BARAYA
Journal of Arabic and World Languages Education
Volume : 02 Issue : 01

THE EFFECTIVENESS OF DUOLINGO AS A SUPPLEMENTARY TOOL IN ARABIC VOCABULARY ACQUISITION AMONG EFL LEARNERS IN HIGHER EDUCATION: A QUALITATIVE CASE STUDY

Muhammad Shodiq[1a] , Khoirotun Nisa[2] , Muhammad Yusfi Azzuhri[3] , Indah Permata Sari[4] , Siti Mufidatul Millah[5]
[1] Al-Falah As-Sunniyah University, Indonesia, 244486232032@uas.ac.id
[2] Al-Falah As-Sunniyah University, Indonesia, 244486232050@uas.ac.id
[3] Al-Falah As-Sunniyah University, Indonesia, 244486232034@uas.ac.id
[4] Al-Falah As-Sunniyah University, Indonesia, 244486232017@uas.ac.id
[5] Al-Falah As-Sunniyah University, Indonesia, 244486232045@uas.ac.id

ABSTRACT

The integration of mobile-assisted language learning (MALL) applications into formal education has gained significant momentum in recent years, particularly in the context of Arabic language instruction. This study investigates the effectiveness of Duolingo as a supplementary learning tool for Arabic vocabulary acquisition among undergraduate students in an Islamic higher education institution in Indonesia. Employing a qualitative case study approach, data were collected through in-depth interviews, classroom observations, and document analysis involving twelveundergraduate students enrolled in an Arabic Language Education program. The findings reveal threekey dimensions: (1) students demonstrated measurable improvement in mufradat (vocabulary) retention when Duolingo was used consistently alongside formal classroom instruction; (2) the application significantly enhanced learner motivation and autonomous learning habits; and (3) several pedagogical and technical limitations were identified, including the app's inadequate coverage of classical Arabic grammar and its disconnect from formal academic Arabic registers. This study contributes to the growing body of literature on MALL in non-Western educational contexts and offers practical recommendations for Arabic language educators seeking to integrate digital tools into their curricula. The implications extend beyond Indonesia, informing global discussions on technology-mediated Arabic language instruction.

Keywords: Duolingo, Arabic, Higher Education

[a] Corresponding author

EFEKTIVITAS DUOLINGO SEBAGAI ALAT PEMBELAJARAN TAMBAHAN DALAM PENGUASAAN KOSAKATA BAHASA ARAB PADA PELAJAR EFL DI PERGURUAN TINGGI: SEBUAH STUDI KASUS KUALITATIF

ABSTRAK

Integrasi aplikasi pembelajaran bahasa berbasis perangkat seluler (*Mobile-Assisted Language Learning*/MALL) ke dalam pendidikan formal telah mendapatkan momentum yang signifikandalam beberapa tahun terakhir, khususnya dalam konteks pengajaran bahasa Arab. Penelitian ini mengkaji efektivitas Duolingo sebagai alat pembelajaran tambahan untuk penguasaan kosakata bahasa Arab di kalangan mahasiswa sarjana di sebuah institusi pendidikan tinggi Islam diIndonesia. Dengan menggunakan pendekatan studi kasus kualitatif, data dikumpulkan melalui wawancara mendalam, observasi kelas, dan analisis dokumen yang melibatkan dua belasmahasiswa sarjana yang terdaftar dalam program Pendidikan Bahasa Arab. Temuan penelitian mengungkapkan tiga dimensi utama: (1) mahasiswa menunjukkan peningkatan yang terukur dalam retensi*mufradat*(kosakata) ketika Duolingo digunakan secara konsisten bersamaan dengan pembelajaran formal di kelas; (2) aplikasi ini secara signifikan meningkatkan motivasi belajar dan kebiasaan belajar mandiri; serta (3) beberapa keterbatasan pedagogis dan teknis berhasil diidentifikasi, termasuk cakupan tata bahasa Arab klasik yang tidak memadai dalam aplikasi tersebut dan ketidaksesuaiannya dengan ragam bahasa Arab akademik formal. Penelitian ini berkontribusi pada khasanah literatur yang terus berkembang mengenai MALL dalam konteks pendidikan non-Barat, serta menawarkan rekomendasi praktis bagi para pengajar bahasa Arab yang ingin mengintegrasikan alat digital ke dalam kurikulum mereka. Implikasi penelitian ini melampaui batas Indonesia dan memberikan kontribusi pada diskusi global mengenai pengajaran bahasa Arab yang dimediasi oleh teknologi.

Kata kunci: Duolingo, Penguasaan Kosakata Bahasa Arab, Pembelajaran Bahasa Berbasis Perangkat Seluler, Mufradat, Pendidikan Tinggi Islam, Indonesia

A. INTRODUCTION

The rapid proliferation of mobile technology has fundamentally transformed the landscape of language education across the globe. Mobile-Assisted Language Learning (MALL) has emerged as a significant pedagogical paradigm, offering learners unprecedented access to language input and practice opportunities that transcend the physical boundaries of the traditional classroom. A systematic review analyzing over 60 studies on MALL and gamification revealed that mobile learning applications enhance language acquisition, particularly in vocabulary and grammar development (Taj et al., 2021). Among the most widely adopted MALL platforms is Duolingo, a gamified language learning application that has become the most downloaded educational app globally, offering free lessons in over 40 languages to more than 500 million users (Blanco, 2021).

One of Duolingo's primary strengths lies in its gamified approach, which is carefully designed to build and sustain learner motivation. The platform builds both extrinsic motivation through features such as streaks, notifications, and leaderboards that encourage daily engagement and intrinsic motivation by inviting learners to set personal goals and interact with content that sparks genuine interest (Duolingo, 2023). This design proves psychologically significant, as research consistently demonstrates that highly motivated learners tend to be more autonomous in foreign language classrooms, exhibiting greater initiative and stronger self-regulation in their learning processes (Jin, 2024). Furthermore, a qualitative study examining language experts who used

Duolingo for third language learning found that MALL aplications such as Duolingo foster learner autonomy and have shown effectiveness in improving motivation and language skills (Puebla et al., 2021).

In the context of Arabic language learning, Duolingo presents both opportunities and distinctive challenges. Users of Duolingo have reported satisfaction in accessing the application to expand their knowledge of foreign language vocabulary, pronunciation, intonation, correct grammar, and vocabulary illustrated by inspirational images (Ritonga et al., 2022). However, research also highlights important limitations. A study investigating Indonesian university students' perceptions of Duolingo-assisted Arabic learning found that one of the major constraints reported by students was the limited availability of vocabulary relevant to academic needs, particularly for writing-related tasks, and that Arabic is rarely used outside academic contexts, which prevents vocabulary learned through Duolingo from being fully internalized (Al-Bariq, 2024).

The challenges facing Arabic language learners are further compounded by the phenomenon of diglossia the coexistence of Modern Standard Arabic (MSA) and numerous colloquial dialects. Research confirms that the Arabic diglossic situation, particularly the functional difference between MSA and colloquial Arabic, creates a significant obstacle for second language learners in terms of social communication in real contexts as well as in the language learning process itself (Alotaibi, 2022). Most university programs teaching Arabic as a foreign language focus exclusively on MSA, which creates difficulties for students who are unable to converse in everyday language with native speakers (Sneed, 2012). This diglossic reality is particularly relevant when evaluating the adequacy of tools like Duolingo, whose Arabic module operates primarily within the MSA register and may not sufficiently address the communicative and classical Arabic demands faced by learners in Indonesian Islamic higher education institutions.

In the Indonesian context, Arabic language education occupies a prominent institutional role. Research exploring Arabic language learning in pesantren across culturally diverse regions of Indonesia reveals that Arabic instruction reflects not only religious but also multicultural dimensions, with regional cultural influences significantly affecting teaching strategies and learner motivations (Haddade et al., 2025). Nevertheless, Arabic education in Indonesian madrasahs continues to face considerable hurdles, as many educators still rely on traditional, rote-based methods that do not effectively foster practical communication skills, and the curriculum often prioritizes memorization over functional language use (Rifa'i et al., 2024). A competency-based Arabic curriculum study further found that integrating learner autonomy principles significantly enhanced students' motivation, enjoyment, and comprehension of Arabic, suggesting the need for more innovative and technology-integrated approaches (Raswan & Muradi, 2025).

Within this context, gamification-based language learning systems represent a potentially transformative supplement to formal Arabic instruction. A comprehensive systematic review of Duolingo literature from 2012 to 2020 confirmed that adding gamification elements to MALL platforms increases language accuracy and learner confidence, though very few studies have adequately assessed learners' vocabulary development, listening comprehension, speaking, and pronunciation within Duolingo specifically (Shortt et al., 2021). Moreover, research on learner autonomy in Arabic language education for non-native speakers highlights that prior studies have not adequately addressed how institutional structures and pedagogical frameworks intersect to influence the development of autonomy in Arabic language instruction, and that research on autonomy has largely focused on Western language contexts with limited attention to non-Western educational settings (Hussein & Haron, 2025). This study therefore seeks to address these gaps by conducting an in-depth qualitative investigation into the experiences of undergraduate Arabic language education students in an Indonesian Islamic higher education institution who used Duolingo as a supplementary learning tool alongside their formal coursework, foregrounding student voices and institutional realities to generate contextually grounded insights for both pedagogical practice and future research.

This study is guided by three primary research objectives. First, this study aims to examine how undergraduate students in an Indonesian Islamic higher education institution perceive and experience the use of Duolingo as a supplementary tool for Arabic vocabulary acquisition. Second, this study seeks to analyze the extent to which Duolingo use influences learner motivation and autonomous learning behavior among Arabic language education students. Third, this study endeavors to identify the pedagogical opportunities and limitations associated with integrating Duolingo as a supplementary resource within the formal Arabic language curriculum at the higher education level.

This study proceeds from the fundamental argument that technology integration in language education must be understood not as a universally applicable solution, but as a contextually situated practice whose effectiveness is mediated by institutional, cultural, linguistic, and individual factors. Duolingo, as a product developed primarily for Western language learning markets, carries embedded assumptions about language learning that may not straightforwardly transfer to the Arabic language learning context in Indonesian Islamic higher education. By subjecting these assumptions to empirical scrutiny through qualitative inquiry, this research aims to produce knowledge that is both theoretically meaningful and practically useful.

The study is anticipated to make several contributions. At the theoretical level, it engages with and extends existing frameworks for understanding MALL effectiveness, learner autonomy, and gamification in language education. At the practical level, it offers concrete, evidence-based recommendations for Arabic language educators, curriculum designers, and educational technology developers seeking to better serve learners in non-Western, Islamic educational contexts. More broadly, this research contributes to a growing scholarly conversation about the decolonization of educational technology the need to critically evaluate digital learning tools whose design logic reflects particular cultural and linguistic assumptions, and to adapt or supplement them thoughtfully for diverse learner populations

B. METHODS

This study adopts an interpretive qualitative approach, grounded in the conviction that understanding human experience requires attention to meaning, context, and subjectivity. The interpretive paradigm holds that social reality is constructed through individuals' interactions with their environments and that the researcher's task is to understand and represent these constructions faithfully and rigorously. This approach is particularly well-suited to the research objectives of this study, which seek to understand how students experience and make meaning of their Duolingo use within the specific institutional and cultural context of Indonesian Islamic higher education. The interpretive approach is complemented by a sociocultural theoretical orientation, drawing on Vygotskian principles of mediated learning and the Zone of Proximal Development. This framework positions technological tools including language learning applications as cultural artifacts that mediate learners' cognitive development and language acquisition processes. Within this orientation, Duolingo is understood not simply as a delivery mechanism for vocabulary input, but as a semiotic tool that shapes and is shaped by learners' existing knowledge, motivations, social practices, and institutional contexts (Cresswell, 2014).

The research design employed in this study is a qualitative case study. Case study methodology is appropriate when the researcher seeks to develop an in-depth understanding of a bounded phenomenon within its real-world context, and when the boundaries between phenomenon and context are not clearly evident. This study treats the use of Duolingo as a supplementary Arabic learning tool within a single Islamic higher education institution as the bounded case, allowing for rich, contextually nuanced data collection and analysis. The case study design enables the triangulation of multiple data sources interviews, observations, and documents thereby enhancing the credibility and trustworthiness of the findings. It also permits the researcher to capture the complexity and particularity of the phenomenon under investigation, moving

beyond generalized claims to produce thick description that illuminates the case in its full contextual richness (Benson, 2011).

The research was conducted at an Indonesian Islamic higher education institution located in a mid-sized city on the island of Java. The institution, referred to throughout this paper as "University Alpha" to protect its identity, offers a Bachelor's degree program in Arabic Language Education (Pendidikan Bahasa Arab) and has a student body of approximately 800 undergraduate students across all programs. The Arabic Language Education program enrolls approximately 120 students per academic cohort. Participants were recruited from the second-semester cohort of the Arabic Language Education program. Purposive sampling was employed to select participants who met the following criteria: (1) currently enrolled in the Arabic Language Education program at University Alpha; (2) had used the Duolingo Arabic course for a minimum of four weeks during the current academic semester; and (3) were willing to participate in interviews and to have their learning activities observed. Twelve students were ultimately selected, comprising eight female and four male participants, consistent with the gender demographics of the program. Their ages ranged from 18 to 22 years. In accordance with ethical commitments to anonymity, all participants are referred to using pseudonyms throughout this paper: Participant A through Participant L. No real names, institutional identifiers, or information that could enable identification of individual participants are disclosed. Two faculty members from the Arabic Language Education program also participated in brief informational interviews to provide contextual background on the curriculum and their perspectives on students' technology use. These participants are referred to as Lecturer 1 and Lecturer 2 (Deterding, 2011).

Data were collected through three primary techniques over a period of eight weeks during the second semester of the academic year. In-depth Semi-structured Interviews: Individual semi-structured interviews were conducted with each of the twelve student participants and both faculty members. Student interviews lasted between 45 and 75 minutes and were conducted in a combination of Indonesian and English, according to participants' preference, with all data subsequently translated into English for analysis. Interview questions addressed participants' patterns of Duolingo use, their perceptions of its impact on vocabulary acquisition, its effects on their motivation, and their views on its alignment with formal coursework. Faculty interviews focused on instructional approaches, curriculum design, and attitudes toward technology integration. Classroom Observation: Six classroom observation sessions were conducted across three different Arabic language courses within the program, with two observations per course. Observations focused on how vocabulary instruction was delivered in formal settings, how students engaged with course materials, and whether any reference was made to digital learning tools including Duolingo. Field notes were recorded during each session and expanded immediately following the observation. Document Analysis: Relevant documents were collected and analyzed, including course syllabi, textbooks used in the Arabic language courses, and, with participants' permission, screenshots of their Duolingo progress dashboards and streak records. These documents provided important contextual data about the formal curriculum and students' actual patterns of Duolingo engagement (Krashen, 1982).

Data analysis followed the principles of thematic analysis as articulated in the qualitative research literature. This involved a recursive, iterative process of reading and re-reading transcripts and field notes, generating initial codes, clustering codes into themes, reviewing and refining themes, and producing coherent thematic descriptions that addressed the research objectives. Analysis proceeded through the following stages. First, all interview transcripts were transcribed verbatim and read multiple times to develop overall familiarity with the data. Second, open coding was conducted, with labels assigned to discrete units of meaning. Third, related codes were grouped into broader categories through axial coding. Fourth, overarching themes were identified that cut across categories and addressed the research objectives. Finally, themes were reviewed against the original data to ensure their validity, and representative quotations were selected to illustrate each theme in the findings section. To enhance rigor and trustworthiness,

member checking was conducted by sharing preliminary thematic summaries with four of the twelve participants, who confirmed that the summaries accurately represented their experiences. Peer debriefing was also employed, with a colleague who was not involved in the study reviewing the coding framework and offering critical feedback. Reflexivity was maintained through the researcher's ongoing engagement with a reflective journal throughout the data collection and analysis process.

This study was conducted in accordance with established ethical principles for qualitative research involving human subjects. Prior to data collection, informed consent was obtained from all participants through written consent forms that clearly explained the study's purpose, procedures, the voluntary nature of participation, the right to withdraw at any time without consequence, and the measures taken to protect confidentiality. Parental consent was not required as all participants were adults (18 years or older) at the time of the study. All identifying information has been anonymized in this paper. University Alpha is a pseudonym for the participating institution. All participant names have been replaced with codes (Participant A through Participant L; Lecturer 1 and Lecturer 2). Data were stored securely in password-protected digital files accessible only to the research team. Interview recordings were deleted following transcription. The study did not involve any deception, and participants were fully informed of the nature of the research prior to their involvement (Kulkuska, 2005).

C. RESULTS *(12pt, Bold) - 25% of manuscript*

1. Student Perceptions and Experiences of Duolingo for Arabic Vocabulary Acquisition

The majority of participants described their initial engagement with Duolingo as driven by curiosity and external recommendation rather than formal institutional guidance. Eight of the twelve participants reported first learning about the application from peers or social media, while the remaining four cited a brief mention by a faculty member during an orientation session. Participant C described her motivation for downloading the application:

"I saw my friend's phone. She had a 100-day streak on Duolingo for Arabic. I thought, if she can do it, I can try. The vocabulary exercises looked different from our textbook more like a game."

Participant G elaborated on how he used the application:

"I usually open Duolingo before sleeping or when I am commuting. I do maybe 15 to 20 minutes. The lessons are short, which I like. In class, we have long sessions with a lot of grammar explanation. Duolingo is different it gives you a word, you see the picture, you hear the pronunciation, and you choose. It is more direct."

When asked about vocabulary retention, nearly all participants expressed the view that Duolingo had helped them remember specific vocabulary items, particularly concrete nouns and everyday phrases. Participant A noted:

"Before I used Duolingo, I kept forgetting words after the exam. But with Duolingo, the same word appears again and again not in a boring way, but in different exercises. I think I remember maybe 200 new words that I can actually use."

However, participants also consistently noted important limitations in the vocabulary coverage provided by the application. Participant D, who was preparing for a course in classical Arabic texts, expressed frustration:

"Duolingo teaches you words like 'apple', 'water', 'house'. These are good for basic Arabic. But in my Qira'ah class, I need words from classical texts, Islamic fiqh, tafsir. Duolingo does not have these words at all."

Classroom observations corroborated this finding. During two observed sessions of a course on classical Arabic reading, the instructor introduced 35 domain-specific vocabulary items related to Islamic jurisprudence. None of these items appeared in participants' Duolingo progress data, and several participants were observed struggling with the terms despite their reported Duolingo use. Participant F offered a nuanced perspective:

"I think Duolingo is like a foundation. It gives you confidence. When I started university, I was embarrassed to speak any Arabic. Now I know many basic words and I am less afraid. But it is not enough for academic Arabic. You still need the textbook and the teacher."

Document analysis of participants' Duolingo dashboards revealed significant variation in usage frequency. Five participants had maintained streaks of 30 days or longer, while three had streaks of fewer than 10 days, suggesting inconsistent engagement. Word review frequency a measure of how often the application prompted users to revisit previously learned vocabulary was consistently high among those with longer streaks, a pattern participants associated with improved retention.

2. Duolingo's Influence on Learner Motivation and Autonomous Learning Behavior

A prominent theme across all twelve interviews was the role of Duolingo's gamified design in sustaining learner motivation. Participants consistently referenced specific design features particularly streaks, experience points (XP), leagues, and immediate feedback as motivationally significant. Participant B described the streak mechanism in detail:

"My streak is now 67 days. Every morning I feel like I have to open Duolingo or the streak will break. I know it is just a game feature, but it makes me study Arabic every day, even on weekends. Before, I only studied before the exam."

Participant J, one of the four participants with shorter streaks, offered a contrasting perspective:

"I lost my streak once because my phone ran out of battery. After that, I felt demotivated for a week. I thought, what is the point? So the streak feature can also be negative if you lose it."

The league and leaderboard features generated mixed responses. Participant H reported that competing with peers was highly motivating:

"I am in the Gold league. I check every day to see my ranking. When I see someone above me, I do more lessons. It is competitive, but in a fun way."

However, Participant K expressed concern that the competitive dimension had led to superficial engagement:

"Sometimes I just rush through exercises to get XP points. I am not really thinking about the meaning. I just want to stay in the league."

Observations of participants' study habits triangulated through interview data and document review indicated that nine of twelve participants had established a daily or near-daily Duolingo practice that was independent of formal coursework requirements or instructor prompting. This represented a notable shift toward autonomous learning behavior for students who, by their own account, had previously studied Arabic only when required by coursework deadlines. Participant E reflected on this shift:

"My Arabic teacher in high school told us what to study and when. Here in university, we have more freedom. Duolingo helps me structure my independent study. I know I need to do at least one lesson every day. It is my own schedule, not the professor's."

Lecturer 1 observed changes in student behavior consistent with participants' self-reports:

"Since students started using these applications, I notice some of them come to class with questions about specific words. They encountered these words in their self-study and want to know more. This is good it shows they are thinking beyond the classroom."

However, Lecturer 2 raised a concern about the quality of autonomous engagement:

"The problem is that students think finishing Duolingo lessons means they have 'done their Arabic study' for the day. But these lessons are very surface-level. They do not develop reading comprehension or writing skills, which are what we assess in formal examinations."

3. Pedagogical Opportunities and Limitations of Duolingo in the Formal Curriculum

The third set of findings addresses the alignment and misalignment between Duolingo's content and pedagogical approach and the formal Arabic language curriculum at University Alpha. Observations and document analysis revealed that the formal curriculum places substantial emphasis on Nahwu (Arabic grammar), Sarf (Arabic morphology), Qira'ah (reading

comprehension of classical texts), and Kitabah (academic writing). Duolingo's Arabic module, by contrast, is structured around conversational vocabulary, listening comprehension, and basic phrase construction. Participants identified several areas where Duolingo complemented formal instruction effectively. Vocabulary introduced in formal classes was sometimes reinforced through Duolingo exercises, even if not systematically. Participant I noted:

"Last week in class we learned words about family relationships father, mother, brother, sister. Duolingo happened to have a lesson on the same topic. Seeing the same words in two different contexts helped me remember them much better."

The application's audio component was also identified as a significant asset by participants, as formal classroom instruction often had limited opportunity for extended listening practice. Participant L stated:

"In class, we hear Arabic from the teacher and from the CD recordings in the textbook. Duolingo gives me more listening practice. The voice is clear and the pronunciation is Modern Standard Arabic, which is what we are learning."

However, significant pedagogical limitations were identified with considerable consistency across all data sources. The most frequently cited limitation was Duolingo's failure to address Arabic grammar in a systematic, academically rigorous manner. Nahwu the study of Arabic grammatical rules is foundational to the University Alpha curriculum and essential for reading classical texts. Participants consistently noted that Duolingo provided insufficient and oversimplified grammatical instruction. Participant A stated plainly:

"Duolingo does not explain Nahwu. It just gives you sentences and asks you to translate. But if you do not understand why a word has Kasrah or Dammah at the end, you cannot read new sentences by yourself. You are just memorizing patterns."

During classroom observations, the researcher noted that instructors regularly engaged students in detailed parsing exercises identifying the grammatical case and function of individual words within Arabic sentences a practice that requires formal grammatical knowledge that Duolingo does not develop. Students with strong Duolingo engagement were not observably more proficient than their peers in these parsing tasks.

A further limitation identified through document analysis was the script and register of Arabic used in Duolingo. The application uses harakat (short vowel diacritics) throughout its Arabic course, which assists beginner learners in pronunciation. However, classical Arabic texts at University Alpha including Quranic texts and Islamic jurisprudence literature are typically presented without harakat, requiring learners to infer vowelization from grammatical and contextual knowledge. Several participants noted that their Duolingo practice had not prepared them for reading unvowelized Arabic text. Participant D summarized the issue:

"Duolingo always shows the harakat. So when I read our textbook without harakat, I feel lost. The skills are different. Duolingo trains you to read with the marks, but real Arabic texts in university do not have these marks."

D. DISCUSSION

1. Duolingo as a Vocabulary Scaffold, Not a Comprehensive Learning System

The first finding that students experienced genuine, though limited, vocabulary acquisition benefits from Duolingo use invites interpretation through the lens of Stephen Krashen's Input Hypothesis, which proposes that language acquisition occurs most effectively when learners are exposed to comprehensible input at a level slightly beyond their current competence, famously formulated as the i+1 principle (Krashen, 1982). Duolingo's spaced repetition algorithm and graduated difficulty structure are broadly consistent with this principle, providing vocabulary input calibrated to the individual learner's demonstrated proficiency. The consistent reports of improved vocabulary retention among participants with regular Duolingo use suggest that the application functions effectively as a scaffolding mechanism for basic lexical acquisition, delivering comprehensible and appropriately leveled input that aligns with the conditions Krashen

identified as conducive to subconscious language acquisition. This interpretation aligns with research conducted in other language learning contexts, where Duolingo has been shown to support vocabulary development among beginner and low-intermediate learners more effectively than among advanced learners, for whom the application's input may fall below the productive challenge threshold. The present findings extend this pattern to the Arabic language learning context, suggesting that Duolingo's vocabulary benefits are most meaningful for students at the early stages of Arabic acquisition.

However, the findings also reveal a significant limitation that existing research has not always adequately foregrounded: the domain specificity of vocabulary needs in specialized educational contexts. The vocabulary prioritized by Duolingo's Arabic module reflecting everyday, conversational communication scenarios diverges considerably from the academic and religious vocabulary that constitutes the core lexical domain of Arabic language education in Indonesian Islamic higher education. This finding resonates with Nation's (2001) model of vocabulary knowledge, which distinguishes between high-frequency general vocabulary, academic vocabulary, technical vocabulary, and low-frequency vocabulary. Duolingo's Arabic course effectively addresses high-frequency general vocabulary a foundational layer but provides minimal support for the academic and technical vocabulary that constitutes the core of University Alpha's curriculum. A learner approaching advanced academic study in Arabic needs to navigate not only the 2,000 most frequent general words but also specialized academic and religious vocabulary that Duolingo's current Arabic module does not systematically address (Nation, 2001). The implication is not that Duolingo is without value, but that its value is bounded and context-specific: it functions effectively as a scaffold for basic lexical knowledge but cannot serve as a substitute for disciplinary vocabulary instruction (Hasan, 2025).

2. Gamification as a Double-Edged Motivational Tool

The second finding that Duolingo's gamified design significantly influenced learner motivation, with both positive and negative dimensions merits interpretation through Self-Determination Theory (SDT), as developed by Ryan and Deci (2020). SDT distinguishes between intrinsic motivation, which is engagement driven by inherent interest and enjoyment, and extrinsic motivation, which is engagement driven by external rewards or pressures. A central premise of SDT is that while extrinsic motivators can initiate and sustain behavior in the short term, they may undermine intrinsic motivation if they become the primary driver of engagement a phenomenon described as the crowding-out effect, whereby external pressure displaces the internal sense of autonomy and competence that sustains deeper learning (Ryan & Deci, 2020). SDT further specifies that for healthy motivational development to unfold, individuals require support for three basic psychological needs: autonomy, competence, and relatedness (Ryan & Deci, 2020). The data reveal a complex interplay of motivational dynamics among participants. The streak mechanism, XP system, and leaderboard features clearly functioned as powerful extrinsic motivators, driving daily engagement among the majority of participants. This aligns with a substantial body of gamification research demonstrating that well-designed reward structures can significantly increase the frequency and consistency of engagement with educational activities (Landers, 2014). The development of daily Arabic study habits among nine of twelve participants a population that had previously studied only under formal institutional requirements represents a meaningful behavioral outcome consistent with what Landers' (2014) theory of gamified learning identifies as a successful mediating process, whereby game elements directly affect learning-relevant behaviors and attitudes (Nugraheni, 2026).

However, the data also reveal concerning signs of motivation quality deterioration in some participants. Participant K's account of rushing through exercises to accumulate XP points without genuine cognitive engagement is a textbook illustration of the crowding-out effect the gamification mechanism has become detached from the underlying learning goal and now serves only the pursuit of game rewards. This is consistent with Landers' (2014) argument that gamification is most educationally effective when reward structures are tightly coupled to

meaningful learning behaviors, and least effective when they can be pursued independently of actual learning outcomes. The finding regarding streak loss and subsequent demotivation, illustrated by Participant J's experience, also warrants attention. Duolingo's streak feature creates a high-investment metric, making learners vulnerable to significant motivational setback when the streak is interrupted a dynamic that suggests a potential design limitation with implications for learner well-being and sustained engagement over longer academic timescales. These findings are further illuminated by research on gamification in online language learning contexts, which confirms that the relationship between gamification and motivation is mediated by learners' experiences of autonomy, competence, and relatedness the three basic needs identified by SDT and that gamification produces fragile motivational benefits when it primarily drives competition and external reward-seeking rather than genuine need satisfaction (Deci & Ryan, 2000). The implication for educational technology designers is that gamified features must be evaluated not merely for their capacity to drive engagement metrics, but for the quality of psychological experience they produce in learners.

3. The Structural Misalignment Between Duolingo and Formal Arabic Curriculum

The third finding that significant structural misalignments exist between Duolingo's pedagogical design and the formal Arabic curriculum at University Alpha invites interpretation through Chapelle's (2001) framework for evaluating CALL tasks, which assesses language learning software along six dimensions: language learning potential, learner fit, meaning focus, authenticity, positive impact, and practicality. Applying this framework to the present data reveals a mixed picture. In terms of language learning potential, Duolingo demonstrates moderate strength for basic vocabulary and listening, but significant weakness in grammar, writing, and classical Arabic competencies. In terms of learner fit, the application's calibration to general conversational Arabic represents a meaningful mismatch with the specialized academic needs of University Alpha students. In terms of authenticity, the conversational scenarios presented by Duolingo do not reflect the authentic language use contexts that University Alpha students will encounter as future Arabic language educators producing and analyzing classical texts, teaching Arabic grammar, and engaging in academic discourse about Islamic sciences. A prior study directly applying Chapelle's criteria to Duolingo's Dutch course similarly found that the platform's tasks questionably comply with the authenticity criterion, suggesting that this limitation is not exclusive to the Arabic module but reflects a broader structural characteristic of the application's design philosophy (Erudio, 2016).

This structural misalignment is not merely a technical limitation but reflects a deeper epistemological tension. Duolingo's Arabic course is designed within a communicative language teaching paradigm, prioritizing functional communication in Modern Standard Arabic. The Arabic language curriculum at Indonesian Islamic higher education institutions, however, is grounded in a classical-textual tradition that foregrounds grammatical analysis, morphological precision, and engagement with pre-modern Arabic literary and religious genres. These represent fundamentally different orientations toward what Arabic proficiency means and what Arabic language education is for a tension that no amount of incremental improvement to Duolingo's Arabic module can fully resolve without a more fundamental reconceptualization of the platform's purpose in this context (Rohmatuszahroh, 2025).

4. Research Significance and Limitation

This study makes several contributions to the intersecting fields of MALL research, Arabic language education, and educational technology in non-Western contexts. Empirically, it provides rich qualitative data on the experiences of Indonesian Arabic language education students using Duolingo a population largely absent from the existing research literature. Theoretically, it demonstrates the value of integrating SDT (Ryan & Deci, 2020), Krashen's (1982) Input Hypothesis, and Chapelle's (2001) CALL evaluation framework in analyzing the multidimensional impact of gamified language learning applications. By foregrounding student voices and classroom realities through qualitative inquiry, this study resists the temptation to

evaluate Duolingo purely on the basis of measurable learning outcomes, and instead attends to the nuanced, lived experiences of learners navigating the intersection of formal institutional instruction and informal mobile-assisted practice. Several limitations of this study must be acknowledged. First, the single-institution case study design limits the transferability of findings, as University Alpha's specific institutional culture, curriculum, and student demographic may not be representative of Islamic higher education institutions across Indonesia or in other national contexts. Second, the study relies primarily on self-reported data regarding Duolingo use and vocabulary acquisition experiences; the absence of independent vocabulary assessment means that claims about learning outcomes are based primarily on participants' perceptions rather than objective measures. Third, the eight-week data collection period may be insufficient to capture longer-term patterns of Duolingo use, motivation, and vocabulary retention, given that research in the MALL field consistently identifies attrition the tendency for technology-mediated learning behavior to decline over time as a significant challenge. Fourth, the researcher's positionality as someone familiar with both Arabic language education and educational technology may have introduced unintended biases into the data collection and analysis process, and while reflexive strategies were employed to mitigate these, they cannot be entirely eliminated.

5. CONCLUSIONS

Provide This study set out to examine three interconnected dimensions of Duolingo use among Arabic language education students in Indonesian Islamic higher education. The conclusions that emerge in response to each research objective are as follows. In response to the first objective to examine student perceptions and experiences of Duolingo for Arabic vocabulary acquisition the study concludes that students broadly perceive Duolingo as a useful supplementary tool for building basic Arabic vocabulary, particularly for high-frequency general vocabulary. The application's spaced repetition, immediate feedback, and graduated difficulty contribute to improved word retention among consistent users. However, students clearly recognize and experience the application's limitations in addressing the specialized academic and classical vocabulary central to their university curriculum. Duolingo functions effectively as a lexical scaffold but is insufficient as a comprehensive vocabulary learning system for the disciplinary demands of Arabic language education at the higher education level.

In response to the second objective to analyze the influence of Duolingo on learner motivation and autonomous learning behavior the study concludes that Duolingo's gamification features produce meaningful, though complex, motivational effects. The majority of participants developed more consistent, self-directed Arabic study habits as a result of engaging with the application, representing a genuine enhancement of learner autonomy. However, the motivational dynamics generated by game elements particularly streaks and leaderboards are fragile and prone to disruption, and in some cases may promote superficial engagement with vocabulary exercises rather than deep learning. The quality of motivation, not merely its quantity, must be a central consideration in evaluating Duolingo's educational value.

In response to the third objective to identify pedagogical opportunities and limitations of Duolingo integration within the formal Arabic curriculum the study concludes that significant structural misalignments exist between Duolingo's CLT-oriented design and the classical-textual pedagogical tradition of Indonesian Islamic Arabic language education. While opportunities exist for complementary vocabulary reinforcement and listening practice, Duolingo's inadequate coverage of Arabic grammar, its reliance on vocalized text, and its absence of classical and religious register vocabulary mean that it cannot serve as a meaningful complement to let alone replacement for formal Arabic instruction in this context. Duolingo is best understood as a motivational entry point and basic vocabulary supplement rather than a curricular resource.

6. FUNDING

This research received no external funding.

7. CONFLICTS OF INTEREST

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. None of the funding sources had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

8. ACKNOWLEDGMENTS

The authors express their sincere gratitude to All

9. REFERENCES

Al-Bariq. (2024). Perceptions of Duolingo-assisted learning in enhancing Arabic vocabulary among university students. *Al-Bariq Journal of Arabic Education*. http://albariq.org/index.php/albariq/article/download/181/71

Alotaibi, A. (2022). Arabic diglossia and its impact on the social communication and learning process of non-native Arabic learners: Students' perspective. *ResearchGate*. https://www.researchgate.net/publication/358172632

Benson, P. (2011). *Teaching and researching autonomy in language learning* (2nd ed.). Pearson Education.

Blanco, M. (2021). Uncurtaining windows of motivation, enjoyment, critical thinking, and autonomy in AI-integrated education: Duolingo vs. ChatGPT. *System*. https://www.sciencedirect.com/science/article/abs/pii/S0023969025000074

Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge University Press.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). *From game design elements to gamefulness: Defining gamification*. ACM Press.

Duolingo. (2023). *The Duolingo method for app-based teaching and learning* [White paper]. https://duolingo-papers.s3.amazonaws.com/reports/Duolingo_whitepaper_duolingo_method_2023.pdf

Erudio. (2016). Assessing Livemocha and Duolingo evaluation tasks: Adapting Chapelle CALL assessment criteria in autonomous learning experience. *Erudio: Journal of Educational Innovation, 5*(1). https://www.researchgate.net/publication/326474777

Haddade, H., et al. (2025). Arabic language learning in a multicultural context at pesantren. *Jurnal Pendidikan Islam*. https://doi.org/10.15575/jpi.v11i1.44104

Hasan, M. F., Rahmi, L., Qoyimah, D., Saputri, M., & Bella, S. (2025). Psikologi Perkembangan Anak Usia SD/MI Teori dan Praktik. Deepublish.

Hussein, F., & Haron, S. (2025). Reframing learner autonomy in Arabic language education for non-native speakers: A theoretical framework of power, control, and motivation. *Frontiers in Education*.

https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1622527/full

Jin, T. (2024). Uncurtaining windows of motivation, enjoyment, critical thinking, and autonomy in AI-integrated education. *System, 131*.

https://www.sciencedirect.com/science/article/abs/pii/S0023969025000074

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

Kukulska-Hulme, A., & Traxler, J. (Eds.). (2005). *Mobile learning: A handbook for educators and trainers*. Routledge.

Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming, 45*(6), 752–768. https://doi.org/10.1177/1046878114563660

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nugraheni, A. S., & Hasan, M. F. (2026). Indonesian language learning adaptation strategies in regional language environment: a phenomenological study in rural primary schools. *Cogent Education*, *13*(1). https://doi.org/10.1080/2331186X.2026.2635805

Puebla, C., et al. (2021). Duolingo, autonomy, and motivation: Perceptions about third language learning by language experts. *Redalyc*. https://www.redalyc.org/journal/5216/521681814051/

Raswan, & Muradi. (2025). Developing a competency-based Arabic curriculum to foster Merdeka Belajar in Indonesian madrasahs. *Jurnal Ilmiah Peuradeun*. https://journal.scadindependent.org/index.php/jipeuradeun/article/view/1930

Rifa'i, I., et al. (2024). Reforming Arabic education in Indonesian madrasahs: A comparative review of KMA Decrees No. 183/2019 and No. 347/2022. *Alsuniyat: Jurnal Penelitian Bahasa, Sastra, dan Budaya Ar ab*. https://ejournal.upi.edu/index.php/alsuniyat/article/download/82533/31541

Ritonga, M., et al. (2022). Duolingo: An Arabic speaking skills' learning platform for andragogy education. *Education Research International*. https://onlinelibrary.wiley.com/doi/10.1155/2022/7090752

Rohmatuszahroh, A. I., Hermawati, K. A., Monita, D., & Hasan, Moh. F. (2025). Inclusive Learning Strategy Based on Braille Modules: Case Study of Tahfidz Classes for Blind Students in Indonesia. *Journal of Disability & Religion*, 1–19. https://doi.org/10.1080/23312521.2025.2531811

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Shortt, M., et al. (2021). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer*

Assisted Language Learning, 36(5). https://www.tandfonline.com/doi/full/10.1080/09588221.2021.1933540

Sneed, C. (2012). Arabic diglossia and Arabic language teaching: Teaching and learning vernacular and standard varieties. *Growing Participator Approach*. https://growingparticipatorapproach.wordpress.com/teaching-msa-and-colloquial-arabic/

Taj, I. H., et al. (2021). Mobile-assisted and gamification-based language learning: A systematic literature review. *PMC/NCBI*. https://pmc.ncbi.nlm.nih.gov/articles/PMC8157183/