



From Gaps to Solutions: Advancing Arabic Language Assessment through the Hamza Academic Test

Alaa Alzahrani ¹, Ebtessam Abdulhaleem ²

¹Independent researcher, Saudi Arabia

²King Salman Global Academy for Arabic, Saudi Arabia

Abstract

With the rapid growth of Arabic language learning worldwide, there is a pressing need for robust and standardized proficiency assessments. Several Arabic proficiency tests have been developed over the past two decades (e.g., Ying, 2023; Banyan, 2024; Al-Kharji, 2024; Dahabi, 2024; Al-Sarie, 2024; Fadl, 2023). However, so far, no comprehensive review has systematically examined the landscape of Arabic assessment to identify gaps and challenges in this evolving field. As such, this paper presents a narrative review of existing Arabic language proficiency assessments, including both global and regionally developed tests. Our review was based on a systematic search and critical analysis of published studies, institutional reports, and test documentation. The aim of this review is to identify key challenges related to test validity, reliability, accessibility, skill coverage, and alignment with international standards. Building on these findings, the paper introduces the Hamza for Academic Purposes Test, a newly developed, standardized Arabic proficiency test designed to address several of the identified challenges. The development process of the Hamza test is briefly outlined, highlighting its innovative features in assessing reading, listening, writing, and speaking skills through authentic academic tasks. The paper argues that the Hamza test represents a model for integrating international best practices with localized adaptations tailored to Arabic's unique linguistic features. By mapping the current state of Arabic language assessment and proposing practical solutions, this study contributes to advancing the field toward more valid globally recognized assessments. This ultimately supports equitable opportunities for Arabic learners worldwide.

Keywords: Arabic assessment, language tests, Hamza test, standardized language proficiency tests

Introduction

The assessment of Arabic language proficiency has gained increasing prominence in recent decades as Arabic continues to expand its role as a global language in education, commerce, diplomacy, and international communication. With over 400 million native speakers and its status as one of the six official languages of the United Nations, Arabic serves as a critical medium for academic scholarship, business transactions, and cross-cultural exchange across the Middle East, North Africa, and diaspora communities worldwide (Ryding, 2013). This growing global significance has generated substantial demand for robust and standardized proficiency assessments that can reliably measure learners' Arabic language competence across diverse contexts and purposes.

In response to this demand, several Arabic proficiency tests have been developed over the past two decades, reflecting varied institutional priorities, theoretical frameworks, and target populations (e.g., Ying, 2023; Banyan, 2024; Al-Kharji, 2024; Dahabi, 2024; Al-Sarie, 2024; Fadl, 2023). These assessments range from internationally recognized certifications aligned with the Common European Framework of Reference for Languages (CEFR) to institution-specific placement tests designed for particular educational contexts. Some tests employ computer-adaptive methodologies to optimize assessment efficiency, while others utilize traditional paper-based formats or hybrid delivery modes. Certain assessments focus exclusively on Modern Standard Arabic, whereas others attempt to address the complex linguistic phenomenon of diglossia by incorporating dialectal variations. This proliferation of assessment instruments demonstrates both the field's vitality and its fragmentation, as stakeholders pursue diverse approaches to measuring Arabic proficiency without unified standards or comprehensive coordination.

However, despite this expansion in Arabic language assessment, no comprehensive review has systematically examined the landscape of Arabic proficiency testing to identify gaps, challenges, and best practices in this evolving field. Existing studies have typically focused on individual tests or specific assessment contexts without providing holistic analysis of the broader assessment ecosystem. This gap in the literature leaves educators, administrators, policymakers, and learners



without clear guidance regarding the relative strengths and limitations of available assessments, hindering informed decision-making about test selection, program evaluation, and credentialing practices. Furthermore, the absence of systematic review obscures opportunities for identifying common challenges and developing collaborative solutions to advance the field collectively.

As such, this paper presents a narrative review of existing Arabic language proficiency assessments, including both globally recognized and regionally developed tests. Our review was based on systematic examination and critical analysis of published studies, institutional reports, test documentation, and professional assessment databases. The review encompasses 18 distinct proficiency tests, examining their structural characteristics, psychometric properties, target populations, accessibility features, and alignment with international proficiency frameworks. The aim of this review is to identify key challenges related to test validity, reliability, accessibility, skill coverage, and alignment with international standards such as CEFR, American Council on the Teaching of Foreign Languages (ACTFL) guidelines, and Interagency Language Roundtable (ILR) scales.

Building on these findings, the paper introduces the Hamza for Academic Purposes Test, a newly developed, standardized Arabic proficiency test designed to address several of the identified challenges in current assessment practices. The development process of the Hamza test is outlined, highlighting its innovative features in assessing reading, listening, writing, and speaking skills through authentic academic tasks that reflect real-world language use in higher education contexts. The paper argues that the Hamza test represents a model for integrating international best practices with localized adaptations tailored to Arabic's unique linguistic features, including its complex morphology, diglossia phenomenon, and rich cultural contexts.

By mapping the current state of Arabic language assessment and proposing practical solutions through the Hamza test development, this study contributes to advancing the field toward more valid, reliable, and globally recognized assessments. This work ultimately supports equitable opportunities for Arabic learners worldwide by facilitating transparent proficiency evaluation, credential portability, and access to educational and professional opportunities requiring demonstrated Arabic language competence.

1. Methodology

1.1 Design

This study employs a narrative review methodology to examine existing Arabic language proficiency tests. A narrative review approach was selected to provide a comprehensive, qualitative synthesis of available proficiency assessments, allowing for descriptive analysis of test characteristics, applications, and limitations without the systematic constraints of meta-analysis (Green et al., 2006). This methodological approach enables the exploration of diverse assessment instruments across varied institutional contexts and facilitates identification of patterns, gaps, and recommendations for Arabic language assessment practices.

1.2 Sample selection

The review encompasses 18 Arabic language proficiency tests currently available or documented in academic and professional literature. These tests represent a diverse range of assessment purposes, including international certification (e.g., CIMA, ACTFL), institutional placement (e.g., King Saud University Test, OSAPT), teacher certification (e.g., TAFL, PTMSA), adaptive assessments (e.g., ALPT, Avant APT), and commercial evaluation services (e.g., ALTA). Test selection criteria included: (a) availability of documented information regarding test structure and purpose, (b) current or recent administration status, and (c) relevance to Modern Standard Arabic proficiency assessment. Tests without verifiable documentation or institutional attribution were included when referenced in Arabic language assessment literature to provide comprehensive coverage of the field. Information for each test was collected from multiple sources including official test websites, institutional documentation, academic publications, and professional assessment databases.

1.3 Analytical framework

Each proficiency test was analyzed according to six critical dimensions:

1. Test Structure and Format: Assessment components (listening, reading, writing, speaking), delivery mode (paper-based, computer-based, online), duration, and task types employed.



2. Alignment with Proficiency Frameworks: Correspondence with internationally recognized standards including the Common European Framework of Reference for Languages (CEFR), American Council on the Teaching of Foreign Languages (ACTFL) guidelines, and Interagency Language Roundtable (ILR) scale.
3. Psychometric Properties: Available evidence regarding reliability, validity, standardization procedures, and quality assurance mechanisms.
4. Target Population and Purpose: Intended test-takers (native speakers, heritage learners, second language learners), assessment objectives (placement, certification, admission), and contextual applications (academic, professional, governmental).
5. Accessibility and Administration: Geographic availability, cost considerations, technological requirements, language varieties assessed (Modern Standard Arabic vs. dialects), and administrative procedures.
6. Recognition and Transferability: International recognition, credential value, institutional acceptance, and applicability across educational and professional contexts.

This analytical framework provides systematic examination of each assessment instrument while enabling comparative analysis across tests to identify strengths, limitations, and best practices in Arabic language proficiency assessment.

1.4 Data analysis

Data were synthesized descriptively for each test, highlighting key characteristics aligned with the analytical framework. Cross-test comparisons were conducted to identify patterns in assessment approaches, recognize gaps in current offerings, and generate recommendations for stakeholders including educators, administrators, policymakers, and test-takers. The narrative synthesis emphasizes both individual test characteristics and broader implications for Arabic language assessment practices.

2. Results

The results are presented in two main components. First, the paper reviews 18 existing Arabic proficiency tests using the analytical framework outlined in the methodology. Second, it introduces the Hamza for Academic Purposes Test as a response to the identified challenges.

2.1 Comparative Analysis

Across the 18 tests, the synthesis highlights shared strengths, recurring limitations, and systematic gaps. Patterns in design and variation in quality are noted, as well as opportunities for advancing Arabic proficiency assessment toward more reliable and equitable practices. These tests are:

- Test 1: CIMA
- Test 2: King Saud University Test
- Test 3: Attanala'l Alarabi
- Test 4: TAFL (Teaching Arabic)
- Test 5: International Institute Test
- Test 6: ALAC
- Test 7: Arabic for Official Purposes
- Test 8: Academic Purposes Test
- Test 9: OSAPT
- Test 10: General Proficiency Examination
- Test 11: Arab Academy ALPT
- Test 12: AL-ARABIYYA-INSTITUTE
- Test 13: ACTFL Arabic Tests
- Test 14: PTMSA
- Test 15: General Institutional Test
- Test 16: APT
- Test 17: ALPA
- Test 18: ALTA



2.2 Diglossia

A critical pattern emerging from this review concerns the tension between Modern Standard Arabic (MSA) and dialectal varieties. The majority of reviewed tests focus exclusively on MSA, reflecting its status as the prestige variety used in formal education, media, and official communication. However, this orientation raises fundamental questions about construct validity and real-world applicability, particularly for learners whose primary communicative needs involve dialectal proficiency. Only a minority of instruments attempt to assess dialectal competence, and none adequately address the diglossic reality that characterizes Arabic language use across most contexts.

2.3 Alignment with International Standards

Approximately two-thirds of the reviewed tests explicitly reference the Common European Framework of Reference for Languages (CEFR), demonstrating an effort to align Arabic assessment with internationally recognized proficiency descriptors. This alignment facilitates comparison with assessments in other languages and supports mobility within educational and professional contexts. However, the adaptation of CEFR to Arabic presents distinctive challenges. The framework was developed primarily for European languages with relatively straightforward relationships between written and spoken forms, whereas Arabic's diglossia, morphological complexity, and script-specific literacy demands require substantial modification of standard descriptors.

Several tests, including the CIMA (Certificat International de Maîtrise en Arabe) and various institutional assessments, have attempted systematic CEFR alignment through calibration studies and expert judgment. Nevertheless, concerns persist regarding the cultural and linguistic appropriateness of applying European frameworks to Arabic without adequate validation research demonstrating equivalence of constructs across proficiency levels.

2.4 Skill coverage

The reviewed assessments demonstrate considerable variability in skill coverage and task design. Most instruments assess reading, writing, listening, and speaking, though the depth and authenticity of these assessments vary substantially. Reading assessments typically employ multiple-choice comprehension questions based on graded passages, though advanced tests increasingly incorporate authentic texts from journalism, literature, and academic discourse. Listening components often rely on recorded dialogues and monologues, with MSA pronunciation that may not reflect the phonological variation learners encounter in naturalistic settings.

Speaking assessments present particular challenges. While some tests employ face-to-face oral interviews with trained raters, resource constraints have driven many programs toward semi-direct assessments using recorded prompts or computer-mediated simulated conversations. The validity of these approaches for Arabic—a language where pragmatic competence and sociolinguistic appropriateness are heavily context-dependent—remains insufficiently researched. Writing assessments typically require production of formal text types (letters, essays, reports) that privilege MSA academic registers, potentially disadvantaging learners with strong functional competence in informal or dialectal writing.

2.5 Psychometric concerns

Critical examination of test documentation reveals significant gaps in psychometric reporting. While many assessments claim high reliability, few provide detailed evidence from classical test theory or item response theory analyses. Validity documentation is particularly sparse, with limited evidence of systematic construct validation, criterion-related validity studies, or consequential validity investigations examining the impacts of test use on teaching and learning.

Inter-rater reliability for productive skills assessments emerges as a persistent concern. Even tests employing detailed scoring rubrics report challenges in achieving acceptable agreement among raters, particularly for speaking assessments where prosodic features, pragmatic appropriateness, and code-switching behaviors must be evaluated. The field would benefit substantially from standardized rater training protocols and ongoing calibration procedures modeled on established practices in other language testing contexts.



2.6 Accessibility and equity issues

Accessibility represents a critical challenge across the Arabic testing landscape. Geographic concentration of test centers, limited availability of computer-based testing options, high examination fees, and inadequate accommodations for candidates with disabilities collectively restrict equitable access. Most tests are administered through institutional partnerships concentrated in urban centers in the Middle East, North Africa, and select Western universities, leaving learners in other regions with limited assessment options.

Language of administration constitutes another accessibility concern. While test instructions are typically available in Arabic, many assessments provide limited or no support materials in learners' first languages, potentially disadvantaging beginning-level examinees who lack sufficient Arabic proficiency to fully understand test requirements. The use of English as the default administrative language for several internationally marketed tests reflects colonial legacies and assumes Western educational contexts that may not align with the backgrounds of many Arabic learners.

2.7 Introduction of the Hamza Test

The second part presents the Hamza for Academic Purposes Test. It outlines its development, innovative features, and Arabic-specific adaptations, while showing alignment with validity and reliability standards.

In recent years, the global landscape of Arabic language education has witnessed remarkable expansion. The number of learners enrolling in Arabic language programs at universities worldwide is increasing at a rate estimated to be ten times higher than that of other foreign language programs (King Salman Global Academy for Arabic Language [KSGAAL], n.d.). This expansion has created an urgent demand for standardized and internationally recognized tools to assess Arabic language proficiency, particularly in academic contexts. Such assessments serve not only educational institutions but also governmental agencies and employers who require reliable evidence of linguistic competence.

The Hamza Academic Test emerges as a strategic response to this demand. Developed by the King Salman Global Academy for Arabic Language, the test is designed to align with international standards, most notably the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001), while also addressing the unique linguistic characteristics of Arabic. The initiative is closely tied to Saudi Arabia's Vision 2030, which emphasizes cultural promotion, national identity, and the global dissemination of the Arabic language as part of a broader societal and educational transformation (Saudi Vision 2030, 2016). Within this vision, standardizing Arabic language proficiency assessment constitutes a foundational step toward enhancing the language's scientific, cultural, and global presence.

The Hamza Academic Test was developed by a team of specialists and experts under the auspices of KSGAAL, with the explicit goal of creating an assessment tool that adheres to the highest international testing standards while remaining faithful to the linguistic structure of Arabic. The development process involved benchmarking against globally recognized language tests and conducting surveys of international best practices to establish clear standards and indicators.

A central pillar of the Hamza Test is its alignment with the CEFR proficiency levels, ranging from A2 (Below Intermediate) to C1 (Advanced). This alignment enables test users to interpret results in a way that is comparable to proficiency in other major world languages. The CEFR provides descriptors for language proficiency in terms of communicative ability, including listening, reading, spoken interaction, spoken production, and writing (Council of Europe, 2001). By adopting this framework, the Hamza Test situates Arabic proficiency within a widely accepted international scale, enhancing transparency and comparability.

The Hamza Academic Test is comprehensive in its assessment of the four core language skills: listening, reading, writing, and speaking. The total duration of the test is 155 minutes, and it comprises 75 items distributed across the skills.

- **Listening Comprehension (30 items):** Candidates listen to diverse audio materials—including monologues, dialogues, and academic discussions—and respond to multiple-choice questions. This section assesses the ability to understand spoken Arabic in both general and academic contexts.
- **Reading Comprehension (40 items):** Examinees read a range of texts and answer multiple-choice questions. Texts are selected to reflect both general topics and academic content, testing the ability to process written Arabic efficiently and critically.

- Writing Section (1 task): Candidates produce a short essay of 200–250 words. This task assesses the ability to organize ideas, use appropriate academic vocabulary and grammar, and express arguments in written Arabic.
- Speaking Section (4 tasks): This component includes answering general questions, discussing topics, describing visual or conceptual prompts, and participating in simulated dialogues. The tasks aim to elicit spontaneous and structured oral language in academic-like situations.

The test is delivered through two modes: computer-based administration at authorized testing centers and a secure virtual computer-based mode for online test-takers. This dual-mode delivery expands accessibility while maintaining standardized administration procedures. The test is offered regularly, at least once every four months, to accommodate candidates worldwide.

The Hamza Academic Test targets a broad range of stakeholders:

1. Universities and Arabic Language Institutes use it in their admission and placement processes for undergraduate and postgraduate programs. The test helps institutions select students with the appropriate linguistic proficiency for academic success.
2. Employers utilize the test to assess the language competence of candidates applying for positions where Arabic proficiency is a requirement, ensuring that language skills are evaluated objectively.
3. Teachers and learners of Arabic benefit from the test as a reliable tool to measure progress, apply for scholarships, or meet program entry requirements.

This broad applicability reflects the test's dual function as both an academic proficiency tool and a credential for professional and institutional contexts.

The Hamza Academic Test represents a significant milestone in Arabic language assessment for several reasons.

First, it standardizes Arabic language proficiency assessment globally, providing a consistent benchmark comparable to tests such as TOEFL or DELF for English and French, respectively. This standardization fills a critical gap in the field of Arabic language education, where the lack of a widely accepted proficiency test has long been a barrier to mobility and recognition.

Second, the test integrates all four language skills in a balanced manner, thereby offering a comprehensive profile of the test taker's abilities. This stands in contrast to older tests that often privileged reading and grammar at the expense of productive skills.

Third, the test embraces digital transformation by offering computer-based and online administration. This innovation enhances global accessibility, reduces logistical barriers, and facilitates regular testing schedules.

Finally, the Hamza Test is directly embedded within national strategic initiatives, particularly Vision 2030, which envisions the Arabic language as a key cultural and educational asset to be promoted internationally (Saudi Vision 2030, 2016). By situating the test within this broader policy framework, KSGAAL has ensured institutional support, sustainability, and international visibility.

3. Discussion

This narrative review highlights both the diversity and the fragmentation of the current Arabic language assessment landscape. Across the 18 reviewed tests, clear patterns emerge regarding scope, methodological approaches, and systemic limitations. Many assessments focus predominantly on Modern Standard Arabic (MSA) and discrete-point testing of grammar and reading, often neglecting productive skills and authentic communicative competence. While several instruments align with international frameworks such as CEFR or ACTFL, adaptation to Arabic's unique linguistic features—particularly diglossia, morphological complexity, and script-specific literacy demands—remains inconsistent. This misalignment raises concerns about construct validity and the ecological relevance of existing tests for real-world Arabic language use.

Psychometric transparency is another key issue. Although many tests claim reliability and standardization, few provide robust evidence of validity, rater consistency, or adaptive measurement quality. Furthermore, access remains uneven. High costs, limited testing sites, and inadequate technological infrastructure exclude many learners, particularly in underrepresented regions. These challenges collectively underscore the absence of a unified, internationally recognized Arabic proficiency standard equivalent to TOEFL, IELTS, or DELF.

The Hamza Academic Test was developed in direct response to these gaps. By aligning with CEFR while integrating Arabic-specific linguistic considerations, it offers a standardized, comprehensive, and internationally interpretable proficiency measure. Its balanced coverage of the four language skills and focus on authentic academic tasks address the limitations of earlier tests that emphasized only

receptive skills. The incorporation of computer-based and online delivery modes also expands accessibility while maintaining standardized administration procedures.

Importantly, Hamza is embedded within broader national and international language policy initiatives, particularly Saudi Arabia's Vision 2030. This strategic positioning provides institutional support and promotes the test's potential to serve as a global benchmark for Arabic proficiency. While further validation studies are necessary to establish its psychometric robustness and cross-contextual applicability, the Hamza Academic Test represents a significant advancement in Arabic language assessment. It offers a coherent response to the field's current fragmentation and sets a foundation for more equitable, reliable, and internationally recognized assessment practices.

Conclusion

This review has mapped the current landscape of Arabic language proficiency assessment, revealing both progress and persistent gaps. Existing tests demonstrate a rich diversity of approaches but lack unified standards, consistent validity evidence, and broad accessibility. The Hamza Academic Test responds to these challenges by integrating international frameworks with Arabic-specific linguistic considerations, offering a comprehensive and standardized tool for academic and professional contexts. As Arabic continues to expand globally, the Hamza Test represents a significant step toward establishing a reliable, internationally recognized benchmark for Arabic proficiency, fostering greater transparency, mobility, and equity for learners worldwide.

REFERENCES

- [1] [A review of the fourth-level Arabic major test in Chinese universities: Its reality and expectations]. 83–55، (1)1، مجلة اختبارات اللغة، <https://doi.org/10.60161/2501-001-001-002>.
- [2] [Evaluation of the placement test in Arabic Linguistics Institute at King Saud University]. 74–25، (2)1، مجلة اختبارات اللغة، <https://doi.org/10.60161/2501-001-002-002>.
- [3] [Content analysis of placement tests for non-native Arabic speakers: A comparative study]. 127–75، (2)1، مجلة اختبارات اللغة، <https://doi.org/10.60161/2501-001-002-003>.
- [4] [Speaking skill tests for non-Arabic speakers in Morocco and Germany: A descriptive and analytical study]. 157–129، (2)1، مجلة اختبارات اللغة، <https://doi.org/10.60161/2501-001-002-004>.
- [5] [Writing test validity in Arabic as a second language: A descriptive-analytical study]. 93–50، (3)2، مجلة اختبارات اللغة، <https://doi.org/10.60161/2501-002-003-002>.
- [6] [The most commonly used objective tests and the levels and outcomes they target]. 170–121، (1)1، مجلة اختبارات اللغة، <https://doi.org/10.60161/2501-001-001-004>.
- [7] ACTFL. (n.d.). Arabic language proficiency tests. American Council on the Teaching of Foreign Languages. <https://www.actfl.org/assessments/actfl-assessments/actfl-postsecondary-assessments>
- [8] AL-ARABIYYA-INSTITUTE. (n.d.). Arabic language proficiency test. <https://www.al-arabiyya-test.org>
- [9] Al-Fanar Media. (n.d.). A new certificate measures Arabic language proficiency. <https://www.al-fanarmedia.org>
- [10] Al-Tanala Al-Arabi Center. (n.d.). Attanala'l Alarabi certification. <https://www.altanala.org>
- [11] ALTA Language Services. (n.d.-a). Language proficiency testing services. <https://www.altalang.com>
- [12] ALTA Language Services. (n.d.-b). Arabic language testing: Dealing with diglossia and other issues. <https://www.altalang.com/arabic-language-testing>
- [13] Arab Academy. (n.d.). ALPT – Arabic Language Proficiency Test online. <https://www.arabacademy.com/alpt>
- [14] AUC School of Continuing Education. (n.d.-a). Career certification in teaching Arabic as a foreign language (CCTAFL). The American University in Cairo. <https://sce.aucegypt.edu/cctafl>
- [15] AUC School of Continuing Education. (n.d.-b). Online Standardized Arabic Proficiency Test (OSAPT). The American University in Cairo. <https://sce.aucegypt.edu/osapt>
- [16] AUC School of Continuing Education. (n.d.-c). Testing services for individuals. The American University in Cairo. <https://sce.aucegypt.edu/testing-services>



- [17] Avant Assessment. (n.d.). Arabic Proficiency Test (APT). <https://www.avantassessment.com/aptarabic>
- [18] Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *Journal of Chiropractic Medicine*, 5(3), 101–117. [https://doi.org/10.1016/S0899-3467\(07\)60142-6](https://doi.org/10.1016/S0899-3467(07)60142-6)
- [19] Institut du monde arabe. (n.d.-a). About CIMA. <https://www.imarabe.org/cima/about>
- [20] Institut du monde arabe. (n.d.-b). CIMA certification. <https://www.imarabe.org/cima>
- [21] Institute of Linguists. (n.d.). Teaching Arabic as a foreign language TAFL. <https://www.iol.org.uk/tafl>
- [22] King Saud University. (n.d.). Arabic language requirements. <https://www.ksu.edu.sa>.
- [23] Practice Aptitude Tests. (2025). Arabic - Arabic Language Proficiency Test (ALPT) assessments: Practice questions. <https://www.practiceaptitudetests.com/arabic-alpt>.
- [24] Ryding, K. C. (2013). *Teaching and learning Arabic as a foreign language: A guide for teachers*. Georgetown University Press.
- [25] The American University in Cairo. (n.d.-a). Teaching Arabic as a foreign language programs. <https://www.aucegypt.edu/tafl>.