Alsina : Journal of Arabic Studies
Vol. 8, No. 2 (2026): 342–355
e-ISSN: 2622-6146; p-ISSN: 2622-6138
DOI: https://doi.org/10.21580/alsina.8.2.32248

Where Alignment Stops: Auditing an Arabic Speaking Syllabus against the IMTA' Proficiency Framework

Irwan Nefri[a*], Nurul Izzati[a], Zubaidah[a], Yelly Putriyani[b]
[a]UIN Sjech M. Djamil Djambek Bukittinggi, Indonesia, [b]Al-Azhar University, Cairo, Egypt
*Corresponding author: irwanabufaqeeb@gmail.com

Article History

Received
22 May 2026
Revised
27 August 2026
Accepted
31 August 2026

Abstract
Arabic is a compulsory component of the curriculum in Indonesian Islamic higher education, yet the syllabus through which it is operationalised is seldom examined against an external proficiency framework, so alignment is asserted in accreditation files rather than demonstrated. This study evaluates the syllabus of a first-semester Arabic speaking course at UIN Sjech M. Djamil Djambek Bukittinggi, Indonesia, against level A1 of the IMTA' framework for teaching Arabic to speakers of other languages. A qualitative content analysis of the course document coded its programme, course, and topic-level outcomes at two separate levels, the substance of the outcome and the operational verb that carries it, and mapped each to the framework's productive indicators. Separating the two levels yields an audit procedure reusable on other course documents, because a document may name the right competences while specifying none in a form that can be taught or assessed; applied before accreditation review, the procedure locates the point at which an international standard reaches the classroom or stops at the level of policy, and it moves the evaluative question from whether a syllabus mentions the standard to whether the standard survives its translation into verbs, activities, and assessment tasks.

Abstrak
Bahasa Arab merupakan komponen wajib kurikulum di perguruan tinggi keagamaan Islam di Indonesia, tetapi silabus mata kuliah yang mengoperasionalkannya jarang diuji terhadap kerangka kemahiran eksternal, sehingga keselarasan lebih sering dinyatakan dalam berkas akreditasi daripada dibuktikan. Penelitian ini mengevaluasi silabus mata kuliah berbicara bahasa Arab semester pertama di UIN Sjech M. Djamil Djambek Bukittinggi terhadap level A1 kerangka IMTA', yaitu kerangka acuan untuk pengajaran bahasa Arab bagi penutur bahasa lain. Penelitian menggunakan analisis isi kualitatif atas dokumen mata kuliah dengan mengodekan capaian pembelajaran program studi, mata kuliah, dan topik pada dua tingkat terpisah, yakni substansi rumusan dan kata kerja operasional yang membawanya, serta memetakan setiap capaian tingkat topik terhadap indikator produktif kerangka tersebut. Pemisahan kedua tingkat itu menghasilkan prosedur audit yang dapat digunakan ulang pada dokumen mata kuliah lain, karena sebuah dokumen dapat menyebut kompetensi yang tepat tanpa merumuskan satu pun dalam bentuk yang dapat diajarkan dan dinilai; jika diterapkan sebelum penilaian akreditasi, prosedur ini menunjukkan titik ketika sebuah standar internasional sampai ke ruang kelas atau berhenti pada tingkat kebijakan, dan menggeser pertanyaan evaluatif dari penyebutan standar menjadi apakah standar itu bertahan ketika diterjemahkan ke dalam kata kerja, aktivitas, dan tugas penilaian.

الملخص
تعد اللغة العربية مكونا إلزاميا فى مناهج التعليم العالى الإسلامى فى إندونيسيا، غير أن الوثيقة التى تترجم هذا المكون إلى مستوى المقرر الدراسى، وهى خطة التعلم الفصلية، قلما تختبر فى ضوء إطار مرجعى خارجى للكفاية، فيكتفى بالإقرار بوجود الموائمة فى ملفات الاعتماد بدل إثباتها. تهدف هذه الدراسة إلى تقويم مقرر مهارة الكلام للمبتدئين فى الفصل الدراسى الأول بجامعة الشيخ محمد جميل جمبيك الإسلامية الحكومية بوكيتينغى، فى ضوء المستوى A1 من الإطار المرجعى إمتاع لتعليم اللغة العربية للناطقين بلغات أخرى. واعتمدت الدراسة تحليل المحتوى الكيفى للوثيقة، مع ترميز مخرجات التعلم على مستوى البرنامج والمقرر والموضوع على مستويين منفصلين: مضمون المخرج، والفعل الإجرائى الذى يحمله، مع مقابلة كل مخرج على مستوى الموضوع بالمؤشرات الإنتاجية فى الإطار. وتتيح الفصل بين مستوى الموائمة إجراء قابلا لإعادة الاستخدام فى تدقيق وثائق المقررات الأخرى، لأن الوثيقة قد تصف الكفايات المناسبة دون أن تصوغ أيا منها صياغة قابلة للتعليم والتقويم؛ واذا طبق هذا الإجراء قبل مراجعة الاعتماد فإنه يبين النقطة التى يصل فيها المعيار الدولى إلى الفصل الدراسى أو يتوقف عند مستوى السياسة، وينقل سؤال التقويم من مجرد ذكر المعيار إلى ما إذا كان المعيار قادرا على البقاء عند ترجمته إلى أفعال وأنشطة ومهام تقويمية.

© 2026 by Alsina : Journal of Arabic Studies. This publication is licensed under a CC BY-NC-SA 4.0

Keywords: Arabic as a foreign language; curriculum alignment; learning outcomes; speaking proficiency; syllabus audit

Cite this article: Nefri, I., Izzati, N., Zubaidah, & Putriyani, Y. (2026). Where alignment stops: Auditing an Arabic speaking syllabus against the IMTA' proficiency framework. Alsina : Journal of Arabic Studies, 8(2), 342–355. https://doi.org/10.21580/alsina.8.2.32248

Introduction

Arabic is more than a means of communication; competence in it conditions access to the Islamic scholarly tradition. In Indonesian Islamic higher education, Arabic is a compulsory component of the curriculum and is expected to serve as a foundational medium for acquiring Islamic knowledge rather than as an isolated subject (Rekan et al., 2025), and research has shown that Arabic instruction strengthens students' understanding of Islam only when objectives, materials, and evaluation are designed around that function (Ritonga et al., 2021). Among the four language skills (*al-mahārāt al-lughawiyyah*), speaking (*mahārat al-kalām*) holds a particularly urgent position, and much recent pedagogical research on Arabic is devoted to developing it (Aburezeq, 2020). It is a productive ability (*al-mahārah al-intājiyyah*) and a tangible indicator of mastery, and it requires an authentic, contextual, and performance-based approach to assessment (Edi & Ibrahim, 2025).

The urgency of *mahārat al-kalām* in Islamic universities rests on two factors. Intellectually, students are expected to discuss Islamic concepts from both classical (*turāth*) and contemporary literature orally. Practically, active communication is needed in a labour market and diplomatic environment where Arabic is increasingly promoted as a language of cultural diplomacy and soft power, not only as a liturgical language (Boulaares, 2025).

However, students' attainment in *mahārat al-kalām* is often still suboptimal. Interviews with Arabic teachers across Indonesian institutions identify low motivation, limited intercultural materials, weak institutional support, and a persistent reliance on grammar-translation methods (Kamal, 2025). Grammar is also the weakest of the speaking components in an Indonesian international class programme, where anxiety depresses oral performance and self-confidence raises it (Bahri et al., 2026). The scarcity of a language environment (*bī'ah lughawiyyah*) compounds both. Strengthening *mahārat al-kalām* is therefore a necessity, so that graduates can not only read texts but also voice moderate and transformative Islamic thought on the world stage.

A problem commonly experienced by second-language learners is speaking. According to Thu'aimah (1986), in the context of Arabic language learning, *mahārat al-kalām* is a core skill that students must master and one of the ultimate goals of foreign language teaching, since it is the ability to communicate oneself to others. One indicator of success in learning a foreign language is therefore the ability to communicate orally (*mahārat al-kalām*), which is the second basic skill after *mahārat al-istimāʿ*. Speaking (*kalām*) is thus a person's expression of what is in their thoughts through the speech organs they possess.

Mahmud Kamil al-Naqah states that the importance of speaking proficiency is evident in the oral aspect of the language itself: speaking is the main aspect of the foreign language curriculum, and most practitioners regard it as the primary goal of foreign language programmes (Al-Naqah, 1985). Tarigan (1991) likewise holds that speaking is, in essence, the ability to pronounce articulated sounds or words to express, state, and convey thoughts, ideas, and feelings. The *Oxford advanced learner's dictionary* defines speaking as "to express or communicate opinions, feelings, ideas, etc by or as talking", involving "psychological, physiological (articulator) and physical (acoustic) stages" (Hornby, 1995, p. 13), and other literature describes it as a "process of building and sharing meaning through the use of verbal and non-verbal symbols, in a variety of contexts" (Chaney & Burke, 1998, p. 13). What these definitions share, and what matters here, is that speaking is defined by what a person does with language rather than by what a person knows about it.

The course syllabus, known in Indonesia as the *Rencana Pembelajaran Semester*, is the operational instrument through which the Indonesian National Qualifications Framework is translated into teaching, and its formulation determines whether the stated outcomes are actually pursued in class (Muhammad & Ariani, 2020). It also serves as a document of accreditation, transparency, and quality audit (Junaidi, 2020). Preparing one is nevertheless difficult: a study of five accredited Arabic language education programmes found that they rarely derived their outcomes independently and instead adjusted existing formulations (Muhammad & Ariani, 2020), and analyses of Indonesian syllabi show that the stated outcomes are frequently misaligned with the intended level of Bloom's taxonomy (Arifiani et al., 2024).

IMTA' is a reference framework published in 2023 by the Educational Center for the Arabic Language of the Gulf States, based in Sharjah, United Arab Emirates (Sharjah 24, 2023). The name is an acronym formed from the Arabic title *al-iṭār al-marjaʿī li-taʿlīm al-lughah al-ʿArabiyyah li-nāṭiqīna bi-lughātin ukhrā*, which translates as "the reference framework for teaching Arabic to speakers of other languages". The framework aims to develop the teaching and learning of Arabic on an educational, scientific, and professional foundation, taking the best technical aspects into account and following international quality standards (Al-Hudaybi, 2023a).

Foreign language teaching therefore needs clear standards for measuring success and attainment. The standards in question are a set of descriptions of the knowledge, skills, and values that learners should acquire as a result of studying one or more parts of a particular curriculum (Al-Hudaybi, 2023b). The adoption of such standards has nevertheless been uneven. The Common European Framework of Reference has been widely accepted as a reference for aligning curricula, learning outcomes, teaching materials, and assessment, yet in Arabic teaching it has been used mainly at the macro level of curriculum and test design and far less at the micro level of classroom teaching and learning (Mohamed, 2023). This asymmetry is where the course syllabus becomes decisive, since it is the point at which a macro-level standard either reaches the classroom or stops at the level of policy.

Method

This study uses a qualitative approach with a descriptive design, since the object of analysis is a single institutional document whose internal structure and formulation are to be described rather than measured. The document examined is the course syllabus for *Mahārat al-Kalām lil-Mubtadi'īn*, the first-semester Arabic speaking course in the undergraduate Arabic Language Education programme at UIN Sjech M. Djamil Djambek Bukittinggi, Indonesia; the Arabic title means "speaking skills for beginners". The syllabus is treated here as the micro-level unit of the curriculum: it is the instrument through which the Indonesian National Qualifications Framework is translated into outcomes, activities, and assessment at course level, and it is therefore the smallest unit at which alignment with an external standard can be observed. Throughout this article the term curriculum refers to this course-level document and not to the study programme as a whole.

The framework against which the document is read is the IMTA' reference framework published in 2023 by the Educational Center for the Arabic Language of the Gulf States. Its entry level is called *al-tahyi'ah*, "the preparatory level", and is designated A1 in the framework's own scheme; that designation is used throughout, since it maps onto the level labels familiar from the Common European Framework of Reference. The course title is retained only as a name, because *lil-mubtadi'īn* describes the course rather than a level of the framework. Level A1 was selected because the course is offered in the first semester to students whose prior exposure to Arabic varies widely, and because it is the level at which the framework states its requirements for oral production most concretely.

Data were collected through documentation study, by systematically examining the content of the syllabus. The analysis proceeded in three stages. The first was descriptive coding, which identified every formulation of the programme, course, and topic-level outcomes and treated each as a unit of analysis. The

second was comparative coding, which assessed each topic-level outcome on two dimensions held separately. On substance, an outcome was conforming when its topic fell within the thematic range of the A1 descriptors and it named a situation in which the topic is to be used, partially conforming when the topic fell within that range but no situation was named, and non-conforming when it fell outside. On operationalisation, the same outcome was conforming when its verb named an observable oral performance, partially conforming when the verb named a performance not observable in speech, and non-conforming when the verb belonged to the cognitive domain. The two dimensions were kept apart because a document may name the right competences while specifying none in teachable form, and a composite judgement would conceal that difference.

For the mapping, only the productive indicators of level A1 were used, that is, those requiring the learner to produce spoken language. Indicators requiring comprehension of heard or written input were recorded but excluded from the conformity judgement, since a speaking outcome cannot be assessed against a listening or reading demand. The classification of each indicator by skill domain is the authors' own coding rather than a category supplied by the framework, and it is shown in Table 3.

The learning activities specified for each topic were recorded and classified as rehearsal (dialogue memorisation, pattern drills) or transfer (a task requiring the learner to speak without a rehearsed script). Because the document lists activities per topic rather than allocating time, the composition is reported as the presence or absence of a transfer task, and no percentage is claimed. The third stage was gap analysis and reconstruction, which set the coded findings against the A1 standard on four curriculum aspects (Table 4). Coding followed a written codebook so that the procedure can be repeated, and the full coding is reproduced in Table 1 rather than summarised. The reconstructed design in Tables 5 and 6 is a proposal: neither the matrix nor the rubric was trialled with students or raters. The document is institutional and contains no personal data.

Results

Analysis of the Syllabus for Arabic Speaking

The programme and course outcomes are quoted verbatim below. Both are formulated as general competence rather than observable performance, which is why the topic-level outcomes carry the analytical weight in what follows.

Programme learning outcome: Students are able to understand and implement themes related to everyday life in communicating in Arabic actively and creatively.

Course learning outcome: This course is designed to develop students' ability to communicate in Arabic well and correctly.

The analysis proceeded by examining the topic-level learning outcomes in the existing Arabic speaking syllabus and comparing them with the A1 indicators. The coding focused on the extent to which each outcome specifies observable communicative performance, including the communicative function, linguistic resources, and situational context required at the A1 level. Table 1 presents the coding results, identifying the alignment between the existing outcomes and the A1 indicators.

Table 1 Coding of the Existing Topic-Level Outcomes against the A1 Indicators

No	Existing topic-level outcome	Verb	Domain	Substance	Operationalisation
1	Able to understand *al-ḥurūf al-ʿarabiyyah*, *al-taʿāruf*, and *al-taḥiyyāt*	understand	Cognitive (C2)	Partially conforming	Non-conforming
2	Able to understand *māddat al-buldān wa-al-lughāt wa-al-safar*	understand	Cognitive (C2)	Partially conforming	Non-conforming
3	Able to understand *māddat al-safar wa-al-muwāṣalāt wa-al-ʿamal*	understand	Cognitive (C2)	Partially conforming	Non-conforming

No	Existing topic-level outcome	Verb	Domain	Substance	Operationalisation
4	Able to understand *māddat al-ittiṣālāt wa-al-ʿilah wa-nashāṭāt* yawmiyyah	understand	Cognitive (C2)	Partially conforming	Non-conforming
5	Able to understand *māddat al-malābis wa-al-ṭaʿām wa-al-wajbāt wa-al-waqt*	understand	Cognitive (C2)	Partially conforming	Non-conforming
6	Able to understand *māddat al-ṣifāt wa-al-manzil wa-al-tasawwuq*	understand	Cognitive (C2)	Partially conforming	Non-conforming
7	Able to understand *māddat al-mustaqbal wa-aqsām al-sanah wa-al-wazn wa-al-ṭūl*	understand	Cognitive (C2)	Partially conforming	Non-conforming
8	Able to understand *māddat al-rīf wa-al-madīnah wa-al-mashāʿir*	understand	Cognitive (C2)	Partially conforming	Non-conforming
9	Able to understand *māddat al-ḥawāss wa-jism al-insān*	understand	Cognitive (C2)	Partially conforming	Non-conforming
10	Able to understand *aḥwāl al-jaww wa-al-ḥayawānāt*	understand	Cognitive (C2)	Partially conforming	Non-conforming
11	Able to understand *māddat al-funduq wa-jawlah siyāḥiyyah wa-al-munāsabāt*	understand	Cognitive (C2)	Partially conforming	Non-conforming
12	Able to understand *al-ṣiḥḥah wa-al-muwāfaqah*	understand	Cognitive (C2)	Partially conforming	Non-conforming

Table 1 reports the coding of the twelve topic-level outcomes. On substance the entries are uniformly partially conforming: every topic falls inside the thematic range the A1 descriptors designate, and none names a situation in which the topic is to be used. On operationalisation none conforms. The same verb, understand, carries all twelve entries, so the document yields one recurring formulation rather than twelve distinct outcomes, and the verb belongs to the cognitive rather than the psychomotor domain (Simpson, 1972). The distribution is therefore 0 conforming, 12 partially conforming, and 0 non-conforming on substance, against 12 non-conforming on operationalisation. Table 2 sets out a revised formulation for each entry; the two columns together explain why the revision is confined to the verb.

Table 2 Revised Topic-Level Outcomes

No	Revised topic-level outcome
1	Able to pronounce *al-ḥurūf al-ʿarabiyyah* and to use greetings and expressions of introduction in *al-taʿāruf* and *al-taḥiyyāt* accurately and fluently
2	Able to hold a short conversation about *al-buldān wa-al-lughāt wa-al-safar* accurately and fluently
3	Able to hold a short conversation about *al-safar wa-al-muwāṣalāt wa-al-ʿamal* accurately and fluently
4	Able to hold a short conversation about *al-ittiṣālāt wa-al-ʿilah wa-nashāṭāt* yawmiyyah accurately and fluently
5	Able to hold a short conversation about *al-malābis wa-al-ṭaʿām wa-al-wajbāt wa-al-waqt* accurately and fluently
6	Able to describe *al-ṣifāt wa-al-manzil wa-al-tasawwuq* in a short conversation accurately and fluently
7	Able to hold a short conversation about *al-mustaqbal wa-aqsām al-sanah wa-al-wazn wa-al-ṭūl* accurately and fluently
8	Able to hold a short conversation about *al-rīf wa-al-madīnah wa-al-mashāʿir* accurately and fluently
9	Able to describe *al-ḥawāss wa-jism al-insān* in a short conversation accurately and fluently
10	Able to hold a short conversation about *aḥwāl al-jaww wa-al-ḥayawānāt* accurately and fluently
11	Able to hold a short conversation about *al-funduq wa-jawlah siyāḥiyyah wa-al-munāsabāt* accurately and fluently
12	Able to hold a short conversation about *al-ṣiḥḥah wa-al-muwāfaqah* accurately and fluently

Note. Each entry retains the topic of the corresponding row in Table 1 and replaces only the cognitive verb with one naming an observable oral performance. Entries 2 to 12 therefore differ from one another only in the topic field.

Table 3 reproduces the achievement indicators of level A1 and classifies each by the skill domain it addresses. Four of the ten indicators require the learner to produce spoken language, one is perceptual, and five require comprehension of input. The mapping in Table 1 uses the four productive indicators only.

Table 3 Achievement Indicators of IMTA' Level A1 (*al-Tahyi'ah*), Classified by Skill Domain

No	Achievement indicator	Skill domain
1	Students are able to pronounce Arabic sounds.	Productive oral
2	Learner-students are able to become accustomed to the pronunciation and discrimination of Arabic sounds, and to use them in communication through short vocabulary and phrases to express basic needs in common situations that have been extensively practiced and memorized. Learners can also control accuracy in those memorized and practiced contexts and can be understood by people who are accustomed to interacting with language learners.	Productive oral
3	Students are able to distinguish between similar sounds and between sounds whose points of articulation (*makhraj*) are close, and to distinguish long vowels (*madd*) from short vowels.	Perceptual
4	Students are able to pronounce/write simple, isolated vocabulary, phrases, and short sentences with the support of some pictures, relating to their personal environment such as their residence, family, and classroom. At this level the language cannot yet be understood, even by listeners or readers who are accustomed to the language of such learners, because they are still in the stage of trying to master the linguistic system of Arabic.	Productive oral
5	Students are able to construct very simple, short sentences using a demonstrative (*ism ishārah*) or a singular pronoun (*ḍamīr mufrad*). They are also able to carry out some simple instructions accompanied by body gestures that clarify them, or by a tone of voice that distinguishes them.	Productive oral
6	Students are able to understand very short, easy, and simple vocabulary, questions, and phrases, provided that they are spoken slowly and clearly and accompanied by body gestures that support comprehension, with repetition if necessary. They are also able to recognize common everyday words or signs, provided that they are presented clearly and slowly in a familiar and specific daily context. In addition, they can recognize numbers, prices, dates, and the days of the week, provided that the information is delivered slowly and clearly in a familiar and specific daily context.	Receptive
7	Students are able to understand short, easy instructions for actions such as: read, write, listen, close the door, and so on. This refers to learners' ability to respond to basic imperative verbs frequently used in classroom situations or very simple everyday interactions.	Receptive
8	Students are able to understand simple, short dialogues (whether heard or read) that are delivered very slowly and repeatedly and contain concrete vocabulary in common everyday-life situations.	Receptive
9	Students are able to use the vocabulary they have learned (concrete vocabulary / vocabulary with closed meaning) to infer basic information from a text.	Receptive
10	At the end of this level, learners are able to follow simple conversations or reading texts that are spoken accurately and clearly and at a speed appropriate to the learner's ability level, paying attention to correct punctuation. They are also able to recognize information related to general topics encountered in everyday life.	Receptive

Note. Adapted from Al-Hudaybi (2023a, p. 79). The domain labels are the authors' classification. Only the indicators marked productive oral were used in the conformity judgement reported in Table 1.

Evaluation against the A1 Standard

Table 4 sets the findings against the A1 standard on four curriculum aspects. Only the first is aligned in substance; the remaining three are where the operational gap becomes visible in activities, materials, and assessment.

Table 4 Evaluation and Reconstruction Matrix of the Arabic Speaking Syllabus

No	Curriculum Aspect	Actual findings in the existing syllabus	Ideal attainment standard of IMTA' (level A1)	Gap Analysis	Direction for development and reconstruction
1	Learning outcomes	The formulation points toward the ability to introduce oneself and to engage in simple daily interaction.	Able to communicate in a limited way on general topics close to daily life.	Conforms in substance. The macro-level conceptual foundation of the syllabus is already aligned with the framework.	Retained. The focus of the outcomes is directed at strengthening indicators of independent language performance.
2	Learning Activities	Dialogue memorisation (*ḥifẓ al-ḥiwārāt*) and pattern drills (*tardīd al-anmāṭ*) for every topic; no transfer task specified.	Requires spontaneous interaction (*al-tafāʿul al-tilqāʾī*) in real contexts.	Methodological gap. The dominance of mechanical instruction limits students' room to produce language adaptively.	Activity redesign. Reduce the share of mechanical drills and add a transfer stage in which the memorised repertoire must be used.
3	Teaching Materials	Fixed on artificial dialogue texts inside the classroom.	Uses an authentic communication approach (*tawāṣuliyyah ḥaqīqiyyah*).	Contextual limitation. The materials do not yet explore dynamic real social situations outside the text.	Material enrichment. Integrate problem-based and real-situation materials that require the memorised repertoire to be used.
4	Assessment Instruments	Tend to test memorization of texts or normative structural tests.	Based on real performance assessment (*al-adāʾ*).	Assessment mismatch. Assessment cannot yet measure students' readiness for the next threshold (A2).	Assessment reorientation. Apply performance-based assessment through semi-scripted communicative tasks assessed with a rubric.

Note. The four aspects correspond to the curriculum components named in the Method. Aligned in substance means the aspect names the required content; the other three judgements concern whether the aspect requires that content to be produced.

Development of the Reconstructed Syllabus Based on IMTA'

The reconstruction proposed here is a conceptual model resulting from a gap analysis between the syllabus and the ideal attainment standard of IMTA' level A1, as presented in Table 4. Its three pillars are design recommendations derived from the coding reported in Table 1, and they have not been tested in classroom practice. The contribution therefore lies in the reconstruction procedure and in the two-dimensional coding that supports it; whether the design raises students' competence toward A2 is left to a further study, preferably a pilot in the same programme.

Developing an Arabic curriculum for non-Arabs must take into account five standards: linguistic, psycholinguistic, educational, knowledge, and cultural (Al-Hudaybi, 2023a). The reconstruction works on the first and the third, since it concerns the formulation of outcomes and the design of the tasks that carry them. The topic-level outcomes were rewritten with operational verbs that name observable speech, so that an outcome oriented toward understanding became one oriented toward performance. The same constraint applies at course level: a verb that cannot be observed in speech cannot be assessed in a speaking course. Psychomotor verbs such as pronouncing, demonstrating, practising, presenting, and creating serve this purpose (Simpson, 1972).

For example: (1) students are able to pronounce *al-ḥurūf al-ʿarabiyyah* in the material on *al-taʿāruf* and *al-taḥiyyāt* well and fluently; and (2) students are able to practice conversations on the topic of *al-buldān wa-al-lughāt wa-al-safar* well and fluently.

To move students past the A1 threshold toward A2, the reconstruction rests on three pillars. First, activity redesign: from rehearsal to graduated transfer. Memorisation is retained, because the level requires it,

but each topic gains a stage at which the memorised repertoire must be used. The stage begins with substitution inside a familiar frame, moves to a semi-scripted exchange in which a situation card fixes the context but not the wording, and ends with a short improvised task. This modifies task-based design rather than abandoning drill: the drill supplies the material, the task requires it to be spent (Richards, 2006).

In application, the third session of each topic is given over to the semi-scripted exchange, and students record a two-minute interaction outside class in which the situation, not the text, is fixed. A speaker of Arabic may be the partner where one is available, otherwise a peer. The condition is that the wording cannot be prepared in advance, so the task tests use rather than recall.

Second, Enrichment of teaching materials: integration of authentic communication materials (*tawāṣuliyyah ḥaqīqiyyah*). The syllabus must update the material component, which has so far been artificial (constructed dialogue texts), with texts or communicative media produced by Arab speakers. Authentic materials at beginner level increase motivation and provide natural exposure, although their comprehension demands scaffolding if the gain in authenticity is not to be offset by a loss of comprehensibility (Gilmore, 2007).

To apply this, lecturers integrate short video clips from Arab vloggers, Arabic-language advertisements, or social media infographics containing everyday conversation, and pair each clip with a comprehension task and a reproduction task, so that the input is used rather than merely shown.

Third, assessment reorientation: applying performance-based assessment (*al-adā'*). Final assessment changes from testing memory to testing productive performance, and the instrument is designed to detect the learner's ability to speak in a situation whose wording has not been rehearsed.

To apply this, the assessment component is replaced with performance-based rubrics. The final examination is not a group dialogue recitation but a semi-scripted role-play in which the student draws a situation card during the examination and must respond within the vocabulary and structures already covered, or a short oral presentation about social identity assessed on fluency, appropriateness of context, and willingness to speak. Assessment resting on observed performance rather than recognition of memorised text is well established (Sanjaya & Hidayat, 2022), and a rubric of this kind makes the transfer stage visible to students before the examination.

Coherence of the Reconstructed Design

The three pillars are stated separately, and their value depends on whether they hold together. Table 5 sets each reconstructed outcome beside the activity that develops it and the task that evidences it, so that the chain from statement to performance can be inspected in one view. It also makes a narrower point visible: the assessment task is where an outcome either becomes operative or remains declarative, since a topic whose outcome names a performance but whose task asks for recall is misaligned at the last step even when the first two are in order.

Table 6 gives the instrument that scores those tasks. Four criteria were retained because they separate what the framework requires from what it leaves open: fluency and accuracy are named in the A1 descriptors, appropriateness follows from their insistence on common situations, and willingness to speak replaces the vaguer notion of courage, naming an observable disposition rather than a trait. Each criterion has four bands. Band 3 is the target, because it restates the framework's own condition that the learner be understood by listeners accustomed to learner language. Band 4 records greater control within the same repertoire rather than movement toward A2, so the scale does not reward the imitation of a higher level.

Two cautions attach. The descriptors are a design proposal, not a validated scale: no estimate of inter-rater agreement is available, so they should be treated as a starting point for moderation rather than as a finished

instrument. The bands also describe speech only, and a course assessing the same outcomes through writing would need a different instrument.

Table 5 Constructive Alignment of the Reconstructed Syllabus

No	Topic in the revised outcome (Table 2)	Transfer stage	Assessment task (situation card)	Criteria
1	*al-ḥurūf al-ʿarabiyyah*, *al-taʿāruf*, *al-taḥiyyāt*	Substitution, then semi-scripted introduction	Introduce yourself to a new classmate (first day of class)	F, A, Ap, W
2	*al-buldān wa-al-lughāt wa-al-safar*	Semi-scripted exchange	Say where you are from and which languages you speak (meeting a visitor)	F, A, Ap
3	*al-safar wa-al-muwāṣalāt wa-al-ʿamal*	Semi-scripted exchange	Ask for and give travel information (at the station)	F, A, Ap
4	*al-ittiṣālāt wa-al-ʿilah wa-nashāṭāt* yawmiyyah	Semi-scripted exchange	Telephone a friend to arrange a meeting (phone call)	F, A, Ap, W
5	*al-malābis wa-al-ṭaʿām wa-al-wajbāt wa-al-waqt*	Semi-scripted exchange	Order a meal and ask the time (at a restaurant)	F, A, Ap
6	*al-ṣifāt wa-al-manzil wa-al-tasawwuq*	Semi-scripted exchange	Describe your home and buy an item (at the market)	F, A, Ap
7	*al-mustaqbal wa-aqsām al-sanah wa-al-wazn wa-al-ṭūl*	Semi-scripted exchange	State next year's plans and describe a person (talking about the future)	F, A, Ap
8	*al-rīf wa-al-madīnah wa-al-mashāʿir*	Semi-scripted exchange	Compare village and city life and name a preference (where would you rather live)	F, A, Ap, W
9	*al-ḥawāss wa-jism al-insān*	Semi-scripted exchange	Describe a symptom to a pharmacist (at the pharmacy)	F, A, Ap
10	*aḥwāl al-jaww wa-al-ḥayawānāt*	Semi-scripted exchange	Make small talk about the weather and an animal (waiting for a bus)	F, A, Ap, W
11	*al-funduq wa-jawlah siyāḥiyyah wa-al-munāsabāt*	Semi-scripted exchange	Check into a hotel and ask about a tour (hotel reception)	F, A, Ap
12	*al-ṣiḥḥah wa-al-muwāfaqah*	Semi-scripted exchange	Accept or decline an invitation and give a reason (invitation to an event)	F, A, Ap, W

Note. F = fluency, A = accuracy, Ap = appropriateness, W = willingness to speak; the criteria are defined in Table 6. Each topic is assessed through the same semi-scripted procedure, in which the situation card fixes the context but not the wording.

Table 6 Analytic Rubric for the Semi-Scripted Speaking Task

Criterion	4	3 (target for A1)	2	1
Fluency	Sustains several connected turns with little hesitation; pauses fall at clause boundaries; no prompting needed.	Completes the exchange with noticeable pauses and some self-repair; occasional repetition or rephrasing required.	Completes the exchange only with prompts from the interlocutor; frequent long pauses and restarts.	Cannot sustain the exchange; responses limited to isolated words or to memorised fragments used out of context.
Accuracy	Uses the structures and vocabulary covered so far with control; errors are few and do not disturb meaning.	Reproduces the memorised patterns accurately and extends them with minor errors that leave meaning clear.	Errors are frequent and occasionally obscure meaning; memorised patterns are reproduced with distortion.	Errors prevent communication; the memorised patterns are not retrievable in usable form.
Appropriateness	Selects the formulaic expressions and register that the situation requires, and adapts them to the interlocutor.	Uses the formulaic expressions expected in the situation, with occasional mismatch.	Uses expressions belonging to a different situation, or produces them without regard to the interlocutor.	No recognisable fit between the utterance and the situation.

Willingness to speak	Initiates turns and volunteers information without prompting.	Responds readily and occasionally initiates.	Responds only when directly addressed, with minimal contribution.	Avoids speaking, or falls silent when the turn passes.

Note. Band 3 restates the A1 condition that the learner be understood by listeners accustomed to learner language (Al-Hudaybi, 2023a, p. 79). Band 4 describes control within the A1 repertoire, not progression to A2. The scale is a design proposal and has not yet been tested with independent raters.

Discussion

The coding reported in Table 1 produces a result that resists a single verdict, so the discussion proceeds in two movements. The first asks whether the document's reliance on memorisation is the defect; the second takes the operational gap as established and asks what it implies for the course's stated destination, the A2 threshold.

Memorisation as Threshold and as Ceiling

The case for the document's design rests on the framework it is measured against. The A1 descriptors reproduced in Table 3 do not treat memorisation as a departure from communicative teaching. Indicator 2 requires learners to express basic needs "in common situations that have been extensively practiced and memorized" and to control accuracy "in those memorized and practiced contexts", while indicator 4 records that the learner's language "cannot yet be understood, even by listeners or readers who are accustomed to the language of such learners". A beginner course built around dialogue memorisation (*ḥifẓ al-ḥiwārāt*) and pattern drills (*tardīd al-anmāṭ*) therefore works with the logic of the level rather than against it. The *pesantren* tradition supplies a second line of support: an integrated curriculum that retains grammatical and textual instruction through one-to-one recitation (*sorogan*) and collective text reading (*bandongan*) alongside dialogue (*ḥiwār*) and speaking drills (*muḥādaṯah*) has been described as developing classical reading and active use together (Mukminin et al., 2025). On this reading the document's method is an inheritance rather than a lag.

The case against is narrower than the charge of being mechanical, and stronger for being narrower. The same descriptors attach memorisation to use: the repertoire is memorised in order to be deployed in "common situations", and accuracy is to be controlled "in those memorized and practiced contexts", so a memorised repertoire is a condition of the level rather than its completion. What the document omits is the step at which the repertoire must leave the script. Krashen's (1982) monitor explains why that step cannot be taken for granted: conscious knowledge of a rule can be consulted only when the speaker has time, is attending to form, and knows the rule, and none of these conditions holds in an unrehearsed exchange. A repertoire that exists only as monitored knowledge does not transfer on its own, and the transfer has to be designed. No topic specifies a task in which the learner must speak without a rehearsed text, and the assessment instruments follow the same pattern, testing recognition rather than performance. The difficulty is therefore not that memorisation occurs but that nothing requires it to be spent.

Once the two readings are set side by side, the apparent paradox dissolves. The document is not conceptually right and operationally adequate, nor wrong on both counts; it is conceptually inside the standard and operationally outside it, which is exactly what the two-dimensional coding records. Substance is partially conforming in twelve of twelve entries because the topics are the everyday-life fields the framework names, and operationalisation is non-conforming in twelve of twelve because a single cognitive verb carries them all. Memorisation is the threshold of A1 and, without a transfer stage, also its ceiling. The defect lies in the missing hinge between the two, and that hinge is a matter of design rather than doctrine.

What the Operational Gap Implies for Progression to A2

The first implication concerns how such documents are audited. Because the syllabus names the framework's competences and covers its thematic fields, it would pass a review that checks for the presence of the required components, which is what accreditation instruments ordinarily do. The gap becomes visible only when the verb is read separately from the topic, and this is why the procedure matters beyond the present

case: a document can be complete in form and inert in function, and the second condition is invisible to a checklist that counts components. Internal audits should therefore test operationalisation, not only coverage.

The second implication concerns assessment, and follows deductively from the first. If an outcome is stated as understanding, an instrument that tests understanding satisfies it, so the assessment misalignment recorded in Table 4 is entailed by the verb-level misalignment rather than independent of it. This matters because assessment is where a document's intentions become visible to students. A course whose outcomes ask for understanding and whose examination asks for recognition of a memorised dialogue is internally consistent and externally misaligned at once. Reorienting assessment without rewriting the outcomes would produce a document that contradicts itself, so the two changes have to be made together. Table 6 makes the consequence concrete: accuracy can be scored on memorised reproduction alone, so the rubric must describe transfer explicitly if it is not to certify the behaviour the outcomes were rewritten to move beyond.

The third implication concerns the route to A2, where the psychological evidence argues for a graduated rather than an abrupt transfer. Anxiety interferes with the cognitive processing stages through which a second language is encoded and retrieved (MacIntyre & Gardner, 1994), so its effect is greatest where retrieval must be improvised. Oral anxiety among learners of Arabic is moderately high and correlates negatively with attainment (Ismail et al., 2023), and in an Indonesian international class programme it has been found to depress speaking performance while self-confidence raises it (Bahri et al., 2026). If the only rehearsed context is the classroom script, the first unrehearsed encounter coincides with the moment of assessment, precisely the condition under which anxiety is highest. A defensible design therefore moves from substitution within a familiar frame, to a semi-scripted exchange, and only then to a short improvised task. Fieldwork of the kind a strong reading of communicative teaching would suggest exceeds what the A1 descriptors support, since indicator 4 states that the learner's speech is not yet comprehensible to ordinary listeners.

Set against the wider literature, this result both agrees and diverges. It agrees with the finding that the CEFR has been absorbed at the macro level of curriculum and test design far more than at the micro level of teaching and learning (Mohamed, 2023), and with the observation that Indonesian Arabic programmes tend to adopt existing outcome formulations rather than derive their own (Muhammad & Ariani, 2020); the document examined here is an instance of both, and it shows what they look like at the level of a single verb. It also sits inside a wider shift in which Indonesian Islamic higher education has begun to align Arabic teaching with externally benchmarked proficiency levels (Musthofa, 2022) while course design in Arabic for specific purposes has moved toward the communicative demands of the target domain (Abdul Ghani et al., 2019). It diverges, however, from accounts that treat the reconstruction of objectives, materials, implementation, and evaluation as a single package (Ritonga et al., 2021), since here the objectives are already adequate in substance and the problem lies downstream. Where objectives have been aligned to a standard but the verbs have not, reform should begin at the verb and the task, not at the statement of competence.

The discussion thus returns to its opening question with an answer. Memorisation is not the defect and its removal is not the remedy; the defect is the absence of a stage at which memorised language must be used, and the remedy is to write that stage into the outcomes, the activities, and the assessment together. Read this way, the two-dimensional result is not a split verdict but a diagnosis of where the work remains to be done.

Conclusion

The main finding is that the misalignment of the Arabic speaking syllabus with IMTA' level A1 is operational rather than conceptual. The document names the competences the framework requires and works within its thematic fields, but it states every one of them through a single cognitive verb and specifies no task in which the memorised repertoire must be used. The course can therefore satisfy its own outcomes without producing the performance those outcomes describe. Memorisation is not the problem; the absence of a stage at which memorised language has to be spent is.

Three limitations bound this claim. The analysis rests on a single document from one programme, so it establishes what one document does rather than how widespread the pattern is. It reads the document and not the classroom, which means the gap identified is a gap in design and nothing is claimed about what lecturers do with it in practice. And the reconstruction proposed here has not been implemented, so its effect on progression from A1 to A2 remains untested.

Two recommendations follow for the programme. The topic-level outcomes should be rewritten with verbs that name observable speech and should attach each topic to a situation in which it is used, and the assessment instruments should be revised in the same operation, so that the document does not state one demand and test another. The rubric in Table 6 is a starting point for that revision rather than a finished instrument, and its descriptors require moderation by two raters before it is used for grading. A further recommendation concerns users of the framework: when IMTA' is applied to a speaking course, productive indicators should be separated from receptive ones before any conformity judgement, since a mixed set will overstate or understate alignment.

Whether the reconstructed design works is an empirical question, and it is the one worth asking next. A semester-long pilot, with oral performance assessed before and after and anxiety measured alongside it, would show whether a transfer stage moves learners toward A2 or merely adds tasks to a crowded syllabus. Replication across programmes would show whether the operational gap is a property of this document or a habit of syllabus writing in Indonesian Islamic higher education more broadly, and of the two findings the second would be the more useful.

Acknowledgments

The authors would like to acknowledge the collaboration and contributions of all co-authors throughout the preparation and completion of this manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of AI Use

The authors used ChatGPT (OpenAI) as an AI-assisted tool to help formulate, revise, and refine the language of the manuscript to improve clarity and align it with academic writing conventions. The use of AI was limited to language-related assistance. The authors reviewed and evaluated the AI-assisted output and take full responsibility for the content and final version of the manuscript.

References

Abdul Ghani, M. T., Wan Daud, W. A. A., & Ramli, S. (2019). Arabic for specific purposes in Malaysia: A literature review. *Issues in Language Studies*, *8*(1), 1–15. https://doi.org/10.33736/ils.1293.2019

Aburezeq, I. M. (2020). The impact of flipped classroom on developing Arabic speaking skills. *The Asia-Pacific Education Researcher*, *29*(4), 295–306. https://doi.org/10.1007/s40299-019-00483-z

Al-Hudaybi, A. A. al-M. (2023a). *Al-Iṭār al-marjaʿī li-taʿlīm al-lughah al-ʿArabiyyah li-nāṭiqīna bi-lughātin ukhrā* [The reference framework for teaching Arabic to speakers of other languages]. Al-Markaz al-Tarbawī lil-Lughah al-ʿArabiyyah li-Duwal al-Khalīj.

Al-Hudaybi, A. A. al-M. (2023b). *Al-Maʿāyīr taʿlīm al-lughah al-ʿArabiyyah li-nāṭiqīna bi-lughātin ukhrā* [Standards for teaching Arabic to speakers of other languages]. Majmaʿ al-Malik Salmān al-ʿĀlamī lil-Lughah al-ʿArabiyyah.

Al-Naqah, M. K. (1985). *Taʿlīm al-lughah al-ʿArabiyyah lil-nāṭiqīn bi-lughāt ukhrā* [Teaching Arabic to

speakers of other languages]. Maʿhad al-Lughah al-ʿArabiyyah.

Arifiani, S., Yunanto, A. A., & Rochimah, S. (2024). Leveraging similarity algorithms for comparative analysis of learning plan documents and standardization against Bloom's taxonomy. In *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)* (pp. 204–209). IEEE. https://doi.org/10.1109/ISRITI64779.2024.10963406

Bahri, R. B., Hula, I. R. N., Pakaya, N. A., Abu Samad, M., & Arrashedy, Y. A. M. G. (2026). The effect of anxiety and self-confidence on Arabic speaking proficiency in the International Class Program (ICP). *Alsina : Journal of Arabic Studies*, *8*(1), 1–34. https://doi.org/10.21580/alsina.8.1.29134

Boulaares, D. M. (2025). Strategic horizons of Arabic: A comparative and geopolitical analysis of Saudi Arabia's role in global language promotion. *Current Issues in Language Planning*, 1–15. https://doi.org/10.1080/14664208.2025.2548644

Chaney, A. L., & Burke, T. L. (1998). *Teaching oral communication in grades K-8*. Allyn & Bacon.

Edi, C., & Ibrahim, N. (2025). Implementasi penilaian non-tes mahārah kalām: Studi kasus multisitus di perguruan tinggi [Implementation of non-test assessment of mahārah kalām: A multi-site case study in higher education]. *Alsina : Journal of Arabic Studies*, *7*(1), 1–28. https://doi.org/10.21580/alsina.7.1.26108

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, *40*(2), 97–118. https://doi.org/10.1017/S0261444807004144

Hornby, A. S. (1995). *Oxford advanced learner's dictionary* (5th ed.; J. Crowther, Ed.). Oxford University Press.

Ismail, Z., Rasit, N., & Supriyatno, T. (2023). Relationship between oral language anxiety and students' Arabic language learning outcomes in Malaysian secondary schools. *International Journal of Language Education*, *7*(1), 143–161. https://doi.org/10.26858/ijole.v1i1.37368

Junaidi, A. (2020). *Penyusunan kurikulum pendidikan tinggi di era industri 4.0 untuk mendukung merdeka belajar-kampus merdeka* [Developing higher education curricula in the industry 4.0 era to support Merdeka Belajar–Kampus Merdeka]. Direktorat Jenderal Pendidikan Tinggi, Kementerian Pendidikan dan Kebudayaan.

Kamal, H. (2025). Teaching Arabic today: Challenges, strategies, and opportunities in Islamic higher education. *International Journal of Learning, Teaching and Educational Research*, *24*(10), 644–659. https://doi.org/10.26803/ijlter.24.10.31

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, *44*(2), 283–305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

Mohamed, S. (2023). The development of an Arabic curriculum framework based on a compilation of salient features from CEFR level descriptors. *The Language Learning Journal*, *51*(1), 33–47. https://doi.org/10.1080/09571736.2021.1923781

Muhammad, I., & Ariani, S. (2020). The development of KKNI-based curriculum at the Arabic language education programs in Indonesian higher education. *Jurnal Ilmiah Peuradeun*, *8*(3), 451–474. https://doi.org/10.26811/peuradeun.v8i3.543

Mukminin, A., Hanun, A., Maskuri, Wassalwa, A., & Sani, D. (2025). Integration of Arabic language curriculum in Salaf and Khalaf. *Alsina : Journal of Arabic Studies*, *7*(1), 143–160. https://doi.org/10.21580/alsina.7.1.27631

Musthofa, T. (2022). CEFR-based policy in Arabic language teaching and cultural dimension in Indonesian Islamic higher education. *Eurasian Journal of Applied Linguistics*, *8*(2), 96–107. https://ejal.info/menuscript/index.php/ejal/article/view/272

Rekan, A. A., Howell, H. R., Mohamad Salleh, M. N., Rosin, N., Tabrani, Z. A., Ma'arif, M. A., & Adnan, M. A. M. (2025). Arabic language curriculum as a foundation for strengthening religious education in

public higher education. *Jurnal Pendidikan Agama Islam*, *22*(1), 97–121. https://doi.org/10.14421/jpai.v22i1.11340

Richards, J. C. (2006). *Communicative language teaching today*. Cambridge University Press.

Ritonga, M., Widodo, H., Munirah, & Nurdianto, T. (2021). Arabic language learning reconstruction as a response to strengthen Al-Islam studies at higher education. *International Journal of Evaluation and Research in Education*, *10*(1), 355–363. https://doi.org/10.11591/ijere.v10i1.20747

Sanjaya, B., & Hidayat, W. (2022). Student speaking skill assessment: Techniques and results. *International Journal of Evaluation and Research in Education*, *11*(4), 1741–1748. https://doi.org/10.11591/ijere.v11i4.22782

Sharjah 24. (2023, November 11). "Al-Tarbawī lil-lughah al-ʿArabīyah" yuṭliqu iṭāran marjiʿīyan ʿālamīyan li-taʿlīm al-lughah al-ʿArabīyah lil-nāṭiqīna bi-ghayrihā [The Educational Centre for Arabic launches a global reference framework for teaching Arabic to speakers of other languages]. Sharjah24. https://sharjah24.ae/ar/Articles/2023/11/11/aa20

Simpson, E. J. (1972). *The classification of educational objectives in the psychomotor domain*. Gryphon House.

Tarigan, H. G. (1991). *Berbicara sebagai suatu keterampilan berbahasa*. Angkasa.

Thu'aimah, R. A. (1986). *Al-Marjaʿ fī taʿlīm al-lughah al-ʿArabiyyah lil-nāṭiqīn bi-lughāt ukhrā* [The reference for teaching Arabic to speakers of other languages]. Jāmiʿat Umm al-Qurā.