



The Effectiveness of the Voice-Based Arabic Conversation Application as an Interactive Learning Innovation for Non-Arab Speakers

Tri Agung Purnama, Muhammad Ikbal Umahani

Program Studi Pendidikan Bahasa Arab, Fakultas Ilmu Tarbiyah dan Keguruan
IAIN Sultan Amai Gorontalo, Indonesia

Volume 6 Nomor 1
Januari 2025: 25-37
DOI: 10.30997/tjpba.v6i1.16559

Article History

Submission: 20-12-2024

Revised: 26-01-2025

Accepted: 26-01-2025

Published: 28-01-2025

Keywords:

Conversation Application, Arabic Language, Non-Arabic Speakers, Speaking Skills, Interactive Learning, T-Test

Keywords:

Conversation applications, Arabic, Non-Arabic Speakers, Speaking Skills, Interactive Learning, T-Test

Correspondence:

(Sella Robi'ah)
(082139427233)
(sellanginayatur@gmail.com)

Abstract: This study aims to test the effectiveness of voice-based Arabic conversation applications in improving the speaking skills of non-Arab speakers. Using a quasi-experimental design with experimental and control groups, this research involved 20 participants. The experimental group used the application for two weeks while the control group used the conventional method. The results showed significant improvements in the experimental group in the aspects of pronunciation, fluency, comprehension and interactive response, compared to the control group. Qualitative data revealed that the app increased participants' comfort, motivation, and self-confidence. The findings show that this sound-based application can be used as an effective learning medium.

Efektivitas Aplikasi Percakapan Bahasa Arab Berbasis Suara Pada Inovasi Pembelajaran Interaktif Bagi Penutur Non Arab

Abstrak: Penelitian ini bertujuan untuk menguji efektivitas aplikasi percakapan bahasa Arab berbasis suara dalam meningkatkan keterampilan berbicara penutur non-Arab. Menggunakan desain eksperimen kuasi dengan kelompok eksperimen dan kontrol, penelitian ini melibatkan 20 partisipan. Kelompok eksperimen menggunakan aplikasi selama dua minggu sedangkan kelompok kontrol menggunakan metode konvensional. Hasil penelitian menunjukkan peningkatan signifikan pada kelompok eksperimen dalam aspek pengucapan, kefasihan, pemahaman dan respons interaktif, dibandingkan dengan kelompok kontrol. Data kualitatif mengungkapkan bahwa aplikasi meningkatkan kenyamanan, motivasi, dan kepercayaan diri partisipan. Temuan menghasilkan



INTRODUCTION

The learning of Arabic for non-Arabic speakers continues to evolve in line with the increasing demand for language skills in a global context (Ruslan et al., 2023). Arabic is not only a language of communication in Arab countries but also a primary language in the study of religion, culture, and education in various parts of the world. (Setiyani et al., 2023).

The main challenges arise, particularly for those learning it as a foreign language, because Arabic has linguistic structures and phonetics that differ from their native languages (Qudrat, 2024). Therefore, effective and interactive learning methods are needed to facilitate the process of understanding and mastering Arabic in a practical manner (Mongi Diodara Christin Natalia, 2015).

With the advancement of technology, innovations in education, particularly in language learning, have undergone significant changes (Chumairoh & Permana, 2023). Voice-based applications have now emerged as one of the alternatives considered

capable of addressing this challenge (Pramuniati et al., 2008). This voice-based learning application allows users to engage in a more practical and direct learning process, as if they were interacting with native speakers (Norlistian Ellen Austin & Sabardila Atiqa, 2024). This technology is supported by increasingly advanced artificial intelligence, enabling users to learn accents, intonation, and language structures accurately, as well as receive instant feedback to identify and correct their pronunciation errors (Ameliyatun Hidayah, 2024).

Based on research conducted by (Kurnia et al., 2023) The title "The Effectiveness of the Utilization of Inspiring-Assisted Learning Multimedia in Improving Motivation and Learning Outcomes in Arabic Language Subjects" can be concluded that the use of multimedia in learning can provide good motivation and learning outcomes for students.

The research conducted by (Zulfa & Huda, 2021) with the title "The Effectiveness of Mentimeter Media in Understanding Arabic Language

Material During the COVID-19 Pandemic Era" can be concluded that using application-based learning media yielded significant results in students' learning outcomes after its implementation.

The analysis results show that the voice-based Arabic conversation application has proven to be effective in learning Arabic for non-Arab speakers. This application not only facilitates mastery and pronunciation but also provides an interactive experience that closely mimics real conversations. Users can develop both speaking and listening skills simultaneously, which are two essential aspects of holistic language mastery. The use of this application has been shown to increase learning motivation, as users can learn with flexibility in terms of time and place, while also receiving accurate corrections through immediate feedback (Hermansyah Hermansyah, 2023).

With these various advantages, the voice-based Arabic conversation

application has great potential to be widely used in Arabic language learning. This technology can serve as an effective learning tool and support the process of language acquisition in a practical and efficient manner, especially among non-Arab speakers who require independent and interactive learning resources. (Sinaga et al., 2021).

METHODS

This study used a quasi-experimental method to test the effectiveness of a voice-based Arabic conversation application in improving speaking skills for non-Arab speakers (Chumairoh & Permana, 2023). This approach was carried out by comparing two groups, namely the experimental group that used the application during learning and the control group that learned using conventional methods without the application (Pera Aprizal, 2021).

The following is a research scheme that the author uses, including the following:

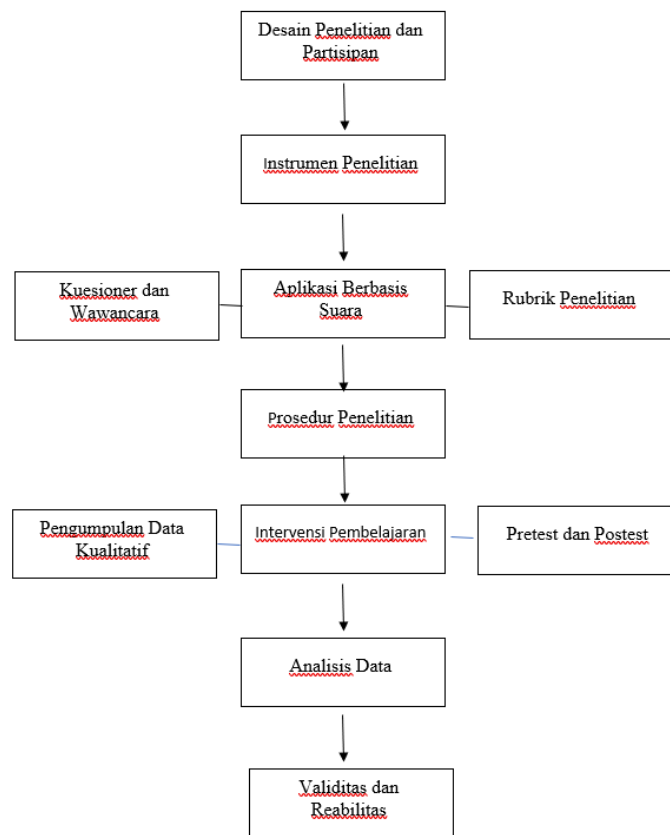


Figure 1 Research Scheme

Based on Figure 1 above, the researcher prepared a research scheme as a guide to designing a systematic and structured research implementation flow.

RESULTS & DISCUSSION

Research Design and Participants

This study used a pre-test and post-test design involving two groups, namely the experimental group and the control group. Each group consisted of 10 participants who were selected purposively based on the criteria of non-Arabic speakers. The purposive

selection of participants was intended to ensure that the samples involved were in accordance with the research objectives, namely to measure the effectiveness of voice-based conversation applications in improving Arabic speaking skills.

a. Experimental Group

This group consisted of 10 participants who used a voice-based conversation application as a learning medium during the duration of the study. The

application was designed to help users improve their pronunciation and increase their speaking fluency through direct feedback. Participants in this group accessed the application according to a predetermined schedule and were asked to practice independently.

b. Control Group

This control group also consisted of 10 participants who followed learning with conventional text and audio teaching materials which are traditional learning methods. They learned using textbooks and audio recordings without the help of sound-based technology. Learning in this group was passive, where participants listened and imitated conversations in Arabic without any automatic correction related to pronunciation or fluency.

The research procedure that the researcher designed was where at the beginning of the study all participants from both groups took a pre-test to measure their speaking skills. After that, the experimental group continued learning with a voice-based application while the control group followed

learning with conventional methods. After the specified learning duration, both groups took a similar post-test to evaluate the development of their speaking skills.

With this design, the study can compare the effectiveness of using voice-based conversation applications against conventional learning methods in improving Arabic speaking skills for non-Arabic speakers.

Research Instruments

To measure the effectiveness of voice-based conversation applications in improving Arabic speaking skills, this study used three main instruments, namely voice-based applications, assessment rubrics and questionnaires, and interviews. Each of these instruments is designed to support quantitative data collection so that it can provide a more complete picture of the effect of the application on participants' speaking skills.

a. Voice Based Applications

The voice-based application is the main instrument used in the experimental group. This application is specifically designed to help non-Arabic speakers practice their Arabic speaking skills.



Figure 2 Voice Based Applications

Based on figure 3.1 above, the main features in this application are voice recognition, conversation simulation and feedback.

b. Assessment Rubric

To assess the participants' speaking ability, this study used a

criteria-based assessment rubric.

This rubric is arranged based on the main aspects such as pronunciation, fluency, comprehension and interactive response. Each aspect is assessed using a Likert scale of 1-5 with the following criteria.

Table 1 Assessment Rubric Table

Aspek Penilaian	Kriteria	Skor 1	Skor 2	Skor 3	Skor 4
Pengucapan	Ketepatan dalam melafalkan kata dan huruf khas bahasa arab	Sangat banyak kesalahan	Banyak kesalahan	Beberapa kesalahan	Sedikit Kesalahan
Kefasihan	Kelancaran berbicara tanpa jeda yang mengganggu	Tidak lancar sama sekali	Banyak Jeda mengganggu	Beberapa jeda	Lancar dengan sedikit jeda
Pemahaman	Kemampuan memahami konteks percakapan dan memberikan tanggapan yang relevan	Tidak memahami sama sekali	Memahami Sebagian kecil	Memahami Sebagian besar	Memahami dengan baik
Respons Interaktif	Kecepatan dan relevansi dalam merespons lawan bicara	Tidak memberikan respons	Respons sangat lambat	Respons cukup lambat	Respon cepat

Based on Table 1 about the assessment rubric table using a Likert

scale of 1-5 for each aspect assessed, namely pronunciation, fluency,

understanding, and interactive response. A score of 1 indicates very low skills, with a need for fundamental improvement while a score of 5 reflects very good skills and meets the expected standards. A score of 2 indicates abilities that are still below standard with many shortcomings while a score of 3 indicates fairly good abilities but requires further development. A score of 4 indicates good abilities with few shortcomings.

This assessment rubric is applied during the pre-test and post-test assessments to evaluate the participants' speaking ability in a structured and objective manner. The

assessments are conducted by two independent raters to minimize bias in scoring. Each participant will be given a score for each aspect based on their performance and the average score of the four aspects is used to calculate the overall development of speaking skills.

c. Questionnaire

The questionnaire was designed to collect data on participants' level of satisfaction with the application, their perceptions of the application's features, and their experience in using the application during the study period.

Table 2 Questionnaire Results Table

Pertanyaan	STS	TS	CS	S	SS	Rata-rata
Aplikasi mudah digunakan	0	0	2	5	5	4,1
Aplikasi efektif memperbaiki pengucapan Bahasa arab	0	1	3	4	2	3,7
Fitur umpan balik instan membantu memperbaiki kesalahan	0	1	4	3	2	3,6
Skenario percakpaan relevan dengan kebutuhan belajar	0	0	2	5	3	4,1
Aplikasi meningkatkan motivasi belajar berbicara Bahasa arab	0	0	3	4	3	4,0

d. Interview

In-depth interviews were conducted to further explore individuals' experiences and motivations in using the application.

Pre-Test Stage

Before the intervention began, all participants took a pre-test to measure their basic speaking skills. The pre-test consisted of a speaking test covering several common Arabic conversation scenarios, which were assessed on

aspects of pronunciation, fluency, comprehension and interactive response.

The initial test was conducted on 20 participants who were divided into two groups.

Table 3 Pre-Test Results

Group	Average Pre-Test Score
Experimental Group	45.8
Control Group	47.2

Based on table 3.3 above, it means that the scores above indicate a similar level of basic speaking skills between the two groups at the beginning.

Overall, the pre-test scores indicated that both groups had fairly equal levels of basic speaking skills.

Learning Interventions

a. Experimental Group

The experimental group used a voice-based conversation application for two weeks with a total of 4 learning sessions. The application used was designed to improve Arabic speaking skills through feedback features, conversation simulations and progress measurements.

b. Control Group

The control group studied using conventional methods for the same duration. Participants studied Arabic conversation texts from the provided module and listened to audio recordings as pronunciation models.

Post - Test

After completing the two weeks, all participants took a final test (post-test) using the same instrument as the pre-test. This test was designed to evaluate changes in Arabic speaking skills such as pronunciation, fluency, comprehension and interactive response.

Table 4 Post-Test Results Table

Kelompok	Rata-rata skor Pre-Test	Rata-rata skor Post-Test	Peningkatan (%)
Eksperimen	45,8	72,3	58%
Kontrol	47,2	50,1	6%

Based on Table 4 above, it can be explained that participants who used the voice-based conversation

application or the experimental group showed an average increase of 26.5 points in the post-test score with a

significant increase in the pronunciation aspect of 65% and interactive response of 58%.

Meanwhile, the increase in scores in the control group was only 2.9 points, indicating that conventional methods are not enough to encourage significant improvement in speaking skills.

Qualitative Data

Qualitative data collection was conducted through structured

interviews and questionnaires given to the experimental group after the learning intervention. The purpose of this data collection was to explore the subjective experiences of participants related to the use of voice-based conversation applications, especially in terms of comfort, learning effectiveness, and learning motivation.

Tabel 5 Qualitative Data

Aspect	Positive Percentage
Comfort	85%
Feedback Effectiveness	95%
Scenario Relevance	80%
Motivation to learn	90%
Confidence	85%

Based on table 3.6 above, it can be explained that the results of qualitative data analysis show that most participants responded to the use of voice-based conversation applications. In terms of convenience, 85% of participants reported a positive experience. The user-friendly interface and intuitive navigation make the application easy to use, even for beginners. The effectiveness of the application's feedback is also highly appreciated with 95% of participants stating that this feature helps them

correct pronunciation errors directly which contributes to increased speaking fluency.

In addition, 80% of participants felt that the conversation scenarios provided by the application were relevant to daily communication needs, making learning more contextual and practical. In terms of motivation, 90% of participants admitted to being more enthusiastic about learning Arabic through this application. The interactive features and challenges provided by the application succeeded in increasing

their involvement in the learning process. Finally, 85% of participants reported increased self-confidence after practicing using the application.

Data analysis

a. Quantitative Analysis

The results of the pre-test and post-test were analyzed using a paired t-test to compare the improvement in mean scores in speaking skills between the experimental group and the control group.

Table 6 Paired t-test of Experimental

Pasangan variabel	Mean	Std.Deviasi	t	df	Sig.(2-tailed)
Pre-Test					
Post-Test	-32.30	1.57	-103.44	9	0.000

Table 7 Paired t-test of Control Group

Pasangan variabel	Mean	Std.Deviasi	t	df	Sig.(2-tailed)
Pre-Test					
Post-Test	-12.40	0.96	-51.56	9	0.000

Based on Figures 3.3 and 3.4 regarding the paired t-test in two groups, it can be explained that the experimental group showed an average increase in post-test scores of 26.5 points from the pre-test with a p value <0.05 indicating a significant difference. While the control group only increased by 2.9 points and the results of the paired t-test showed a p value > 0.05 indicating an insignificant increase.

b. Qualitative Analysis

Most participants found the app easy to use despite some technical issues such as accent recognition. Instant feedback and relevant

scenarios were the main benefits reported by participants, as well as a significant motivational impact on learning.

Validation Test and Reliability Test

Instrument validation testing in this study was conducted by involving experts to assess the quality of the assessment rubric and application content.

Reliability testing was conducted using the internal consistency method through Cronbach's alpha to ensure that the assessment rubric used can provide consistent and stable results when tested on various participant samples.

Table 8 Validity test of the Assessment Rubric (Expert Review)

Instrumen	Keterangan	Validitas
Rubrik penilaian pengucapan	Diperiksa oleh 3 ahli Bahasa Arab dan 2 ahli teknologi pembelajaran	Valid
Rubrik penilaian kefasihan	Diperiksa oleh 3 ahli Bahasa Arab dan 2 ahli teknologi pembelajaran	Valid
Rubrik penilaian pemahaman	Diperiksa oleh 3 ahli Bahasa Arab dan 2 ahli teknologi pembelajaran	Valid
Rubrik penilaian respons interaktif	Diperiksa oleh 3 ahli Bahasa Arab dan 2 ahli teknologi pembelajaran	Valid

Based on Table 8 above, all assessment rubric instruments used in this study have been validated by a team of experts in the field of Arabic language and learning

technology. This validity confirms that the assessment rubric is appropriate for measuring Arabic speaking skills in accordance with the objectives of the study.

Table 9 Reliability Test (Cronbach's Alpha)

Instrumen	Cronbach's Alpha	Interpretasi
Rubrik Penilaian Pengucapan	0.91	Sangat Reliabel
Rubrik Penilaian Kefasihan	0.89	Sangat Reliabel
Rubrik Penilaian Pemahaman	0.88	Reliabel
Rubrik Penilaian Respons Interaktif	0.90	Sangat Reliabel

Based on Table 9 above, the Cronbach's Alpha value is greater than 0.80 indicating that the assessment rubric has very good

reliability. Thus, this instrument can be relied on to measure Arabic speaking skills consistently in research.

Table 10 Initial Trial (Pilot Test)

Instrumen	Nilai Pre-Test	Nilai Post-Test	Perbedaan (Mean)
Rubrik penilaian pengucapan	56.2	72.1	15.9
Rubrik penilaian kefasihan	58.4	74.0	15.6
Rubrik penilaian pemahaman	57.3	73.5	16.2
Rubrik penilaian respons interaktif	59.1	75.4	16.3

Based on Table 10 above, the initial trial showed a significant increase in pre-test and post-test scores in all aspects tested. The difference in average scores between the pre-test and post-test

indicates that the assessment instrument is effective in measuring Arabic speaking skills after using the voice-based conversation application.

CONCLUSION

Based on the research conducted, several explanations can be drawn regarding the effectiveness of the voice-based Arabic conversation application as a learning medium for non-Arabic speakers.

The effectiveness of the application has been proven in enhancing Arabic speaking skills among non-Arabic speakers. This is evidenced by a significant increase in the post-test scores of the experimental group compared to the control group using conventional methods. The experimental group that used the application achieved an average post-test score of 72.3, while the control group only reached 50.1.

Interviews and questionnaires revealed that most participants felt comfortable using the application, especially due to its user-friendly interface and instant feedback filters that helped improve pronunciation and speaking fluency. The application also provided a more relevant learning experience through the conversation scenarios it offered, while enhancing participants' motivation and self-confidence.

Validity tests conducted by experts and reliability tests using Cronbach's alpha indicated that the research instruments, namely the speaking skills assessment rubrics, had a very high level of validity and reliability. All assessment rubrics were proven to consistently and accurately measure aspects of Arabic speaking skills.

Based on quantitative and qualitative data analysis, it can be concluded that the voice-based application is effective in improving Arabic speaking skills among non-Arabic speakers, particularly in the aspects of pronunciation, fluency, comprehension, and interactive responses. Additionally, the application also boosts participants' learning motivation and self-confidence.

Overall, this study provides evidence that the voice-based conversation application can be an effective and engaging alternative for learning Arabic, particularly for non-Arabic speakers who wish to improve their speaking skills in a more interactive and enjoyable manner.

ACKNOWLEDGMENT

We express our deepest gratitude to both of our parents for their prayers,

encouragement, and support, both morally and materially. We also express our gratitude to all parties who have been involved in the process of making this journal, so that this journal can be can be resolved well.

REFERENCES

- Ameliyatun Hidayah, S. (2024). *SAJAK Analisis Penggunaan Hiponim dan Hipernim dalam Koran Riau Pos Edisi Agustus 2020*. 3(1). <https://journal.uir.ac.id/index.php/sajak>
- Chumairoh, U., & Permana, H. (2023). Polisemi Dalam Media Online Di CNN Arabic. *A Jamiy: Jurnal Bahasa Dan Sastra Arab*, 12(2), 337. <https://doi.org/10.31314/ajamiy.12.2.337-346.2023>
- Hermansyah Hermansyah. (2023). *KAIDAH WADIH AL_DILALAH:HIERARKI DAN URGENSINYA DALAM PENAFSIRAN AL-QURAN*.
- Kurnia, N., Darmawan, D., Program, P., Teknologi, S., Konsentrasi, P., Pembelajaran, T., Pascasarjana, S., Program, P., Teknologi, S., Universitas, P., Indonesia, P., Program, P., Teknologi, S., Konsentrasi, P., Pembelajaran, T., Pascasarjana, S., & Indonesia, U. P. (2023). *EFEKTIVITAS PEMANFAATAN MULTIMEDIA PEMBELAJARAN*. 8(2003), 458-468.
- Mongi Diodara Christin Natalia. (2015). *RELASI LEKSIKAL DAN PROPERTI MAKNA DALAM NASKAH FILM THE BEST EXOTIC MARIGOLD HOTEL*.
- Norlistian Ellen Austin, & Sabardila Atiqa. (2024). *Kohesi Gramatikal dan Leksikal dalam Keterangan Instagram Folkative*.
- Pera Aprizal, A. (2021). Urgensi Pembelajaran Bahasa Arab dalam Pendidikan Islam. *Jurnal Pendidikan Guru*, 2(2), 181-191. <https://doi.org/10.47783/jurpendigu.v2i2.232>
- Pramuniati, I., Bahasa, F., & Seni, D. (2008). *SEMANTIK LEKSIKAL, SEMANTIK KALIMAT, MAKNA DAN KONTEKS BAHASA ACEH BESAR*.
- Qudrat, Y. (2024). *SHAVKAT RAHMON ASARLARIDA ANTONIM VA ANTITEZA MASALALARI*.