh Nurjati Cirebon, Indonesia

¹ rachmatulaziz@syekhnurjati.ac.id, ² Durtam@uinssc.ac.id, ³ suciramadhini95@gmail.com

Abstract

This systematic literature review examines the challenges and opportunities associated with the use of artificial intelligence (AI) in Arabic language learning and teaching. Following the PRISMA 2020 protocol and a Population-Concept-Context framework, 98 records were identified through an AI-assisted multi-database search layer and citation snowballing; 52 studies published between 2016 and 2026 met the eligibility criteria and were synthesized thematically. The review finds that AI applications, including generative chatbots, intelligent tutoring systems, automatic speech recognition, and text-to-speech tools, create opportunities for personalized feedback, expanded practice of listening and speaking skills, and greater learner autonomy. Persistent challenges include Arabic's dialectal diversity and morphological complexity, inconsistent accuracy of AI-generated content, unequal digital infrastructure, and gaps in teacher training and assessment policy. The review concludes that pedagogically guided, teacher-mediated integration is necessary to realize AI's potential in Arabic language education responsibly.

Keywords artificial intelligence; Arabic language learning; systematic literature review; PRISMA; educational technology

مستخلص البحث

تستعرض هذه الدراسة، القائمة على منهج المراجعة الأدبية المنهجية وفق بروتوكول بريزما 2020، التحديات والفرص المرتبطة باستخدام الذكاء الاصطناعي في تعلم اللغة العربية وتعليمها. باستخدام إطار السكان والمفهوم والسياق، تم حصر 98 دراسة عبر طبقة بحث متعددة القواعد بمساعدة الذكاء الاصطناعي وتتبع الاستشهادات المرجعية، واستوفت 52 دراسة نُشرت بين عامي 2016 و2026 معايير الأهلية وخضعت للتحليل الموضوعي. تُظهر النتائج أن تطبيقات الذكاء الاصطناعي، مثل روبوتات الدردشة التوليدية وأنظمة التعليم الذكية والتعرف الآلي على الكلام وتحويل النص إلى كلام، تتيح فرصًا للتغذية الراجعة الشخصية وتوسيع ممارسة مهاراتي الاستماع والكلام وتعزيز استقلالية المتعلم. أما التحديات المستمرة فتشمل تنوع اللهجات العربية وتعقيد الصرف، وتفاوت دقة المحتوى المُولّد آليًا، وعدم تكافؤ البنية التحتية الرقمية، والفجوات في تدريب المعلمين وسياسات التقييم. وتخلص الدراسة إلى ضرورة الدمج التربوي الموجّه بإشراف المعلم لتحقيق إمكانات الذكاء الاصطناعي في تعليم اللغة العربية بمسؤولية.

2Muhammad Rachmatul Aziz, Durtam, Suci Rahmadhini, Challenges and Opportunities of Artificial Intelligence Use in Arabic Language Learning: A Systematic Literature Review

الذكاء الاصطناعي؛ تعلم اللغة العربية؛ المراجعة المنهجية؛ برزما؛ التكنولوجيا التعليمية

كلمات أساسية

Artificial intelligence (AI) has moved from a peripheral experiment to a mainstream feature of language classrooms worldwide, and Arabic language education is no exception. Chatbots built on large language models, automatic speech recognition engines, intelligent tutoring systems, and text-to-speech platforms are now used to support listening, speaking, reading, and writing instruction for learners of Arabic as a first, second, heritage, or foreign language (Moustafa et al., 2026; Zubair, 2026). The stakes of this shift are considerable. Arabic is spoken by more than 422 million people and serves as the liturgical language of over a billion Muslims worldwide (Al Moaiad et al., 2024), yet it is also linguistically demanding to teach and learn: it combines a non-concatenative root-and-pattern morphology, a diglossic gap between Modern Standard Arabic and regional dialects, and an orthography that typically omits short-vowel diacritics (Farghaly & Shaalan, 2009; Boulesnam & Boucetti, 2025). These features that make Arabic linguistically distinctive also make it comparatively difficult for AI systems, most of which were developed and tuned primarily on English-language data, to process reliably (Alayba, 2025). Understanding how AI is actually being taken up in Arabic classrooms, and where it helps or hinders learning, is therefore both a pedagogical and a technical question.

A growing body of empirical work has begun to answer that question. Studies published over the last decade, and concentrated heavily in the last three years, report that generative chatbots and large language model tools such as ChatGPT are used by Arabic teachers to prepare materials, generate practice items, and provide first-pass feedback (Alkaabi & Almaamari, 2025; Nasaruddin, 2024), while automatic speech recognition and text-to-speech systems have been trialled for pronunciation training and listening comprehension (Shao et al., 2022; Tolba et al., 2024; Hijriyah, 2025). Learner-facing studies from Indonesia, Saudi Arabia, the Gulf, and other Arabic-teaching contexts report gains in motivation, engagement, and learner autonomy alongside AI use (Sumiarni et al., 2026; Zaimah et al., 2023; Faqeeh et al., 2024). At the same time, a parallel and equally consistent strand of findings documents difficulties: teachers describe AI systems as unreliable when handling regional dialects and culturally specific content, students report inconsistent or inaccurate Arabic output, and institutions report gaps in infrastructure, teacher training, and assessment policy (Alkaabi & Almaamari, 2025; Adawiyah, 2025; Ridho et al., 2025).

This dual pattern, real gains reported alongside real limitations, is the field-level gap that motivates this review: individual studies each describe a slice of the challenge-opportunity picture, generally within a single country, institution, or AI tool, but no consolidated account yet exists of how these findings add up across the field as a whole for Arabic specifically. A recent and closely related systematic review by Hady et al. (2026) synthesized 15 studies drawn from Scopus and Google Scholar between 2023 and 2025 and reported that AI supports the four language skills through adaptive feedback and conversational agents while facing resource and accuracy constraints; that review, however, was restricted to a two-year publication window and to two search sources. The present review extends that line of work by broadening the search across a longer time window (2016 to 2026), drawing on an AI-assisted multi-database discovery layer supplemented by citation snowballing, and by organizing the synthesis explicitly around challenges and opportunities as parallel, comparably weighted themes rather than around AI applications alone. This is the research gap the present review addresses: a broader, more current, and more symmetrically balanced synthesis of the challenge-opportunity landscape of AI in Arabic language education.

Accordingly, this review aims to (1) map the range of AI technologies that have been applied to Arabic language learning and teaching, (2) synthesize the opportunities these technologies are reported to create for learners and teachers, and (3) synthesize the challenges that limit or complicate their use, drawing only on claims that are traceable to studies retrieved during the course of the review. Its contribution is threefold: it offers Arabic language teacher educators and curriculum developers an evidence-based map of where AI currently helps and where it currently struggles; it gives researchers a structured account of gaps that remain open, including the shortage of studies with rigorous experimental designs and the near-absence of dialect-focused evaluation; and it provides a transparent, reproducible search and screening record that later reviews can build on or challenge.

The remainder of the article is organized as follows. The Method section describes the review protocol, search strategy, eligibility criteria, and screening process, and reports the PRISMA 2020 flow of records. The Results and Discussion section presents the characteristics of the included studies and synthesizes findings under the two organizing themes of opportunities and challenges. The Conclusion summarizes the review's main findings and proposes directions for future research.

طريقة البحث

This review followed the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting protocol, adapted for a thematic rather than a meta-analytic synthesis, because the review question concerns qualitatively different challenges and opportunities rather than a single pooled quantitative effect.

Review question and framework

The review question was framed using the Population-Concept-Context (PCC) framework recommended for scoping-style systematic reviews. Population was defined as learners and teachers of Arabic language, at any education level and in any setting (first-language, second-language, heritage, or foreign-language instruction). Concept was defined as the challenges and opportunities associated with applying artificial intelligence technologies, including generative chatbots and large language models, intelligent tutoring systems, automatic speech recognition, text-to-speech and speech synthesis, machine translation, and adaptive or personalized learning platforms. Context was defined as Arabic language learning and teaching settings, whether formal (schools, universities, Islamic boarding schools) or informal (self-directed or community learning).

Eligibility criteria

Records were eligible if they (a) discussed an AI, machine learning, natural language processing, or generative AI technology applied specifically to Arabic language learning, teaching, or acquisition; (b) had an explicit educational or pedagogical focus involving learners, teachers, or classroom-level application, rather than being purely computational or engineering work with no educational application; (c) reported, discussed, or permitted inference of at least one challenge, limitation, benefit, or opportunity of AI use in this context; and (d) were published as a peer-reviewed journal article, conference proceeding paper, book chapter, or systematic/scoping review. Records were excluded if they concerned AI applied to a language other than Arabic, including English-as-a-foreign-language instruction of Arabic-speaking students (a different population and a different target language); if they were purely computational or engineering papers on Arabic natural language processing with no learner- or teacher-facing evaluation or discussion; or if they were preprints, unpublished theses, or records from venues that could not be verified as peer-reviewed.

Search strategy

Search terms were built from the PCC concept blocks, combining Arabic-language-education vocabulary in English, Indonesian, and Arabic transliterated terms (for example, “AI”, “artificial intelligence”, “generative AI”, “ChatGPT”, “chatbot”, “intelligent tutoring system”, “speech recognition”, and “text-to-speech” combined with “Arabic language”, “bahasa Arab”, “maharah”, and “pembelajaran bahasa Arab”). Live queries to subscription bibliographic databases such as Scopus and Web of Science were not available in this review session; this is stated explicitly as a methodological limitation rather than left implicit. Records were instead identified through two complementary channels. The first channel was an AI-assisted multi-database discovery layer (Consensus), which draws on Semantic Scholar, Scopus-indexed records, PubMed, and arXiv, queried with nine concept-block query variants (Table 1). The second channel was citation and reference-list snowballing from records retrieved in the first channel. This combined search yielded 98 uniquely titled records (92 from the discovery layer, 6 from snowballing); no duplicate titles were found across the two channels. Because the discovery-layer channel is relevance-ranked rather than a reproducible Boolean query against a single named database, the identification counts reported here should not be read as equivalent to an exhaustive Scopus or Web of Science export; they represent a transparent, reproducible-in-principle search of the sources actually queried.

Table 1. Search query themes

No.	Query theme	Channel	
1	Artificial intelligence, Arabic language learning and teaching	AI-assisted layer	discovery
2	Chatbot, Arabic language education, learners	AI-assisted layer	discovery
3	Automatic speech recognition, Arabic pronunciation learning	AI-assisted layer	discovery
4	ChatGPT, generative AI, Arabic language teaching, higher education	AI-assisted layer	discovery
5	Teacher perceptions, AI, Arabic language instruction, challenges	AI-assisted layer	discovery
6	Arabic NLP, low-resource, dialect, diacritics, morphology (background/context)	AI-assisted layer	discovery
7	Intelligent tutoring system, adaptive learning, Arabic grammar/vocabulary	AI-assisted layer	discovery

No.	Query theme	Channel
8	Machine translation, AI writing feedback, Arabic learners, academic integrity	AI-assisted discovery layer
9	Artificial intelligence, Arabic reading comprehension, listening skill	AI-assisted discovery layer
10	Artificial intelligence, Arabic language learning, motivation, engagement	AI-assisted discovery layer
11	Citation and reference-list snowballing from retrieved records	Snowballing

Selection process

One reviewer, assisted by an AI-based research tool for record retrieval and initial classification, screened all 98 records against the eligibility criteria in two stages. Because a second independent human screener was not involved in this iteration, this is reported as a limitation, and inter-rater reliability was not computed. At the title-and-abstract stage, 41 records were excluded: 27 for concerning a language other than Arabic or the wrong learner population, 13 for lacking any educational or pedagogical focus, and 1 for not being confirmed as an AI-specific technology, leaving 57 reports assessed in full for eligibility. At the eligibility stage, a further 5 reports were excluded: 3 for inappropriate publication type (an arXiv preprint, a record in an unverifiable venue, and an unpublished doctoral thesis), 1 for a construct mismatch discovered on closer reading (the tool it evaluated was explicitly described as not using artificial intelligence), and 1 because it was a systematic review of near-identical scope to the present one (Hady et al., 2026), which is discussed as related work in the Introduction and Discussion rather than double-counted as a primary source. Fifty-two studies, published between 2016 and 2026, were included in the final synthesis. Figure 1 presents the PRISMA 2020 flow diagram.

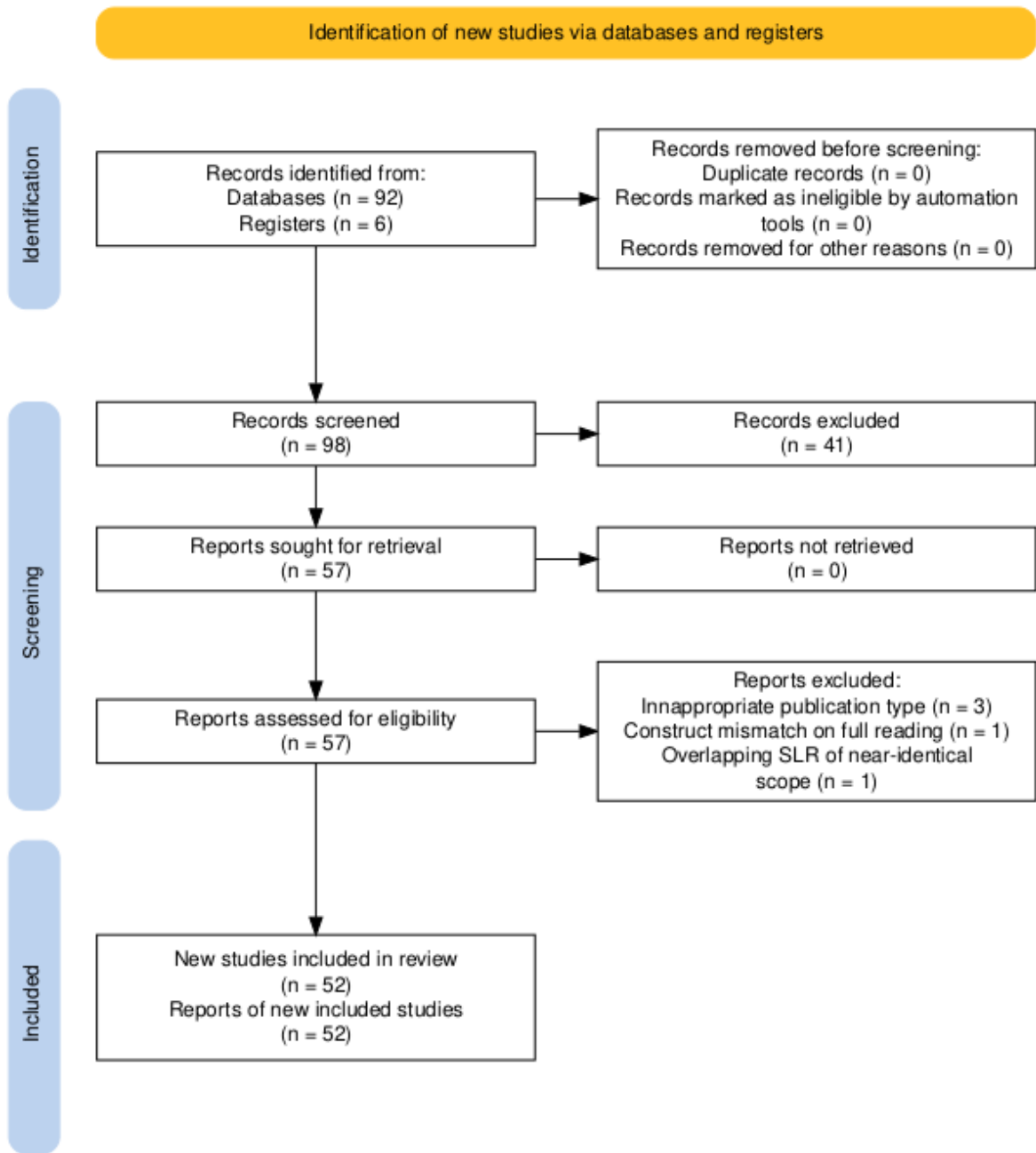


Figure 1. PRISMA 2020 flow diagram of record identification, screening, and inclusion

Data extraction and synthesis

For each included study, the following information was extracted where reported: first author and year, AI technology or tool, educational context and country or region, and study design. Table 2 summarizes these characteristics across all 52 included studies. Because the included literature combines experimental, survey, qualitative case-study, and conceptual or literature-review designs addressing a common but non-comparable set of outcomes, the studies were synthesized narratively using thematic analysis rather than statistically pooled.

Two organizing themes, Opportunities and Challenges, were used as the primary coding frame, consistent with the review question; within each theme, sub-themes were derived inductively from patterns recurring across at least three independent studies. Full multi-author lists could not be independently verified for every included record beyond first-author-level indexing metadata; where this applies, citations use “et al.” and this is noted here as a further transparency limitation rather than a claim of complete bibliographic verification.

Table 2. Characteristics of included studies (n = 52)

Author, Year	AI technology	Educational context	Study design
Mahmoud et al., 2016	Intelligent tutoring system (rule-based)	Primary/elementary grammar, Egypt	System development + testing
Mahmoud, 2018	Multi-agent intelligent tutoring system	Primary/elementary grammar, Egypt	System development + testing
Mohammdi et al., 2021	AI-based sign-language translation system	Deaf learners, Arabic Sign Language	System development + user evaluation
Natheer, 2022	Agent-based intelligent tutoring system	Higher education grammar	System development
Shao et al., 2022	AI speech tutor (ASR + deep learning)	University, Moroccan dialect	System development + evaluation
Simon, 2023	Generative AI (conceptual)	Arabic language education, general	Qualitative literature review
Yatri et al., 2023	Multiple AI tools (ChatGPT, Grammarly)	English & Arabic classes, mixed	Qualitative, learner/teacher interviews
Zaimah et al., 2023	LLM-based chatbot	Higher education, Indonesia	Mixed methods, quasi-experimental
Anon. (Al-Ta'rib), 2024	AI tools, general	Arabic language education, literature-based	Qualitative literature review
Berliani et al., 2024	ChatGPT	Arabic language learners, general	Descriptive-qualitative
Faqeeh et al., 2024	AI applications, general	Higher education	Quantitative survey
Garba et al., 2024	Multiple apps (ArabicPod101, Lingbe, Duolingo, Scribe)	Non-Arabic speakers, Nigeria	Descriptive survey
Nasaruddin, 2024	ChatGPT	Arabic as foreign language, teachers	Descriptive/conceptual
Sa'idah et al., 2024	AI tools (PLS-SEM evaluated)	Schools/university, Indonesia	Quantitative, PLS-SEM
Samin et al., 2024	AI integration planning	Higher education, Indonesia	Qualitative case study
Sarif et al., 2024	AI text-to-speech	Reading skills, general	Descriptive/conceptual
Tolba et al., 2024	AR + AI speech recognition	Phonetics education, adult learners	Experimental user study (n=83)

10 Muhammad Rachmatul Aziz, Durtam, Suci Rahmadhini, Challenges and Opportunities of Artificial Intelligence Use in Arabic Language Learning: A Systematic Literature Review

Author, Year	AI technology	Educational context	Study design
Adawiyah, 2025	Multiple AI tools (teacher-reported)	Islamic higher education	Qualitative case study
Alasmari, 2025	AI & m-learning (bibliometric)	Arabic countries, higher education	Bibliometric analysis
Albab et al., 2025	AI curriculum content tools	Higher education	True experimental (single group)
Aleedy et al., 2025	Multi-agent chatbot (GPT-based)	Higher education, translation learners	System development + user study
Alkaabi & Almaamari, 2025	Generative AI (teaching assistant)	Higher education instructors	Qualitative, interviews (n=15)
Alrashoudi et al., 2025	Mispronunciation detection (transformer)	Non-native learners	System development + human evaluation
El Zahraa, 2025	AI & digital tools, general	Sociolinguistic competence	Qualitative/conceptual
Fanoush et al., 2025	Mispronunciation detection (SSL model)	Primary school, Saudi Arabia	System development + dataset evaluation
Fredlina et al., 2025	ChatGPT	Higher education, Indonesia	Quantitative descriptive survey (n=152)
Hijriyah, 2025	SUNO.AI (listening)	Higher education, Indonesia	R&D, pre/post-test (n=60)
Madwi, 2025	AI applications, general	Arabic language education	Qualitative literature review
Mohd Yusof et al., 2025	AI chatbots (conceptual framework)	Arabic as foreign language	Conceptual/theoretical
Mufid et al., 2025	AI speaking tools (Google Assistant, FunEasyLearn)	Higher education, Indonesia	Qualitative, phenomenological
Muttaqin et al., 2025	ChatGPT	Higher education, Indonesia	Qualitative, interviews
Nasution et al., 2025	ASR (phonetic error analysis)	Higher education, Indonesia	Descriptive-analytical
Rafi'atun et al., 2025	ChatGPT (writing)	Higher education, Indonesia	Qualitative field study
Ridho et al., 2025	AI tools, general	Higher education, Indonesia (Lombok)	Qualitative
Ridwan et al., 2025	AI, general (literature review)	Arabic acquisition, general	Qualitative literature review
Samiya, 2025	Multiple AI tools, general	Arabic language education, general	Narrative/conceptual review
Setyawan et al., 2025	AI writing assistance	Higher education, Indonesia	Qualitative case study
Shalihah et al., 2025	AI chatbot vs traditional method	Arabic language learners	Comparative qualitative
Sismat et al., 2025	ChatGPT	Non-native Arabic learners	Empirical study (book chapter)

Author, Year	AI technology	Educational context	Study design
Supriatna et al., 2025	Multiple AI tools, general	Islamic university, Indonesia	Descriptive-analytical literature review
Syed et al., 2025	AI + Metaverse (GNT)	Conceptual system design	System/framework design
Al-Jarf, 2026	AI-narrated YouTube videos (TTS)	Foreign learners, listening	Descriptive-analytical
Humairoh et al., 2026	ChatGPT (autonomous learning)	Arabic as foreign language	Qualitative conceptual + simulation
Kamal, 2026	Multiple AI assessment tools	Gifted students, Malaysia	Mixed methods (n=80)
Moustafa et al., 2026	Generative AI, general	Arabic language education, global	Systematic literature review
Ni'mah, 2026	AI role-play (MemBot/Memrise)	Higher education, Indonesia	Classroom action research (n=40)
Oluwatoyin, 2026	Multiple AI tools (ChatGPT most used)	Higher education, Nigeria	Quantitative descriptive survey (n=60)
Sa'diyah et al., 2025	Suno.Ai (listening)	Islamic boarding school, Indonesia	Qualitative case study
Sofian et al., 2026	AI microlearning + vlog (listening)	Higher education	Quasi-experimental (n=60)
Sumiarni et al., 2026	Multiple AI tools, general	Higher education, teachers & students	Qualitative, interviews
Yuliani et al., 2025	AI text-to-speech (listening)	Arabic language learning, general	Qualitative literature review
Zubair, 2026	Multiple AI tools (synthesis of 4 studies)	Higher education, India	Systematic review/synthesis

النتائج والمناقشة

Of the 52 included studies, the majority (n = 38) were published in 2024 or later, confirming that empirical interest in AI and Arabic language learning is a very recent phenomenon that has accelerated sharply since the wider availability of generative AI tools such as ChatGPT. Geographically, Indonesia is the single most represented research context (n = 19 studies), followed by Saudi Arabia and other Gulf states (n = 9), with the remainder distributed across Malaysia, Nigeria, Egypt, Jordan, the United States, and multi-country or conceptual studies without a single fieldwork site. Methodologically, qualitative case studies and interview-based designs predominate (n = 24), followed by system-development and technical-evaluation studies (n = 11), quantitative surveys (n = 9), quasi-experimental or

experimental designs ($n = 6$), and literature reviews or conceptual papers ($n = 8$; some studies combine categories). Only a small minority of studies used a true experimental or quasi-experimental design with a comparison condition, which is itself a methodological finding discussed further below.

Opportunities of AI in Arabic language learning

Personalized feedback and adaptive practice. Intelligent tutoring systems built specifically for Arabic grammar demonstrate that AI can track an individual learner's error patterns and adjust the difficulty and sequencing of exercises accordingly, a capability first demonstrated for elementary-level Arabic grammar instruction in Egypt and since extended to higher-education contexts (Mahmoud et al., 2016; Mahmoud, 2018; Natheer, 2022). More recent generative-AI studies report a similar pattern for open-ended practice: teachers use large language models to generate differentiated materials and assessment items matched to individual student levels, and students report that AI-based platforms adjust pacing and content to their proficiency in ways that traditional textbooks cannot (Alkaabi & Almaamari, 2025; Sa'idah et al., 2024; Albab et al., 2025). Several studies link this personalization directly to measurable outcomes: a quasi-experimental evaluation of an AI curriculum-content tool reported significant pretest-to-posttest gains for participating students (Albab et al., 2025), and a structural-equation-modelling study across Indonesian schools and universities found a significant positive path coefficient between AI usage and learning outcomes when teacher training was also present (Sa'idah et al., 2024).

Expanded practice of listening and speaking skills. A cluster of studies converges on AI's contribution to skills that are traditionally the hardest to practice outside the classroom: listening comprehension and oral production. AI-based listening platforms and AI-narrated or AI-generated audio content, including tools such as SUNO.AI and AI text-to-speech engines, have produced statistically significant gains in listening comprehension scores in pre-test/post-test and quasi-experimental designs (Hijriyah, 2025; Sofian et al., 2026; Sa'diyah et al., 2025; Yuliani et al., 2025). For speaking and pronunciation, automatic speech recognition systems that detect mispronunciation at the phoneme level give learners immediate, private feedback that does not depend on a teacher's availability, which several studies describe as reducing the anxiety associated with speaking practice in front of peers (Shao et al., 2022; Tolba et al., 2024; Fanoush et al., 2025; Nasution et al., 2025). An AI-driven role-play model built on a chatbot feature within a commercial

vocabulary application produced measurable increases in both motivation and oral proficiency in a classroom action-research design (Ni'mah, 2026).

Conversational practice, motivation, and learner autonomy. Generative chatbots are consistently reported to increase learners' willingness to initiate practice outside class time, because they are available on demand and are perceived as a lower-stakes conversational partner than a human interlocutor (Zaimah et al., 2023; Aleedy et al., 2025; Mohd Yusof et al., 2025). Several qualitative studies frame this as supporting learner autonomy specifically: students describe using ChatGPT to rehearse and revise independently, at their own pace, before bringing work to a teacher for validation (Humairoh et al., 2026; Berliani et al., 2024; Sismat et al., 2025). Motivation gains are reported quantitatively as well as qualitatively; one interview-based study of teacher and student experience reported an 85 percent increase in self-reported motivation and engagement alongside AI adoption (Sumiarni et al., 2026), and a descriptive survey of AI awareness found consistently high average scores for perceived effectiveness across a large student sample (Faqeeh et al., 2024).

Support for teachers and reduced preparation burden. From the teacher's side of the classroom, AI is most consistently valued as a material-generation and first-pass-assessment tool. Instructors report using generative AI to draft reading passages, comprehension questions, and vocabulary exercises matched to a specific lesson topic, freeing time that would otherwise go to material preparation (Nasaruddin, 2024; Alkaabi & Almaamari, 2025; Samin et al., 2024). Several studies frame this as a shift in the teacher's role from information deliverer to facilitator of personalized learning, with teachers increasingly positioned as curators and validators of AI output rather than sole content creators (Sumiarni et al., 2026; Zubair, 2026).

Access and inclusion. A smaller but important set of studies documents AI's contribution to learners who are otherwise underserved by conventional Arabic instruction. An AI-based translation system between spoken Arabic and Arabic Sign Language was shown to significantly improve deaf students' ability to learn Arabic content, with a 99 percent sign-recognition rate in system testing (Mohammdi et al., 2021). Mobile and AI-supported learning is also reported to extend access in regions with dispersed or under-resourced Arabic-teaching institutions (Alasmari, 2025; El Zahraa, 2025), and one assessment study found AI-based tools particularly well suited to differentiating instruction for gifted and talented Arabic learners, who benefited from faster feedback and more challenging, individually calibrated material (Kamal, 2026).

Challenges of AI in Arabic language learning

Linguistic and technical limitations specific to Arabic. The single most frequently reported challenge across the corpus concerns the mismatch between AI systems' underlying training data, still overwhelmingly weighted toward English and Modern Standard Arabic, and the linguistic reality of Arabic instruction, which routinely involves regional dialects, diacritical ambiguity, and morphologically dense word forms. Instructors in one interview study explicitly identified AI accuracy in dialect processing as their most significant obstacle, alongside difficulty maintaining cultural authenticity in AI-generated content (Alkaabi & Almaamari, 2025). A descriptive-analytical study of AI-narrated Arabic audio found systematic pronunciation errors tied specifically to unmarked short-vowel diacritics, producing confusions between grammatically distinct word forms that a native listener would not make (Al-Jarf, 2026). Speech-recognition and mispronunciation-detection systems report similar constraints arising from the scarcity of annotated Arabic speech data relative to English (Alrashoudi et al., 2025; Fanoush et al., 2025).

Accuracy, reliability, and academic integrity concerns. Beyond dialect-specific errors, multiple studies report that AI-generated Arabic content is inconsistent in quality even for Modern Standard Arabic, which erodes learner and teacher confidence in using it unsupervised. A mixed-methods study of a large-language-model chatbot found that while learners' motivation increased, formal statistical tests did not show a significant improvement in language-learning effectiveness, and participants raised explicit concerns about the consistency of chatbot responses (Zaimah et al., 2023). Survey research on ChatGPT use among Arabic-education students found the tool strong for receptive skills such as reading simplification and vocabulary expansion but comparatively weak for productive skills such as writing correction and speaking practice, with grammar inaccuracy cited by nearly two-thirds of respondents (Fredlina et al., 2025). Academic-integrity concerns recur across writing-focused studies: students and teachers alike report ambiguity around plagiarism, difficulty verifying whether AI-assisted writing reflects genuine student competence, and, in one case study, confusion regarding intellectual attribution when AI is used across the stages of thesis-proposal writing (Rafi'atun et al., 2025; Setyawan et al., 2025; Muttaqin et al., 2025).

Unequal digital infrastructure and access. Several studies situate AI's promise against a backdrop of uneven technological readiness. A qualitative study conducted in Lombok, Indonesia, found that infrastructure gaps, limited teacher

training, and variable student digital readiness constrained what AI could realistically achieve in under-resourced regions, even where teachers and students held positive attitudes toward the technology in principle (Ridho et al., 2025). Similar infrastructure and access concerns are reported in Nigerian and Indian higher-education contexts (Oluwatoyin, 2026; Zubair, 2026), suggesting that this is not a problem confined to any single national setting but a structural feature of AI adoption in Arabic language education wherever institutional resourcing is limited.

Teacher readiness and professional development gaps. Closely related to infrastructure, a recurring theme is that Arabic language teachers frequently lack structured training in how to use AI tools pedagogically rather than merely operationally. Case-study research in Islamic higher-education settings found lecturers were broadly positive about AI's potential but constrained by their own limited familiarity with the tools, the absence of institutional guidelines, and uncertainty about how to adapt existing curricula (Adawiyah, 2025; Samin et al., 2024; Mufid et al., 2025). One study framed this explicitly as a policy gap rather than merely an individual skills gap, calling for institution-level AI literacy and ethics guidelines specific to Arabic instruction rather than generic AI-in-education policy (Alkaabi & Almaamari, 2025).

Cultural and contextual authenticity. A distinctive challenge in the Arabic case, less prominent in AI-in-education literature focused on other languages, concerns the cultural and religious register that much Arabic instruction is embedded in, particularly where Arabic is taught for access to Islamic textual sources. Several studies report that AI-generated content, however grammatically fluent, can miss register, cultural nuance, or religious context that a human teacher would supply automatically, and that this gap is felt most acutely in productive tasks such as writing and free conversation rather than in narrowly bounded drills (Muttaqin et al., 2025; Berliani et al., 2024; El Zahraa, 2025).

Assessment and evaluation strategy gaps. Finally, several studies note that assessment practices have not kept pace with AI adoption: institutions report using conventional rubrics to evaluate work that may have been partly AI-assisted, without a clear framework for distinguishing AI contribution from student competence, and instructors describe uncertainty about how to design authentic assessment tasks that remain meaningful once AI tools are freely available to students (Alkaabi & Almaamari, 2025; Sumiarni et al., 2026).

Synthesis and comparison with prior reviews

Read together, the opportunities and challenges identified in this review are not independent of one another; they largely track the same underlying tension. The same generative flexibility that lets a chatbot produce endless, on-demand practice material is also what makes its output inconsistent across dialects and registers; the same adaptivity that personalizes an intelligent tutoring system is only as good as the (currently thin) Arabic-specific training and assessment data behind it. This pattern is broadly consistent with the earlier review by Hady et al. (2026), which similarly reported that AI supports the four Arabic language skills while facing resource and accuracy constraints. The present review extends that finding in two respects. First, drawing on a larger and more recent corpus, it shows that infrastructure, teacher training, and assessment-policy gaps are reported at least as frequently as the purely linguistic and technical limitations that dominated earlier discussion, suggesting the field's central bottleneck may now be institutional and pedagogical readiness rather than the underlying AI technology alone. Second, the studies identified here skew heavily toward descriptive and qualitative designs; only a small subset used experimental or quasi-experimental designs with a comparison condition, and rigorous dialect-focused evaluation remains rare. This is itself a finding of the review: claims about AI's effectiveness for Arabic language learning currently rest on a comparatively thin base of controlled evidence, even where they are directionally consistent across many studies.

الاستنتاج والخلاصة

This systematic review synthesized 52 studies, identified through an AI-assisted multi-database search and citation snowballing and screened against explicit PRISMA 2020 criteria, to map the challenges and opportunities of artificial intelligence use in Arabic language learning and teaching. The evidence base shows that AI, particularly generative chatbots, intelligent tutoring systems, automatic speech recognition, and text-to-speech tools, offers genuine and recurring opportunities: personalized and adaptive feedback, expanded low-stakes practice of listening and speaking skills, increased learner motivation and autonomy, reduced material-preparation burden for teachers, and improved access for underserved learner groups. These opportunities are consistently reported alongside a comparably weighted set of challenges rooted in Arabic's dialectal and

morphological complexity, inconsistent output accuracy, unequal digital infrastructure, insufficient teacher training and institutional policy, gaps in cultural and religious authenticity, and unresolved assessment practices. Neither picture is complete without the other, and reviews or policy recommendations that report only the opportunities, or only the obstacles, misrepresent a field in which both are simultaneously and robustly documented.

Further research should prioritize four directions suggested by the gaps identified in this review. First, more experimental and quasi-experimental studies with comparison conditions are needed to move the evidence base beyond descriptive and perception-based designs. Second, dialect-focused evaluation of AI tools, rather than evaluation limited to Modern Standard Arabic, would address the specific technical limitation most frequently reported by teachers and learners. Third, institution-level studies of teacher training models and AI-specific assessment policy would help translate the individual-classroom findings synthesized here into scalable practice. Fourth, given that this review's search was conducted through an AI-assisted discovery layer and citation snowballing rather than direct manual export from subscription databases such as Scopus and Web of Science, a subsequent review with full manual multi-database access, independent dual screening, and complete author-list verification would strengthen confidence in the precise counts reported here, even though the substantive thematic pattern is unlikely to change with a larger corpus. With these caveats, the overall conclusion of this review is that AI holds substantial and well-documented promise for Arabic language education, provided its integration remains pedagogically guided and teacher-mediated rather than treated as a stand-alone substitute for human instruction.

كلمة الشكر والتقدير

The author(s) would like to express their sincere gratitude and appreciation to the Department of Arabic Language Education (PBA) at Universitas Islam Negeri Siber Syekh Nurjati Cirebon (UIN SSC) for providing the academic support and resources necessary to complete this research. Special thanks are extended to the lecturers of the PBA Department for their valuable insights, guidance, and constructive feedback throughout the research process. The author(s) are also deeply grateful to the students of the PBA Department who actively participated in this study, provided assistance during data collection, and contributed to the proofreading and language refinement of the manuscript. Their support and collaboration were invaluable to the success of this work

- Adawiyah, R. (2025). Implementing AI in Arabic language learning: Challenges and insights from Islamic higher education. *AL-ISHLAH: Jurnal Pendidikan*.
- Al Moaiad, Y., et al. (2024). Challenges in natural Arabic language processing. *Edelweiss Applied Science and Technology*.
- Al-Jarf, R. (2026). Are Arabic YouTube videos narrated by artificial intelligence suitable for training foreign students in listening skills? *Frontiers in Computer Science and Artificial Intelligence*.
- Al-Ta'rib. (2024). Innovative Arabic language learning: Redefining language archetypes using artificial intelligence. *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, 12(2), 293-308. <https://doi.org/10.23971/altarib.v12i2.8323>
- Alasmari, T. (2025). Artificial intelligence and m-learning in Arabic countries: Innovations, trends, and regional perspectives. *International Journal of Interactive Mobile Technologies*.
- Alayba, A. M. (2025). Arabic natural language processing (NLP): A comprehensive review of challenges, techniques, and emerging trends. *Computers*, 14.
- Albab, U., et al. (2025). Exploring artificial intelligence in Arabic language learning in higher education. *Thariqah Ilmiah: Jurnal Ilmu-Ilmu Kependidikan dan Bahasa Arab*.
- Aleedy, M., et al. (2025). A multi-agent chatbot architecture for AI-driven language learning. *Applied Sciences*.
- Alkaabi, M. H., & Almaamari, A. S. (2025). Generative AI implementation and assessment in Arabic language teaching. *International Journal of Online Pedagogy and Course Design*, 15(1), 1-18. <https://doi.org/10.4018/IJOPCD.368037>
- Alrashoudi, N. A., et al. (2025). Improving mispronunciation detection and diagnosis for non-native learners of the Arabic language. *Discover Computing*.
- Berliani, D., Azhari, I., & Taufiqurrochman, R. (2024). Exploring the potential of ChatGPT as a technology-based tool for Arabic language learning: Benefits and risks. *Kitaba*, 2(2), 62-69, <https://doi.org/10.35891/kitaba.v2i2.4695>
- Boulesnam, I., & Boucetti, R. (2025). Arabic language characteristics that make its automatic processing challenging. *International Arab Journal of Information Technology*, 22(4), 814-831, <https://doi.org/10.34028/iajit/22/4/18>
- El Zahraa, F. (2025). Leveraging artificial intelligence and digital technologies to enhance sociolinguistic competence and Arabic language skills. *Proceeding International Collaborative Conference on Multidisciplinary Science*.

- Fanoush, T., et al. (2025). Mispronunciation detection and diagnosis for young Arabic learners using transfer learning. IEEE Access.
- Faqeeh, M. H. A., et al. (2024). The students' awareness degree of the effectiveness of artificial intelligence applications in learning the Arabic language. Research Journal in Advanced Humanities, 5.
- Farghaly, A., & Shaalan, K. (2009). Arabic natural language processing: Challenges and solutions. ACM Transactions on Asian Language Information Processing, 8(4), 1-22, <https://doi.org/10.1145/1644879.1644881>
- Fredlina, S. F., et al. (2025). The role of ChatGPT as an AI-assisted learning tool in teaching Arabic to students at higher education institution Malang. Proceeding of International Conference on Islamic Education (ICIED).
- Garba, M. A., et al. (2024). Use of AI in learning Arabic language by non-Arabic speakers. International Journal of Research and Innovation in Social Science, 8.
- Hady, Y., et al. (2026). Artificial intelligence-based Arabic language learning: A systematic study of the development and challenges of pedagogical innovation. Online Learning in Educational Research (OLER).
- Hijriyah, U. (2025). How effective is SUNO.AI in enhancing Arabic listening skills? An evaluation of AI-based personalized learning. International Journal of Information and Education Technology.
- Kamal, M. H. (2026). Assessment of Arabic language skills using AI tools among gifted talented students. Malaysian Journal of Social Sciences and Humanities.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. RELC Journal, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Madwi, F. H. M. (2025). Integrating artificial intelligence in Arabic language education: Challenges and opportunities. Dzil Majaz: Journal of Arabic Literature.
- Mahmoud, M. H. (2018). A multiagents based intelligent tutoring system for teaching Arabic grammar. International Journal of Education.
- Mahmoud, M. H., et al. (2016). An intelligent tutoring system for teaching the grammar of the Arabic language. Journal of Electrical Systems and Information Technology, 3(1), 1-12, <https://doi.org/10.1016/j.jesit.2015.12.002>
- Mohammdj, H. M., et al. (2021). An intelligent system to help deaf students learn Arabic Sign Language. Interactive Learning Environments, 31(6), 3591–3608. <https://doi.org/10.1080/10494820.2021.1943689>

- Mohd Yusof, M. A., Ismail, N., & Wan Daud, W. A. A. (2025). Exploring Arabic language learning in the age of AI chatbots: A conceptual framework. *SIBAWAYH Arabic Language and Education*, 6(2).
- Moustafa, A. T. H., et al. (2026). Generative AI for learning and teaching Arabic: Opportunities, challenges, and ethical considerations. *IEEE Access*.
- Mufid, M., et al. (2025). Integrating artificial intelligence in speaking skill development: Teachers' perceptions at an Indonesian Islamic university. *Alsinatuna*.
- Muttaqin, C., et al. (2025). Students' perceptions of ChatGPT as a learning aid in Arabic language education at Universitas Islam KH. Ruhiat Cipasung. *Borneo Journal of Language and Education*.
- Nasaruddin, N. (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>
- Nasution, S., et al. (2025). Analyzing phonetic errors among non-native Arabic learners through artificial intelligence. *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*.
- Natheer, S. (2022). Agent-based intelligent tutoring system for Arabic grammar. *Journal of King Abdulaziz University: Computing and Information Technology Sciences*.
- Ni'mah, M. (2026). Artificial intelligence-driven role-playing communicative learning: A transformative model to enhance learners' speaking skills in an Arabic language course. *Al-Muyassar: Journal of Arabic Education*.
- Oluwatoyin, A. (2026). Integrating artificial intelligence into Arabic language education: Pedagogical strategies and learning outcomes. *University of Arusha Academic Journal*.
- Rafi'atun, R., et al. (2025). The role of ChatGPT in enhancing academic Arabic writing skills among students of Arabic language education. *LISANIA: Journal of Arabic Education and Literature*.
- Ridho, R., Perkasa, M. R., & Maujud, F. (2025). Exploring the application of artificial intelligence in Arabic language learning: Perspectives of students and teachers. *Academia Open*, 10(2). <https://doi.org/10.21070/acopen.10.2025.12886>
- Ridwan, M., Jezy, A., Hawary, M. S. R., Apip, M., & Efendi, A. (2025). Pemerolehan bahasa Arab di era artificial intelligence: Studi pustaka tentang potensi, strategi, dan tantangan pembelajaran berbasis AI. *Lisanul Arab: Journal of Arabic Learning and Teaching*, 14(2), 148-160. <https://doi.org/10.15294/laa.v14i2.24233>

- Sa'diyah, K. K. H., et al. (2025). The utilizing Suno.Ai technology to enhance Arabic listening skills in foreign language education. *Journal of Arabic Language Teaching*.
- Sa'idah, M. A., et al. (2024). Enhancing Arabic language teaching through artificial intelligence: Assessing effectiveness and educational implications. In 2024 3rd International Conference on Creative Communication and Innovative Technology (ICCIT).
- Samin, S. M., et al. (2024). Integrating artificial intelligence into the Arabic language teaching plan at higher education. *SHS Web of Conferences*.
- Samiya, R. (2025). Artificial intelligence in Arabic language education: Current applications, challenges, and future perspectives. *Studies in Education Sciences*.
- Sarif, S., et al. (2024). Efektivitas artificial intelligence text to speech dalam meningkatkan keterampilan membaca. *Jurnal Naskhi: Jurnal Kajian Pendidikan dan Bahasa Arab*.
- Setyawan, C. E., et al. (2025). Role of AI assistance in academic writing: A case study of Arabic language students from prompt to publication. *Alsinatuna*.
- Shalihah, S., Fradana, H., & Aljabr, B. N. (2025). Traditional method vs AI chatbot-assisted in Arabic learning. *LISANIA: Journal of Arabic Education and Literature*, 9(1), 1-14. <https://doi.org/10.18326/lisania.v9i1.1-14>
- Shao, S., et al. (2022). AI-based Arabic language and speech tutor. In 2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA).
- Simon, A. S. (2023). Prospek pembelajaran bahasa Arab di era generative artificial intelligence. *Assuthur: Jurnal Pendidikan Bahasa Arab*.
- Sismat, A., Abdullah, R., Abdul Kahar, K., Abdullah, R., & Jarudin, A. (2025). Exploring the impact of ChatGPT on non-native learners of Arabic language: Challenges and opportunities. In V. P. H. Pham, A. Lian, & J. White (Eds.), *Empowering educators: Integrating AI tools for personalized language instruction* (pp. 281-305). Springer Nature. https://doi.org/10.1007/978-3-032-01348-4_12
- Sofian, G. Y., et al. (2026). The effectiveness of AI based microlearning platforms through vlog content to improve Arabic listening skills. *Arabiyatuna: Jurnal Bahasa Arab*.
- Sumiarni, N., et al. (2026). Transforming Arabic language learning with artificial intelligence: Teacher perspectives and student experiences. *Ijaz Arabi: Journal of Arabic Learning*, 9(1), 237-252. <https://doi.org/10.18860/ijazarabi.v9i1.35113>
- Supriatna, A., et al. (2025). Enhancing Arabic language proficiency among Islamic university students in Indonesia through artificial intelligence in the digital

- age. Tadris Al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban.
- Syed, T. A., et al. (2025). Revolutionizing Arabic learning: GNT and AI-enhanced Metaverse environments. *Frontiers in Computer Science*.
- Tolba, R., et al. (2024). Interactive augmented reality system for learning phonetics using artificial intelligence. *IEEE Access*.
- Yatri, D., et al. (2023). Artificial intelligence (AI) in language learning (English and Arabic class): Students' and teachers' experience and perceptions. *Transformational Language Literature and Technology Overview in Learning (TRANSTOOL)*.
- Yuliani, S. Y., et al. (2025). Integration of AI-based text-to-speech technology in Arabic listening skills learning. *Aphorisme: Journal of Arabic Language, Literature, and Education*.
- Zaimah, N. R., et al. (2023). Acceptability and effectiveness analysis of large language model-based artificial intelligence chatbot among Arabic learners. *Mantiqu Tayr: Journal of Arabic Language*, 4(1), 7. <https://doi.org/10.25217/mantiquatayr.v4i1.3951>
- Zubair, K. (2026). Artificial intelligence in Arabic language education: Applications, perceptions, and pedagogical integration. *Journal of Digital Learning and Distance Education*.