



Vol 5
No 15

Proceeding Of International Student Conference
on Arabic Language and Teaching 2026
Hosted By Student Association of Arabic Literature
Department Faculty of Letters Universitas Negeri
Malang

September

2026

Page: ----

Dhou: A Concept of Developing an Arabic Language Learning Application for Fusha and 'Ammiyyah in the Age of Artificial Intelligence

Dhou: تصور لتطوير تطبيق لتعلم اللغة العربية الفصحى والعامية في عصر الذكاء الاصطناعي

Fariz Mirza Abdillah^{1*}

¹Universitas Abdul Chalim, Indonesia

*farizmirza.fm@gmail.com

Abstract: This study employs a design thinking approach in the UX design process, focusing on the stages of empathize, define, ideate, prototype, and testing. The result of this study is a conceptual UI/UX design for an Arabic language learning application that provides separate yet interconnected learning pathways for *fusha* and *'ammiyyah* in Arabic. The proposed features include structured learning materials, vocabulary and everyday expressions, contextual exercises, and AI-powered conversation practice that enables users to simulate Arabic communication in various situations. The proposed design demonstrates how AI can be integrated as a supporting component within a broader digital learning environment rather than as the sole focus of the application. This concept is expected to provide a foundation for the future development and evaluation of a more flexible, interactive, and contextually relevant Arabic language learning application.

Keywords: Arabic Language Learning; *Fusha* and *'Ammiyyah*; Mobile Application; Conversational AI; Design Thinking

تستخدم هذه الدراسة منهجية التفكير التصميمي (Design Thinking) في عملية تصميم تجربة المستخدم (UX)، مع التركيز على مراحل: التعاطف (Empathize)، وتحديد المشكلة (Define)، وتوليد الأفكار (Ideate)، والنموذج الأولية (Prototype)، والاختبار (Testing). وتتمثل نتيجة هذه الدراسة في تصميم تصوري لواجهة وتجربة المستخدم (UI/UX) لتطبيق لتعلم اللغة العربية، يوفر مسارات تعليمية منفصلة ومتداخلة في الوقت نفسه لكل من العربية الفصحى والعامية. وتشمل الميزات المقترحة مواد تعليمية منظمّة، ومفردات وتعبيرات يومية، وتمارين سياقية، إضافة إلى ممارسة المحادثة المدعومة بالذكاء الاصطناعي، بما يتيح للمستخدمين محاكاة التواصل باللغة العربية في مواقف متنوعة. ويظهر التصميم المقترح كيفية دمج الذكاء الاصطناعي بوصفه عنصرًا هامًا ضمن بيئة تعليمية رقمية أوسع، وليس باعتباره المحور الوحيد للتطبيق. ومن المتوقع أن يشكل هذا التصور أساسًا لتطوير التطبيق وتقييمه مستقبلًا، بما يسهم في توفير تجربة تعلم أكثر مرونة وتفاعلية وارتباطًا بالسياق.

تعلم اللغة العربية؛ العربية الفصحى والعامية؛ تطبيق الهاتف المحمول؛ الذكاء الاصطناعي للمحادثة؛ التفكير التصميمي.

Arabic language learning presents a distinctive challenge for non-native learners because competence in formal Arabic does not necessarily result in readiness for everyday communication. Formal education commonly emphasizes *fusha*, particularly Modern Standard Arabic, which remains essential for literacy, academic communication, formal discourse, news, and written materials. Everyday communication across Arabic-speaking societies, however, frequently involves regional spoken varieties generally categorized as *'ammiyyah*. This sociolinguistic relationship has long been discussed through the concept of diglossia, in which related varieties perform different social and communicative functions (Albirini, 2016; Ferguson, 1959).

For learners who later study, travel, or temporarily reside in Middle Eastern countries, the difference between formal and everyday Arabic can become practically significant. Previous knowledge of *fusha* provides an important linguistic foundation, but participation in daily communication may require familiarity with vocabulary, pronunciation, expressions, interactional patterns, and communicative conventions that receive less attention in formal learning environments. Research on Arabic diglossia similarly demonstrates that the relationship between standard and spoken Arabic has important educational consequences (Saiegh-Haddad, 2023; Thomure et al., 2025).

This does not mean that *fusha* should be replaced by *'ammiyyah*. Formal Arabic remains central to literacy, academic communication, and communication across Arabic-speaking regions. The educational challenge concerns the transition between formal knowledge and communicative adaptation. Learners may possess a grammatical and lexical foundation in *fusha* while still requiring additional exposure to the spoken varieties used in everyday interaction. Putri and Thoyyibah (2024) similarly emphasize the relevance of both formal and spoken Arabic to communicative adaptability.

Mobile-assisted language learning creates new possibilities for responding to this gap. Mobile applications can combine audiovisual input, short exercises, listening practice, speech recording, repetition, feedback, progress tracking, and personalized review within one environment. Research on mobile-assisted and gamified language learning suggests that digital practice can support engagement and continuity when technological features remain aligned with meaningful educational objectives (Ishaq et al., 2021; Luo, 2023).

Existing digital language-learning applications demonstrate several useful approaches. Duolingo is widely associated with short progressive lessons, visible learning paths, immediate feedback, repeated practice, and gamification (Duolingo, n.d.; Shortt et al., 2023). The Cake app provides a complementary model through contextual audiovisual content, practical expressions, listening, pronunciation, repetition, and speaking-oriented activities (Cake, n.d.; Hardiana & Lubis, 2024; Paramita et al., 2022; Sujana et al., 2024). Memrise is also relevant because its Arabic learning environment combines structured language learning, native-speaker video, and AI-supported conversation (Memrise, n.d.). These applications provide useful benchmarks, but Dhou is not intended to reproduce their proprietary interfaces.

Generative artificial intelligence can create an additional opportunity for the transition from knowledge to communication. Conversational AI can provide flexible interaction, immediate responses, and repeated practice without requiring another human conversation partner to remain continuously available. Reviews of generative AI in language education identify conversation, feedback, personalization, and interactive language production among its important affordances (Law, 2024; Yang & Li, 2024). However, generative systems may also produce inaccurate or inconsistent language, including unmarked switching among Arabic varieties.

Dhou therefore conceptualizes AI as a practice environment rather than the primary source of instructional content. Human-reviewed audiovisual materials, vocabulary, learning activities, pronunciation models, and variety labels provide the verified instructional foundation, while AI provides a flexible environment for communicative practice.

Based on these considerations, this study proposes Dhou, an Arabic language-learning application organized around three principal components: *video learning*, a *gamified learning path*, and *conversational AI*. A review function supports learning continuity across these components. The architecture can be summarized as video learning → gamified learning path → conversational AI → review, or pedagogically as contextual input → structured practice → communicative production →

reinforcement. The objective of this study is to develop the conceptual UI/UX architecture of Dhou using design thinking and to explain how learner needs, product benchmarking, audiovisual learning, gamified practice, Arabic-variety awareness, and conversational AI can be integrated coherently in a single educational-technology app.

Methods

This study applies the *Design Thinking* framework as the primary product-development approach. Design Thinking was selected because educational-product development requires systematic understanding of learner problems before technological solutions are defined. The framework is organized around five iterative stages: Empathize, Define, Ideate, Prototype, and Testing (Razzouk & Shute, 2012), as illustrated in Figure 1.

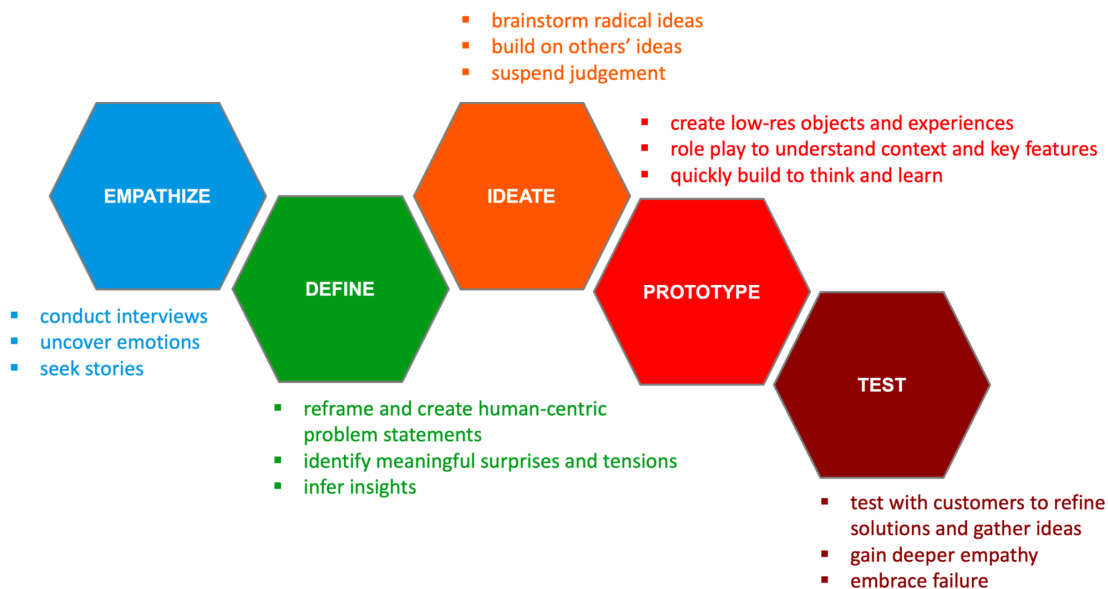


Figure 1. Design Thinking Process

Source: <https://www.designreview.byu.edu/>

In this study, the Empathize stage contains two complementary research activities: *unmoderated questionnaire-based user research* and *product benchmarking*. The first examines the experiences and needs of potential users, while the second examines how comparable digital products currently address related learning problems. The Define stage synthesizes evidence from both activities into a central design problem. The Ideate stage transforms the problem into the major components of Dhou. The Prototype stage establishes the conceptual information architecture of the app, while the Testing stage was conducted through moderated usability research.

Empathize

The Empathize stage aims to understand the problem from two perspectives: the learner perspective and the existing product

environment. The first perspective identifies what users experience, while the second identifies existing interaction patterns and opportunities in contemporary digital language learning.

Unmoderated Questionnaire-Based User Research

The first activity used unmoderated user research through an online questionnaire. In an unmoderated process, participants respond independently without the continuous presence of a researcher or facilitator. This data-gathering format allows participants to describe their experiences according to their own schedules and supports participation across different locations.

The target respondents consisted of students and university students who had previous experience studying, visiting, or temporarily living in Middle Eastern countries and who had previously received Arabic-language education. The questionnaire explored previous Arabic learning, perceived competence in *fusha*, experience communicating outside formal learning environments, exposure to spoken Arabic, difficulties experienced during adaptation, learning resources used, and perceived needs for additional support.

The responses indicated a recurring gap between previous formal Arabic learning and everyday communicative experience. Participants generally possessed previous knowledge of *fusha*, which supported formal learning, reading, and academic activities. Nevertheless, they reported greater difficulty when adapting to everyday interaction in environments where regional spoken Arabic was frequently used.

The difficulty did not necessarily result from an absence of Arabic knowledge. Instead, participants encountered vocabulary, expressions, pronunciation patterns, speech rates, and communicative conventions that differed from the Arabic emphasized in their previous learning experiences. Respondents also perceived learning resources for *fusha* as relatively structured and available, whereas resources designed specifically for practical learning of relevant *'ammiyyah* varieties were considered less accessible.

The responses therefore suggested a need for more contextual listening exposure, opportunities to repeat spoken material, pronunciation practice, and spaces in which learners could rehearse communication before or alongside real-world interaction. The present paper reports this recurring qualitative pattern without claiming statistical generalization; the final empirical report should include the actual respondent count, recruitment process, questionnaire instrument, data-collection period, and detailed analysis procedure.

Product Benchmarking

The second research method was product benchmarking or competitive research. Its purpose was not to determine which application is universally superior or to reproduce a particular interface. Instead, benchmarking was used to understand how established language-learning products address structured progression, contextual exposure, pronunciation practice, review, and conversational interaction. Duolingo, Cake, and Memrise were selected because they represent complementary patterns relevant to the learner needs identified through the questionnaire. Table 1 summarizes the comparison across nine design dimensions.

Table 1. Comparison of Duolingo, Cake, and Memrise and Opportunities for Dhou

Design Dimension	Duolingo	Cake	Memrise	Opportunity for Dhou
Structured progression	Progressive learning path with short lessons and review	Level- and topic-based learning	Scenario- and level-oriented learning	Progressive Arabic-specific learning path
Gamification	Strong progress and motivational mechanisms	Present but less central	Progress and learning continuity	Gamification supporting regular Arabic practice
Contextual video	Present in selected experiences	Core emphasis on short real-life videos	Strong use of native-speaker videos	Core contextual Video Learning environment
Interactive exercises	Reading, listening, writing, speaking, recognition, review	Listening, quizzes, sentence activities, pronunciation , speaking	Vocabulary, listening, review, pronunciation , speaking	Wide variety of Arabic learning activities
Voice/pronunciation	Speaking practice available	Strong recording and pronunciation practice	Speaking practice available	Recording, shadowing, and pronunciation practice
Conversational AI	Available in selected AI-supported experiences	Not the central model	AI language partner available	Core communicative-practice component
Arabic focus	Arabic course available	Arabic not a principal current target language	Arabic course available	Designed specifically around Arabic needs

Design Dimension	Duolingo	Cake	Memrise	Opportunity for Dhou
Arabic-variety awareness	Not the central architecture	Not applicable to current principal target languages	Not central to Arabic-course architecture	Clear labeling and selection of supported Arabic varieties
Learning integration	Depends on available course features	Strong relation between video and practice	Connects vocabulary, video, and AI conversation	Video → Learn → Conversation → Review

The benchmarking reveals complementary strengths. Duolingo provides a useful reference for progressive practice and gamification; Cake demonstrates the potential of short contextual audiovisual learning and pronunciation-oriented activities; and Memrise provides an important reference for connecting structured learning, native-speaker video, and AI-supported communication. The comparison also shows that Dhou cannot define its contribution merely by combining existing features. Instead, those interaction patterns need to be reorganized around the Arabic-specific learner problem identified through the questionnaire.

Define

The Define stage synthesizes the findings obtained during Empathize. The questionnaire indicates that some learners possess a foundation in *fusha* but continue to experience difficulty when adapting to everyday spoken Arabic. The benchmarking analysis shows that contemporary applications already provide established interaction patterns for structured learning, audiovisual input, pronunciation practice, repeated exercises, and conversational AI.

The central design problem is therefore formulated as follows: *How can a mobile learning application support learners who possess a foundation in fusha but require greater contextual exposure, structured practice, and communicative opportunities for everyday Arabic?*

The design requirements derived from this problem are contextual audiovisual exposure; clear identification of the Arabic variety being presented; structured and repeatable practice; varied activities across vocabulary, listening, sentence construction, pronunciation, comprehension, and active recall; opportunities for flexible language production; conversational AI that maintains the selected variety as consistently as possible; and a review mechanism that returns difficult material to later practice.

Ideate

The Ideate stage transforms these requirements into a simplified learning architecture consisting of three principal components: *video learning*, a *gamified learning path*, and *conversational AI*. A *review* function operates across the three components as a reinforcement mechanism. Video Learning responds to the need for contextual exposure; the Gamified Learning Path responds to the need for systematic practice and continuity. Conversational AI responds to the need for flexible communication, and Review supports consolidation.

The concept deliberately limits itself to these major functions so that the application does not become a collection of disconnected features. The underlying learning model is contextual input → structured practice → communicative production → reinforcement.

Prototype

The Prototype stage translates the learning model into a conceptual information architecture with four primary destinations: *Video*, *Learn*, *AI Conversation*, and *Review*. Video contains contextual audiovisual learning. Learn contains the gamified path and structured activities. AI Conversation contains interactive conversational practice, and Review consolidates material requiring reinforcement.

An Arabic variety functions as a system-wide attribute rather than a separate module. Video content identifies its variety, learning activities identify the variety being practiced, conversational practice begins with the relevant variety selection, and review retains the variety information associated with saved or difficult material. Information can also move among components: language encountered through video learning can become structured practice, while difficult items from Learn or AI conversation can return through Review.

Testing: Moderated Usability Research

The Testing stage was conducted through moderated usability research involving 20 participants who had previous experience learning Arabic. The sessions were conducted remotely through Zoom, allowing the researcher to observe participants while they interacted with the Dhou prototype and completed predetermined tasks representing the application's major user flows. The tasks included navigating the gamified learning path, identifying the selected Arabic variety, accessing video-based learning materials, completing structured learning activities, initiating an AI-supported conversation, and locating the Review function. During each session, the moderator observed task completion, navigation behavior, hesitation, repeated actions, misunderstanding of interface elements, and participant responses to the Arabic-variety labels.

A think-aloud protocol was also used to encourage participants to verbalize their expectations and decision-making while interacting with

the prototype. After completing the tasks, participants were asked short semi-structured questions concerning interface clarity, perceived usefulness, difficulty, and areas that required improvement. The resulting observation notes, task outcomes, participant comments, and interview responses were analyzed descriptively and qualitatively to identify recurring usability patterns.

Results

Conceptual Architecture of Dhou

The primary result of the design thinking process is a conceptual learning architecture consisting of three principal learning environments supported by review. The visible product architecture remains simple, while the pedagogical sequence moves from exposure to practice, production, and reinforcement. Table 2 summarizes these components, their primary functions, and the learner processes they support.

Table 2. Principal Components of Dhou and Their Learner Processes

Component	Primary Function	Learner Process
Video Learning	Contextual language input	Observe, listen, notice, and understand
Gamified Learning Path	Structured language practice	Recognize, retrieve, construct, pronounce, and reinforce
Conversational AI	Communicative language production	Formulate, respond, interact, and maintain communication
Review	Learning reinforcement	Revisit, correct, repeat, and consolidate

The architecture can be summarized as Video Learning → Gamified Learning Path → Conversational AI → Review. This sequence does not require a completely rigid path for every user. Instead, it describes the pedagogical relationship among the major functions: Video provides language in context; the learning path transforms that input into deliberate practice, conversational AI provides a space for active language use, and review supports consolidation.

The Dhou mockups (Figures 2–4) illustrate the application’s three main learning environments: video learning, gamified learning path, and conversational AI. Video learning provides contextual Arabic exposure through audiovisual materials; the gamified learning path supports structured and progressive practice, while conversational AI allows learners to apply their knowledge through interactive communication with artificial intelligence. Together, these interfaces represent Dhou’s core learning flow from contextual input and structured practice to active language production and continuous reinforcement.



Figure 2. Gamified Learning Path Interface

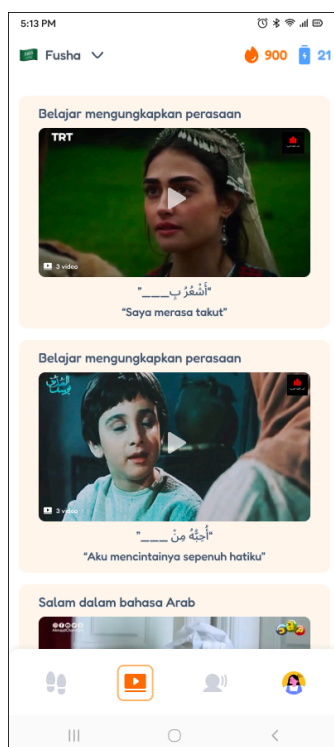


Figure 3. Video Learning Interface



Figure 4. Conversational AI Interface

Video-Based Contextual Learning

Video Learning provides audiovisual exposure to Arabic within communicative contexts rather than presenting language exclusively through isolated words and sentences. Potential sources include original educational productions, appropriately licensed audiovisual material, authorized excerpts, and other legally usable media relevant to language learning. Film- and music-related content may also be used where licensing requirements are addressed appropriately.

The pedagogical value of audiovisual material lies in the combination of linguistic and contextual information. Learners can hear pronunciation, speech rhythm, intonation, connected speech, and communicative timing while observing the context in which language is used. Each learning item should clearly indicate whether the material uses *fusha* or a supported regional *'ammiyyah* variety. Table 3 summarizes the principal functions supporting Video Learning and their pedagogical purposes.

Table 3. Video Learning Functions and Their Pedagogical Purposes

Function	Pedagogical Purpose
Arabic subtitles	Connect spoken language with Arabic orthography
Optional translation	Provide meaning support when needed
Interactive transcript	Support focused examination of spoken content
Vocabulary extraction	Identify important lexical items

Function	Pedagogical Purpose
Expression extraction	Identify useful communicative expressions
Sentence-level replay	Support focused listening and repetition
Playback-speed control	Provide additional listening support
Listening comprehension	Evaluate understanding of audiovisual input
Expression recognition	Develop contextual language recognition
Subtitle on/off control	Encourage increasing listening independence
Shadowing	Develop rhythm, fluency, and pronunciation
Voice recording	Encourage active spoken production
Record-and-compare	Support reflection on learner pronunciation
Pronunciation feedback	Provide guidance when technically reliable
Save to Review	Transfer useful or difficult language to practice later
Arabic-variety label	Identify <i>fusha</i> or the relevant supported <i>'ammiyyah</i>
Related practice	Connect contextual material to structured activities

These functions transform audiovisual input into an active learning resource. Learners requiring greater support can use subtitles, translation, slower playback, and repeated segments. As comprehension develops, those supports can be reduced. Language encountered through contextual video can also be reinforced later through vocabulary, listening, sentence construction, pronunciation, or review activities.

Gamified Learning Path

The gamified learning path organizes language practice progressively and makes learner progress visible to users. Short activities allow users to practice regularly without requiring long sessions. Gamification provides motivational support through progress visualization, completion indicators, learning continuity, and achievements, but progression remains tied to meaningful learning activity rather than game mechanics alone (Dehghanzadeh et al., 2021; Luo, 2023).

A central design requirement is activity variety. The path should not depend excessively on one recognition-based quiz format because language knowledge involves recognition, comprehension, retrieval, listening, sentence construction, pronunciation, and oral production. Table 4 lists the proposed activity types and their primary learning functions.

Table 4. Gamified Learning Path Activity Types and Primary Learning Functions

Activity Type	Primary Learning Function
Vocabulary learning	Introduce and reinforce lexical knowledge
Word–meaning matching	Develop semantic recognition
Fill in the blank	Develop contextual lexical and structural selection
Sentence arrangement	Develop sentence-construction awareness
Word pairing	Reinforce lexical and structural relationships
Listen and choose	Develop auditory recognition
Audio-to-text dictation	Connect listening comprehension with written production
Text-to-audio recognition	Connect written forms with spoken forms
Controlled translation	Reinforce cross-language meaning recognition
Contextual response selection	Develop pragmatic understanding
Dialogue completion	Develop understanding of interactional sequence
Pronunciation repetition	Support spoken production through model imitation
Voice recording	Enable learners to monitor their own speech
Record-and-compare	Support reflection on pronunciation
Shadowing	Develop rhythm, fluency, and pronunciation
Listening comprehension	Evaluate understanding of extended spoken input
Active recall	Strengthen retrieval of previously learned language
Mixed-skill practice	Integrate several language skills
Personalized error review	Return recurring difficulties to structured practice
Spaced review	Reinforce material over increasing intervals

The proposed activities move learners through different levels of engagement with language. Recognition activities create familiarity, retrieval requires stronger memory access, listening connects linguistic knowledge with auditory input, sentence-building requires active organization, and recording or shadowing moves learning toward oral production. Arabic-variety information remains visible throughout the learning path so that users know whether they are practicing *fusha* or a supported *‘ammiyyah* variety.

Conversational AI

Conversational AI provides opportunities to use Arabic interactively after or alongside structured practice. Conventional digital exercises typically operate within a limited response space, whereas conversation requires learners to interpret incoming language, formulate responses, and maintain interaction when the next turn is not predetermined.

The Dhou interaction is designed to remain simple. Before starting a conversation, the learner selects the Arabic variety to be practiced, such as *fusha* or one of the supported ‘*ammiyyah*’ varieties. The learner is not required to select a specific AI character or complete a long configuration process. Conversation may proceed freely; when useful, the AI may briefly suggest a situational conversation and establish the context automatically, but this remains part of the general conversation experience rather than a separate product feature. Table 5 summarizes the principal functions of the conversational AI component.

Table 5. Conversational AI Functions and Their Purposes

Function	Purpose
Arabic-variety selection	Allows practice in <i>fusha</i> or a supported ‘ <i>ammiyyah</i> ’ variety
Interactive conversation	Provides active language-production practice
Contextual conversation suggestion	Allows AI to introduce a situational context when useful
Variety consistency	Maintains the selected Arabic variety
Adaptive language level	Keeps interaction reasonably appropriate to learner ability
Learning support	Provides assistance when needed
Feedback	Identifies areas requiring additional practice
Review connection	Returns relevant difficulties to later learning

The most important Arabic-specific requirement is variety consistency. A generative system may produce grammatically acceptable Arabic while switching unexpectedly from *fusha* to a regional variety or combining forms from different dialects. AI evaluation should therefore consider not only grammaticality and relevance but also consistency with the variety selected by the learner. Human-reviewed vocabulary, transcripts, pronunciation material, learning activities, and variety labels remain the primary instructional reference, while generated interaction provides flexibility for communicative practice.

Review and Learning Continuity

Review forms the reinforcement layer of Dhou. Vocabulary or expressions saved during video learning can be revisited; errors in the learning path can return for additional practice; pronunciation activities can be repeated; and difficulties identified during conversational interaction can inform later review. Items associated with repeated difficulty may appear more frequently, while successfully retrieved material can return at longer intervals. The complete cycle is therefore

Contextual Exposure → Structured Practice → Communicative Production → Review → Repeated Use.

Conceptual Positioning of Dhou

The benchmarking analysis indicates that individual Dhou functions are not technologically unprecedented. Video learning, gamified progression, speech recording, pronunciation practice, and conversational AI already exist in contemporary products. The conceptual contribution therefore lies in how these functions are organized around an identified Arabic-learning problem. Preliminary user research suggests that some learners enter Arabic-speaking environments with previous *fusha* knowledge but remain insufficiently prepared for everyday spoken communication. Dhou responds by combining contextual Arabic input, structured gamified practice, variety-aware conversational AI, and continuous review.

Moderated Usability Testing Results

The moderated usability testing with 20 participants indicated generally positive usability outcomes across the principal Dhou user flows. Eighteen participants (90%) successfully completed the main navigation tasks without moderator assistance, while 17 participants (85%) correctly understood the distinction between the *fusha* and *'ammiyyah* variety labels. Nineteen participants (95%) successfully accessed and understood the Video Learning interface, 18 participants (90%) completed the activities within the Gamified Learning Path, and 17 participants (85%) successfully initiated and maintained interaction with the Conversational AI feature. In the post-task evaluation, 18 participants (90%) described the overall interface as easy to understand and perceived the combination of Video Learning, progressive learning activities, and AI-supported conversation as a clear learning flow. Nevertheless, several recurring usability issues were identified. Participants occasionally experienced difficulty locating the Review function, interpreting several interface icons, and understanding the feedback presented after completing an activity. These findings indicate that the principal Dhou learning architecture was generally understandable to participants while also identifying specific interface elements requiring refinement.

Discussion

Learner Problem: Fusha and Everyday Communication

The findings suggest that the central problem addressed by Dhou is not simply access to Arabic-learning materials but the gap between formal competence and communicative readiness. The preliminary questionnaire indicates that learners may already possess a foundation in *fusha* while still encountering difficulty in everyday spoken interaction. This shifts the design problem away from teaching Arabic entirely from

the beginning toward supporting the transition from previously acquired formal knowledge to more flexible communicative competence.

This interpretation is consistent with the sociolinguistic characteristics of Arabic. Diglossia has long been used to describe the functional differentiation of Arabic varieties (Ferguson, 1959), while contemporary work emphasizes a more dynamic relationship shaped by context, identity, education, and communicative purpose (Albirini, 2016). From an educational perspective, competence in *fusha* remains highly valuable but does not automatically prepare learners for every spoken form encountered in local daily communication.

The questionnaire findings, therefore, should not be interpreted as evidence that conventional *fusha* education is ineffective. Instead, they point to a mismatch between the communicative situations emphasized during prior study and those encountered after learners enter Arabic-speaking environments in real life. This interpretation is compatible with research emphasizing the educational consequences of Arabic diglossia and the importance of communicative adaptability across varieties (Putri & Thooyibah, 2024; Saiegh-Haddad, 2023; Thomure et al., 2025).

Empathize Findings and Product Benchmarking

Using unmoderated questionnaire research and product benchmarking within Empathize provides two complementary forms of evidence. User research identifies what learners struggle with, while benchmarking (competitive analysis) shows how comparable products currently address related learning needs. This combination reduces the risk of developing a product based only on user requests or only on competitor features.

If development relied only on benchmarking, Dhou could become a collection of familiar interaction patterns without a clearly defined learner problem. If it relied only on questionnaire responses, the design might ignore useful patterns already established in mobile language learning. The synthesis of both sources allows the Define stage to identify an opportunity that is simultaneously user-centered and product-aware.

The benchmark products provide complementary references. Duolingo demonstrates the value of short progressive practice and visible learning structures; Cake illustrates contextual audiovisual learning, and Memrise demonstrates the increasingly common combination of native-speaker video, structured learning, and AI conversation. These similarities also prevent exaggerated novelty claims. Dhou's contribution must therefore be explained through the Arabic-specific problem it addresses and the way its components are integrated.

Video Learning as Contextual Input

Video Learning responds primarily to insufficient contextual exposure. Written materials can represent vocabulary and grammar efficiently, but spoken communication also includes pronunciation, rhythm, intonation, connected speech, timing, speaker relationships, and contextual meaning. Audiovisual material can expose learners to several of these dimensions simultaneously.

For learners identified during Empathize, this exposure may be especially important because their difficulty concerns the applicability of previous formal learning to everyday communication activities. Video, therefore, should not function as passive entertainment. Subtitles, transcripts, vocabulary extraction, replay, listening tasks, shadowing, recording, and review transform audiovisual material from passive exposure into guided input.

This design principle is consistent with studies of Cake, which emphasize contextual audiovisual material, pronunciation, listening, and speaking-oriented practice (Hardiana & Lubis, 2024; Paramita et al., 2022; Sujana et al., 2024). Dhou adapts the pedagogical value of that interaction pattern to Arabic while adding explicit variety labeling. Such labeling is important because learners may otherwise encounter unfamiliar spoken forms without understanding their regional or functional status.

Contextual video also creates a balance between authenticity and instructional control. Authentic or semi-authentic material can expose learners to realistic language but may also overwhelm them with rapid speech or unfamiliar forms. Interface support can reduce difficulty without removing context entirely. Content selection should therefore consider level, clarity, communicative usefulness, cultural appropriateness, variety, and learning objectives for users.

Gamified Learning Path as Structured Practice

Contextual exposure alone is insufficient for systematic language development. Learners need repeated opportunities to retrieve vocabulary, process listening input, construct sentences, and produce spoken language. The gamified learning path, therefore, acts as the structured-practice layer of Dhou.

Research on mobile-assisted learning indicates that regular digital practice can support continuity (Ishaq et al., 2021), while gamification can contribute to engagement when game mechanisms remain aligned with pedagogical activity (Dehghanzadeh et al., 2021; Luo, 2023). Duolingo provides a useful benchmark because its learning model demonstrates how short sequential activities, feedback, and visible progress can organize large amounts of language material into manageable sessions (Shortt et al., 2023).

For Dhou, however, the varied activity types are more important than gamification itself. Recognition, comprehension, retrieval, sentence construction, pronunciation, recording, shadowing, and review involve different levels of processing. A learner who can recognize a word is not necessarily able to retrieve it spontaneously; a learner who understands a written sentence may still fail to recognize the same language in connected speech. The learning path therefore aims to move from recognition toward controlled production rather than allowing visual progress to substitute for linguistic progress.

This relationship also explains why video learning and the gamified learning path should remain connected. Video answers how language is used in context, while the learning path provides repeated opportunities to process and retain that language. Together they combine contextual authenticity with deliberate practice.

Conversational AI as Communicative Practice

Conversational AI extends the learning sequence from controlled activity toward freer, active language use. Structured exercises generally operate within predetermined response spaces, while conversation requires learners to interpret incoming language, retrieve relevant vocabulary, formulate responses, and maintain interaction dynamically. This transition is particularly relevant to the Empathize findings because everyday communicative difficulty cannot be addressed entirely through additional recognition exercises.

Generative AI offers repeated interaction without requiring another person to be continuously available, and recent research identifies conversation and individualized feedback as important potential uses in language learning (Law, 2024; Yang & Li, 2024). Nevertheless, Dhou positions AI as a practice environment rather than a curriculum. Curated audiovisual materials, vocabulary, activities, and pronunciation models remain the stable instructional reference, while conversational AI provides flexibility for production.

Arabic introduces a specific quality requirement: variety consistency. A generated response may be grammatically understandable while switching unexpectedly between *fusha* and regional varieties. Such output may be recognizable to experienced speakers but confusing to learners. Dhou, therefore, uses explicit variety selection and treats consistency with that selection as an AI evaluation criterion. The objective is not to deny natural variation in Arabic, but to provide sufficiently stable scaffolding for learners before they are expected to interpret complex sociolinguistic variation independently.

Integration and Contribution of Dhou

The strongest conceptual contribution of Dhou lies in the relationship among its components rather than in any individual technology. Video learning, gamification, speech recording, and conversational AI already exist in contemporary products. The contribution of Dhou is therefore better expressed as an integrated learning cycle: Contextual Input → Structured Practice → Communicative Production → Reinforcement.

At a more detailed pedagogical level, the sequence can be understood as Exposure → Noticing → Recognition → Retrieval → Controlled Production → Communicative Production → Feedback → Review. This structure reduces the risk of creating a feature-rich but pedagogically fragmented application. Language encountered through video learning can become structured practice; difficulties in learning can enter review; and problems revealed during AI conversation can return to more controlled practice.

Benchmarking also clarifies Dhou's position relative to existing products. Duolingo is a reference for progressive gamified practice, Cake for contextual audiovisual learning, and Memrise for the integration of native-speaker video and AI-supported communication. Dhou should therefore not be described merely as a combination of these products. A more academically defensible positioning is that Dhou is a variety-aware Arabic mobile learning environment designed to support the transition from contextual exposure to structured practice and active communication.

Arabic-variety awareness operates throughout the system rather than as an independent module. Video items are labeled, learning activities identify the variety being practiced, AI conversation uses a selected variety, and review retains that information. This reduces interface complexity while acknowledging that variation influences listening, vocabulary, pronunciation, speaking, and interaction across the complete learning process.

The architecture also has pedagogical implications beyond individual self-study. Dhou can complement formal *fusha* instruction rather than replace it. Teachers may use contextual audiovisual materials, structured practice, or AI-supported conversation as supplementary learning outside limited classroom time, while higher-stakes assessment and linguistic explanation remain under human supervision.

Limitations and Future Development

Although the prototype has undergone initial moderated usability testing with 20 participants, the findings should be interpreted within the scope of an early-stage usability evaluation. The participant group remains limited and does not represent the diversity of all Arabic-language learners. Furthermore, the usability test primarily evaluated navigation, interface comprehension, task completion, Arabic-variety labeling, and

interaction with the major application components rather than longitudinal learning outcomes. Therefore, the present study does not claim that Dhou has been empirically demonstrated to improve listening, speaking, vocabulary, pronunciation, or long-term communicative competence; it evaluates only the ease of use and user experience of the app.

The usability findings nevertheless provide several directions for subsequent prototype refinement. Future development should improve the visibility of the Review function, increase the clarity of selected interface icons, and provide more explicit feedback following activity completion. Subsequent testing should evaluate the revised prototype with a broader and more diverse participant group and examine whether the usability improvements remain consistent across different learner backgrounds and levels of Arabic proficiency. Later empirical studies should also investigate learning effectiveness over longer periods and evaluate the linguistic reliability, variety consistency, and pedagogical usefulness of Dhou's conversational AI.

The scope of spoken Arabic should also remain deliberately limited during the initial Minimum Viable Product (MVP). Dhou is proposed to support *fusha* alongside Egyptian Arabic and Levantine Arabic rather than attempting to include numerous regional varieties simultaneously. This limited scope allows the linguistic content, pronunciation models, conversational output, and variety labeling to be reviewed and validated more systematically before additional Arabic varieties are introduced into Dhou.

Conclusion

This study proposes Dhou as an educational technology for learning Arabic that integrates contextual video learning, structured gamified practice, conversational AI, and review within a learning environment that distinguishes between *fusha* and supported *'ammiyyah* varieties. Moderated usability testing with 20 participants indicated that the proposed interface and learning flow were generally understandable and usable, while also revealing several areas requiring further refinement.

Future research should focus not only on improving interface clarity and validating Arabic content and AI-generated responses but also on developing a more systematic and pedagogically structured approach for learning materials. This includes organizing content according to learner proficiency levels, communicative needs, progression between *fusha* and *'ammiyyah*, and clear learning objectives to ensure that the learning pathway remains coherent, measurable, and relevant to real-world Arabic communication.

Acknowledgment

The author acknowledges Universitas Abdul Chalim (UAC) for providing an academic environment that supported the conceptual development of Dhou. Appreciation is also extended to the students and university students whose questionnaire responses and experiences contributed to identifying the Arabic-learning and adaptation challenges addressed in this study, as well as to educators and colleagues whose discussions contributed to the development of the concept.

References

- Albirini, A. (2016). *Modern Arabic sociolinguistics: Diglossia, variation, codeswitching, attitudes and identity*. Routledge. <https://doi.org/10.4324/9781315683737>
- Cake. (n.d.). *Learn real-life English & Korean*. Retrieved August 23, 2026, from <https://cake.day/>
- Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaei, E., & Noroozi, O. (2021). Using gamification to support learning English as a second language: A systematic review. *Computer Assisted Language Learning*, 34(7), 934–957. <https://doi.org/10.1080/09588221.2019.1648298>
- Duolingo. (n.d.). *How to use Duolingo for language learning*. Retrieved August 23, 2026, from <https://blog.duolingo.com/duolingo-101-how-to-learn-a-language-on-duolingo/>
- Ferguson, C. A. (1959). Diglossia. *Word*, 15(2), 325–340. <https://doi.org/10.1080/00437956.1959.11659702>
- Flinn, A. (2026). Surveying native speakers to find the proportions of registers used in Levantine Arabic. *Register Studies*, 8(1), 70–93. <https://doi.org/10.1075/rs.25002.fli>
- Hardiana, N. D., & Lubis, Y. (2024). Cake application as a tool for boosting vocabulary mastery in EFL education. *RETORIKA: Jurnal Ilmu Bahasa*, 10(2), 594–602. <https://doi.org/10.22225/jr.10.2.2024.594-602>
- Ishaq, K., Mat Zin, N. A., Rosdi, F., Jehanghir, M., Ishaq, S., & Abid, A. (2021). Mobile-assisted and gamification-based language learning: A systematic literature review. *PeerJ Computer Science*, 7, e496. <https://doi.org/10.7717/peerj-cs.496>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, 5, 100156. <https://doi.org/10.1016/j.caeai.2023.100156>
- Law, L. H. L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature

- review. *Computers and Education Open*, 6, 100174.
<https://doi.org/10.1016/j.caeo.2024.100174>
- Luo, Z. (2023). The effectiveness of gamified tools for foreign language learning (FLL): A systematic review. *Behavioral Sciences*, 13(4), 331.
<https://doi.org/10.3390/bs13040331>
- Memrise. (n.d.). *Arabic language course*. Retrieved August 23, 2026, from <https://www.memrise.com/en/learn-arabic/arabic-course>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Paramita, I. G. A. D., Ardika, I. W. D., Setyono, E. Y., Yuliantini, N. N., & Suciani, N. K. (2022). The use of Cake application to improve speaking ability. *Journal of Applied Studies in Language*, 6(2), 214–220.
<https://doi.org/10.31940/jasl.v6i2.345>
- Putri, A. H., & Thoyyibah, A. (2024). Exploring Arabic diglossia in learning kalām: Bridging the gap in daily communication. *Alsinatuna*, 9(2), 184–198. <https://doi.org/10.28918/alsinatuna.v9i2.7076>
- Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? *Review of Educational Research*, 82(3), 330–348.
<https://doi.org/10.3102/0034654312457429>
- Saiegh-Haddad, E. (2023). Embracing diglossia in early literacy education in Arabic: A pilot intervention study with kindergarten children. *Oxford Review of Education*, 49(1), 48–68.
<https://doi.org/10.1080/03054985.2022.2090324>
- Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). 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(3), 517–554.
<https://doi.org/10.1080/09588221.2021.1933540>
- Sujana, I. K. B. A., Santosa, M. H., & Mahayanti, N. W. S. (2024). Improving students' pronunciation skill using Cake application. *Journal of English Language Learning*, 8(2).
<https://doi.org/10.31949/jell.v8i2.10148>
- Telaumbanua, Y. A., Harefa, S. T. H., Zebua, M. R., Zega, S. M., Zalukhu, J., & Mendrofa, R. (2024). Comparison of the effectiveness of Duolingo and Cake applications in improving students' English language proficiency. *Jurnal Penelitian Inovatif*, 4(4), 2395–2400.
<https://doi.org/10.54082/jupin.841>
- Thomure, H. T., Taha, H., Sabella, T., & Saleh, R. (2025). A systematic review of teaching and learning in the context of Arabic diglossia.

- Social Sciences & Humanities Open*, 12, 102281.
<https://doi.org/10.1016/j.ssaho.2025.102281>
- Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System*, 124, 103351.
<https://doi.org/10.1016/j.system.2024.103351>
- Zaidan, O. F., & Callison-Burch, C. (2014). Arabic dialect identification. *Computational Linguistics*, 40(1), 171–202.
https://doi.org/10.1162/COLI_a_00169