

A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ulum

 <https://doi.org/10.31004/jele.v11i3.2376>

*Arini Roihati Farha, Muhammad Hifdil Islam, Zainuddin^{abc} 

¹²³Universitas Islam Zainul Hasan Genggong Probolinggo Indonesia

Corresponding Author: ariniroihati@gmail.com

A B S T R A C T

Accurate pronunciation is fundamental to the development of oral proficiency, as it directly governs the clarity and comprehensibility of verbal interaction; nevertheless, many EFL learners continue to encounter phonetic hurdles stemming from insufficient practice, a lack of corrective feedback, and interference from their native language. While existing scholarship has widely acknowledged the advantages of Mobile-Assisted Language Learning (MALL), most studies remain centered on broad language acquisition, leaving a gap in understanding how specific applications are utilized to resolve pronunciation challenges within localized educational settings. Consequently, this study examines the integration of the CAKE application in English pronunciation instruction and its role in mitigating students' speech difficulties at MA Mirqotul Ulum through a descriptive qualitative framework. Involving an English instructor and their students, data were gathered via classroom observations, semi-structured interviews, phonetic assessments, and documentation, then processed through an interactive model of data analysis. The findings demonstrate that CAKE fosters a self-directed learning environment through its speech recognition and 'Listen and Repeat' functionalities, which offer instantaneous evaluation and lead to observable progress in both segmental and suprasegmental elements, alongside bolstered student motivation and confidence. Ultimately, this research underscores the practical utility of mobile based interventions in pronunciation pedagogy and enriches the discourse on MALL by providing empirical evidence of a specialized application's efficacy in an EFL context.

Keywords: *Pronunciation, MALL, CAKE Application, EFL Students*

Article History:

Received 13th April 2026

Accepted 13th May 2026

Published 19th May 2026



INTRODUCTION

In an increasingly integrated era of globalization, English has solidified its position as an essential universal communication tool across various sectors, ranging from education and technology to cross cultural diplomacy (Swami, 2025). For students, mastering English is not merely a curricular requirement but a fundamental competency for accessing global literacy and participating actively in international dialogue (Kumar, 2023). Consequently, the integration of English instruction within the national education system particularly in Indonesia, where it is taught as a foreign language carries strategic significance.

Among the four fundamental language skills, speaking is often regarded as the most crucial competency, as it serves as a direct representation of an individual's ability to articulate ideas (Akhter, 2021). This skill demands more than just a theoretical understanding of the language system; it requires communicative competence that harmonizes fluency, accuracy, and articulatory quality. Without adequate pronunciation, the effectiveness of communication will be hindered, even if the speaker possesses an extensive command of vocabulary and grammar (Mutali, 2025).

The component of pronunciation is a vital element in the process of acquiring English as a foreign language. This is because articulatory precision significantly influences the success of message transmission, ensuring it is clearly understood by the listener. As emphasized by

A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ullum
 Celce-Murcia, Brinton, and Goodwin, deficiencies in pronunciation frequently trigger semantic distortion and misunderstandings, even when the sentence structure adheres to grammatical standards (Celce-Murcia et al., 2010).

Furthermore, contemporary discourse regarding pronunciation instruction emphasizes the aspect of intelligibility rather than the pursuit of absolute native-like mimicry. This paradigm stresses that the primary focus of pronunciation pedagogy is to assist students in producing speech that is clear and functional in real-world communicative situations. Thus, achieving clarity of meaning during interaction becomes the primary objective for learners (Jenkins, 2015).

Despite its importance, mastering the English sound system often poses a significant obstacle for foreign language learners (Oxána & Krechkovska, 2024). A determining factor for this difficulty is the discrepancy between the English phonological system and the students' native language. Such structural differences in sound often make it challenging for students to produce specific phonemes that are either unfamiliar or non-existent in their native linguistic repertoire.

Technically, the obstacles faced by students encompass both segmental features, such as vowels and consonants, and suprasegmental features, including stress, rhythm, and intonation. Gilakjani argues that a failure to internalize these elements can cause a student's speech to sound unnatural and difficult to comprehend (Gilakjani, 2017). Without a balanced mastery of both aspects, meaningful and effective oral communication remains difficult to achieve in linguistic practice.

Unfortunately, pronunciation teaching practices in many EFL (English as a Foreign Language) classrooms remain confined by the limitations of conventional approaches (Tiwari, 2023). Teachers' instructional focus tends to be largely absorbed by reinforcing grammar and vocabulary, while pronunciation often receives a minimal portion of attention (Pouresmaeil & Vali, 2023). Even when pronunciation materials are provided, the methods used are typically more theoretical than practical, leaving students with little room to hone their motor speech skills intensively.

This issue is further exacerbated by the absence of immediate corrective feedback on students' articulatory errors. Without instant evaluation, students tend to repeat mispronunciations continuously without realizing where their mistakes lie. This condition, according to Derwing, Munro, and Wiebe, potentially triggers the fossilization of errors, where incorrect pronunciation patterns become permanently ingrained and increasingly difficult to rectify in the future (Derwing et al., 1998).

Given the complexity of these challenges, the need for effective and interactive learning media has become increasingly urgent to accommodate (Mayarisa, 2023). Such media are expected to provide a platform for students to engage in independent repetitive practice while receiving an instant correction mechanism. The presence of such learning tools is believed to assist students in improving their pronunciation accuracy while building better phonological awareness of the errors they commit (Awadh et al., 2024).

The rapid evolution of digital technology has fundamentally reshaped language instruction through the adoption of Mobile Assisted Language Learning (MALL). This methodology leverages mobile devices as versatile and reachable instruments, empowering students to interact with language materials regardless of their physical location or the time of day (Alisoy & Sadiqzade, 2024).

Current academic discussions suggest that MALL has transitioned into a highly interactive, student-focused model that encourages learner autonomy and provides immediate evaluative feedback (Metruk, 2024). Furthermore, the use of mobile-based educational applications has been linked to higher levels of student motivation and significant improvements in productive competencies, particularly regarding oral communication and phonetic accuracy (N. Nasim et al., 2022). These technological breakthroughs offer novel avenues for pronunciation pedagogy, enabling users to engage with authentic phonetic models and utilize automated speech recognition for real time assessment. Consequently, MALL serves as a vital tool in addressing the shortcomings of conventional teaching methods,

A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ulum especially within EFL environments where access to natural, native like input is often restricted.

One digital platform that stands out in facilitating pronunciation reinforcement is the CAKE application. This application presents authentic audio input and engaging video-based learning materials, equipped with speech recognition technology to evaluate the user's articulatory quality. Through consistent linguistic exposure and structured feedback systems, students are expected to gradually enhance their speaking proficiency and build confidence in communication.

A number of previous literatures have provided empirical evidence that the integration of digital media and mobile based learning can positively impact students' pronunciation abilities and motivation (Metruk, 2024). These findings confirm that technology creates a more interactive and student centered learning environment compared to traditional pedagogical methods, which tend to be one-dimensional. Nevertheless, the majority of existing studies currently focus on general EFL contexts or merely explore students' perceptions of technology use at a macro level (Rezai et al., 2024). There remains a significant research gap regarding how a specific application like CAKE is implemented practically to mitigate pronunciation difficulties within educational institutions that possess specific socio cultural characteristics.

This specific challenge is highly relevant to the context of MA Mirqotul Ulum, a boarding-based Islamic senior high school that implements an integrative education system between academics and religion. Students in this environment often face limitations in obtaining authentic English exposure and opportunities for personalized articulatory practice. Based on these conditions, this study aims to investigate the implementation patterns of the CAKE application in supporting pronunciation instruction and its effectiveness in addressing students' articulatory problems at MA Mirqotul Ulum.

METHOD

This study employed a qualitative research approach using a descriptive design. Qualitative research is intended to understand and interpret phenomena in natural settings by focusing on meaning, process, and participants' experiences. This approach was considered appropriate because the study aimed to explore how the CAKE application was implemented in teaching pronunciation and how it addressed students' pronunciation difficulties rather than measuring the outcomes statistically (Creswell, 2014).

The research was conducted at MA Mirqotul Ulum, a boarding-based Islamic senior high school. The participants of this study consisted of students and an English teacher involved in pronunciation learning using the CAKE application. The students had different levels of English proficiency, particularly in pronunciation ability. Most of them still experienced difficulties in producing accurate English sounds, including vowels, consonants, stress, and intonation. The researcher acted as the primary instrument who directly collected and analyzed the data during the research process (Siyoto & Sodik, 2015).

Data Sources

The data of this study were obtained from both primary and secondary sources (Ajayi, 2023). Primary data were collected directly from students and the English teacher at MA Mirqotul Ulum. These data included students' pronunciation performance, their responses during the learning process, and their experiences in using the CAKE application. The teacher also provided information regarding students' pronunciation difficulties and the instructional strategies applied in the classroom.

Secondary data were obtained from supporting documents such as lesson plans, photos, videos, and students' learning records during the implementation of the CAKE application (Shek et al., 2023). These data were used to support and strengthen the findings obtained from primary data sources.

Data Collection Techniques

To obtain comprehensive data, several data collection techniques were employed, namely observation, interviews, pronunciation tests, and documentation. Observation was

A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ullum conducted to examine how the CAKE application was implemented in the classroom. The researcher focused on students' participation, engagement, and pronunciation practice activities during the learning process. Observation is considered effective in capturing real classroom situations and students' behavior (Ary et al., 1972).

The researcher utilized semi-structured interviews to investigate the participants' perspectives and personal experiences with the CAKE application. While a predefined guideline featuring open-ended questions about pronunciation challenges and the app's efficacy was employed, the researcher maintained the flexibility to pose spontaneous follow-up inquiries. This approach was designed to elicit more profound insights and comprehensive details from the respondents (Creswell, 2014).

These sessions, which involved the English instructor and specific students, took place directly in the classroom. To maintain high data integrity, each 10–15 minute session was audio-recorded with prior permission from the participants. Furthermore, the researcher supplemented these recordings with field notes to document significant observations and key responses in real-time.

Pronunciation tests were administered to identify students' pronunciation difficulties before and after using the CAKE application. The tests focused on both segmental features (vowels and consonants) and suprasegmental features (stress and intonation). These tests were used to analyze students' pronunciation performance and to identify improvements after the implementation of the application (Fulcher & Davidson, 2012).

Documentation was also used as a supporting technique. It included photos, videos, and students' work during the learning process. Documentation was used to provide additional evidence and to support the data obtained from observation and interviews.

Data Analysis

The data were analyzed using the interactive model proposed by (Miles et al., 2014), which consists of three steps: data reduction, data display, and conclusion drawing. In the data reduction stage, the researcher selected and simplified the data obtained from observation, interviews, tests, and documentation by focusing only on relevant information related to the research objectives.

In the data display stage, the data were organized and presented in descriptive form to make them easier to understand. The pronunciation test results were categorized based on pronunciation aspects, such as vowels, consonants, stress, and intonation, to identify patterns of students' difficulties.

In the conclusion drawing stage, the researcher interpreted the data and drew conclusions regarding the implementation of the CAKE application and its role in addressing students' pronunciation difficulties. The conclusions were continuously verified to ensure their validity.

Research Validity and Trustworthiness

To ensure the trustworthiness of the data, this study applied triangulation techniques. Triangulation was conducted by comparing data obtained from different sources and methods, including observation, interviews, pronunciation tests, and documentation. This process helped ensure the consistency and credibility of the findings. In addition, the researcher carefully reviewed and rechecked the data during the analysis process to maintain the accuracy and reliability of the results. The use of multiple data sources and techniques strengthened the validity of the study and ensured that the findings accurately reflected the real conditions in the classroom (Schlunegger et al., 2024).

FINDINGS AND DISCUSSION

The Use of CAKE Application in Teaching English Pronunciation

The findings of this study revealed that the implementation of the CAKE application in teaching pronunciation created a more interactive and engaging learning environment. Based on classroom observation and interviews, the application was used as a medium for listening, repeating, and practicing pronunciation through its features such as Listen and

A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ullum
Repeat and speech recognition. These features allowed students to actively participate in pronunciation practice rather than passively receiving explanations from the teacher.

The English teacher's interview pointed out that students initially felt a sense of reluctance when they began utilizing the CAKE platform. the teacher stated:

"The students are actually used to using cell phones, but they were surprised that an app could evaluate their voices. At first, they were shy about speaking because they were afraid of getting a bad score or a 'C,' but eventually they started competing with their classmates".

These insights suggest that the application's speech recognition capabilities played a significant role in shaping the students' affective domain. Although there was an initial presence of learner anxiety, it eventually shifted toward self-confidence and healthy competition, which fostered a more active involvement in their pronunciation drills.

Furthermore, classroom observation showed that students were actively involved in pronunciation practice. They repeatedly listened to the audio model and imitated it until they achieved better scores. The availability of instant feedback encouraged students to correct their own pronunciation independently. This supports the idea that immediate feedback is essential in pronunciation learning because it helps learners recognize and repair their errors directly (Derwing et al., 1998).

In addition, the learning process became more student centered. Students were not only dependent on the teacher but also actively explored the application features. The teacher's role shifted into a facilitator who guided students during the learning process. This finding is consistent with the principles of Mobile-Assisted Language Learning (MALL), which promote autonomous and flexible learning (Kukulska-Hulme & Shield, 2008).

Students' Pronunciation Difficulties Before and After Using CAKE Application

The findings also showed that students experienced several pronunciation difficulties before using the CAKE application, particularly in segmental and suprasegmental aspects. Based on observation, students often mispronounced certain consonants and vowels that do not exist in their first language, especially influenced by their local dialect (Madurese). In addition, students had difficulties in applying correct stress and intonation, which affected the clarity of their speech.

During the observation, the researcher noted that students struggled with unfamiliar sounds and often pronounced words exactly as they were written. This confirms that differences between the first language and English sound systems are a major source of pronunciation difficulties (Gilakjani, 2017).

Data from the interview with the English instructor indicated that learners initially experienced a degree of apprehension when they first engaged with the CAKE app. As noted by the teacher:

"The students are actually used to using cell phones, but they were surprised that an app could evaluate their voices. At first, they were shy about speaking because they were afraid of getting a bad score or a 'C,' but eventually they started competing with their classmates".

This observation implies that the integration of speech recognition technology significantly impacted the students' psychological responses. While the learners initially struggled with performance anxiety, this feeling gradually evolved into a sense of assurance and competitive drive, which ultimately led to more persistent and enthusiastic pronunciation practice (Celce-Murcia et al., 2010).

In addition, students became more aware of their pronunciation errors through the scoring system provided by the application. This awareness is important because it helps learners focus on specific aspects that need improvement. As a result, students showed better pronunciation performance after continuous practice using the application.

Students' Experiences in Using CAKE Application

A participant shared their personal experience using the platform as follows:

"During the dialogue practice section, Miss, sometimes I feel like I'm saying it right, but the app still says 'Wrong.' That makes me curious, so I listen to

A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ullum
the original audio again slowly, then I repeat it until I get an 'A' (Student 1 Interview).

This statement illustrates that the feedback mechanism within the app stimulated iterative practice and piqued students' curiosity instead of causing frustration. Furthermore, it demonstrates how real-time evaluation fosters learner autonomy and persistence, driving students to continuously refine their pronunciation.

Another student shared their favorable view of the platform, noting:

"This is better, Miss. Since there's a video, you can see the expressions. Plus, with a book, you don't know if your pronunciation is correct or not, but with CAKE, you can immediately see which word you got wrong" (Student 2 Interview).

This perspective indicates that the integration of audiovisual materials alongside real-time correction significantly improves students' phonetic awareness. Furthermore, it highlights that multimedia elements are essential in creating a more purposeful and engaging learning environment compared to traditional printed resource (Joseph & Uther, 2009).

In addition, classroom observation revealed interesting social interactions among students. Students frequently asked their peers about their scores and shared strategies to achieve better results, such as asking, "How do I pronounce this word to get an "Excellent"?" This indicates that the use of the CAKE application also fostered collaborative learning among students.

However, the implementation of the application also faced some challenges. One of the main issues observed was classroom noise, which sometimes interfered with the accuracy of the speech recognition feature. Students needed to bring their phones closer to their mouths to ensure better detection. Despite this limitation, students continued to engage actively in the learning process. Overall, the findings indicate that the CAKE application not only helped students improve their pronunciation but also increased their motivation, engagement, and learning autonomy. The integration of technology in pronunciation learning provides a more effective and interactive alternative compared to traditional methods (S. M. Nasim et al., 2022).

CONCLUSIONS

This study investigated the use of the CAKE application in teaching English pronunciation at MA Mirqotul Ullum and its role in reducing students' phonetic difficulties. The findings indicate that features such as "Listen and Repeat" and automated speech recognition create an interactive learning environment that supports autonomous practice and provides immediate feedback. As a result, students showed improvement in both segmental aspects, including vowels and consonants, and suprasegmental features such as stress and intonation. The application also increased students' awareness of pronunciation errors, confidence, motivation, and participation in learning activities. This study suggests that Mobile-Assisted Language Learning (MALL) applications like CAKE can serve as effective tools for pronunciation instruction, particularly in EFL contexts with limited exposure to authentic English. Theoretically, this research contributes to MALL literature by demonstrating the potential of mobile applications to improve pronunciation skills and learning engagement. Future research should examine long-term impacts using broader research designs and participants.

REFERENCES

- Ajayi, V. O. (2023). A review on primary sources of data and secondary sources of data. Available at SSRN 5378785.
- Akhter, S. (2021). Exploring the significance of speaking skill for EFL learners. *Sjesr*, 4(3), 1-9.
- Alisoy, H., & Sadiqzade, Z. (2024). Mobile-assisted language learning (MALL): Revolutionizing language education. *Luminis Applied Science and Engineering*, 1(1), 60-72.

- A Descriptive Qualitative Study on the Use of Cake Application to Address Students' Pronunciation Difficulties at MA Mirqotul Ulum
Ary, D., Jacobs, L. C., & Razavieh, A. (1972). *Introduction to research in education*.
- Awadh, F. M. A. M., Putro, N. H. P. S., Pohan, A. E., & Alsamiri, Y. A. (2024). Improving English pronunciation through phonetics instruction in Yemeni EFL classrooms. *JOLTT Journal of Languages and Language Teaching*, 12(2), 930–940.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide*. Cambridge University Press.
- Creswell, J. W. (2014). *Qualitative, quantitative and mixed methods approaches*. Sage.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393–410.
- Fulcher, G., & Davidson, F. (2012). *The Routledge handbook of language testing*. Routledge New York, NY.
- Gilakjani, A. P. (2017). English pronunciation instruction: Views and recommendations. *Journal of Language Teaching and Research*, 8(6), 1249–1255.
- Jenkins, J. (2015). Repositioning English and multilingualism in English as a Lingua Franca. *Englishes in Practice*, 2(3), 49–85.
- Joseph, S. R. H., & Uther, M. (2009). Mobile devices for language learning: Multimedia approaches. *Research and Practice in Technology Enhanced Learning*, 4(01), 7–32.
- Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271–289.
- Kumar, R. (2023). English language and digital literacy: Navigating the information age. *Journal of International English Research Studies (JIERS)*, ISSN: 3048-5231, 1(2), 25–31.
- Mayarisa, D. (2023). Analytical study on the urgency of media learning in Islamic religious education. *International Journal Education and Computer Studies (IJECS)*, 3(3), 76–80.
- Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, 29(13), 16255–16282.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis*. sage.
- Mutali, A. (2025). The Role And Place Of Pronunciation In Teaching Oral And Written Communication. *Yessenov Science Journal*, 2(51), 148–154.
- Nasim, N., Sandeep, I. S., & Mohanty, S. (2022). Plant-derived natural products for drug discovery: current approaches and prospects. *The Nucleus*, 65(3), 399–411.
- Nasim, S. M., AlTameemy, F., Ali, J. M. A., & Sultana, R. (2022). Effectiveness of digital technology tools in teaching pronunciation to Saudi EFL learners. *FWU Journal of Social Sciences*, 16(3), 68–82.
- Oxána, K., & Krechkovska, O. (2024). *Difficulties in language learning in the elementary classes in Transcarpathia*.
- Pouresmaeil, A., & Vali, M. (2023). The effects of incidental focus on form on learning vocabulary, grammar, and pronunciation. *Language Teaching Research*, 13621688231185420.
- Rezai, A., Soyooof, A., & Reynolds, B. L. (2024). Ecological factors affecting students' use of informal digital learning of English: EFL teachers' perceptions. *Teaching and Teacher Education*, 145, 104629.
- Schlunegger, M. C., Zumstein-Shaha, M., & Palm, R. (2024). Methodologic and data-analysis triangulation in case studies: A scoping review. *Western Journal of Nursing Research*, 46(8), 611–622.
- Shek, D. T. L., Wong, T., Li, X., & Yu, L. (2023). Use of instructional videos in leadership education in higher education under COVID-19: A qualitative study. *Plos One*, 18(9), e0291861.
- Siyoto, S., & Sodik, M. A. (2015). *Dasar metodologi penelitian*. literasi media publishing.
- Swami, N. (2025). The Role Of English In Global Communication: Challenges And Opportunities. *Journal of Advanced Interdisciplinary Studies-p-ISSN 3050-8924 En e-ISSN 3050-8932*, 1(1), 1–9.
- Tiwari, H. P. (2023). Challenges in teaching pronunciation: Secondary level English teachers' perspectives. *Journal of Linguistics and Language in Education*, 17(2), 1–29.