Journal of Practice Learning and Educational Development
https://digitalpress.gaes-edu.com/index.php/jpled
E-ISSN: 2809-1205
Vol 6, No. 2 (2026) 2187-2196
DOI: 10.58737/jpled.v6i2.1698

Submitted: 13-03-2026
Revised: 23-04-2026
Accepted: 26-05-2026

Claude in Arabic Language Learning: Critical AI Literacy, Self-Regulation, and Process-Oriented Pedagogy

Ahmad Aziz Fuadi[1], Dul Saiin[2], Muhammad Fatkul Munir[3]
STAI Ma'arif Kendal Ngawi[1,2,3]
E-mail: fuadiaaf@gmail.com[1]

Abstract

Generative artificial intelligence in Arabic language education requires learning designs that preserve student agency rather than merely accelerate answer production. This article formulates a pedagogical model for using Claude as a bounded cognitive scaffold in university-level Arabic learning, with emphasis on critical AI literacy, self-regulated learning, and process-oriented assessment. The study applies an integrative literature review combined with an affordance analysis of Claude. Sources were purposively selected from 2023-August 2026 publications on generative AI, language learning, Arabic education, assessment, and human-AI collaboration, supplemented by foundational Arabic linguistics and pedagogy and official Anthropic documentation. The synthesis identifies five principles: source-grounded reading, layered grammatical verification, reason-based writing revision, a PLAN-ASK-TEST-REVISE-REFLECT self-regulation cycle, and assessment of process traces. Claude is most useful when it supports text interrogation, comparison of linguistic alternatives, and feedback that students must verify. Its risks include fluent but inaccurate output, cognitive offloading, unequal paid access, and opaque AI assistance. The resulting framework positions human verification, task design, and accountable linguistic reasoning as the conditions for pedagogically responsible use.

Keywords: Claude, Critical Ai Literacy, Arabic Language Learning, Self-Regulated Learning, Process-Oriented Assessment

Introduction

Debates about generative artificial intelligence in higher education are moving beyond the simple question of whether students should be allowed to use AI. A more consequential question is what kinds of learning activities still require students to read, compare, test, and write with their own academic responsibility when automated assistance is continuously available. This shift is particularly important in Arabic language education. Linguistic competence is not developed merely by receiving sentences that appear correct on the surface. It is built through recognizing patterns, understanding relations between form and meaning, testing usage in context, and repeatedly revising errors. Generative language models therefore need to be embedded in pedagogical designs that deliberately preserve students' cognitive work rather than simply accelerate task completion.

Claude is a paid plan for the Claude conversational AI environment. According to Anthropic's documentation, the Pro plan provides higher usage capacity than the free service, priority access during periods of high traffic, model selection, and features such as Projects and knowledge bases (Anthropic, 2026a). These characteristics make the service relevant to academic activities involving long texts, collections of course materials, writing drafts, and multi-step conversations. Yet broader functionality

does not automatically create pedagogical value. The easier it becomes to upload materials and receive coherent responses, the more important it is to define what students must read independently, what AI may assist with, and what students must personally justify.

Anthropic presents Claude for Education as a tool intended to support exploration, understanding, and responsible AI use rather than simply provide answers (Anthropic, 2026b). This orientation is consistent with a growing emphasis on human-in-the-loop learning. Saarela, Gunasekara, and Kumarage (2026), in a meta-analysis of 51 studies, report positive effects of generative AI on language proficiency and affective-cognitive outcomes, but the magnitude and form of the effects vary according to context, learner characteristics, targeted skills, and intervention design. The pedagogical question is therefore not only whether AI can be effective, but for which tasks, under what verification mechanisms, and for which learners.

Research on Arabic language learning in Indonesia shows a similar movement from general digital support toward more specific assistance for reading, writing, translation, listening, and instructional design. Nasaruddin (2024) demonstrates that generative AI can support the preparation of Arabic texts, questions, and teaching materials, while Bustam, Mahmudah, and Hidayat (2024) argue that technological innovation becomes educationally meaningful only when it is connected to clear pedagogical needs and media design. Mustar et al. (2025) further emphasize the importance of bridging established Arabic pedagogy with AI proportionally. Such findings are directly relevant to Claude because Arabic morphology, i'rab, gender and number agreement, register variation, and culturally embedded expressions continue to require human verification.

Arabic-specific quality concerns are illustrated by Allaithy and Zaki (2025), who evaluated AI-generated Arabic reading-comprehension materials and identified patterns of linguistic unnaturalness, including structural dominance, verb and pronoun problems, unsystematic lexical repetition, and influence from English structures. Their findings support a fundamental distinction: linguistic fluency in an AI response should not be equated with linguistic naturalness or accuracy. This distinction is especially important for learners whose current competence may be insufficient to detect subtle errors in grammatically complex or culturally marked Arabic.

Arabic linguistic scholarship provides an additional reason not to judge AI output solely by surface fluency. Tammam Hassan conceptualizes Arabic structure through the relation of mabna, ma'na, and contextual qara'in, whereas Ahmad Mukhtar 'Umar demonstrates that meaning depends on semantic relations and the environment in which words are used (Hassan, 1998; 'Umar, 1993). These perspectives imply that apparently fluent AI-generated sentences still need to be evaluated at morphological, syntactic, semantic, and contextual levels. In pedagogy, Rushdi Ahmad Tu'aymah relates learning objectives, learner needs, language skills, materials, methods, and culture, while 'Ali Ahmad Madkur emphasizes directed and staged practice across the arts and skills of Arabic (Tu'aymah, 1989; Madkur, 2008). Technology should therefore remain subordinate to the goals of Arabic language learning.

Recent work also highlights the distribution of responsibility between humans and AI. Anders and Speltz (2025) show that generative AI literacy can be developed through planning, iteration, evaluation, and human-in-the-loop practice. In Arabic writing, Zubaidi et al. (2025) propose a tridimensional human-AI collaboration framework in which humans remain conceptualizers, validators, and executors while AI serves an assisting function. This division of roles is compatible with self-regulated learning because students are expected to plan their use of AI, inspect its output, and consciously make final decisions. Salhab and Aboushi (2025) likewise connect AI literacy and broader twenty-first-century skills with students' acceptance of generative AI, indicating that competent adoption involves more than technical familiarity.

The same affordances that make generative AI attractive may also create cognitive offloading. In writing tasks, students may move from using AI to identify a problem to asking it to generate a complete paragraph; in translation, they may accept a polished rendering without examining the source; and in grammar, they may adopt an i'rab analysis without testing the rule. Zubaidi et al. (2025) and

Mustar et al. (2025) both underline the need for human validation and proportional integration. Consequently, prompts such as 'write this for me,' 'translate this,' or 'give the answer' should not by themselves be regarded as instructional innovation. Pedagogical value emerges when students continue to compare, reject, revise, and justify.

Assessment also needs to change when polished AI-generated products can be obtained with minimal visible effort. If instructors assess only the final text, high-quality language may receive a high grade even when the student cannot explain it. By contrast, assessment that includes an initial draft, revision notes, justification of linguistic choices, source verification, and oral explanation makes the learning process visible. Such an approach is particularly important in Arabic because competence is demonstrated not only by producing a correct form, but by being able to explain why that form is appropriate in relation to grammar, meaning, register, and context.

Against this background, this article does not treat Claude primarily as an automated tutor or conversational partner. It conceptualizes Claude as a bounded cognitive scaffold: a workspace that can expand access to explanations, alternatives, and feedback while treating every AI output as a linguistic hypothesis that must be tested. The study therefore aims to identify Claude affordances relevant to Arabic learning, synthesize linguistic and cognitive risks that require control, formulate procedures that preserve verification and self-regulation, and design process-oriented assessment that makes AI use transparent. Its main contribution is a five-principle integration framework and a PLAN-ASK-TEST-REVISE-REFLECT cycle applicable to reading, grammar and morphology, translation, and writing.

Methods

This study employed a qualitative integrative literature review combined with a technology-affordance analysis. The integrative design was selected because the purpose was not to calculate the experimental effect size of Claude, but to synthesize findings across generative AI, language learning, Arabic language pedagogy, self-regulation, linguistic quality, and assessment into a coherent pedagogical framework. Affordance analysis was used to identify possible learning actions enabled by Claude features while explicitly distinguishing these possibilities from claims of effectiveness that have not been directly tested with Arabic-language students.

The source corpus was assembled through targeted searches of journal websites and academic publisher pages for work published from 2023 through August 2026. Priority topics were generative AI in education and language learning, AI-assisted Arabic learning, reading, writing, translation, listening, grammar, assessment transformation, human-AI collaboration, and AI literacy. Indonesian literature was prioritized from accredited Sinta 1 and Sinta 2 journals represented in the source manuscript, including Izdihar, Indonesian Journal of Applied Linguistics, Arabiyat, Arabiyatuna, and LISANIA. To keep the analysis rooted in Arabic studies, four foundational works by Tammam Hassan, Ahmad Mukhtar 'Umar, Rushdi Ahmad Tu'aymah, and 'Ali Ahmad Madkur were included. Official Anthropic documentation was used only to describe Claude characteristics and educational positioning, not as evidence of learning effectiveness.

Sources were included when they offered directly relevant evidence, conceptual frameworks, or documented product characteristics. Popular opinion pieces, promotional materials without academic metadata, and sources that could not be linked to the analytical questions were excluded. The review procedure consisted of four stages. First, each source was read for meaning units related to AI functions, learning activities, risks, and control mechanisms. Second, these units were coded using categories such as source-grounding, contrastive checking, revision loop, cognitive offloading, self-monitoring, process evidence, teacher mediation, and access equity. Third, related codes were clustered into themes specifically applicable to Arabic language learning. Fourth, the themes were converted into design principles and conceptually tested against four common task types: reading Arabic texts, nahwu-sharf analysis, translation, and writing.

Conceptual triangulation was used to strengthen the synthesis. Claims about generative AI benefits were compared with literature on dependency and cognitive offloading; general language-

learning findings were compared with studies evaluating Arabic-specific linguistic quality; and assessment recommendations were checked against research on human validation, transparency, and institutional responsibility. The study did not collect interviews, focus-group data, classroom observations, or student achievement scores. Accordingly, the term 'results' in this article refers to analytical synthesis rather than field or experimental outcomes. This limitation is stated explicitly so that the proposed model is not interpreted as direct evidence that Claude has already improved student competence in a particular institution.

Results and Discussions

1. Results

The integrative synthesis and affordance analysis produced five principles for using Claude in Arabic language learning. The principles are organized according to the learning work that should remain visible: students begin from sources, test linguistic forms, revise with reasons, regulate the sequence of assistance, and present evidence of process during assessment. The resulting model therefore places students, Arabic texts, curricular objectives, and accountable academic judgment at the center; Claude remains a supporting environment rather than the center of instruction.

a. Source-Grounded Reading and Translation

The first result is a source-grounded use pattern. Claude features such as Projects and knowledge bases can organize course texts, glossaries, validated examples, assignment criteria, and related materials in a shared working context (Anthropic, 2026a). The pedagogical value lies not in automatic summarization but in constraining inquiry to selected learning materials. A reading task can therefore require students to annotate a text before using AI, ask targeted questions about specific clauses or vocabulary, and then verify the response against dictionaries, grammatical references, commentaries, or instructor explanations.

This pattern also changes the role of translation. Instead of asking Claude for a final translation, students first produce a provisional translation and then use AI to identify a limited number of ambiguous points or alternative lexical choices. Students compare the options and justify the final selection. In this design, the system expands the space of alternatives while semantic judgment remains with the learner. Table 1 maps the main Claude affordances to recommended Arabic tasks, predictable risks, and pedagogical controls.

Table 1.
Mapping Claude affordances to Arabic language tasks

Affordance	Recommended task	Main risk	Pedagogical control
Projects / knowledge bases	Source-grounded qira'ah using texts and glossaries	Summaries replace actual reading	Require initial annotation before AI use
Multi-step dialogue	I'rab analysis and structural comparison	Fluent answers are assumed to be correct	Request alternatives and verify with references
Text feedback	Staged kitabah revision	Students accept every correction	Require a reason for each accepted revision
Variation generation	Vocabulary, sentence-pattern, and transformation exercises	Examples may be unnatural or register-inappropriate	Provide instructor criteria and validated examples

Higher Pro usage capacity	Repeated exploration and longer projects	Overuse and unequal paid access	Keep core tasks possible without paid access

b. Layered Grammatical Verification

The second result is a three-layer verification procedure for nahwu and sharf. Layer one is the student's own analysis: identifying word classes, syntactic relations, provisional i'rab, roots, or morphological patterns. Layer two is AI critique: Claude is asked to identify the most questionable points in the student's analysis rather than generate a complete analysis from the beginning. Layer three is academic adjudication: the student checks rules and evidence in traceable Arabic references and records the final decision.

The procedure is designed to turn AI answers into testable hypotheses. Prompts can be structured to expose uncertainty, for example by asking for two possible i'rab analyses when the context allows them, the conditions for each analysis, and the additional contextual information needed to decide between them. The same logic can be applied to morphology by asking for alternative roots or patterns and requiring dictionary verification. The outcome of the task is therefore evaluated by the quality of the student's decision and evidence, not by the quantity of AI-generated analysis.

c. Reason-Based Writing Revision

The third result is a revision-studio model for kitabah. The initial draft is written without text generation so that the student's actual starting competence remains visible. Claude is then used for limited diagnosis, such as cohesion, agreement, word choice, or potentially literal transfer from Indonesian. Students decide which feedback to accept or reject and provide a brief reason for each decision. Revision is completed by the student, not by accepting a fully rewritten paragraph.

The final stage compares the initial and revised drafts and requires a short reflection on the most important changes, rejected AI suggestions, and the basis for the final linguistic choices. Table 2 summarizes the five-stage revision sequence and the evidence collected at each stage.

Table 2.
Reason-based kitabah revision workflow with Claude

Stage	Student activity	Claude role	Evidence collected
1. Initial draft	Write using current ability	Not used for text generation	Draft version 1
2. Diagnosis	Define a specific problem focus	Identify limited problem areas	Prompt and response
3. Decision	Accept or reject suggestions with reasons	Clarify feedback when requested	Decision notes
4. Revision	Rewrite independently	Compare versions when needed	Draft versions 2 / 3
5. Reflection	Explain learning and AI boundaries	No final-answer generation	150-200 word reflection

d. PLAN-ASK-TEST-REVISE-REFLECT Self-Regulation Cycle

The fourth result is a five-phase self-regulation cycle. PLAN requires students to define the learning objective, the point of difficulty, and an indicator of success before opening Claude. ASK requires the smallest useful form of assistance rather than a request for a final product. TEST requires comparison of the AI response with the source text, dictionary, grammar, instructor example, or

previously learned knowledge. REVISE requires the student to make a decision after verification. REFLECT records what changed in understanding, which errors were nearly accepted, and which strategy should be used next.

The sequence intentionally interrupts the prompt-output-submit pattern. It can be simplified into a worksheet for beginners or extended to classical-text analysis and academic translation for advanced students. Table 3 presents the phases, guiding questions, Arabic-learning examples, and observable indicators of self-regulation.

Table 3.
PLAN-ASK-TEST-REVISE-REFLECT cycle for Arabic learning

Phase	Key question	Arabic-learning example	Self-regulation indicator
PLAN	What is my learning objective?	Decide whether the focus is meaning, i'rab, or style	A specific objective exists before AI use
ASK	What is the minimum assistance I need?	Ask about one construction rather than request a full translation	The prompt is narrow and purposeful
TEST	What evidence supports this response?	Compare with a dictionary, grammar reference, or parallel text	A verification source is documented
REVISE	What is my decision after verification?	Change or retain the analysis with reasons	The student makes the final decision
REFLECT	What changed in my understanding?	Record errors, strategies, and AI limitations	Metacognitive reflection is present

e. Process-Oriented Assessment and Staged Implementation

The fifth result is an assessment design that treats process traces as evidence of competence. In qira'ah, students can submit pre-AI annotations, questions asked, problematic responses, and their corrections. In grammar, they can defend an i'rab analysis orally. In kitabah, the progression from draft to revision is assessed. In translation, students can present two alternatives and justify the final choice. A concise AI-use declaration records the tool, purpose, and type of assistance. This approach does not prohibit AI; it makes AI assistance visible and prevents the final product from becoming the only evidence of learning.

The model can be implemented in four stages. First, students learn that AI output is a hypothesis by comparing conflicting explanations. Second, Claude is restricted to one function per activity: clarification in reading, critique in grammar, diagnosis in writing, and alternatives in translation. Third, students complete a small project using selected source materials and document key AI interactions. Fourth, summative assessment combines AI-permitted and AI-free tasks so that instructors can observe both critical collaboration with technology and independent Arabic competence.

2. Discussions

a. Claude as a Bounded Cognitive Scaffold

The results support a reframing of Claude from automated tutor to bounded cognitive scaffold. This interpretation is consistent with research showing that technology becomes pedagogically valuable through instructional positioning rather than technological sophistication. Arifin et al. (2023) found that smart technology can make Arabic learning easier and more interactive but also produces positive and negative consequences that require management. Bustam et al. (2024) similarly connect innovation to deliberate instructional design. These studies do not establish Claude

effectiveness directly, but they support the principle that affordances must be converted into bounded learning actions.

A bounded scaffold also addresses a central problem identified in generative AI research: fluency can conceal error. Allaithy and Zaki (2025) show that AI-generated Arabic reading materials may remain linguistically unnatural despite coherent presentation. Alkaabi and Almaamari (2025) likewise emphasize implementation and assessment considerations in Arabic language teaching. Accordingly, Claude should not occupy the epistemic position of a dictionary, grammar, edited text, or expert teacher. Its output is most useful when it gives learners something to interrogate rather than something to submit.

b. Arabic-Specific Verification and Linguistic Authority

The layered verification result is particularly relevant to Arabic because i'rab and morphological interpretation are frequently context-sensitive. Maulidiya, Abdurrahman, and Saleh (2025) compared generative AI systems in Arabic grammar analysis and found differences in model performance while maintaining the need for manual correction. This finding reinforces the proposed three-layer procedure: students first analyze, AI then critiques, and academic sources adjudicate. The procedure preserves retrieval practice and provides a visible record of why one analysis was selected over another.

The theoretical works of Hassan (1998) and 'Umar (1993) deepen this argument. If structure depends on interacting linguistic clues and meaning depends on semantic relations and context, a single confident AI label is insufficient evidence. The learner must identify the qara'in, structural relations, lexical meanings, and contextual constraints that justify an interpretation. The same principle applies to translation. Rozak, Umbar, and Albantani (2024) show that AI can expand possibilities in Arabic translation learning, but critical attention to lexical equivalence and context remains necessary. In this sense, the educational value of AI lies in multiplying hypotheses while human reasoning determines which hypothesis is defensible.

Reading and listening studies also suggest that AI should be tied to a specific maharah rather than generalized as an all-purpose language authority. Salam et al. (2023) report the use of AI to support reading skills, while Khuluq et al. (2025) combine AI-supported reading materials with structured instructional facilitation. Baharun et al. (2026) describe AI-assisted transcription as scaffolding for Arabic listening comprehension, and Sofian et al. (2026) report the effectiveness of AI-based microlearning when paired with vlog content and an instructional design. Nasution et al. (2025) use AI to analyze phonetic errors among non-native Arabic learners. Together, these studies suggest that AI effectiveness is function- and skill-specific, which supports restricting Claude to clearly defined roles in each learning activity.

c. Human-AI Collaboration, Self-Regulation, and Kitabah

The reason-based writing model is supported by recent Arabic writing research. Qomariah et al. (2025) report that generative AI can help students obtain ideas, translate, and develop academic Arabic writing, but they also identify concerns related to accuracy, cultural context, source verification, and plagiarism. Zubaidi et al. (2025) more explicitly retain human authority by assigning the human the roles of conceptualizer, validator, and executor. These findings justify a design in which Claude diagnoses limited problems and students remain responsible for rewriting and explaining the final text.

The PLAN-ASK-TEST-REVISE-REFLECT cycle also extends human-in-the-loop research into a memorable classroom procedure. Anders and Speltz (2025) connect generative AI literacy with planning, iteration, evaluation, and self-regulation. Salhab and Aboushi (2025) show that AI literacy and broader skills shape student acceptance of generative AI. The proposed cycle therefore defines critical AI literacy not as sophisticated prompting alone, but as the ability to decide when AI is needed, minimize the scope of assistance, test the result, make an independent decision, and reflect on the process. A technically complex prompt that merely outsources thinking would not satisfy this definition.

This interpretation also aligns with Arabic pedagogy. Tu'aymah (1989) and Madkur (2008) place language skills, materials, methods, exercises, and learner needs within an intentional pedagogical relationship. The technology should therefore be selected according to the maharah and learning objective. Clarification may be appropriate in qira'ah, diagnostic feedback in kitabah, alternative analyses in nahwu, and contrastive choices in tarjamah. The model resists using one AI feature as a universal solution for all aspects of Arabic learning.

d. Assessment Transparency, Academic Responsibility, and Equity

Process-oriented assessment responds to the decreasing reliability of take-home products as sole evidence of competence. Mustar et al. (2025) and Zubaidi et al. (2025) both emphasize human mediation and validation in AI-supported Arabic learning. Ravi et al. (2025) further show that university staff and students may not share identical understandings of competent and ethical AI use. Explicit assessment rules are therefore necessary. A declaration such as 'Claude was used to identify possible pronoun disagreement and suggest alternative diction; the final decisions and revisions were made by the author' is more informative than a generic statement that AI was used.

Rubrics should likewise reward reasoning that remains visible. A course may allocate marks to final language quality, linguistic justification, revision quality, source verification, and reflection rather than placing the entire grade on the polished product. The exact percentages should remain course-specific, but the principle is stable: evidence of thinking must carry meaningful assessment weight. Oral defense and spontaneous tasks can complement written products when instructors need evidence that a student can reproduce or explain competence without AI assistance.

Equity is a further consideration because Claude is a paid service. A required assignment that can only be completed through Pro-specific capacity would disadvantage students without subscriptions. The results therefore support two options: paid features remain optional, or the institution provides access when they are essential to the curriculum. Core learning outcomes should remain achievable through academic sources and ordinary course resources. This ensures that technological innovation does not create a new access hierarchy within the classroom.

e. Linguistic, Epistemic, Cognitive, and Institutional Limits

Four limitations define the boundaries of responsible use. The linguistic limit concerns naturalness, register, dialect, pragmatics, and ambiguous structures. Arabic generated by a model may be fluent in one context and inappropriate in another. The epistemic limit concerns authority: a model generates responses from patterns and available context, whereas academic claims require traceable evidence. Students should therefore distinguish between a plausible explanation and a claim supported by a dictionary, grammar, edited text, or scholarly publication.

The cognitive limit arises when convenience removes productive struggle. Saarela et al. (2026) find that generative AI can support language outcomes, but its benefits depend on context and design. Nasaruddin (2024) and Mustar et al. (2025) likewise position AI as support rather than a replacement for pedagogy. Some parts of language learning should remain deliberately effortful: inferring meaning from context, attempting sentence construction, retrieving a rule, and making errors before receiving feedback are all part of competence development.

The institutional limit concerns cost, privacy, account management, data handling, and academic policy. Students should not be required to upload personal data, grades, sensitive documents, or restricted manuscripts without clear institutional guidance. Policies should distinguish between individual practice, assessed coursework, research, and sensitive data processing. Finally, Claude should not be used as the sole high-stakes evaluator. Model outputs may differ across versions and prompts, and Arabic register variation makes human oversight even more important. AI can support formative diagnostic comments, but academic grading decisions should remain with instructors using transparent rubrics.

Conclusion

This integrative analysis indicates that Claude is most defensible in Arabic language education when it is treated as a bounded cognitive scaffold rather than an automated linguistic authority. The synthesis

produced five design principles: source-grounded reading, layered grammatical verification, reason-based writing revision, self-regulation through the PLAN-ASK-TEST-REVISE-REFLECT cycle, and process-oriented assessment. Together, these principles deliberately interrupt the prompt-output-submit pattern and redirect attention from the apparent fluency of an AI product to the quality of students' linguistic judgment. Claude can support qira'ah, nahwu-sharf, tarjamah, and kitabah when instructors require verification, provide traceable references, restrict AI functions to defined stages, and assess process evidence. Its unstructured use may instead encourage cognitive offloading, dependency, unequal access, and academic products that do not represent actual student competence. Arabic linguistic and pedagogical scholarship remains the authority for evaluating form, meaning, skill development, and context. Because this study is an integrative review rather than a field experiment, the framework should next be tested empirically. Future research may compare structured-cycle classes with less regulated AI use and examine grammatical verification, writing development, self-regulation, and students' ability to defend linguistic decisions orally.

References

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

Allaithy, A., & Zaki, M. (2025). Evaluation of AI-generated reading comprehension materials for Arabic language teaching. Computer Assisted Language Learning, 1-33. https://doi.org/10.1080/09588221.2025.2474037

Anders, A. D., & Speltz, E. D. (2025). Developing generative AI literacies through self-regulated learning: A human-centered approach. Computers and Education: Artificial Intelligence, 9, 100482. https://doi.org/10.1016/j.caeai.2025.100482

Anthropic. (2026a). What is the Pro plan? Anthropic Help Center.

Anthropic. (2026b). Claude for Education. Anthropic.

Arifin, Z., Desrani, A., Ritonga, A. W., & Ibrahim, F. M. A. (2023). Arabic language learning approach using smart technology in higher education. Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature, 6(1), 1-12. https://doi.org/10.22219/jiz.v6i1.25011

Baharun, S., Hanifansyah, N., Habib, M. T., Mahmudah, M., & Batrisya, N. A. (2026). Exploring AI-assisted transcription as scaffolding in Arabic listening comprehension. Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature, 9(1), 57-76. https://doi.org/10.22219/jiz.v9i1.42312

Bustam, B. M. R., Mahmudah, K. N. L., & Hidayat, K. (2024). Arabic language learning innovation based on technology utilization. Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban, 11(1), 72-88. https://doi.org/10.15408/a.v11i1.34876

Hassan, T. (1998). Al-lughah al-'Arabiyyah: Ma'naha wa mabnaha [The Arabic language: Its meaning and structure]. 'Alam al-Kutub.

Khuluq, M., Ainin, M., Rosyidi, A. W., & Imamah, N. (2025). The development of reading skill teaching materials based on Prezi Artificial Intelligence. Arabiyatuna: Jurnal Bahasa Arab, 9(1), 327-346. https://doi.org/10.29240/jba.v9i1.10886

Madkur, 'A. A. (2008). Tadris funun al-lughah al-'Arabiyyah [Teaching the arts of the Arabic language]. Dar al-Fikr al-'Arabi.

Maulidiya, A. R., Abdurrahman, M., & Saleh, N. (2025). Exploring AI capabilities in Arabic grammar: Comparative analysis of ChatGPT and Gemini. Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban, 11(2), 160-174. https://doi.org/10.15408/a.v11i2.42671

Mustar, M., Raswan, R., Chabib Uluum, D., Tamami, K., & Bin Mohamad, M. R. (2025). Bridging traditional language pedagogy and AI: Lessons from Arabic language programs in Indonesia. Arabiyatuna: Jurnal Bahasa Arab, 9(2), 770-785. https://doi.org/10.29240/jba.v9i2.13708

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., Asari, H., Al-Rasyid, H., & Khalilah, Z. (2025). Analyzing phonetic errors among non-native Arabic learners through artificial intelligence. Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban, 12(2), 229-242. https://doi.org/10.15408/a.v12i2.47342

Qomariah, R. N., Awad, A., Fikri, S., & Lazreque, S. (2025). The role of ChatGPT in enhancing academic Arabic writing skills among students of Arabic Language Education. LISANIA: Journal of Arabic Education and Literature, 9(1), 152-173. https://doi.org/10.18326/lisania.v9i1.152-173

Ravi, M., Kaur, K., Wright, C., Bawn, M., & Cutillo, L. (2025). University staff and student perspectives on competent and ethical use of AI: Uncovering similarities and divergences. International Journal of Educational Technology in Higher Education, 22, 55. https://doi.org/10.1186/s41239-025-00557-7

Rozak, A., Umbar, K., & Albantani, A. M. (2024). The model of Arabic learning translation using artificial intelligence in social-media. Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban, 11(1), 89-99. https://doi.org/10.15408/a.v11i1.39757

Saarela, M., Gunasekara, S., & Kumarage, P. (2026). A meta-analysis of generative AI effects on language proficiency and affective-cognitive outcomes in language learning. Discover Computing, 29, 116. https://doi.org/10.1007/s10791-026-10015-1

Salam, M. Y., Mahayuni, M., Taman, M., & Mudinillah, A. (2023). Using artificial intelligence for education in the Education 5.0 era to improve reading skills. Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban, 10(2), 149-162. https://doi.org/10.15408/a.v10i2.35487

Salhab, R., & Aboushi, M. M. (2025). Influence of AI literacy and 21st-century skills on the acceptance of generative artificial intelligence among college students. Frontiers in Education, 10, 1640212. https://doi.org/10.3389/feduc.2025.1640212

Sofian, G. Y., Sutaman, Taufiqurrochman, R., Hidayat, A. F. S., & Nukman, N. (2026). The effectiveness of AI based microlearning platforms through vlog content to improve Arabic listening skills. Arabiyatuna: Jurnal Bahasa Arab, 10(1), 39-62. https://doi.org/10.29240/jba.v10i1.16265

Tu'aymah, R. A. (1989). Ta'lim al-'Arabiyyah li-ghayr al-natiqin biha: Manahijuhu wa asalibuhu [Teaching Arabic to non-native speakers: Its curricula and methods]. Islamic Educational, Scientific and Cultural Organization (ISESCO).

'Umar, A. M. (1993). 'Ilm al-dalalah [Semantics]. 'Alam al-Kutub.

Zubaidi, A., Munip, A., Widodo, S. A., & Zerrouki, T. (2025). Enhancing Arabic writing skills using Chat GPT-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