

Musthofa Sha'ban's Perspective on Arabic Listening Assessment for Beginners

**Lalu Rahmat Sugiara¹, R. Umi Baroroh², Amy Fitriani Siregar³, Khaerul Muttakin⁴,
Ilham Dhya'ul Wahid⁵**

^{1,2,3,4}Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia

⁵Azhar University, Egypt

✉ Email: 23204022017@student.uin-suka.ac.id¹, umi.baroroh@uin-suka.ac.id²,
amyfitriani2000@gmail.com³, 24204022031@student.uin-suka.ac.id⁴,
ilhamdhyal4@gmail.com⁵

Article Information:

Received July 12, 2025

Revised July 22, 2025

Accepted September 27, 2025

Published November 8, 2025

Keywords: *Listening skills test, Arabic language learners, Musthofa Sya'ban*

Kata Kunci: *tes keterampilan menyimak, pembelajar bahasa Arab, Musthofa Sya'ban*

Abstract:

The listening comprehension test (*maharah istima'*) instrument for basic-level Arabic learners is still limited and has not yet been able to measure language competence comprehensively. Previous studies have emphasized literal understanding but have paid less attention to phonetic, cognitive, and visual dimensions, resulting in suboptimal instruments. This study aims to analyze and formulate the standardization of the *maharah istima'* test based on Musthofa Sya'ban's perspective, which is relevant to the needs of beginners. The method used is a library research approach, with primary data from *Ma'ayir Lughawiyah* and secondary data from books, journals, and scientific articles. The analysis was conducted through identifying key concepts, categorizing themes, and comparing them with previous research findings to develop a conceptual model. This study offers a listening comprehension test model based on *Musthofa Sya'ban's* standards suitable for beginners. It thus can serve as a reference in developing Arabic language evaluation instruments at the basic level.

Abstrak:

Instrumen tes pemahaman menyimak (*maharah istima'*) bagi pembelajar bahasa Arab tingkat dasar masih terbatas dan belum mampu mengukur kompetensi berbahasa secara komprehensif. Penelitian-penelitian sebelumnya lebih menekankan pada pemahaman literal, namun kurang memperhatikan dimensi fonetik, kognitif, dan visual, sehingga menghasilkan instrumen yang kurang optimal. Penelitian ini bertujuan untuk menganalisis dan merumuskan standarisasi tes *maharah istima'* berdasarkan perspektif Musthofa Sya'ban yang relevan dengan kebutuhan pemula. Metode yang digunakan adalah pendekatan penelitian kepustakaan, dengan data primer yang bersumber dari *Ma'ayir Lughawiyah* dan data sekunder dari buku, jurnal, serta artikel ilmiah. Analisis dilakukan melalui identifikasi konsep-konsep kunci, pengelompokan tema, serta perbandingan dengan temuan penelitian terdahulu untuk mengembangkan model konseptual. Penelitian ini menawarkan model tes pemahaman menyimak berdasarkan standar Musthofa Sya'ban yang sesuai untuk pembelajar tingkat dasar. Dengan demikian, hasil penelitian ini dapat menjadi rujukan dalam pengembangan instrumen evaluasi pembelajaran bahasa Arab pada tingkat dasar.

Correspondent Author: 23204022017@student.uin-suka.ac.id (Lalu Rahmat Sugiara)

How to cite:	Rahmat Sugiara, L., Baroroh, R. U., Siregar, A. F., Muttakin, K., & Wahid, I. D. (2025). Musthofa Sha'ban's Perspective on Arabic Listening Assessment for Beginners. <i>Journal of Arabic Language Teaching</i> , 5(2), 249-262. https://doi.org/10.35719/arkhas.v5i2.2373
Publisher:	Arabic Language Education Department, Postgraduate of UIN Kiai Haji Achmad Siddiq Jember

Introduction

Listening skills can be understood as an active process that involves receiving and responding to both spoken and unspoken signals. DeVito describes listening as an efficient process for taking in information, processing it, and responding appropriately to what is heard (Ma, 2024). The skill of *istima'* falls under receptive skills, so students should master it first before other skills (Gunarti, 2020). As a receptive skill, *maharah istima'* plays a crucial role in the process of understanding and internalizing the sound system, vocabulary, and structure of the Arabic language (Baroroh & Rahmawati, 2020). These skills become increasingly vital when applied to elementary learners who are still in the stage of forming a basic understanding of the Arabic language system.

However, basic-level Arabic language learning has its own complexities that require a specialized approach in the evaluation and assessment process (Wulan, 2023). Learners at this level face various challenges, ranging from limited exposure to Arabic sounds and difficulty in distinguishing similar-sounding letters to limited vocabulary and obstacles in understanding simple utterances (Haniah, 2018). These specific characteristics necessitate an evaluation system that not only accurately measures abilities but also considers the psychological aspects of learners (Suyitno, 2021). These challenges align with global trends in foreign language assessment, which are now shifting toward competency-based evaluation, as seen in the CEFR framework. This shift means that assessment no longer merely measures literal comprehension, but also integrative and communicative language proficiency (Handriawan, 2021). The absence of appropriate evaluation instruments causes difficulties for teachers in identifying learners' listening comprehension levels and tracking their progress (Jamil, 2023).

In this context, Musthofa Sya'ban, an expert in Arabic language learning, has developed a perspective on listening comprehension testing (Rohmat, 2025). His perspective emphasizes the importance of considering aspects such as phonetic understanding, vocabulary, sentence structure, and context in evaluating Arabic listening skills. The approach offered by Musthofa Sya'ban provides a systematic and measurable evaluation framework that can be adjusted according to the learners' proficiency level. Academically, this approach contributes to enriching the field of receptive skill assessment in Arabic, which remains underexplored at the elementary level. Practically, this model can serve as both a conceptual reference and a technical guide for educators in designing listening evaluation instruments that are more aligned with the characteristics of beginner-level learners.

Based on a review of previous studies, several obstacles were found in the evaluation system of *maharah istima'* for the elementary level (Firdaus et al., 2023). First, there is no standardization of tests that specifically considers the characteristics of beginner learners (Rika Noverma & Hikmah, 2024). Second, there is a lack of evaluation instruments that comprehensively integrate phonetic aspects and understanding. Third, the limited assessment models are progressive and gradual according to the learners' developmental abilities. Fourth, there are still insufficient evaluation instruments that consider the psychological aspects of elementary-level learners (Upiyani, 2024). Aizzatin Habibah and Syihabuddin examined

"Evaluating Listening Skills by Utilizing the Whispering Chain Game in Arabic Language Learning," which found that this evaluation helps students identify aspects they have not mastered. Teachers understand aspects that have not been taught (Habibah & Syihabuddin, 2021). Zulpina's research on "Quizizz, an Alternative Media for Online Arabic Language Learning for *Madrasah Ibtidaiyah* Students" found that the features available in this application can be used for evaluation and learning activities in the Arabic language such as listening skills, speaking skills, reading skills, speaking skills, and writing skills (Ashari et al., 2024; Hakim et al., 2025; Soleha et al., 2025; Zulpina, 2022).

Although these studies have made significant contributions, several limitations remain. Previous research has tended to emphasize the technical aspects of evaluation without considering the specific characteristics of beginner-level learners, as proposed by Musthofa Sya'ban. Therefore, the purpose of this study is to conduct an in-depth analysis and formulate the standardization of *maharah istima'* tests in Arabic language learning using Musthofa Sya'ban's perspective as the main theoretical framework, as well as to identify the various challenges faced by beginner-level learners.

This study hypothesizes that Musthofa Sya'ban's framework provides a more appropriate and comprehensive standardization for basic-level listening skills tests, as it integrates phonetic, cognitive, and visual dimensions in accordance with the characteristics of beginner learners. Thus, this research is expected to make an academic contribution to the development of Arabic language evaluation theory as well as a practical contribution in the form of a conceptual model for educators in designing more effective listening skills test instruments.

Method

This study employs a qualitative descriptive method with a library research and content analysis approach (Creswell, 2014). The primary source of the research is *Ma'ayir Lughawiyah* by Musthofa Sya'ban, which serves as the primary reference for exploring the principles of evaluating listening skills. In contrast, the secondary sources consist of books, articles, and scholarly journals discussing the development of Arabic language tests, particularly the evaluation of *maharah istima'*. The selection of literature was carried out selectively by considering its relevance to the research topic, its connection to the assessment of basic-level listening skills, and its contribution to the development of competency-based evaluation theory.

The data collection technique involved literature documentation, where primary and secondary sources were systematically gathered and reviewed (Astuti et al., 2022). Data analysis was conducted using a content analysis approach in two main stages. First, the researcher identified key concepts in Musthofa Sya'ban's work that are relevant to listening skills testing. Second, these concepts were categorized into central themes, including phonetic aspects, vocabulary, sentence structure, context, and the psychological dimensions of learners. The results of this categorization were then synthesized to formulate a conceptual model for standardizing *maharah istima'* tests in accordance with the characteristics of beginner learners.

Through these stages, this study develops a conceptual design of listening skills tests based on five main evaluation criteria: validity, reliability, objectivity, practicality, and discriminatory power, so that it can serve as a reference for developing Arabic test instruments more suited to the needs of beginner learners.

Results and Discussion

Result

Basic Level Learning in Listening Skills

Elementary School is the initial stage of formal education aimed at developing students' basic potential as preparation for continuing to higher levels of education. This education is designed to equip children with solid skills and foundations to participate in social and community life. Therefore, in elementary school, a learning process is implemented that supports all subjects, including Arabic language lessons.

The ideal Arabic language learning should be based on several important principles, one of which is paying attention to the characteristics of children. These principles include: (1) learning that is appropriate to the child's world, (2) starting from things that the child already knows or things that are close and easily accessible to them, (3) connecting learning with things that interest the child, (4) organizing learning materials based on the knowledge the child already possesses, (5) providing tasks oriented towards activities or actions, (6) using teaching materials that combine elements of fiction and non-fiction, and (7) adjusting learning materials to the age of the students.

Listening skills are the process of carefully listening to verbal symbols, understanding, appreciating, and interpreting them to obtain information, grasp the main points of the conversation, and comprehend the meaning of communication that the speaker conveys indirectly through words or spoken language (Muhammad Fathoni, 2018). Meanwhile, Madkur stated that listening skills encompass knowledge of linguistic formulas, which involves distinguishing sounds and understanding the meaning conveyed through them, followed by a review of what has been heard and concluding.

The process of learning *istima* requires complete concentration to develop listening skills. According to asy-Syunthy, there are three levels in *istima'*: *Sima'* (listening unintentionally, like the sound of birds), *istima'* (listening intentionally and understanding its meaning), and *inshot* (the highest level of *istima'*). Generally, *istima'* is passive-receptive, where the initiative of communication comes from others, while the listener's role is to listen and understand the content of the message. Good listening skills will enhance a person's ability to absorb the information or knowledge they receive. In the learning process, this skill is essential for every student. Students who are less skilled in listening tend to experience obstacles during learning, while those who are skilled will find it easier to understand and complete their studies. Thus, listening skills are closely related to the success of the learning process, as they are a fundamental language skill that precedes the mastery of other language skills.

Musthofa Sya'ban's Perspective on the Basic Level Listening Skills Test

Although there have been changes in teaching methods, the content of foreign language learning, and the facilities and technologies used, these changes have not been accompanied by similar developments in listening comprehension testing methods. Even in the context of a functional approach, most of the focus still relies on oral tests to measure comprehension of heard speech. In principle, two types of criteria must be met by language skills tests, including listening tests: basic criteria that apply to every test regardless of the topic, and additional criteria that are specifically applied to each test (هاني اسماعيل رمضان, ٢٠١٨).

The basic criteria are general guidelines that establish the primary specifications to be considered when designing various types of tests, including the skills measured by those tests.

Ta'ima has established these five basic criteria.

1. Validity (Warsah, 2022). Validity refers to the extent to which a test can measure what it is supposed to measure. Based on this, a listening test filled with difficult words and expressions that do not match the students' proficiency level cannot be considered valid, as it not only measures listening skills but also assesses vocabulary mastery.
2. Reliability refers to the consistency of a test's results when administered to the same group of individuals under the same conditions and within a short time frame. Therefore, a test that shows significant changes in student outcomes when repeated cannot be considered reliable.
3. Objectivity refers to the absence of personal influence from the evaluator in the grading or assessment of students on tests. To achieve objectivity, students must clearly understand the test instructions, and the questions and expected answers must have uniform interpretations. Additionally, the physical and psychological conditions of the students during the test must be ensured to be supportive (Mustopa et al., 2021).
4. Practicality. Practicality means that the test does not require significant effort from the teacher in its preparation, implementation, or assessment. For example, some tests require students to use specific types of pens or paper, or to be conducted under certain time and place conditions, or involve complex assessment and interpretation procedures.
5. Discrimination Ability (Discrimination). Discrimination ability is the extent to which a test can differentiate between strong and weak students. To achieve this, a range of questions is needed that covers difficulty levels from easy to hard, so that the scores obtained by students can be distributed proportionally between the highest and lowest scores.

The five basic criteria in test design play an important role in determining the quality of measurement being carried out. Validity ensures that the test truly measures the ability or competence being targeted, so that the results are relevant to the instructional goals (Arbeni et al., 2025) An invalid test, such as a listening test filled with difficult terms beyond the students' capacity, actually obscures the results because it measures other aspects not intended. Reliability, on the other hand, functions to maintain consistency of test results when administered under the same conditions. When a test is unreliable, fluctuating outcomes can raise doubts about the accuracy of the measurement.

Objectivity is equally important, as it prevents the assessor's subjective influence on test results. To achieve this, clear questions and unambiguous instructions are required, along with a supportive testing environment. In addition, practicality becomes an important criterion that emphasizes efficiency. A test that is complex and time-consuming not only burdens teachers but can also hinder the assessment process on a larger scale. Finally, discrimination ability is crucial to ensure that the test can distinguish between students with different proficiency levels. This allows the test results to provide a more detailed picture of individual strengths and weaknesses.

According to Mushtofa Sya'ban, listening tests for the basic level should be designed in such a way that they can measure students' ability to distinguish various phenomena accompanying the listening process. This is important because at the basic level, listening skills are more focused on recognition, comprehension, and differentiation between basic elements in spoken language. Listening tests at this level not only measure the ability to understand meaning but also sensitivity to phonetic aspects, intonation, word stress, and characteristic

rhythm patterns in the target language (هاني اسماعيل رمضان, ٢٠١٨). The criteria are as follows:

1. The ability to distinguish similar sounds. The test must measure the students' ability to distinguish similar sounds that have slight differences, such as "ba' " and "mim", "ta' " and "dal", or "sin" and "sad". Additionally, students must also be able to differentiate between long and short vowels. (*mad* and short *harakat*).
2. The ability to distinguish sound combinations. The test should assess students' ability to understand sound blending, such as when *hamzah washal* connects with the following word (*iqra' wa sam'a*), or the omission of the lam sound in the sun letters, as in the word *al-shams*.
3. The ability to distinguish sounds without written letters. The test needs to measure students' ability to recognize sounds that are not written, such as *tanwin* or sound changes in words like "lakin" and "walakin".
4. The ability to distinguish changes in termination (*waquf*). The test should train students to distinguish stopping rules in words, such as when stopping at a noun with *tanwin* (*jaa'a taalibun*) or one ending with *ta' marbuta* (*hiya taalibah*).
5. The ability to distinguish between two consonants that meet. The test should also include questions to recognize changes when two consonants meet, such as in "uktub al-waqjib" or "qadd antaqa al-qithar".
6. Ability to interact physically. This test measures students' understanding of simple movement-related instructions, such as "Sit down," "Quick," or "Open your book".
7. The ability to interact with images. The test must involve images. Students are asked to match the sound or word they hear with the corresponding image, or to answer whether the teacher's description of the image is correct or incorrect.

To understand the main difficulties faced by exam question setters in assessing listening skills at the basic level, Musthofa Sya'ban designed an exam and demonstrated through this design the difficulties faced by students and how to overcome them by choosing the types of questions for the exam. The listening test model suitable for the *maharah istima'* for beginner learners includes, *first*, questions to distinguish sounds. *Second*, questions to differentiate body movement interactions. *Third*, questions to distinguish image descriptions.

1. Question to distinguish sounds.

Arabic has a wealth of phonetics that requires precision in listening and pronunciation. In learning this language, one often encounters letters that have similar sounds but carry very different meanings.

Table 1

Distinguishing similar and nearly similar sounds

Letter	Example 1	Example 2	Example 3
ع vs أ	سؤال / سعال	أمل / عمل	أين / عين
ه vs ح	حائل / هائل	نحر / نحر	ساجر / شاهر
ط vs ت	فتور / مطور	تابع / طابع	تبار / طيار
غ vs ح	خير / غير	خاب / غاب	خلاف / غلاف

س VS ث	ثُمَّ / سَمَّ	حَارِثُ / حَارِسَ	التَّائِرُ / السَّائِرُ
ص VS س	سَعِيدُ / صَعِيدُ	السَّيْفُ / الصَّيْفُ	السُّورَةُ / الصُّورَةُ
ظ VS ذ	نَذِيرُ / نَظِيرُ	فَذُّ / فَظُّ	ذِرَاعُ / ظِرَارُ
ز VS ذ	ذَكِيَّ / زَكِيَّ	نَذِيرُ / نَزِيرُ	ذِمَامُ / زِمَامُ

Note. (هاني اسماعيل رمضان, ٢٠١٨)

Table 2

Differences between short and long vowels

Short Vowel	Short Vowel
فَتَحَ	فَاتَحَ
ضَرَبَ	ضَارَبَ
كَتَبَ	كَاتَبَ

Note. (هاني اسماعيل رمضان, ٢٠١٨)

Table 3.

Hamzah wasal and its combination

Initial Sentence	After Being Merged
إِقْرَأْ وَ سَمِعْ	إِقْرَأْ وَ اسْمِعْ
يَوْمَ لَثْنَيْنِ	يَوْمَ الْإِثْنَيْنِ
بَعْدَ نَتَهَاءِ دَرَسِ	بَعْدَ انْتِهَاءِ دَرَسِ

Note. (هاني اسماعيل رمضان, ٢٠١٨)

2. Questions about physical body interactions

This type of question is marked when trained during the study because it can provide students with several words and sentences that have the characteristics of imperative sentences (imperatif). This enhances their ability to absorb such verbs and improves their understanding of expressions that require them to carry out specific instructions related to certain body movements, connecting their limbs with external objects. The goal is for them to perform movements or describe actions based on the spoken words they hear.

3. Questions to differentiate image descriptions.

This type of exercise has a unique aspect in the learning stage because it helps to revisit students' memories of sounds, words, and sentences they have previously learned. In practice, teachers can use a series of images, where students are asked to listen to a specific sound or word, and then match it with the corresponding image. Additionally, teachers can bring simple or combined images and then describe each one. The students' task is to determine whether the description is correct or incorrect. The teacher can also read specific sentences and ask the students to choose the image that best matches the sentence from the available options. This

exercise aims to connect sentences or speech with images. With the increasing variety of images used, students will become more trained in overcoming various difficulties in this type of question. Therefore, in the exam, such questions are usually divided into five main categories. (هاني اسماعيل رمضان, ٢٠١٨).

a. Questions to connect sound with images, Example:

- 1) For example, students listen to the sound (ث), and then are asked to choose the letter that corresponds to that sound from the available images.

ث س ش

Students must choose the picture of the letter ث.

- 2) For example, students listen to the sound of the letter ṭā' (ط) and are then asked to choose the word containing that letter from the available pictures.



Source: <https://www.canva.com>

In this case, the students must choose picture number (1) and number (3) (airplane – طائرة – محطة القطار)

b. Questions to connect words with images, Example:

- 1) Students listen to the word and then are asked to mark (correct) in front of each picture that corresponds to the word; for example, listening to the word (children/ أطفال).



Source: <https://www.google.com/search?q=gambar>

In this case, the student must mark the correct answer in front of picture numbers (3) and (4).

- 2) Alternatively, the students listen to (the number) and then are asked to mark (correct) in front of each picture that corresponds to that number; for example, listening to the number (two).

a) Questions that involve commenting on an image with a simple description listened to by students, then asked to evaluate the description, such as:



هذه شجرة فالعبارة هنا صحيحة

- b) Questions that involve commenting on an image with a more complex description that is listened to by students, who are then asked to evaluate the description, such as:



هذه امرأة وسيمة تضحك

- c) Questions that involve matching sentences with images, where students listen to a sentence and are asked to choose the image that corresponds to the sentence they heard, such as:

[A man tastes the food.] رجل يتذوق الطعام

Image No. (2) is the image that was heard.



Note. <https://www.google.com/search?q=gambar>

Discussion

Listening skills are regarded as a primary skill that must be mastered, as they form a crucial foundation for acquiring other language skills, such as speaking, reading, and writing. The process of listening involves the ability to hear, understand, appreciate, and interpret spoken messages effectively. Additionally, listening also requires complete concentration to grasp the essence of the conversation and understand the implied meaning in communication. As Tampubolon stated, listening skills are considered a fundamental skill in acquiring the Arabic language, serving as a foundation for other language skills such as speaking, reading, and writing (Muhammad Syafii Tampubolon a, 2024; Soleha et al., 2025). This involves the ability to listen, understand, appreciate, and interpret spoken messages effectively (Muhammad Fathoni, 2018). To improve listening skills, various methods and teaching media can be used. For example, podcasts on Spotify provide a flexible and practical way for students to enhance their listening skills at any time and from anywhere (Mumtaz & Abdurrahman, 2022).

When developing a listening skills test, several important criteria must be considered, including validity, reliability, objectivity, practicality, and discrimination ability. According to Sihabudin, the development of practical language tests, particularly for Arabic listening skills, requires careful consideration of the main criteria (Ma, 2024). These criteria include validity, reliability, objectivity, practicality, and discriminative power (Sihabuddin, 2023). To ensure validity, test developers must follow specific procedures, including determining learning objectives, selecting appropriate materials, creating indicators, and preparing test blueprints.

Analysis of test items using programs like *Anates* can provide insights into validity, reliability, discrimination power, and difficulty level (Gusmizain, 2022).

The basic level listening skills test (*maharah istima'*) is designed to measure students' ability to distinguish elements of Arabic phonetics. Sample questions include differentiating similar letter sounds, such as *ba'* and *mim*, as well as recognizing long and short vowels. According to Musyarofah, the development of listening skills assessment instruments (*maharah istima'*) in Arabic language learning is very important for evaluating students' abilities. At the basic level, this test is designed to measure students' ability to distinguish phonetic elements, including similar letter sounds and vowel lengths (Musyarofah, 2024; Qomaruddin & Haq, 2022). The development of this instrument aims to create a structured and systematic test that accurately measures students' Arabic listening skills, which are fundamental for effective communication (Pahlefi, 2022). The appropriate assessment tools are essential to determine the level of students' ability in listening comprehension of the Arabic language.

In addition, the questions also train students to understand sound combinations, such as *hamzah washal* or stopping rules (*waquf*). Another type of question is physical interaction, which involves understanding simple commands, and image-based questions, which ask students to match words or sounds with images, assess the accuracy of descriptions, or select images that correspond to the statements heard. According to Fathoni, listening exercises in Arabic language learning can be designed to enhance students' understanding of sound combinations (Imawan et al., 2023; Muhammad Fathoni, 2018). Visual aids, especially image-based questions, play an important role in enhancing student engagement and learning outcomes, particularly in Arabic listening skills. By presenting relevant images, students can more easily connect the audio information they hear with the visual context, thereby improving their understanding of the material (Fahrur Rosikh et al., 2022). Thus, the design of the *Istima* test not only functions as a measuring tool but also as a learning tool that forms comprehensive listening competence.

When compared with previous research, the evaluation of basic-level listening skills still has various limitations. Firdaus et al. (2023) found the absence of standardized tests that take into account the characteristics of beginners and the integration of phonetic and comprehension aspects. Upiyani (2024) noted that the existing evaluation model overlooks students' gradual development and psychological aspects. Habibah & Syihabuddin (2021) through the whispering chain game succeeded in improving listening skills, but the instrument was limited to games, without providing a comprehensive evaluation framework (Ahmad Wahyudi, 2024). Zulpina (2022) uses the Quizizz application to test language skills, including *istima'*, but the focus is more on online media rather than instrument standardization. In contrast, this research presents a comprehensive approach based on Musthofa Sya'ban's theory, which integrates phonetic, cognitive, visual, and humanistic aspects to develop a test model that is more structured, valid, and tailored to the needs of elementary-level learners.

Based on these findings, this research confirms several important recommendations. First, Arabic teachers should utilize audio-visual media in listening tests so that students get used to connecting phonetic input with visual context. Second, the preparation of evaluation instruments should be based on indicators of validity, reliability, objectivity, and distinguishing power to ensure the quality of test results. Third, a humanistic approach needs to be integrated into the evaluation process to maintain student motivation and participation. The main contribution of this research lies in the formulation of an *istima'* test model that is

comprehensive and appropriate for the characteristics of beginners, making it a conceptual and methodological reference for developing evaluations of Arabic listening skills at the basic level.

Conclusion

This research confirms that listening skills (*maharah istima'*) are the primary foundation in learning Arabic, especially for elementary level students. The main findings show that the *istima'* test design, developed based on Musthofa Sya'ban's perspective, not only functions to assess literal understanding but also encourages the development of more complex linguistic abilities, including phonetic, cognitive, visual, and psychological aspects. Thus, this instrument can accommodate the needs of novice learners more thoroughly than previous test models.

The contribution of this research can be seen in three dimensions. Conceptually, this research emphasizes the urgency of standardizing holistic *istima'* tests, which do not just focus on measuring learning outcomes but also on the process of forming basic competencies. Theoretically, this research expands Arabic language evaluation studies by integrating the principles of Musthofa Sya'ban into the context of beginner-level learners, a perspective that is still rarely studied. Methodologically, this research provides a framework based on a literature study and content analysis, which can serve as a starting point in developing an *istima'* test instrument that is more valid, reliable, and appropriate for the characteristics of beginners.

However, this research has several limitations. The study remains theoretical in nature, so empirical research is needed to test the validity, reliability, and effectiveness of the instrument in learning practice. In addition, further research can be directed at developing *istima'* tests for intermediate and advanced levels, as well as integrating digital technology as an adaptive evaluation medium. This recommendation will open up opportunities for the birth of listening test instruments that are more applicable, innovative, and relevant to the needs of Arabic language education in the digital era.

Acknowledgment

Alhamdulillah, we thank Allah SWT for giving His grace, Taufik, and guidance, so that the writing of this research can be completed properly. On this occasion, we would like to express our sincere gratitude to all parties who have provided their support and contributions during this research process.

First of all, we would like to express our gratitude to Sunan Kalijaga State Islamic University Yogyakarta, which has provided the means and opportunity to carry out this research. Thank you also to all parties who have contributed to creating an academic atmosphere that supports the growth of science and education at this university. We would also like to express our highest appreciation to the supporting lecturers who have guided and provided guidance during this research process. Especially to Dr. R. Umi Baroroh, S. Pd. I, M. Si, who has provided very valuable guidance, encouragement, and advice in the preparation of this research. Without his guidance and support, this research would not have been able to be carried out properly.

In addition, we would like to thank all parties who have provided moral and material support, both directly and indirectly. Hopefully, all the kindness and assistance that has been given will get a fair reward from Allah SWT. In the end, hopefully, this research can provide benefits for the development of science and educational practice, especially in learning Arabic.

Thank you to all who have contributed to this research.

References

- Ahmad Wahyudi. (2024). Use of Picture Card Media to Improve Beginning Reading Skills of Elementary Students. *Journal of Childhood Development*, 4(1), 154–160. <https://doi.org/10.25217/jcd.v4i1.3971>
- Arbeni, W., Bilbina, A., Intan, Y., Karokaro, B., Wijaya, D., Ginting, I. R., Irma, A., & Harahap, Y. (2025). Planning Strategies in Test Development: A Systematic Approach to Educational Evaluation. *Jurnal Nasional Holistic Science*, 5(1), 40–43.
- Ashari, M. Y., Huda, M. M., & Mahfudhoh, R. (2024). Enhancing Arabic Writing Skills Through the Genre-Based Approach in Senior High School. *Journal of Arabic Language Teaching*, 4(2), 121–134. <https://doi.org/10.35719/arkhas.v4i2.2137>
- Astuti, S., Afrianni, L., Fathyyah Cahyani, I., & Education Program STKIP Kusuma Negara Jakarta, E. (2022). A LIBRARY RESEARCH: THE IMPLEMENTATION OF GENRE BASED APPROACH IN INDONESIAN VOCATIONAL HIGH SCHOOLS. In *JOLADU : Jurnal Of Language and Education* (Vol. 1, Issue 2).
- Baroroh, R. U., & Rahmawati, F. N. (2020). Metode-Metode Dalam Pembelajaran Keterampilan Bahasa Arab Reseptif. In *Urwatul Wutsqo: Jurnal Studi Kependidikan dan Keislaman* (Vol. 9, Issue 2, pp. 179–196). STIT Al-Urwatul Wutsqo Jombang. <https://doi.org/10.54437/urwatulwutsqo.v9i2.181>
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). Sage Publications.
- Fahrur Rosikh, Zumrotus Sholihah, Dyah Putri Larasati, & Ach Nurhamid Awalluddin. (2022). Penggunaan Media Gambar untuk Meningkatkan Minat dan Hasil Belajar Bahasa Arab. *Kilmatuna: Journal Of Arabic Education*, 2(2), 293–302. <https://doi.org/10.55352/pba.v2i2.80>
- Firdaus, D., Ainin, M., Muashomah, & Siregar, H. H. (2023). Menakar Al Maharah Al Lughawiyah dalam KMA 183 dengan Standar ACTFL. In *Shaut al Arabiyyah* (Vol. 11, Issue 1, pp. 1–18). Universitas Islam Negeri Alauddin Makassar. <https://doi.org/10.24252/saa.v11i1.29255>
- Gunarti, T. T. (2020). Pengembangan Media Pembelajaran Bahasa Arab Berbasis Audio Visual Untuk Meningkatkan Maharah Istima' Pada Siswa-Siswi Madrasah Ibtidaiyah. *Awwaliyah: Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, 3(2), 122–129.
- Gusmizain, A. (2022). Karakteristik Butir Soal Tes Mata Kuliah Matriks & Ruang Vektor Mahasiswa Matematika. *Jurnal Evaluasi Pendidikan*, 13(2), 127–130. <https://doi.org/10.21009/jep.v13i2.28328>
- Habibah, A., & Syihabuddin. (2021). Evaluasi Ketrampilan Menyimak Dengan Memanfaatkan Permainan Berbisik Berantai Dalam Pembelajaran Bahasa Arab. *Al-Ittijah: Jurnal Keilmuan Dan Kependidikan Bahasa Arab*, 12(2), 97–106.
- Hakim, T. L., Haris, A., & Huda, M. M. (2025). Assessment of Speaking Skills (Maharah Kalam) in Arabic Language Lesson at Madrasah Ibtidaiyah Riyadlus Sholihien Jember. *Journal of Science and Education (JSE)*, 5(2), 599–605. <https://doi.org/10.58905/jse.v5i2.485>
- Handriawan, D. and N. M. (2021). Evaluasi Pembelajaran Bahasa Arab. In *Sanabil Publishing*. Penerbit Adab.

- Haniah, H. (2018). Analisis Kesalahan Berbahasa Arab Pada Skripsi Mahasiswa Jurusan Bahasa Dan Sastra Arab. *Arabi : Journal of Arabic Studies*, 3(1), 23. <https://doi.org/10.24865/ajas.v3i1.62>
- Imawan, Y., Rahmatan, M., Hania, I., & Alimudin, A. (2023). Ashwat's Teaching Strategies and Their Implications In The Learning of Maharah Istima'. In *International Journal of Education and Teaching Zone* (Vol. 2, Issue 1, pp. 13–24). Yayasan Nurul Yakin Bunga Tanjung. <https://doi.org/10.57092/ijetz.v2i1.55>
- Jamil, M., Nawawi, M., Rohmaniah, S., & Andrianto, D. (2023). Desain Tes Bahasa Arab Menggunakan Aplikasi Wondershare Quiz Creator Untuk Meningkatkan Maharah Istima'. *Attractive : Innovative Education Journal*, 5(3), 1–12.
- Ma, X. (2024). Enhancing language skills and student engagement: investigating the impact of Quizlet in teaching Chinese as a foreign language. *Language Testing in Asia*, 14(1). <https://doi.org/10.1186/s40468-024-00275-3>
- Muhammad Fathoni. (2018). Maharah Istima '. *Jurnal Komunikasi Dan Pendidikan Islam*, 1, 2020.
- Muhammad Syafii Tampubolon a, R. I. (2024). The priority of istima' skills in arabic learning. *Amorti : Jurnal Studi Islam Interdisipliner*, 3(2), 22–26.
- Mumtaz, F., & Abdurrahman, M. (2022). Pembelajaran Maharah Al-Istima' Menggunakan Media Podcast pada Aplikasi Spotify. In *JURNAL PENDIDIKAN BAHASA* (Vol. 12, Issue 2, pp. 41–45). STKIP Taman Siswa Bima. <https://doi.org/10.37630/jpb.v12i2.1005>
- Mustopa, A., Jasim, J., Basri, H., & Barlian, U. C. (2021). Analisis Standar Penilaian Pendidikan. *Jurnal Manajemen Pendidikan*, 9(1), 24–29. <https://doi.org/10.33751/jmp.v9i1.3364>
- Musyarofah, A. (2024). Pengembangan Instrumen Tes Keterampilan Mendengar Untuk Pembelajar Bahasa Asing. *Afeksi: Jurnal Penelitian Dan Evaluasi Pendidikan*, 5(3), 431–441. <https://doi.org/10.59698/afeksi.v5i3.284>
- Pahlefi, M. R. (2022). Pengembangan Instrumen Penilaian Keterampilan Menyimak (Mahārah al-Istima') dalam Pembelajaran Bahasa Arab. *Uktub: Journal of Arabic Studies*, 2(2), 68–84. <https://doi.org/10.32678/uktub.v2i2.6458>
- Qomaruddin, F., & Haq, M. A. (2022). Intervensi Fonetik Bahasa Arab dan Bahasa Indonesia (Studi Kontrastif Bahasa Arab dan Indonesia Pada Tataran Pelafalan Kata). In *Borneo Journal of Language and Education* (Vol. 2, Issue 1, pp. 71–83). UINSI Samarinda. <https://doi.org/10.21093/benjole.v2i1.6215>
- Rika Noverma, Hikmah, H. D. (2024). Standardization of Arabic Language Test Questions For Class Xii Man 1 Pekanbaru Based On The Perspective Of Actfl. *ELMAKTABAH : International Journal Of Educational Research*, 15(1), 37–48.
- Rohmat, B., Maulana, A., & Huda, M. M. (2025). THE ROLE OF DIGITAL LITERACY IN ARABIC LANGUAGE LEARNING IN THE ERA OF INDUSTRIAL REVOLUTION 4.0. *Islamic Management: Jurnal Manajemen Pendidikan Islam*, 8(01), 541–550. <https://doi.org/10.30868/im.v8i01.8197>
- Sihabuddin. (2023). Prosedur Penyusunan Tes Berbasis Hots Pada Empat Keterampilan Berbahasa Arab. *LEARNING : Jurnal Inovasi Penelitian Pendidikan Dan Pembelajaran*, 3(1), 40–47. <https://doi.org/10.51878/learning.v3i1.2035>
- Soleha, S., Bukhori, E. M., & Huda, M. M. (2025). *Transforming Student 's Arabic Writing Skills through Word Square Media : Examining its Effectiveness Transformasi*

- Keterampilan Menulis Bahasa Arab Siswa melalui Media Word Square : Menguji Efektivitasnya*. 08, 89–107. <https://doi.org/10.30762/asalibuna.v8i02.5231>
- Suyitno, S. (2021). Penerapan Kompetensi Psikologi Guru dalam Peningkatan Motivasi Belajar Siswa. *Jurnal Basicedu*, 6(1), 58–65. <https://doi.org/10.31004/basicedu.v6i1.1900>
- Upiyani, I. C. (2024). Prinsip-Prinsip Dasar Evaluasi dan Implikasi dalam Menilai Instruksi Iman Islam di Lembaga Pendidikan. *Jurnal PAI Raden Fatah*, 15(1), 37–48. <https://doi.org/10.19109/pairf.v6i1>
- Warsah, I. (2022). Evaluasi Pembelajaran (Konsep . Fungsi dan Tujuan). *Jurnal Kajian Pendidikan Islam*, 1, 190.
- Wulan, T. A. (2023). *Model Evaluasi Countance Stake Dalam Pembelajaran Bahasa Arab Di Smp It Ruhul Jadid Kecamatan Tigaraksa Kabupaten* 22.
- Zulpina, Z. (2022). Quizizz, Media Alternatif Pembelajaran Bahasa Arab Online Bagi Siswa Madrasah Ibtidaiyah. *Al-Madrasah: Jurnal Pendidikan Madrasah Ibtidaiyah*, 6(3), 775. <https://doi.org/10.35931/am.v6i3.1089>
- هاني اسماعيل رمضان. (٢٠١٨). معايير مهارة اللغة العربية للناطقين بغيرها. المتتدى العربي التركي.