

From Recognition to Production: Diagnosing Arabic Language Skill Imbalances among Non-Native Learners at the Islamic University of Maldives through a Task-Based Framework

Mohamed Uwais Mohamed Ameen

Islamic University of Maldives

DOI: <https://doi.org/10.51244/IJRSI.2026.1309000016>

Received: 08 September 2026; Accepted: 21 September 2026; Published: 29 September 2026

ABSTRACT

This paper presents a theoretically grounded pedagogical proposal informed by a directional diagnostic pattern among non-native undergraduate learners at the Islamic University of Maldives (IUM). The underlying project involved a quantitative phase with 120 learners selected through convenience sampling and a qualitative phase with 30 purposively selected participants. It used pre- and post-tests across listening, speaking, reading, and writing, together with classroom observations, questionnaires, and semi-structured interviews; the questionnaire was reported to have acceptable internal consistency (Cronbach's $\alpha = .87$). The available records indicate an uneven four-skill profile: reading appears comparatively stronger, while listening and speaking emerge as the most pronounced areas of difficulty; writing is also constrained by weak independent sentence construction, vocabulary retrieval, grammatical accuracy, cohesion, and organization. Because complete numerical and qualitative outputs are not available in the current records, the paper treats this pattern as directional rather than as a complete statistical account. Drawing on recent research from 2018–2026, it proposes a six-stage TBLT cycle integrating activation, input, language focus, collaborative task performance, individual output, feedback, and re-performance. The framework prioritizes listening, oral interaction, structured sentence production, guided-to-independent writing, pronunciation, and vocabulary activation while using reading as scaffolding rather than as the dominant instructional endpoint. The paper contributes a context-responsive pedagogical model and identifies the empirical evidence required for rigorous future evaluation.

Keywords: Arabic as a foreign language; Islamic University of Maldives; listening; speaking; writing; sentence construction; Task-Based Language Teaching; communicative competence

INTRODUCTION

Arabic occupies an important position in Islamic higher education because it functions not only as a language of communication but also as an academic and disciplinary resource for the study of Islamic texts. At IUM, non-native learners are expected to read, understand, discuss, and write in Arabic for academic, religious, and professional purposes. Such expectations require more than declarative knowledge of vocabulary and grammar. They require learners to process spoken input in real time, retrieve linguistic resources rapidly, construct well-formed sentences, sustain interaction, and organize written ideas. Recent Arabic-language research emphasizes that communicative competence depends on coordinated development across listening, speaking, reading, and writing rather than isolated mastery of forms (Aljanadbah & Al Sharef, 2024; Almelhes, 2024; Damsing & Che Haron, 2026).

The diagnostic concern motivating the present study is that the four skills do not appear to develop equally. The present study indicates that learners appear comparatively stronger in reading but experience greater difficulty when required to understand connected speech, respond spontaneously, retrieve vocabulary, construct sentences, or produce independent writing. This pattern is pedagogically important because receptive recognition can mask limitations in usable language. A student may identify a word or sentence when it remains visible on the page yet fail to recognize the same item in continuous speech or retrieve it

quickly enough for speaking. Modality research similarly shows that spoken and written L2 production place different lexical and phraseological demands on learners (Kim & Crossley, 2023).

Arabic presents additional complexity because learners may encounter variation between Modern Standard Arabic and spoken varieties, creating sociolinguistic demands beyond grammar and vocabulary alone. Recent work on L2 Arabic shows that bidialectal and multidialectal practices can shape learners' pragmatic development and communicative repertoires (Al Masaeed, 2020, 2022; Nassif & Al Masaeed, 2022). Diglossia may also affect how learners perceive, access, and practise spoken Arabic (Hashem, 2022). Therefore, a curriculum that privileges reading and explanation without systematic listening and interaction may not provide sufficient conditions for communicative development.

TBLT is especially relevant to this problem because it organizes learning around purposeful language use. In contemporary formulations, tasks require learners to mobilize available linguistic resources to achieve a meaningful outcome while form-focused support is provided when performance reveals a need (Ellis et al., 2020). Recent studies have demonstrated the value of collaborative, technology-mediated, mobile-supported, role-play, and repeated tasks for expanding interaction and production (Daif-Allah & Al-Sultan, 2023; Fang et al., 2021; González-Lloret, 2020; Suzuki & Hanzawa, 2022). Rather than adding more explanation to an already receptive-heavy environment, TBLT offers a mechanism for converting known language into performance through repeated cycles of input, interaction, output, feedback, and re-performance.

Accordingly, this paper has two linked purposes: first, to interpret the available directional diagnostic pattern in the IUM context; and second, to translate that pattern into a practical task-based pedagogical framework. Its contribution is the explicit connection between a documented skill imbalance and instructional design. The study asks: (1) What are the major areas of Arabic language weakness among non-native learners at IUM? (2) How do listening, speaking, and writing compare with reading in the available diagnostic evidence? (3) What difficulties characterize structured Arabic sentence production? (4) What learner and classroom factors contribute to these weaknesses? and (5) Which task-based activities are most appropriate for addressing the diagnosed needs?

LITERATURE REVIEW AND CONCEPTUAL BASIS

Integrated skill development and the recognition-to-production gap

Contemporary language competence is multidimensional. The CEFR Companion Volume emphasizes reception, production, interaction, and mediation as interdependent modes of language activity rather than treating language proficiency as knowledge of discrete forms (Council of Europe, 2020). For Arabic learners, this means that reading performance alone cannot represent communicative ability. Assessment research also highlights the importance of defining and evaluating communicative Arabic proficiency rather than relying exclusively on native-speaker norms (Alabar, 2019). The present study therefore distinguishes between recognition, in which linguistic information is externally available and can be revisited, and production, in which learners must retrieve, assemble, and monitor language independently.

The distinction is particularly useful for explaining why reading may be stronger than listening, speaking, or writing. Reading provides stable visual input and allows rereading, whereas listening is transient and time constrained. Speaking adds conceptualization, lexical retrieval, grammatical encoding, articulation, monitoring, and social pressure. Writing allows more processing time but still requires learners to generate content and language without copying the model. Kim and Crossley (2023) demonstrate that L2 spoken and written production differ in lexical and phraseological characteristics, reinforcing the need to diagnose modality-specific performance rather than assume that knowledge transfers automatically across skills.

Listening and speaking in L2 Arabic

Listening is often difficult because learners must segment continuous speech, recognize known vocabulary in altered phonological forms, maintain information in working memory, and construct meaning while the

input continues. These demands may intensify in Arabic when learners move between carefully articulated instructional language and naturally occurring speech. Research on L2 Arabic diglossia and bidialectal practice indicates that learners benefit from pedagogies that acknowledge linguistic variation rather than treating one variety as the only legitimate communicative resource (Al Masaeed, 2022; Nassif & Al Masaeed, 2022). A diagnostic listening programme should therefore examine whether difficulty arises from vocabulary knowledge, speed, phonological recognition, sentence processing, limited exposure, or dependence on written support.

Speaking requires accessible vocabulary, sentence frames, pronunciation, interactional strategies, and confidence. Classroom research with non-native Arabic learners highlights the value of purposeful classroom activities that increase opportunities for speaking practice and learner participation (Abu Shaikha & Hamzeh, 2023). Role-play can provide a structured communicative reason to negotiate meaning, ask questions, respond, and sustain dialogue, and recent Arabic research reports positive effects of role-play on dialogue skills (Daif-Allah & Al-Sultan, 2023). Repetition also matters: repeating a task can reduce planning demands and support fluency, although the scheduling of repetition should avoid mechanical performance (Suzuki & Hanzawa, 2022).

Writing, sentence construction, and language activation

Writing exposes the difference between recognizing a correct structure and producing one independently. Non-native Arabic learners may know a grammatical rule yet struggle to apply agreement, word order, verb forms, connectors, and lexical choices while developing an idea. Sentence construction is therefore a productive bottleneck that affects both speech and writing. Recent classroom research on syntactic priming in L2 Arabic suggests that repeated exposure to and production of targeted structures can support syntactic development (Alzahrani & Almalki, 2024). This provides a rationale for sentence expansion, substitution, combining, picture-to-sentence tasks, and reformulation before learners are expected to produce longer texts without scaffolding.

Vocabulary and pronunciation also require activation. Knowing the meaning of a word does not ensure that it can be retrieved rapidly in speech or incorporated accurately into writing. Tasks should therefore recycle target vocabulary across modalities and require original use. Digital pronunciation support may add individualized feedback; recent Arabic-language work demonstrates continuing advances in automatic mispronunciation detection and diagnosis for non-native learners (Alrashoudi et al., 2025). Technology, however, is most useful when it serves a defined communicative task rather than becoming an isolated activity.

TBLT as the instructional bridge

TBLT provides the central pedagogical mechanism in this study. Ellis et al. (2020) describe tasks as meaning-oriented activities with an outcome that require learners to draw on their linguistic resources. This makes TBLT well suited to a recognition-to-production gap: learners can begin with comprehensible input and scaffolded language, then use that language to solve a problem, exchange information, perform a role, retell content, or produce a written outcome. Beginner-level research also demonstrates that tasks can be designed with sufficient input and support for learners who do not yet possess extensive productive resources (Erlam & Ellis, 2018).

Technology can broaden the range and frequency of task performance. Mobile-supported TBLT has been associated with gains in linguistic achievement and conversational interaction (Fang et al., 2021), while collaborative online tasks can promote spoken and written output (González-Lloret, 2020). Smith and González-Lloret (2021) caution, however, that technology-mediated TBLT requires principled alignment among task goals, technological affordances, complexity, sequencing, and context. For IUM, this supports the use of short audio, learner voice recording, collaborative documents, and mobile practice only where they increase meaningful exposure, interaction, feedback, or re-performance.

METHODOLOGY

Research design and participants

This paper draws on an underlying mixed-methods diagnostic project in which quantitative performance evidence was complemented by qualitative evidence about learner difficulties, participation, confidence, and instructional needs. Mixed methods is appropriate when numerical patterns alone cannot explain why learners perform differently across skills and qualitative perceptions alone cannot establish the relative profile of those skills (Creswell & Creswell, 2018). The quantitative phase comprises 120 undergraduate non-native Arabic-speaking learners selected through convenience sampling. The qualitative phase comprises 30 purposively selected participants with relevant experience of Arabic instruction. These sampling strategies fit an institution-based diagnostic study but limit statistical generalization beyond the study context.

Instruments and data collection

The underlying project used pre- and post-tests covering listening, speaking, reading, and writing, together with classroom observations, questionnaires, and semi-structured interviews. The instruments target listening comprehension, spontaneous oral production, reading, independent writing, sentence construction, vocabulary retrieval, classroom participation, confidence, and perceived instructional challenges. The available study documentation reports expert review of the study materials and questionnaire reliability of Cronbach's $\alpha = .87$, indicating acceptable internal consistency for the reported questionnaire. A rigorous journal submission should additionally provide the test blueprint, scoring ranges, example tasks, speaking and writing rubrics, rater training procedures, and inter-rater reliability where performance was scored by human judges.

Analysis and integration

The underlying project was designed to compare performance across the four skills and, where the same learners completed pre- and post-tests, to evaluate change over time. A complete inferential report should include sample size per analysis, means, standard deviations, confidence intervals, appropriate paired comparisons or models, exact p values, and effect sizes after checking relevant assumptions. The available study records do not provide those numerical outputs; consequently, the present paper reports the diagnostic pattern only at the level supported by the available evidence and does not introduce statistics that cannot be verified. Qualitative evidence is analysed thematically to identify recurring patterns in learner difficulty and instructional experience. Braun and Clarke's (2021) approach provides a transparent basis for familiarization, coding, theme development, review, definition, and reporting. Integration occurs by comparing skill-performance patterns with observation, questionnaire, and interview evidence to identify instructional priorities.

Ethics and research transparency

Ethics approval, informed-consent procedures, data-retention arrangements, and anonymization procedures are not documented in the records currently available to the authors. These details must be verified from the original project records and inserted before submission. The final version should also report whether pre- and post-test observations are paired, how missing data were handled, whether assessors were blinded to test occasion, and how the qualitative and quantitative samples overlap.

Results: Diagnostic Profile

The available diagnostic records suggest an uneven four-skill profile rather than a uniform lack of Arabic knowledge. Reading appears comparatively stronger, while listening and speaking emerge as the clearest areas of difficulty. Writing also appears to require further development, particularly when learners must construct sentences independently, retrieve vocabulary, maintain grammatical and syntactic accuracy,

connect ideas, and extend sentences into coherent paragraphs. Table 1 consolidates the diagnostic pattern reported in the present study and links each area to the immediate pedagogical need.

Area	Diagnostic pattern	Likely performance demand	Priority response
Listening	Difficulty processing connected speech; familiar words may not be recognized in real time; reliance on repetition or written support.	Rapid segmentation, phonological recognition, retention, meaning construction.	Short graded audio, repeated listening, information-gap listening, key-word and note-taking tasks.
Speaking	Hesitation, limited spontaneous speech, slow vocabulary retrieval, short responses, pronunciation/fluency problems, reduced confidence.	Rapid retrieval, sentence assembly, articulation, interaction management.	Pair tasks, role-play, retelling, information gaps, repeated performance and low-stakes speaking.
Writing	Weak independent sentence production, organization, agreement/word order, connectors, and paragraph development.	Independent lexical/grammatical selection, organization, revision.	Guided sentence building, expansion, picture-based writing, collaborative drafting, gradual removal of scaffolds.
Reading	Comparatively stronger and supported by stable visual input.	Recognition and self-paced processing.	Use reading as input/scaffolding for subsequent listening, speaking, and writing rather than over-practising it.

Listening and speaking weaknesses

Listening problems involve processing connected Arabic, identifying known vocabulary in speech, locating key information, and responding immediately after listening. The reported dependence on repetition and written support may indicate that part of the difficulty lies in converting visually accessed lexical knowledge into phonological processing. Speaking displays a parallel retrieval problem: learners hesitate, produce short or disconnected responses, and have difficulty sustaining interaction without teacher support. Fear of grammatical mistakes, pronunciation difficulty, and limited fluency further restrict oral participation. Taken together, these patterns may point to limited experience with real-time processing and low-risk communicative practice, rather than reflecting grammatical knowledge alone.

Writing and structured sentence production

Writing problems are concentrated at the point of independent generation. Learners have difficulty producing complete sentences, ordering words, maintaining agreement, selecting connectors, retrieving vocabulary, and organizing ideas into a paragraph. The present study identifies sentence construction as an important area of difficulty because it affects both writing and speaking. Learners may recognize a correct sentence yet struggle to create a parallel sentence without a model. This finding supports staged production: controlled but meaningful sentence manipulation should precede increasingly independent writing and speaking, with feedback directed to errors that impede clarity or reflect the current target structure.

Reading as a strategic resource

The relative strength in reading should not lead to further concentration of instructional time on reading. Instead, reading can function as a scaffold. A short text can supply topic knowledge, vocabulary,

collocations, sentence patterns, and ideas; those resources can then be recycled through listening, pair discussion, retelling, role-play, sentence transformation, and a short written response. The diagnostic implication is therefore redistribution rather than abandonment: preserve reading as a source of comprehensible input while increasing the proportion of lesson time in which learners must process or produce Arabic without continuous visual support.

Task-Based Pedagogical Framework for IUM

The framework translates the directional diagnostic pattern into a proposed six-stage weekly cycle. Its design follows the TBLT principle that language is strengthened through purposeful use while allowing explicit attention to vocabulary, grammar, and pronunciation when learners need support (Ellis et al., 2020). It also incorporates task repetition, which can lower processing demands and support fluency when re-performance has a communicative purpose (Suzuki & Hanzawa, 2022).

Stage	Core activity	Primary skill focus	Example
1. Activate	Topic prediction and vocabulary retrieval	Speaking/listening	Picture prediction; semantic map; quick pair questions.
2. Input	Short reading and/or graded audio	Reading/listening	Listen for gist, then key details; compare with a short text.
3. Focus	Sentence pattern, vocabulary, pronunciation	Grammar/lexis/form	Noticing, substitution, sentence expansion, model-record-compare.
4. Task	Meaningful pair/group outcome	Speaking/listening	Information gap, role-play, problem solving, sequencing, interview.
5. Individual output	Independent transfer	Speaking/writing	One-minute retell, picture description, message, mini-paragraph.
6. Feedback + re-performance	Selective feedback followed by a second attempt	Integrated skills	Repeat with new partner/topic; revise paragraph; record again.

Targeted activities

For listening, activities should progress from prediction and gist to specific-information and response tasks: listen-and-select, listen-and-order, information-gap listening, keyword identification, short note-taking, repeated listening with gradually reduced textual support, and dictogloss. For speaking, teachers can use pair interviews, role-play, picture description, question chains, information gaps, story retelling, guided discussion, short presentations, and conversation circles. These formats are intended to increase the number of turns each learner produces and may reduce the communicative pressure associated with speaking before the whole class (Abu Shaikha & Hamzeh, 2023; Daif-Allah & Al-Sultan, 2023).

Sentence construction should connect grammar directly to meaning. Useful activities include word-order reconstruction, sentence expansion, substitution tables, question-answer transformation, sentence combining, correction of learner-generated sentences, picture-to-sentence production, and syntactic priming

through repeated narrative patterns (Alzahrani & Almalki, 2024). Writing should then move from guided sentences to paragraph completion, picture-based writing, short summaries, messages, academic mini-paragraphs, collaborative drafting, and independent revision. Vocabulary should be pre-activated and then required in original output; pronunciation can be addressed through short shadowing and model-record-compare cycles, with digital diagnosis used where appropriate (Alrashoudi et al., 2025).

DISCUSSION

The available evidence is best interpreted as suggesting a recognition-to-production imbalance. Learners appear able to access some Arabic knowledge when forms remain visually available but have greater difficulty when language is transient, interactive, or must be generated independently. This interpretation explains why reading can be comparatively stronger while listening, spontaneous speaking, and sentence production remain weak. It is also consistent with evidence that L2 modality changes lexical and phraseological demands (Kim & Crossley, 2023) and with Arabic-specific research emphasizing the need for communicative, interactive, and multidialectally aware instruction (Al Masaeed, 2022; Aljanadbah & Al Sharef, 2024; Nassif & Al Masaeed, 2022).

The findings therefore suggest that a pedagogical response based primarily on additional explanation may not be sufficient. Grammar and vocabulary remain essential, but knowledge that is repeatedly recognized without retrieval may remain inaccessible during communication. The proposed framework treats form as a resource for task performance: learners notice a pattern, use it to achieve a communicative outcome, receive focused feedback, and perform again. This sequence is consistent with contemporary TBLT theory and with evidence supporting beginner tasks, mobile-supported interaction, collaborative online work, role-play, and carefully designed task repetition (Ellis et al., 2020; Erlam & Ellis, 2018; Fang et al., 2021; González-Lloret, 2020; Suzuki & Hanzawa, 2022).

A second implication concerns Arabic variation. Listening and speaking curricula should prepare learners to cope with authentic variation rather than equating communicative competence solely with decontextualized grammatical accuracy. Research on translanguaging, bidialectal development, and multidialectal practice shows that L2 Arabic learners draw on complex repertoires when negotiating meaning (Al Masaeed, 2020, 2022; Nassif & Al Masaeed, 2022). IUM can maintain Modern Standard Arabic as an academic target while sensitizing learners to common spoken variation through carefully selected listening input and pragmatic comparison. The objective is not uncontrolled dialect mixing but more realistic comprehension and sociolinguistic awareness. A third implication is that confidence should be treated as an outcome of structured successful performance rather than addressed only through encouragement. Low-stakes pair work, planning time, useful sentence frames, repeated tasks, selective error correction, and opportunities to observe progress may help reduce the communicative pressure associated with mistakes. Technology can extend these opportunities outside class through brief listening, voice-recording, and collaborative-writing tasks, provided that tool choice remains subordinate to the communicative objective (Smith & González-Lloret, 2021).

Implications for Curriculum, Assessment, and Teacher Development

Curriculum planning should begin with a short diagnostic profile at the start of each course and allocate instructional time according to demonstrated need. In the present context, listening and speaking should receive increased systematic practice, while writing should follow a planned progression from sentence construction to connected paragraphs. An IUM task bank could classify activities by level, topic, skill, language function, expected outcome, and required scaffolding. Reading texts already used in courses can be redesigned as launch points for integrated listening, speaking, and writing tasks rather than treated as endpoints.

Assessment should report the four skills separately so that a strong reading score cannot compensate for severe productive weakness. Speaking criteria can include task completion, comprehensibility, fluency, vocabulary access, interaction, pronunciation, and grammatical control; writing can include content, organization, sentence accuracy, vocabulary, cohesion, and task fulfilment. Performance tasks should

resemble classroom goals and include repeated low-stakes assessments. Teacher development should focus on task design, scaffolding, managing pair/group interaction, selecting audio, providing selective feedback, and using digital tools for purposeful practice. Such alignment between diagnosis, teaching, and assessment is important if communicative development is to be evaluated systematically rather than treated as an aspirational goal.

LIMITATIONS AND FUTURE RESEARCH

The paper has several limitations. Convenience sampling in the quantitative phase and purposive sampling in the qualitative phase restrict generalizability. More importantly, the available study documentation does not include the numerical pre/post score tables, inferential tests, effect sizes, confidence intervals, participant demographics, qualitative coding frequencies, representative interview excerpts, or the ethics approval identifier. These elements should be recovered from the original dataset and study documentation before the paper can support full empirical scrutiny. The present results therefore remain a directional diagnostic profile rather than a complete statistical account.

Future research should test the proposed framework through a transparent intervention design. A quasi-experimental or randomized design, where feasible, could compare task-based instruction with current practice using common pre-, post-, and delayed measures. Speaking and writing should be double rated with reported inter-rater reliability, while qualitative process data can examine confidence, willingness to communicate, perceived usefulness, and teacher implementation. Longitudinal work should determine whether gains transfer beyond rehearsed tasks and whether greater exposure to spoken varieties improves comprehension without weakening the programme's academic focus on Modern Standard Arabic.

CONCLUSION

This paper interprets an available directional diagnostic pattern of Arabic language skills among non-native undergraduate learners at the Islamic University of Maldives through a task-based pedagogical perspective. The available diagnostic evidence suggests that learners' language development is not uniformly weak; rather, it appears to be characterized by an imbalance between relatively stronger reading performance and more difficult listening, speaking, and independent writing. Difficulties in processing connected speech, retrieving vocabulary, constructing sentences spontaneously, maintaining grammatical accuracy, and organizing written ideas suggest that learners may possess a degree of receptive or recognition-based knowledge without being able to consistently transform that knowledge into independent language production.

The findings therefore highlight the importance of moving beyond a predominantly recognition-oriented approach to Arabic instruction. Knowing a word, grammatical structure, or sentence pattern is not sufficient if learners cannot retrieve and use that knowledge during real-time communication. For this reason, the proposed framework places listening, speaking, sentence construction, writing, pronunciation, vocabulary activation, and interaction at the center of classroom practice. Reading is not excluded; instead, it is repositioned as a source of comprehensible input and linguistic scaffolding that prepares learners for subsequent productive tasks.

TBLT offers a potentially practical bridge between linguistic knowledge and communicative performance. The six-stage cycle proposed in this study—activation, input, language focus, collaborative task performance, individual output, and feedback with re-performance—offers a structured way to transform learners from passive recipients of language into active users of Arabic. The framework also allows teachers to provide explicit attention to grammar, vocabulary, and pronunciation when these become necessary for successful task completion. Thus, task-based instruction in this context should not be understood as replacing language form with communication, but as connecting language form to meaningful use.

The study also has implications for Arabic language education in Islamic higher education. Since Arabic at IUM serves academic, religious, and professional purposes, learners need to develop the ability to understand and produce language across different communicative situations. Greater attention to spoken

input, interaction, repeated oral practice, sentence-level production, and guided-to-independent writing may help reduce the gap between what learners recognize and what they can use. At the same time, exposure to appropriately selected spoken Arabic can help learners develop greater confidence in understanding authentic communication while maintaining Modern Standard Arabic as an important academic target.

From a curriculum perspective, the findings suggest that skill development should be deliberately balanced rather than assuming that progress in reading automatically leads to equivalent progress in speaking, listening, or writing. Assessment should therefore provide separate and meaningful evidence of learners' performance in each skill, particularly productive skills. Teachers should also be supported in designing communicative tasks, providing appropriate scaffolding, managing pair and group interaction, and using selective feedback and task repetition to encourage more independent production.

The contribution of this paper is consequently both diagnostic and pedagogical. It identifies a specific recognition-to-production problem within the IUM context and translates that problem into a context-responsive instructional framework. However, the framework should be regarded as a theoretically and diagnostically informed proposal rather than as evidence of an already established intervention effect. The available study record does not provide sufficient statistical detail to claim that the proposed framework has itself produced significant learning gains. Future research should therefore implement the framework systematically and evaluate its effectiveness through pre-, post-, and delayed post-tests, appropriate statistical analyses, speaking and writing performance measures, inter-rater reliability, classroom observations, and learner perspectives.

Ultimately, improving Arabic proficiency among non-native learners requires a shift from asking “What Arabic do learners know?” to also asking “What can learners do with the Arabic they know?” The central pedagogical challenge is not simply increasing linguistic knowledge, but creating repeated, meaningful opportunities for learners to retrieve, negotiate, produce, receive feedback on, and reuse that knowledge. A task-based approach offers a promising pathway for making this transition and for developing more balanced, confident, and communicatively capable Arabic learners at IUM and in comparable higher-education contexts.

REFERENCES

1. Abu Shaikha, A. M., & Hamzeh, M. (2023). Using classroom activities in teaching speaking skills for non-native Arabic learners: Teacher's perspective. *Journal of Language Teaching and Research*, 14(6), 1524–1534. <https://doi.org/10.17507/jltr.1406.10>
2. Al Masaeed, K. (2020). Translanguaging in L2 Arabic study abroad: Beyond monolingual practices in institutional talk. *The Modern Language Journal*, 104(1), 250–266. <https://doi.org/10.1111/modl.12623>
3. Al Masaeed, K. (2022). Bidialectal practices and L2 Arabic pragmatic development in a short-term study abroad. *Applied Linguistics*, 43(1), 88–114. <https://doi.org/10.1093/applin/amab013>
4. Alabar, R. (2019). An assessment norm of communicative Arabic proficiency. *Arab Journal of Applied Linguistics*, 4(2), 87–106.
5. Aljanadbah, A., & Al Sharef, A. (2024). Improving the language competencies of non-native Arabic speakers through the communicative approach. *International Journal of Learning and Teaching*, 10(3), 418–427. <https://doi.org/10.18178/ijlt.10.3.418-427>
6. Almelhes, S. (2024). Enhancing Arabic language acquisition: Effective strategies for addressing non-native learners' challenges. *Education Sciences*, 14(10), 1116. <https://doi.org/10.3390/educsci14101116>
7. Alrashoudi, N., Al-Khalifa, H., & Alotaibi, Y. (2025). Improving mispronunciation detection and diagnosis for non-native learners of the Arabic language. *Discover Computing*, 28, Article 1. <https://doi.org/10.1007/s10791-024-09489-8>
8. Alzahrani, A., & Almalki, H. (2024). Syntactic priming in the classroom: Using narratives to prime L2 Arabic speakers. *Asian-Pacific Journal of Second and Foreign Language Education*, 9, Article 68. <https://doi.org/10.1186/s40862-024-00292-y>
9. Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.
10. Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing.

11. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
12. Daif-Allah, A. S., & Al-Sultan, M. S. (2023). The effect of role-play on the development of dialogue skills among learners of Arabic as a second language. *Education Sciences*, 13(1), 50. <https://doi.org/10.3390/educsci13010050>
13. Damsing, N., & Che Haron, S. (2026). Developing the four Arabic language skills and their role in enhancing communicative competence among non-native learners. *International Journal of Modern Education*, 8(29), 651–664. <https://doi.org/10.35631/IJMOE.829039>
14. Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press. <https://doi.org/10.1017/9781108643689>
15. Erlam, R., & Ellis, R. (2018). Task-based language teaching for beginner-level learners of L2 French: An exploratory study. *Canadian Modern Language Review*, 74(1), 1–26. <https://doi.org/10.3138/cmlr.3831>
16. Fang, W.-C., Yeh, H.-C., Luo, B.-R., & Chen, N.-S. (2021). Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. *ReCALL*, 33(1), 71–87. <https://doi.org/10.1017/S0958344020000208>
17. González-Lloret, M. (2020). Collaborative tasks for online language teaching. *Foreign Language Annals*, 53(2), 260–269. <https://doi.org/10.1111/flan.12466>
18. Hashem, R. (2022). A qualitative investigation into the impact of diglossia on the self-learning of an Arabic spoken dialect by Arabic as a foreign language learners. *Journal of Arts, Literature, Humanities and Social Sciences*, (80), 112–126. <https://doi.org/10.33193/JALHSS.80.2022.685>
19. Kim, M., & Crossley, S. A. (2023). Lexical and phraseological differences between second language written and spoken opinion responses. *Frontiers in Psychology*, 14, 1068685. <https://doi.org/10.3389/fpsyg.2023.1068685>
20. Nassif, L., & Al Masaeed, K. (2022). Supporting the sociolinguistic repertoire of emergent diglossic speakers: Multidialectal practices of L2 Arabic learners. *Journal of Multilingual and Multicultural Development*, 43(8), 759–773. <https://doi.org/10.1080/01434632.2020.1774595>
21. Smith, B., & González-Lloret, M. (2021). Technology-mediated task-based language teaching: A research agenda. *Language Teaching*, 54(4), 518–534. <https://doi.org/10.1017/S0261444820000233>
22. Suzuki, Y., & Hanzawa, K. (2022). Massed task repetition is a double-edged sword for fluency development. *Studies in Second Language Acquisition*, 44(2), 536–561. <https://doi.org/10.1017/S0272263121000358>