



Arabic Vocabulary Test Standards for Beginners According to Common European Framework of Reference for Languages (CEFR)

Ihwan Mahmudi, Syifa Firdausi Putri, Diana Fitri Islami,
Khansa Nabila, Raihana Nur Azizah

Universitas Darussalam Gontor, Indonesia

Corresponding E-mail: 452024817043@student.unida.gontor.ac.id

Abstract

Vocabulary is a fundamental component in second language learning, especially for beginners. However, Arabic vocabulary assessment has not been fully standardized according to the Common European Framework of Reference for Languages (CEFR). This study aims to identify vocabulary test criteria for Arabic at the beginner level aligned with CEFR. This research applied a descriptive qualitative method with a literature study approach. Primary data were collected from vocabulary test items in *al-Ikhtibārāt al-Lughaniyyah* by Ali Khuly, supported by secondary data on CEFR and language assessment. The results that A1 learners should master approximately 500–600 high-frequency words within themes such as family, time, numbers, daily routines, colors, animals, clothing, weather, and food. Effective test formats include fill-in-the-blank, picture-based multiple-choice, definition-based questions, synonym/antonym tasks, and word usage in context. An ideal proportion of 60% receptive and 40% productive tasks is recommended for beginner-level learners. It is concluded that Arabic vocabulary tests must prioritize validity, reliability, and practicality, while aligning with CEFR for international comparability.

Keywords: *assessment, beginner level, CEFR, Arabic vocabulary test*

Introduction

Vocabulary plays a central role in foreign language learning, as it forms the foundation of all communicative skills (Setiyadi, et al., 2025). Mastering vocabulary is one of the fundamental aspects of language learning (Jumatiah & Helalsah, 2021). Vocabulary serves as the main base for language skill development, as it directly supports the development of listening, speaking, reading and writing skills (Hasanah et al., 2024). For beginner learners, vocabulary is the primary gateway to understanding the meaning of speech, forming simple sentences, and interacting in daily communicative (Fatati & Sutarjo, 2021). Without sufficient vocabulary, learners are unable to express ideas effectively or comprehend spoken and written texts (Solehudin et al., 2025). A limited vocabulary range often results in restricted comprehension and difficulty expressing even simple ideas, which hinders overall language acquisition.

In the context of Arabic as a foreign language, vocabulary acquisition is particularly challenging due to the diglossia situation between Modern Standard Arabic and colloquial varieties, the non-Latin script, and its complex root-and-pattern morphology (Rizqia et al., 2024). These features often make Arabic vocabulary learning more demanding compared to other languages (Lee et al., 2022). Therefore, mastering sufficient vocabulary at the early stages of learning is widely acknowledged as the foundation upon which the other language skills are built, making it an essential focus of both teaching and assessment (Brown & Abeywickrama, 2019).

Vocabulary in language learning has been defined as the collection of words and lexical units that learners must acquire to achieve communicative competence (Hendra & Rufaidah, 2021). Scholars distinguish between receptive vocabulary, which refers to the words learners can recognize in listening and reading, and productive vocabulary, which refers to the words they can actively use in speaking and writing (Mahmudi et al., 2025). At the beginner stage, receptive vocabulary typically exceeds productive vocabulary, emphasizing the importance of balancing recognition and usage when designing both instruction and assessment. In Arabic learning, this distinction becomes even more critical given the complex orthographic and morphological features of the language (Zaid et al., 2025).

In the context of learning evaluation, vocabulary tests measure the extent to which learners have mastered the target vocabulary taught. The design of a good test requires clear, measurable criteria that are in line with internationally recognized reference standards. A widely used reference is the Common European Framework of Reference for Languages (CEFR), which provides language competency descriptors at various levels, from Beginner Level (A1) to Proficient Level (A2) (Rohman & Rosyadi, 2021). The CEFR not only regulates general language skills (Dianto et al., 2022), but also provides guidance on the range and complexity of vocabulary that should be mastered at each level (Pratama et al., 2024).

For Arabic learners at the beginner level (A1) and elementary level (A2), the CEFR recommends mastery of basic vocabulary related to everyday communicative needs, such as introducing oneself, describing objects around them, mentioning time, or explaining daily routines. However, there are still challenges in translating CEFR descriptors—which were originally developed for European languages—into the context of Arabic, which has different morphological, phonological, and semantic characteristics. As a result, many Arabic vocabulary tests are not yet fully aligned with CEFR parameters, in terms of word selection, level of difficulty, and question types (Fahri & Supriadi, 2023).

Previous studies further highlight the importance of aligning Arabic vocabulary testing with CEFR. In Arabic-specific contexts, discussed the challenges posed by diglossia when applying CEFR descriptors to Arabic (Khalil, 2024). This research collectively demonstrate the growing relevance of CEFR in Arabic pedagogy, but they also reveal the absence of practical, standardized frameworks for beginner-level vocabulary assessment in Arabic.

The need for the development of Arabic vocabulary test criteria that align with CEFR is becoming increasingly urgent, especially to ensure standardization in the

evaluation of vocabulary proficiency across educational institutions. With clear criteria, teachers and test developers can create evaluation instruments that are valid, reliable, and relevant to learning objectives. This article aims to identify Arabic vocabulary indicators at the beginner level based on the CEFR as a reference in evaluating Arabic language learning.

Method

This study uses descriptive qualitative research. Qualitative research is research that has qualitative data, both verbal and written, and it can be observed in detail to obtain the implied meaning in the document (Sholahudin et al., 2025; Alimudin et al., 2023). In terms of data sources, this study is a literature study. A literature study is research that is synonymous with text analysis to obtain accurate facts (Amalia et al., 2024). This study has two data sources: primary and secondary (Husna et al., 2023). Primary data is the main data obtained from original sources to answer the research questions (Aziz et al., 2025). The primary data source for this study is the vocabulary test questions from the book *al-Ikhtibârât al-Lughawiyah* by Ali Khuly. Secondary data is supporting data obtained from books and other documents in the literature review. The secondary data sources for this study are books and articles discussing language tests and the CEFR.

Result and Discussion

Description of A1-Level Vocabulary

The CEFR offers an internationally recognized model for describing language proficiency, dividing it into six levels (A1–C2). At the A1 beginner level, learners are expected to recognize and use a limited range of everyday vocabulary related to personal information, family, daily routines, food, numbers, and classroom objects (Mohamed, 2023). Their vocabulary use remains restricted, relying on memorized expressions and basic phrases for immediate needs (Islami et al., 2025). A summary of CEFR descriptors for A1 level vocabulary is presented in Table 1, which outlines learners' expected range, control, comprehension, production, and thematic domains. These descriptors provide a benchmark for aligning Arabic vocabulary teaching and testing with internationally accepted standards (Soliman & Familiar, 2024).

Table 1. CEFR Descriptors for Vocabulary at A1 Level

Aspect	CEFR Descriptor (A1)	Example in Arabic Context
Vocabulary Range	Has a very basic repertoire of isolated words and phrases related to personal details, everyday needs, and concrete situations.	بيت (house), كتاب (book), مدرسة (school), أكل (eat), شرب (drink)
Vocabulary Control	Can use memorized phrases and basic expressions but may struggle with accuracy and flexibility.	Greetings: السلام عليكم, مرحباً

Comprehension	Can recognize familiar words and very basic phrases when spoken slowly and clearly.	نعم (yes), لا (no), شكراً (thank you)
Production	Can produce simple words and expressions to describe immediate needs or surroundings.	أنا طالب (I am a student), هذا كتاب (This is a book)
Thematic Domains	Vocabulary limited to family, numbers, food, daily routines, classroom, time, and basic adjectives.	أسرة (family), واحد-عشرة (numbers), صباح (morning), كبير/صغير (big/small)

According to the CEFR, learners at level A1 are described as having the ability to understand and use everyday expressions to interact in routine and concrete situations with clear pronunciation (Pratama et al., 2023). CEFR itself does not specify an exact number of vocabulary items to be mastered at each level. However, several studies provide estimations for A1 vocabulary size. For instance, Natalie et al. found that A1 learners master approximately 1,000 lemmas based on the Lex test (Finlayson et al., 2024), while Zhiqing Li indicated that the number of A1-level lemmas could be fewer than 1,500, depending on the metric and the target language (Li et al., 2024).

The descriptors for level A1 further highlight that learners should be able to understand and use everyday expressions and simple phrases aimed at meeting immediate needs. In the context of Arabic language teaching for beginners, a target of 500–600 words pragmatically adopted as it is considered sufficient for basic communication within the A1 curriculum design. Which enables them to perform essential communicative functions such as introducing themselves, asking and answering simple questions, telling the time, shopping for basic items, and giving or asking for directions. These abilities demonstrate how vocabulary knowledge at A1 supports basic interaction in real-life contexts. Aligning vocabulary tests with CEFR therefore ensures that assessments reflect these communicative goals, making them standardized, reliable, and comparable across different educational settings while also guiding curriculum planning and instructional design.

For Arabic beginners, assessment should emphasize high-frequency and functional vocabulary (Azizah et al., 2024). Indicators include recognition of common nouns (*kitâb* “book”, *bayt* “house”), understanding of verbs of daily activities (*kataba* “to write”, *dhababa* “to go”), basic adjectives (*kabîr* “big”, *shaghîr* “small”), and functional expressions such as greetings and self-introductions. Different test formats such as receptive tasks (matching, multiple choice), productive tasks (gap-filling, short writing), and integrated tasks (dialogue role-play) can be employed to capture these abilities. The vocabulary used in this research instrument is grouped into several main themes adapted to the daily communication needs of A1 learners, namely: (1) family, (2) numbers, (3) time, (4) places, (5) daily activities, (6) colors, (7) animals, (8) clothing, (9) weather, and (10) food. The selection of such themes follows the principle of high-frequency

vocabulary, which facilitates beginner learners in understanding and using essential vocabulary for basic communication (Rahmani et al., 2022).

Vocabulary Test Criteria

Effective vocabulary assessment must adhere to the fundamental principles of validity, reliability, and practicality (Bachman & Palmer, 2010). Validity ensures that the test measures learners' vocabulary knowledge accurately at the intended level. For instance, a vocabulary test designed for beginner learners (A1) should focus on high-frequency and concrete words rather than abstract or advanced terms. Reliability guarantees consistency of results, meaning that the test should yield similar outcomes when administered under similar conditions. Practicality, on the other hand, refers to the feasibility of conducting the assessment in classroom settings, including considerations such as time allocation, clarity of instructions, and ease of scoring. Together, these principles form the foundation of effective vocabulary evaluation across contexts.

For Arabic vocabulary testing, these principles gain additional significance due to the unique orthographic and morphological features of the language. The Arabic writing system requires learners to distinguish between consonants, short vowels, and diacritical marks, while its root-and-pattern morphology produces multiple words from the same root, such as *كتب - كتاب - كاتب - مكتبة*. This complexity poses challenges for test designers in determining whether items should measure recognition of isolated words or broader morphological awareness. For example, a fill-in-the-blank test that expects learners to complete the sentence “ذهب أحمد إلى ____ صباحًا” with “المدرسة” remains valid for the A1 level because it reflects familiar, concrete vocabulary. Conversely, including derived forms or rare lexical items would risk reducing the validity of the test.

A wide variety of test formats can be employed to measure both receptive and productive vocabulary knowledge (Nation, 2013; Schmitt, 2014). Receptive tasks include picture-based multiple-choice items, such as identifying the image of a lion with “أسد”, or definition-based items, such as selecting “المدرسة” for the definition “a place for learning and teaching.” These tasks measure learners' ability to recognize and comprehend vocabulary. Productive tasks, by contrast, require learners to actively retrieve and use vocabulary in context, for instance: “استخدم كلمة 'قلم' في جملة: كتب الطالب ____ بالقلم.” Synonym and antonym tests (ثري = غني; طويل ↔ قصير) also encourage learners to deepen their semantic networks, while matching tasks strengthen recognition of collocations and thematic word fields. Together, these formats ensure a balanced approach that addresses both recognition and active application of vocabulary.

By aligning test types with CEFR descriptors, teachers and test developers can ensure that assessments are standardized, reliable, and pedagogically appropriate for beginner learners. At the A1 level, vocabulary tests should emphasize picture identification, simple fill-in-the-blank exercises, and definition-based multiple-choice items, which are accessible and engaging for novice learners. At the A2 level, however, learners can begin to handle tasks that involve semantic relations, such as synonym/antonym matching and collocations, which require deeper lexical

understanding. Table 1 summarizes the main types of vocabulary tests, their examples, guiding principles, and dominant CEFR levels.

Effective vocabulary assessment must adhere to the fundamental principles of validity, reliability, and practicality (Bachman & Palmer, 2010). Validity ensures that the test items accurately reflect the vocabulary knowledge expected at the beginner level, while reliability guarantees consistency of results across different test administrations. Practicality ensures that assessments can be feasibly implemented in classroom contexts, considering time, scoring, and learner accessibility. These principles have been highlighted by several studies. For instance, Nation (2013) emphasized that vocabulary tests must differentiate between receptive and productive dimensions to maintain construct validity. Schmitt (2014) further argued that reliability in vocabulary testing can only be achieved through clear item design and standardized scoring procedures. In the context of Arabic, Al-Husban (2021) stressed the importance of aligning assessment practices with CEFR descriptors to ensure international comparability. Similarly, Ismail (2019) demonstrated that vocabulary tests developed with CEFR-based indicators enhanced both validity and practicality, as they guided teachers to design level-appropriate items. Taken together, these studies confirm that well-constructed vocabulary assessments not only measure lexical knowledge effectively but also provide reliable data for instructional decision-making.

When assessing vocabulary, several things need to be considered, including:

1. Word Difficulty Level

In designing the vocabulary test, words are selected progressively, starting from highly common and concrete to more specific terms. For example, the word “بيت” (house) is introduced earlier than “مستشفى” (hospital). This approach aligns with the high-frequency vocabulary focus, which states that beginner learners should first master high-frequency words before moving to low-frequency ones (Webb et al., 2023).

2. Assessment Formats

The vocabulary assessment is designed to assess both receptive vocabulary knowledge and productive vocabulary knowledge in a balanced manner. The formats used are:

Table 2. Assessment Formats

Test Format	Description	Example Item
Fill in the Blank	Complete the missing word in a sentence	أعلى نقطة في الجبل تُدعى _____ إنه رجل ثري
Synonym (MCQ)	Select the closest meaning	أ. غني ج. كريم ب. فقير د. بخيل

Test Format	Description	Example Item
Picture and Word Choice	Select the word matching the picture	 أ. أسد ب. نمر ج. فيل د. سمكة
Definition and Word Choice	Definition and Word Choice	الشخص الذي يسلّم النقود من المشتري هو: أ. المحاسب ب. السمسار ج. أمين الصندوق د. المالك
Word Meaning in Context	Word Meaning in Context	ذهب الطفل إلى الحديقة ليركض ويلعب. ما معنى الحديقة؟ أ. غرفة النوم ب. مكان فيه أشجار وألعاب
Matching Words	Matching Words	سعيد – فرح كبير – ضخم
Sentence Production	Sentence Production	واسع → بيتي واسع
Completion with Letter Clues	Completion with Letter Clues	أذهب إلى ____ (م) لشراء الطعام → متجر

3. Receptive and Productive Proportion

After reviewing the beginner-level test instrument aligned with CEFR standards, the composition of items was determined to be 60% receptive and 40% productive (Sulistyaningrum & Purnawati, 2021). This is consistent with the findings of Gonzalez-Fernandez et al., who reported that beginner learners perform better when exposed to more receptive vocabulary before being required to produce it actively (Xia et al., 2023).

4. Analysis of test Format Allignment with A1 Level

Based on the test instrument formats described above, it can be concluded that all formats meet the vocabulary learning characteristics of A1 level according to CEFR. First, the items are concrete and visual, which facilitates learners' ability to connect word form and meaning. Second, the sentence structures are simple, in line with beginner-level competence. Third, the contexts are relevant to learners' daily experiences, enhancing engagement and comprehension. Fourth,

the difficulty level is measured and progressive, with items moving from easy to more challenging while remaining within the A1 range.

With this consistent approach, it is recommended that beginner vocabulary evaluation utilize visual stimuli and simple contexts to effectively connect form and meaning. Furthermore, these findings align with Krashen's Comprehensible Input principle, which states that material should be understandable to learners with a slight challenge to optimally promote language development (Hajimia et al., 2020).

Therefore, the criteria and test formats designed in this study can serve as a reference for developing Arabic vocabulary assessment instruments for beginners that are consistent with the CEFR framework.

Example of Blueprint Test

A test blueprint is a document that outlines the key elements of a test, including the content to be assessed, the emphasis allocated to each content area, and other essential features that support the test's validity and reliability (Ezenwaka & Adinna, 2025). The process of creating a test blueprint typically involves several steps: identifying the main competencies to be assessed, specifying clear assessment objectives, selecting appropriate assessment methods, and determining item distribution based on content weight (Waluyo et al., 2024). A well-structured test blueprint improves the validity of test results and also serves as a guide for focused instructional and assessment planning.

The purpose of developing a test blueprint is to assist teachers in designing valid and reliable tests aligned with predetermined learning objectives. A test blueprint serves as a comprehensive plan that outlines content coverage, learning goals, and the cognitive levels to be assessed. With a clear blueprint, teachers can develop fair, systematic, and competency-based evaluation instruments. Moreover, the blueprint helps ensure proportional question distribution, balanced representation across topics, and clarity and consistency in assessment. A well-constructed test blueprint also promotes fairness and accessibility for all test-takers while reinforcing the validity and reliability of the assessment outcomes.

Table 3. Example of Blueprint Test

Competence (A1 - CEFR)	Indicator	Theme	Question Type	Keywords	Cognitive Level
Uses numbers in basic Arabic expressions.	Identify and complete sentences with number- related words.	Numbers	Fill in the Blank	واحد، اثنان، ثلاثة	Understanding
Recognizes similar meanings of common time- related words.	Choose synonyms of given words.	Time	Synonym (MCQ)	صباح، نهار، ليل	Understanding

Matches places with corresponding vocabulary.	Match pictures to appropriate words.	Places	Picture and Word Choice	مدرسة، مسجد، بيت	Remembering
Identifies vocabulary based on definitions.	Choose the correct word from a definition.	Daily Activities	Definition and Word Choice	يأكل، ينام، يكتب	Understanding
Recognizes and identifies basic Arabic vocabulary related to family.	Identify the correct vocabulary related to family.	Family	Matching Words	أب، أم، أخ، أخت	Remembering

CEFR Compatibility Analysis of the Vocabulary Used in the Test

1. Fill in the Blank

Question:

أكمل الجملة بالكلمة العربية الصحيحة التي تعبر عن الرقم!

أعطيتني أمي _____ تفاحات

This item assesses students' ability to use numerical vocabulary in the context of a simple sentence. This competence aligns with the CEFR A1 descriptors, which emphasize mastery of basic expressions related to personal information and quantity. According to English Radar, learners at this level are expected to use numbers, days, months, and understand quantity and prices in every situation (Gazali & Saefuloh, 2023). Therefore, the format and content of the item are considered appropriate for A1 level, as they focus on the recognition and application of basic numerical vocabulary in meaningful and communicative contexts.

When compared with the question formats in *Ikhtibârât Lughawīyyah*, there is a noticeable difference in cognitive complexity. The fill-in-the-blank items in the book require learners to complete sentences using conceptual vocabulary, such as geographical or scientific terms, for example, "_____ أعلى نقطة في الجبل تُدعى". These items are productive and objective in nature, demanding deeper semantic understanding and background knowledge beyond basic everyday language, thus making them more appropriate for A2 or B1 levels. In contrast, the item developed in this study focuses on the use of basic numerical vocabulary within simple sentence structures, aligning with CEFR A1 descriptors that emphasize functional language for daily communication.

2. Synonym

Question:

اختر الكلمة الأقرب في المعنى إلى الكلمة: يوم

أ. صباح ☒ ب. ليل ج. مساء د. فجر

This item assesses learners' understanding of time-related vocabulary through synonym recognition, specifically focusing on the word "نهار" and its closest meaning "صباح". According to the CEFR Companion Volume, A1 learners are expected to relay simple and predictable information about times and places, particularly when the input is delivered slowly and clearly (Negishi et al., 2012). This suggests that temporal vocabulary, including basic words like "daytime" and "morning", falls within the expected lexical scope for this level. Therefore, the design of this item aligns well with CEFR A1 descriptors, emphasizing familiar, everyday expressions essential for basic communicative competence.

When compared with a similar synonym-based item such as "إنه رجل ثري" (with "غني" as the correct answer), a slight difference in lexical familiarity is evident. While both items assess synonym recognition, the use of time-related vocabulary in this item is generally more concrete and functionally accessible to A1 learners. Thus, this task is well-aligned with A1 descriptors, focusing on essential and immediately applicable language.

3. Picture and Word Choice

Question:

انظر إلى الصورة. اختر الكلمة الصحيحة التي تطابقها!



أ. مسجد ب. بيت ج. مدرسة ☒ د. مطبخ

This question invites learners to recognize place-related vocabulary through visual cues, such as identifying "مدرسة" from a depicted image. CEFR A1 reading descriptors highlight that learners at this level can recognize familiar words accompanied by pictures, such as in illustrated menus or simple picture books, indicating that the combination of visual stimulus and high-frequency vocabulary is appropriate for beginners (Erlina, 2022).

Both this item and the animal recognition task from *Iktibārât Lughawīyyah* (e.g., "أ. أسد، ب. نمر، ج. فيل، د. سمكة") rely on visual cues to elicit recognition of concrete vocabulary. Although the thematic content differs—places in the current task and animals in the reference item—the cognitive demands are similar. Each requires learners to match a picture with a familiar, high-frequency noun without

involving abstract reasoning or complex sentence processing. Such tasks are well-suited for CEFR A1 learners, who are expected to identify common objects, locations, and animals encountered in daily life. The inclusion of image support further reduces processing load, making both items effective tools for assessing foundational vocabulary recognition at the beginner level (Iskandar et al., 2021).

4. Definition and Word Choice

Question:

اقرأ التعريف واختر الكلمة الصحيحة.

"الفعل الذي يقوم به الشخص عندما يضع الطعام في فمه"

أ. ينام ب. يأكل ☒ ج. يكتب د. يشرب

This item presents a simple definition describing an everyday action and asks learners to select the corresponding verb. In this case, the definition "The action someone does when they put food in their mouth" clearly refers to the verb "يأكل". Tasks of this kind promote understanding of core vocabulary through short, familiar explanations, an approach that aligns well with CEFR A1 expectations. At this level, learners are expected to comprehend and respond to highly frequent language related to daily routines and personal needs. According to the CEFR Companion Volume, A1 learners can understand short, simple descriptions and follow familiar vocabulary connected to basic needs and habits (Sawaguchi, 2025).

The item from *Iktibârât Lughawīyyah* asks learners to identify a profession based on the definition: "الشخص الذي يسلّم النقود من المشتري هو ...", with the correct answer "أمين الصندوق". While both this and the current task use definition-based formats, the lexical content varies. Identifying أمين الصندوق requires familiarity with occupational terms, which are more suited for A2 level. In contrast, verbs like يأكل, ينام, يكتب represent high-frequency vocabulary tied to daily routines. This makes the current item more appropriate for A1 learners, as it reflects CEFR descriptors focusing on simple and essential language for everyday communication.

5. Matching Words

Question:

زاوج الكلمات العربية مع معانيها!

Arabic Word	Meanings Word
أب	Sister
أم	Brother
أخ	Father
أخت	Mother

This item requires students to match basic Arabic family vocabulary with their meanings. According to the CEFR level A1, family-related vocabulary is part of the essential lexical inventory that beginners are expected to master, as it falls under the themes of personal information and family and friends. Thus, the selection of vocabulary in this item can be considered appropriate for A1 descriptors, since it emphasizes the recognition of core vocabulary that supports simple, everyday communication.

When compared with the test formats in *Iktibârât Lughawīyyah*, a difference in complexity is evident. The book requires learners to match words based on semantic relations such as synonymy, antonymy, or collocation. For example, in one item, students are asked to match “سعيد” with “فرح”, which illustrates the demand for recognizing word relationships beyond surface meaning. Such tasks are more suitable for A2 level and above, as they demand an understanding of relational meaning rather than simple word identification. Accordingly, this first item confirms its position as an appropriate instrument for the A1 level, as it focuses on the direct identification of basic family vocabulary without requiring more complex semantic analysis.

The findings highlight that effective vocabulary tests must adhere to the principles of validity, reliability, and practicality. Validity requires that the test items accurately reflect the vocabulary knowledge expected at the beginner level, while reliability ensures that the results remain consistent across different administrations. Practicality is equally important, ensuring that the tests are feasible to implement within classroom contexts. The specific characteristics of Arabic, particularly its orthographic and morphological system, further underline the need for careful test design that balances receptive (recognition, comprehension) and productive (recall, application) vocabulary knowledge.

Furthermore, the classification of test formats ranging from fill-in-the-blank and picture-based multiple choice to synonym/antonym matching and word-in-context tasks demonstrates how each type can serve a specific purpose in measuring vocabulary competence. At the A1 level, learners benefit most from concrete and visually supported tasks such as picture identification or simple fill-in-the-blank exercises. At the A2 level, however, assessments can gradually include tasks that measure deeper lexical relations, such as semantic networks and collocations. Aligning test design with CEFR descriptors thus ensures that vocabulary assessments remain both pedagogically appropriate and internationally standardized.

Conclusion

This study emphasizes the crucial role of vocabulary assessment in Arabic language learning, particularly at the beginner (A1) level. Vocabulary is not only the foundation for the development of other language skills: listening, speaking, reading, and writing but also a key determinant of learners' ability to communicate meaningfully. By referring to the CEFR framework, which provides internationally recognized descriptors for language proficiency, it becomes possible to establish clear, standardized, and comparable criteria for Arabic vocabulary testing.

Based on the findings of this study, several recommendations can be proposed for teachers, curriculum designers, and future researchers. For teachers and practitioners, it is essential to employ a balanced combination of receptive and productive tasks in vocabulary assessment so that both recognition and active use of words can be measured comprehensively. Classroom tests should integrate vocabulary items related to real-life communicative contexts, such as daily routines, family, and school activities, while also aligning with CEFR descriptors to ensure international comparability. For curriculum designers and institutions, the development of teaching materials and assessment frameworks that explicitly reference CEFR levels particularly A1 and A2 is strongly recommended to enhance the global recognition of Arabic language programs. This effort should be supported by teacher training on the principles of validity, reliability, and practicality, with particular emphasis on addressing the morphological and orthographic challenges of Arabic. In addition, the use of formative assessment tools, such as picture-based quizzes or digital vocabulary games, can complement summative tests to improve learner engagement and retention.

For future researchers, it is recommended to extend the scope of study beyond A1 to higher CEFR levels in order to construct a more comprehensive framework for Arabic vocabulary assessment. Further research should also investigate how vocabulary testing can be integrated with other skills, such as reading comprehension or oral communication, to provide a holistic view of language competence. Finally, exploring technology-enhanced assessment tools, including computer-adaptive testing and mobile-based applications, would improve the practicality, scalability, and authenticity of vocabulary testing in the digital age.

REFERENCES

- Alimudin, A., Dania, I., & Saputri Hidayat, A. A. (2023). Analisis Bahan Ajar Mahârah Qirâ'ah di SMP Muhammadiyah 3 Yogyakarta Berbasis CEFR. *Al-Hikmah: Jurnal Agama dan Ilmu Pengetahuan*, 20(1).
- Amalia, N. N., M. Abdul Hamid, Zakiyah Arifa, & Asbarin. (2024). Analysis of the Arabic Language Textbook for Junior High School from the Perspective of the Education, Standards, Curriculum, and Assesment Agency (BSKAP). *Alibbaa': Jurnal Pendidikan Bahasa Arab*, 5(2).
- Aziz, A. M., Kosim, N., & Wahyudin, D. (2025). Analysis of The Content Alignment of The Grade VII Arabic Textbook Published by Erlangga with the Merdeka Curriculum. *Tadris Al- 'Arabiyyah*, 4(1).
- Azizah, R. N., Zarkasyi, A. H., Sifaurohmah, A., & Hasyim, N. I. H. (2024). Penerapan Media Educandy untuk Meningkatkan Hasil Belajar dalam Materi Bahasa Arab pada Kelas 5B MI Qurrota A'yun Yogyakarta. *Al Mahara: Jurnal Pendidikan Bahasa Arab*, 10(2).
- Brown, H. D., & Abeywickrama, P. (2019). *Language Assessment Principles and Classroom Practices (Third Edition)*. Pearson Education.

- Dianto, T. N., P, N. J., Fatoni, A., & Kalita, S. (2022). CEFR-Based Beginner Arabic Reading and Writing Curriculum Design in Indonesia/ Desain Kurikulum Membaca dan Menulis Bahasa Arab Pemula Berbasis CEFR di Indonesia. *Ijaz' Arabi Journal of Arabic Learning*, 5(3).
- Erlina. (2022). Kaidah Penyusunan Tes Bahasa Arab (Pilihan Ganda). *El-Jaudah : Jurnal Pendidikan Bahasa dan Sastra Arab*, 3(2).
- Ezenwaka, C. M., & Adinna, D. P. I. (2025). Role of Test Blueprint in Test Construction Among Secondary School Teachers in Onitsha Education Zone of Anambra State. *Journal of Educational Research and Development*, 8(1).
- Fahri, A. J., & Supriadi, R. (2023). Implikasi CEFR Pada Buku Ajar Sahabatku Indonesia untuk Penutur Bahasa Arab Level 1. *Madab: Jurnal Bahasa dan Sastra*, 14(1).
- Fatati, A., & Sutarjo, J. (2021). Implementasi Metode Mimicry Memorization (Mim-Mem) dalam Pembelajaran Mufradat. *An Nabighoh*, 23(1).
- Finlayson, N., Marsden, E., & Hawkes, R. (2024). Creating and Evaluating Corpus-Informed Word Lists for Adolescent, Beginner-to-Low-Intermediate Learners of French, German, and Spanish. *Language Teaching Research*, 1.
- Gazali, E., & Saefuloh, H. (2023). Development of an Arabic Receptive Proficiency Test Instrument Based on the Common European Framework of Reference for Languages. *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab LAIN Palangka Raya*, 11(2).
- Hajimia, H., Chethiyar, S. D. M., & Singh, M. K. S. (2020). Second Language Acquisition: Krashen'S Monitor Model and the Natural Approach. *PEOPLE: International Journal of Social Sciences*, 6(3).
- Hasanah, M., Hasibuan, R., & Jundi, M. (2024). Elevating Arabic Vocabulary Learning: Integrating Teams Games Tournament and Show & Tell Method. *Arabiyat : Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 11(1).
- Hendra, F., & Rufaidah, M. F. (2021). Compiling Vocabulary Lists for Corpus-Based Arabic for Toutrism Teaching. *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 8(1).
- Husna, I., Azkiya, H., Febrianti, S. R., Rahmawati, & Kustati, M. (2023). The Analysis of Arabic Language Textbook for Class II Elementary School. *Lectura: Jurnal Pendidikan*, 14(7).
- Iskandar, I., Efendi, A. S., & Santosa, I. (2021). The Model of Listening CEFR-Based Test Specification for Higher Education. *Linguist: Journal Of Linguistics and Language Teaching*, 7(2).
- Islami, D. F., Yasin, A., Nur, I. R., Saidah, U., Alfin, M. F. S., Education, A., Program, S., Darussalam, U., & Ponorogo, G. (2025). Fā'ālīyāh Istikhdam Wāsā'il āt-Tā'lim ās-Sāwtīyyāh wāā āl-Mār'īyyāh Sparkol Video Scribe 'ālā Itqānī āl-

- Mufrādāt āl-‘Arābiyyāh lī āl-Mādrāsāh āl-Ibtidā’iyyāh. *Mantiqū Tayr: Journal of Arabic Language*, 5(2).
- Jumatiah, J., & Helalsah, I. (2021). Developing the Arabic Vocabulary Test Material from Cognitive and Skill Domain Perspective. *Jurnal Al-Maqayis*, 8(2).
- Khalil, A. (2024). Integrating CEFR Standards in Arabic Language Education: Challenges and Opportunities. *IJAS: International Journal of Arabic Studies*, 1(2).
- Lee, J. H., Ahn, J. J., & Lee, H. (2022). The Role of Motivation and Vocabulary Learning Strategies in L2 Vocabulary Knowledge: A Structural Equation Modeling Analysis. *Studies in Second Language Learning and Teaching*, 12(3).
- Li, Z., Li, J. Z., Zhang, X., & Reynolds, B. L. (2024). Mastery of Listening and Reading Vocabulary Levels in Relation to CEFR: Insights into Student Admissions and English as a Medium of Instruction. *Languages*, 9(7).
- Mahmudi, I., Nisa, S. S., Widyanti, L. N., Jariah, A., & Rini, S. (2025). Namūdzu al-Ikhtibārātī al-Qowāidī li al-Lughotī al-‘Arabiyyati Muassasan ‘alā an-Nazhariyyati ACTFL. *Lisanan Arabiya: Jurnal Pendidikan Bahasa Arab*, 9(1).
- Mohamed, S. (2023). The Development of an Arabic Curriculum Framework Based on A Compilation of Salient Features From CEFR Level Descriptors. *Language Learning Journal*, 51(1).
- Negishi, M., Tono, Y., & Fujita, Y. (2012). A Validation Study of the CEFR Levels of Phrasal Verbs in the English Vocabulary Profile. *English Profile Journal*, 3.
- Pratama, A., Mahmudah, K., Sastra Sumardi, M., Putri, R., Fitria, W., Rozal, E., Nirwana, I., & Sakunti, S. R. (2023). Teaching English At a1 Movers Class and a2 Flyers With Common European Framework Reference (Cefr) At Junior High School Jambi City. *Abdi Dosen: Jurnal Pengabdian pada Masyarakat*, 7(4).
- Rahmani, A., Asadi, V., & Xodabande, I. (2022). Using Mobile Devices for Vocabulary Learning Outside the Classroom: Improving the English as Foreign Language Learners’ Knowledge of High-Frequency Words. *Frontiers in Psychology*, 13.
- Rizqia, A. S., Sopian, A., Ridlo, A., & Abdulkhaleq, Al-Rawafi. (2024). Using Mimicry Memorization Method in Improving Student’s Learning Outcomes on the Lesson of Arabic Language. *Arabi: Journal of Arabic Studies*, 9(1).
- Rohman, H., & Rosyadi, F. I. (2021). Pengembangan Bahan Ajar Bahasa Arab Berbasis Common European Framework of Reference (CEFR) Untuk Meningkatkan Keterampilan Bahasa Arab Siswa Development of Arabic Teaching Materials Based on the Common European Framework of Reference (CEFR) to Improve. *Al Mahāra Jurnal Pendidikan Bahasa Arab*, 7(2).
- Sawaguchi, R. (2025). Developing a CEFR-based Diagnostic Test to Assess Japanese University Students’ Productive Knowledge of Lexical Bundles. *Language Testing in Asia*, 15(1).

- Setiyadi, A. C., Anhar, A., & Muttaqin, H. (2025). Utilizing Comics as a Learning Media to Improve Students' Arabic Speaking Skills. *Arabiyatuna: Jurnal Bahasa Arab*, 9(1).
- Sholahudin, S., Al Munawar, A. H., & Aburrahman, M. (2025). Analysis of Arabic Textbook for Grade VII Merdeka Curriculum Based on Textbook Praxeological Analysis. *Didaktika Jurnal Kependidikan*, 14(1).
- Solehudin, M., Hanifansyah, N., & Izzuddin, I. F. (2025). Enhancing Arabic Listening and Vocabulary Acquisition through AI-Powered Music: A Study on Suno AI. *Arabiyat : Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 11(2).
- Soliman, R., & Familiar, L. (2024). Creating a CEFR Arabic Vocabulary Profile: A Frequency-based Multi-Dialectal Approach. *Critical Multilingualism Studies*, 11(1).
- Sulistyaningrum, S. D., & Purnawati, P. (2021). Incorporating CEFR Bands and ICT-Competences in Grammar Syllabuses of English Language Education Study Program in Indonesia. *Journal on English as a Foreign Language*, 11(2).
- Waluyo, B., Zahabi, A., & Ruangsung, L. (2024). Language Assessment at a Thai University: A CEFR-Based Test of English Proficiency Development. *REFlections*, 31(1).
- Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How Effective is Second Language Incidental Vocabulary Learning? A Meta-Analysis. *Language Teaching*, 56(2).
- Xia, T., Chen, X., Parsaei, H. R., & Qiu, F. (2023). An Intelligent Vocabulary Size Measurement Method for Second Language Learner. *Language Testing in Asia*, 13(1).
- Zaid, A. H., Putri, S. F., & Anggraini, R. M. (2025). Analysis of Nahwu Errors in Hamzah Washl And Hamzah Qath'i Class 5 KMI At Gontor for Girls. *Ijaz Arabi Journal of Arabic Learning*, 8(1).