



Enhancing English Grammar Education Through a Mobile-Assisted Language Learning Application: A Case Study Using the Formula 33 Approach

Sri Mulyati^{1*}, Nizamuddin Sadiq², Astri Hapsari², Galuh Ihsan Nurkhois¹

¹ Department of Informatics, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta 55281, Indonesia

² Department of English Language Education, Faculty of Psychology and Sociocultural Sciences, Universitas Islam Indonesia.

Received: December 16, 2025

Revised: February 05, 2026

Accepted: March 25, 2026

Published: March 31, 2026

Corresponding Author:

Sri Mulyati

mulya@uii.ac.id

DOI: [10.29303/jppipa.v12i3.13915](https://doi.org/10.29303/jppipa.v12i3.13915)

 Open Access

© 2026 The Authors. This article is distributed under a (CC-BY License)



Abstract: The integration of technology into language learning has created new opportunities to enhance effectiveness and learner engagement. However, existing grammar learning applications often lack sufficient contextual scaffolding, provide minimal adaptive feedback, and rely on chatbots that are not grounded in curated learning sources, limiting their effectiveness for structured language acquisition. This study focuses on developing a mobile-assisted English grammar-learning application using the Formula 33 approach and an RAG-based chatbot. The research uses a prototyping methodology, enabling iterative development and continuous improvement informed by user feedback. The application was built using Android Studio (Kotlin), integrated with Firebase Realtime Database, and features a Retrieval Augmented Generation (RAG)-based chatbot that uses the Formula 33 book as its main self-learning source. The implementation results show that all main features, including formula learning, practice exercises, quizzes, and personalized feedback, are present and function effectively. System evaluation using Black Box Testing confirmed functional reliability, while the System Usability Scale (SUS) test yielded an average score of 83.125, categorized as “Excellent” and “Acceptable” for usability. These findings demonstrate that the Formula 33 method, implemented as a pedagogical framework integrated with RAG technology based on curated textbook sources and evaluated through SUS standards, effectively supports users in progressively and contextually understanding English grammar structures. The proposed model offers strong potential as a replicable framework for the development of future technology-assisted language learning applications.

Keywords: Formula 33; Prototyping; RAG Chatbot; Technology assisted language learning

Introduction

Technology has revolutionized the world of education, fundamentally transforming classroom dynamics, teaching methods, and curriculum design (Abidin et al., 2017). The rapid advancement of digital infrastructure and the widespread adoption of mobile devices have created new and exciting opportunities in education, as documented in various studies across

diverse educational contexts. For example, Badia (2015) identifies a range of informal learning environments enabled by technology, including virtual, game-based, and mobile learning environments, each of which extends the boundaries of traditional education. These developments reflect a global shift toward more technology-oriented education, where learners can access educational materials from anywhere and at any time, fostering a more flexible and self-directed learning

How to Cite:

Mulyati, S., Nizamuddin, N. S., Hapsari, A., & Nurkhois, G. I. (2026). Enhancing English Grammar Education Through a Mobile-Assisted Language Learning Application: A Case Study Using the Formula 33 Approach. *Jurnal Penelitian Pendidikan IPA*, 12(3), 172-183. <https://doi.org/10.29303/jppipa.v12i3.13915>

culture (Mohammed, 2020). Scholars emphasize that this "anytime-anywhere" capability fosters a ubiquitous learning environment, where educational boundaries are blurred, allowing students to seamlessly bridge formal classroom instruction with informal daily practice (Crompton & Burke, 2018).

Studies across secondary and higher education contexts demonstrate that smartphones and digital technology broadly impact educational outcomes: they extend students' attention spans, enhance cognitive capabilities, improve academic performance, and strengthen learner collaboration (Sabron et al., 2020). At the higher education level in particular, technology has improved student collaboration, backlog management, and access to educational materials, while also accelerating the delivery of more personalized and flexible learning pathways (Mohammed, 2020). Furthermore, utilizing mobile devices in familiar, authentic environments – such as local communities and campuses provides learners with opportunities for daily encounters and real communication, which effectively bridges the gap between formal instruction and real-world application (Shadiev et al., 2020). Previous studies also indicate that the integration of technology in education requires appropriate pedagogical strategies to ensure that learning objectives are achieved effectively. These findings collectively underscore the importance of designing technology-supported learning environments that carefully balance accessibility, engagement, and instructional effectiveness.

Within the broader landscape of educational technology, language learning applications have emerged as an increasingly important tool, providing easier and more flexible access to learning resources while offering a more interactive and engaging learning method (Refat et al., 2020; Sadiq, 2016). Technology is becoming a key element in modern language education, prompting further research into its effectiveness and influence on the learning process (Refat et al., 2020). The changing needs and preferences of users in learning English, especially due to technological developments, are essential considerations in the design of effective language learning tools (Ahmadi, 2018). Mobile-Assisted Language Learning (MALL) has been extensively studied as a means of supporting authentic, real-world language practice, allowing learners to bridge formal classroom instruction with informal daily language use (Crompton & Burke, 2018). Language learning applications have evolved significantly: originally designed as simple tools for beginners, they have now evolved into complex platforms catering to users from beginner to advanced levels, featuring conversation simulation, adaptive content, and grammar-focused instruction tailored to individual

progress (Xu et al., 2019). Intuitively designed learning apps have been shown to increase students' motivation to learn, and applications such as Duolingo have become essential tools in modern language teaching, providing flexible access to resources and allowing learners to study at their own pace (Mohammed, 2020). Recent studies further confirm that this ubiquity not only supports personalized learning paths but also reduces cognitive load, allowing students to process information more effectively compared to rigid classroom schedules (Ozer & Kiliç, 2018).

Among the pedagogical frameworks suitable for technology-mediated grammar instruction, Formula 33 stands out as an innovative English learning strategy that has been tested and proven to produce positive results, especially in improving students' ability to create accurate grammar in English sentences. Formula 33 adopts the SCORE paradigm, Simple, Clear, Original, Relevant, and Fun (Sadiq, 2015, 2016) enabling learners to master grammar structures through a gradual and structured learning process. Research on the effectiveness of translating words from Indonesian to English emphasizes the importance of word structures such as contractions, semi-definite expressions, and tenses (Maulidiyah & Garnida, 2021), areas where Formula 33 provides strength. One of the key features of a learning program based on Formula 33 is the way it approaches instruction by utilizing concepts closely related to students' daily lives, allowing students to understand the concepts being taught slowly and deliberately through three domains: instruction, delivery, and student assistance (Sadiq, 2016). This approach is highly relevant to constructivist learning theory, which emphasizes the importance of scaffolding and the Zone of Proximal Development (ZPD), enabling teachers to provide adequate support for students to achieve higher understanding independently (Sadiq, 2015). Furthermore, Formula 33 is aligned with the principles of Contextual Learning Theory (CTL), creating meaningful and relevant learning experiences for English learners. By combining rule-based decision logic with forward chaining, Formula 33 offers a structured, personalized, and engaging framework, making it not only a learning method but a comprehensive framework for transforming the way students learn English grammar.

Despite the growing adoption of MALL and grammar-focused learning applications, several critical gaps remain unaddressed. First, while Formula 33 has demonstrated positive outcomes for grammar acquisition in classroom settings (Sadiq, 2015, 2016) its integration into mobile-based digital applications has not yet been systematically explored, limiting its reach and scalability beyond face-to-face instructional settings. Research by Xu et al. (2019) shows that rule-based

technology helps adult learners improve the quality of their writing by integrating proper grammatical structure, and Sadiq (2016) further confirms that the integration of grammatical rules through technology-based systems yields positive results, yet a dedicated mobile application implementing Formula 33 in full remains absent from the literature. Second, most existing language learning applications rely on conventional Large Language Models (LLMs) to provide feedback and explanations; however, these models are prone to generating hallucinations, plausible but factually incorrect or fabricated information, which poses significant risks to the integrity of grammar instruction and student learning (Danyaro et al., 2025).

Technology integration in education continues to face challenges related to access gaps, pedagogical adaptation, and the psychological impact arising from dependence on technology (Husni et al., 2021). The increasing complexity of learning needs demands the development of educational technology that presents adaptive digital learning according to individual needs, yet the risk of inaccurate LLM output undermines trust and academic integrity. With the increasing role of technology in the learning process and the various challenges that accompany it, a structured and adaptive approach is needed that can be effectively integrated into the technology-based learning process. These gaps collectively highlight the need for a grammar learning solution that is not only mobile-based and pedagogically structured, but also capable of delivering reliable, consistent, and contextually grounded feedback to support English grammar mastery.

To address these limitations, recent advances in Retrieval-Augmented Generation (RAG) technology offer a promising and innovative solution. RAG Chatbot technology has emerged as an innovative approach in digital learning by combining real-time information retrieval capabilities, as noted by Swacha et al. (2025) with the generative ability of language models, providing more accurate and reliable responses to users. By functioning as an "open book system" that retrieves data from verified external sources, RAG ensures that the educational content generated is accurate and compatible with verified knowledge. Recent research demonstrates that RAG chatbots are highly effective in online learning environments, where they provide tailored conceptual explanations and examples based on specific course materials, resulting in high student satisfaction regarding the helpfulness and relevance of the explanations (Lang & Gürpınar, 2025). Furthermore, by leveraging external information, RAG can simplify complex reasoning tasks, allowing the model to address problems with higher complexity than standard LLMs (Gao et al., 2024). However, despite these advances, the integration of RAG with a structured pedagogical

framework, particularly Formula 33, for English grammar instruction within a mobile application has not yet been explored in prior research.

The novelty of this study, therefore, lies in the integration of Formula 33, Mobile-Assisted Language Learning (MALL), and a document-based Retrieval-Augmented Generation (RAG) chatbot within a single mobile application, enabling adaptive and consistent grammar feedback, an approach not yet explored in previous research. This approach combines constructivist learning theory with technology-based pedagogy, operationalizes Formula 33 in a scalable digital format, mitigates the hallucination risks of conventional LLMs, and supports learners with real-time, reliable, and personalized feedback. Therefore, the purpose of this study is to develop and evaluate a mobile-assisted English grammar learning application based on Formula 33 with an integrated RAG chatbot, with usability assessed through the System Usability Scale (SUS). This research is urgent given the increasing demand for reliable and adaptive language learning tools in the Indonesian educational context, where technology-enhanced English grammar instruction remains critically needed to support learners in achieving communicative competence.

Method

Research Design

This research adopts a prototyping development methodology, an iterative approach widely recognized in software engineering for its efficacy in developing software technology systems. This development model is particularly advantageous as it permits early functional models to be evaluated directly by end-users, facilitating refinement through a continuous iteration process and feedback loops (Carey, 1989). In the context of learning application development, the prototyping approach is considered highly effective because it accommodates the evolving nature of user requirements, specifically the pedagogical needs of learners, and mitigates the risk of misaligned system specifications during the final implementation phase.

Prototyping Method Stages

The research methodology used is prototyping, which enables developers to create early versions of language-learning applications and obtain feedback directly from users. Through user needs analysis, developers can design an appropriate application design and implement the necessary functionality. Furthermore, after the initial implementation, an evaluation is conducted to identify the application's strengths and weaknesses (Kortum & Sorber, 2015). This evaluation process, often involving end-users in the

testing of preliminary models, has been proven to significantly improve institutional visibility and user satisfaction by ensuring the final product accurately reflects the users' expectations and technical needs (Grieve, Braaten, MacPherson, Liu, & Jung, 2025).

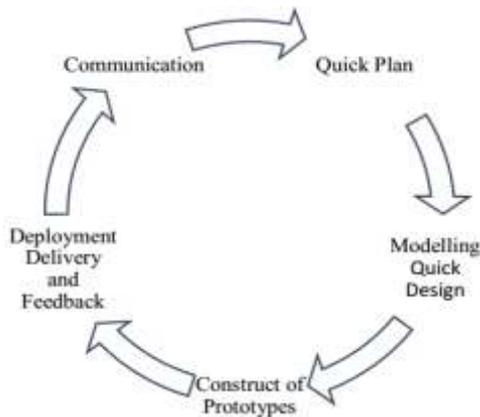


Figure 1. Prototyping method

Communication

The communication stage is conducted to identify system requirements through discussions with stakeholders, including English learning experts and target users. At this stage, the main problems in learning grammar are explored, along with the need for a structured digital learning flow based on Formula 33. Stakeholders provide input on essential features, including formula learning, practice questions, quizzes, and the integration of an AI-based chatbot for self-learning support.

This stage also confirms the requirement for an expert website to manage question banks, formulas, and user learning data. The outcomes of this communication process form the initial requirements list, which serves as the foundation for planning, quick design, and prototype development in the next stages.

Quick Plan

The planning stage is carried out to determine the scope at each stage of development, determine the priority scale of system needs, and set quality targets that need to be achieved. The initial stage of system development is focused on developing features related to the user's learning flow, starting from learning formulas, practice questions, and quizzes. Application development also focuses on improving the user interface and the implementation of AI-based chatbots.

Every piece of data used in the system will be stored in the Realtime Database on Firebase. The database will also be connected to the expert website as an admin system that functions in question bank management, formulas, and user score monitoring. System development, especially for applications, will be used by

Kotlin with Android Studio, while expert website development uses Bootstrap and JavaScript. Furthermore, each system's needs will be compiled into a list of needs organized in the development priority scale. This list of needs is a reference for the development of the Formula 33 English MALL system.

Modelling Quick Design

The quick design stage is a stage to translate the needs of the system into an initial design of the system that can be tested. This design is organized into artifacts consisting of a use case diagram, activity diagram, database design, chatbot workflow, and system wireframe interface. This wireframe design is used to visualize the layout of each component in a single feature page. The design of the wireframe user interface prioritizes updates to the interface display so that it is easier for users to navigate and, with the addition of the application usage instructions feature, it is expected to make it easier for users to use the application.

Meanwhile, in terms of system architecture, the Firebase Realtime Database Service is used as a database to store every component used in the system, such as question content, formulas, and user scores. This database is integrated with Android apps and expert websites as an admin system.

As a self-learning support tool, the system is designed to implement an RAG-based chatbot. Utilizing the Formula 33 book embedded into a vector, the chatbot is only limited to learning needs. All artefacts of the system design become guidelines for the next stage of development. Wireframes are a reference for the development of system interfaces, data schemas are used as a reference for database development, chatbot workflows are used as a reference for chatbot development and use cases, and activity diagrams are a reference for business flows and system usage.

Build Prototype

This stage is the realization of the previously prepared system design, which includes feature design, interface layout, and database structure. At this point, the English MALL application based on Formula 33 is developed using Android Studio with the native Kotlin programming language. The implementation focuses on core features – Learn Formula, Practice Questions, and Quizzes – all of which are integrated with the Firebase Realtime Database to ensure that question data, formulas, and quiz results are stored and retrieved seamlessly through the quiz score history module.

The chatbot component is implemented using the Gemini LLM, enhanced with a Retrieval-Augmented Generation (RAG) mechanism. The Formula 33 book is embedded as the primary knowledge source, enabling

the chatbot to deliver context-specific grammar explanations and provide adaptive learning support.

In parallel, a supporting expert website is developed using Bootstrap and JavaScript. The website includes essential features such as a dashboard, question bank management, formula management, and user score monitoring. Through the integration of the mobile application and the expert website, the system forms a structured learning ecosystem aligned with the principles of Formula 33 while enabling experts to manage learning content flexibly and efficiently.

Deployment Delivery and Feedback

The deployment stage is the process of implementing the system into a real environment so that it can be used directly by users. At this stage, the mobile application that has been built will be exported in the form of an APK file and installed on Android devices, while the expert website will be deployed through a web hosting service with a special authentication mechanism to maintain access security. The provision of these two platforms is aimed at ensuring that the integration between the user-used learning system and the expert-managed admin system runs according to plan.

Delivery is carried out by distributing applications to respondents who are target users, such as students and university students. Respondents will be asked to install applications, try out key features, and access expert websites according to the specified role. After deployment and delivery, the next stage is to collect feedback through a testing process. Feedback is obtained using two approaches, namely functional testing using the Black Box Testing method to verify the suitability of the input-output flow, and evaluation of the usability level using the System Usability Scale (SUS) method to

measure user experience of the application. These two mechanisms are used to assess whether the prototype meets the needs and provide a basis for improvement in subsequent development.

The prototype method is essential in the development of learning applications to ensure that challenging English language learning can be well and effectively integrated (Husmildin et al., 2020). The focus of the application development aspect with the prototyping method is the analysis, design, implementation, and evaluation that can be carried out in repeated iterations. This method allows developers to receive feedback directly from users and integrate changes and improvements gradually. With this approach, language learning apps can be designed to better meet the needs of users and provide a more effective learning experience. There are learning applications in Indonesia related to translation, such as Kamus Pro and Oxford Dictionary of English. In the Dictionary Pro app, users can easily search for translations of words or phrases in different languages and get additional information such as synonyms, antonyms, and examples of the use of those words.

System Architecture

The system architecture of the English Mobile Assisted Language Learning (MALL) application is designed to integrate three main components: the Android-based Formula 33 learning application, the expert website for content management, and a Retrieval Augmented Generation (RAG) chatbot for self-learning support. The architecture adopts a client-server model, where the mobile application and expert website function as clients connected to Firebase Realtime Database as the centralized backend service.

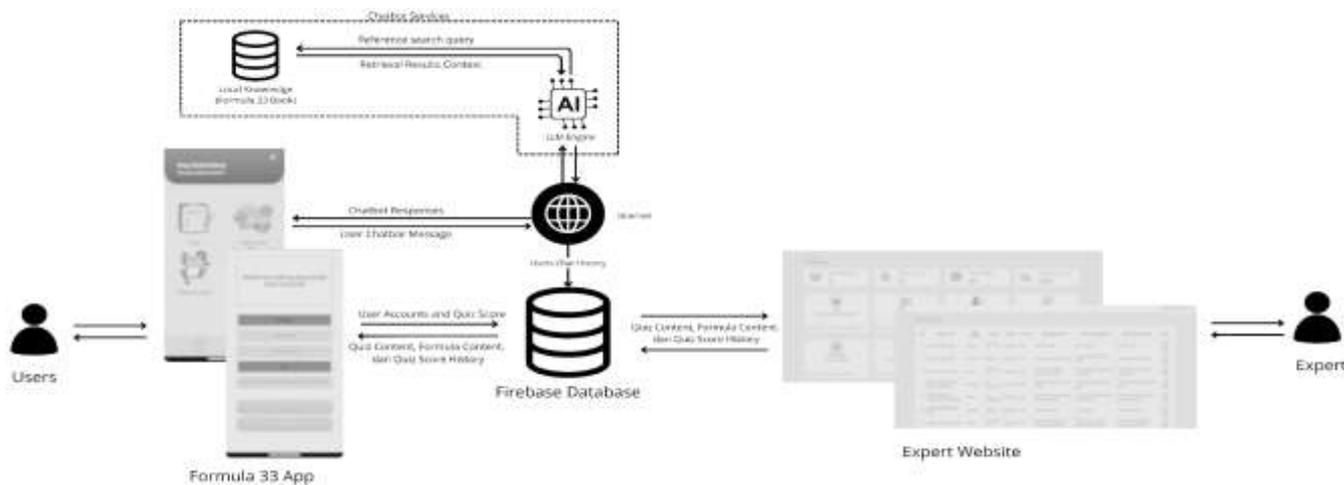


Figure 2. System architecture

The Formula 33 application is developed using Kotlin and Android Studio. It provides core learning

functionalities such as formula learning, practice exercises, quizzes, score history, and chatbot services for

self-learning. All learning data, including formula structures, quiz questions, and quiz results, are retrieved in real time from Firebase, enabling seamless synchronization across devices. For the expert website, it is developed with Bootstrap and JavaScript. This website serves as an administrative panel that allows English learning experts to manage formula content, question banks, and user assessment data. Both the expert website and mobile application access the same Firebase backend to ensure data consistency and maintain a unified learning ecosystem.

The system also incorporates a RAG-based chatbot module that supports self-learning for users. The Formula 33 textbook is converted into vector embeddings and stored in a local knowledge base in the application. When a query is submitted, the retriever component searches for the most relevant text segments in the vectored Formula 33 book, which are then passed to Gemini LLM to generate contextually accurate explanations. This knowledge from Formula 33 minimises hallucinations and ensures that responses remain aligned with the instructional content of Formula 33.

To complement these components, the system incorporates a chatbot developed with a Retrieval-Augmented Generation (RAG) framework. Adhering to the system workflow, the chatbot relies on the Formula 33 textbook as its primary knowledge corpus; this text is transmuted into vector embeddings and stored locally to facilitate efficient context retrieval. Upon receiving a user query, the Google Gemini LLM engine initiates a retrieval process using cosine similarity to identify the most semantically relevant segments within the embedded knowledge base.

This retrieved context is then synthesised with the user’s prompt, enabling the LLM to construct a structured response that strictly adheres to Formula 33 grammar protocols. Prior to being displayed, the output is processed by a response handler to guarantee coherence and readability. This architectural design effectively mitigates hallucination risks, ensuring the chatbot functions as a dependable instrument for independent learning.

Dataset & Knowledge Source for RAG

The Retrieval Augmented Generation (RAG) component in this research utilises a knowledge source derived exclusively from the Formula 33 textbook as the primary reference for grammar explanations and sentence analysis. The content of the book is first converted into digital text and then segmented into smaller, coherent units that represent specific rules, formula descriptions, and example sentences. After segmentation, each text unit is cleaned and normalised before being transformed into a numerical embedding vector using a text embedding model. All embeddings and their corresponding text segments are stored in a JSON-based vector repository that includes the raw text, vector representation, metadata, and a unique identifier. These vectors form a local knowledge base that stores the structured representations of all Formula 33 learning materials.

During chatbot interaction, user queries are embedded into vectors and compared against the stored knowledge base using cosine similarity to identify the most relevant instructional segments. The retrieved text chunks are then supplied as contextual input to the generative model, allowing the chatbot to produce responses that remain aligned with the Formula 33 methodology. Through this approach, the system maintains accuracy, prevents model hallucination and deviation from the intended curriculum, and ensures that all explanations strictly reflect the rules and learning structure defined in the textbook.

Evaluation Methods

The system evaluation utilized a dual approach, incorporating both functional and usability assessments to ensure comprehensive validation. Functional verification was conducted via black box testing, which examined whether the system’s components aligned with specific requirements and produced expected input-output behaviors without inspecting the underlying source code. This process covers all features of the mobile application and expert website, such as education content management, quiz features, practice modules, and chatbot interaction. By utilizing test cases that mimicked user interactions with the system, the testing assessed the accuracy of system responses (both

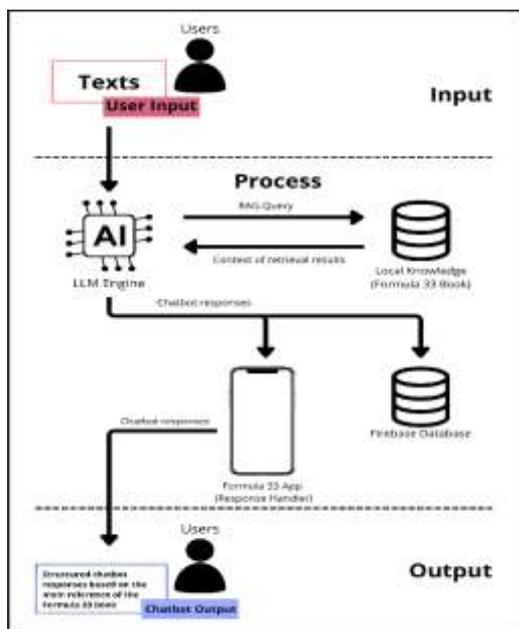


Figure 3. Chatbot workflow

positive and negative responses), error management, and interface workflow, ultimately confirming that every module functioned reliably according to operational goals.

Complementing the functional tests, a usability evaluation was performed using the System Usability Scale (SUS) to gauge user perception regarding the application's clarity and ease of use. The SUS method utilizes a ten-item questionnaire with a Likert scale to generate a standardized usability score. During this phase, participants installed the application, navigated through key features, completed a quiz, and subsequently rated their experience. The resulting data provided a quantitative metric of the system's intuitiveness, efficiency, and acceptability, offering insights into user satisfaction that technical validation alone could not uncover.

Result and Discussion

Sentence Translation Process Using Formula 33

The development of a grammar learning application model begins with an identification process. The results of the identification provide an overview of the learning model to be built. In general, the process of translating sentences from Indonesian to English follows certain rules. To translate a sentence from Indonesian to English, the first step is to understand the meaning of the sentence. This includes identifying the type or character of the sentence, then determining the time aspect of the sentence. After these steps, the appropriate translation formula is determined.

The process of translating Indonesian sentences into English requires a systematic and structured approach to ensure that the resulting translation accurately reflects the original meaning. This process involves several carefully sequenced stages: (1) understanding the meaning of sentences within their context; (2) identifying the type of sentence, whether it contains verbs, nouns, or passive forms; (3) identifying the time aspect, whether present, past, or future; (4) identifying the grammatical aspect, including ongoing, completed, future-oriented, or empty aspects; (5) selecting the appropriate sentence rules; (6) double-checking to ensure accuracy; and (7) choosing the most appropriate translation. At each of these stages, the Formula 33 method can be systematically applied, as it has been proven effective in improving students' ability to construct grammatically accurate English sentences. By following this structured sequence, learners are guided not only toward grammatical precision but also toward translations that clearly and accurately convey the intended meaning of the original text.

The testing results show that users applied the Formula 33 stages with greater consistency than

conventional translation patterns, indicated by a decrease in structural errors in the application's early exercises. This suggests that the digitalization of the Formula 33 analysis flow helps users understand grammar patterns in a more directed way.

System Development Stages

The development stages of the English MALL application with the Formula 33 method were carried out with 3 iterations of gradual development that adopted a prototyping approach (Isal et al., 2021). This approach is chosen so that each stage of development can be tested, evaluated, and improved according to the needs of the system and the user. Each iteration carried out brings improvements both in terms of functionality and user experience.

The initial stages of development are focused on the implementation of the core features of the application, namely quizzes, practice questions, and learning formulas, which are commonly used as effective components in Mobile-Assisted Language Learning applications (Chaib et al., 2023; Jabbar, 2024). This initial development stage is also the beginning of development for each initial interface appearance of the application and the integration of the application with real-time databases, as well as user registration and authentication mechanisms. The development goal of this first iteration is to ensure the app can be used for learning flows based on the Formula 33 method. The first iteration resulted in a high-fidelity prototype that already functions according to the basic requirements of the system.

Based on the results of testing and evaluation in the initial prototype, the second iteration is directed to develop supporting features in the application to improve the user experience. Features developed in the second iteration include a tutorial page as a guide to using the app for users, a feedback quiz page to display the results and evaluation of user quiz results, as well as improvements to the user interface design. The feedback given to users is personalized based on the results of the quiz so that users can receive a clearer evaluation of learning outcomes. With these supporting features, the app becomes more user-friendly for new users and can provide an intuitive learning experience.

Furthermore, the development of the chatbot is carried out by utilising Gemini LLM to answer user questions (Wiboolyasarini et al., 2025). Chatbots are designed as a means of self-learning for users; the answers from chatbots come from the main reference, the Formula 33 book, and chatbots are only used in the context of English learning.

The last stage of development is focused on the development of a support system in the form of an expert website. This website functions as an application content management platform, including quizzes and

practice questions as well as formula validation, as well as statistical monitoring of user scores. The integration of expert websites with mobile applications ensures that the questions and formulas used as learning content by users are in accordance with the rules in the Formula 33 concept. In addition to presenting a mobile application to support user learning, the system also provides a means for experts to manage and ensure that the learning content used is appropriate.

Based on iterative testing throughout each prototyping cycle, it was found that usability improved significantly in the second and third iterations, particularly following the addition of quiz feedback features and interface refinements. This pattern suggests that support features, such as tutorials and personalised feedback, play a crucial role in enhancing the effectiveness of application usage. The integration of an expert website within the development workflow enables continuous content quality control, and testing demonstrated that updates to formulas and question banks could be immediately reflected in the application without system disruption. These findings reinforce the function of the Formula 33 learning ecosystem as an adaptive and easily updatable model.

System Implementation Results

The Formula 33 mobile application was successfully developed with Kotlin on Android Studio by integrating the Firebase Realtime Database as the system's database. Implementation of the main features in the application, which include a learning formula page, a question practice pages and a quiz page. The results of the implementation show that all features are functionally running well.



Figure 4. Formula 33 quiz features

The interface has been successfully updated to make it more intuitive; this is done based on user needs through surveys from previous versions of the application. In addition, in the development of this latest version of the application, an application instructions page was also added to help users understand the flow of use of each application feature. Also, a personalised feedback feature from quiz results was added, to provide evaluation of quiz results and user understanding.

One of the innovations in the development of the second version of the Formula 33 application is the implementation of an AI-based English learning chatbot with a Retrieval Augmented Generation (RAG) approach. Chatbots are designed as a means of self-learning for users; the answers from chatbots come from the main reference, the Formula 33 book, and chatbots are only used in the context of English learning.

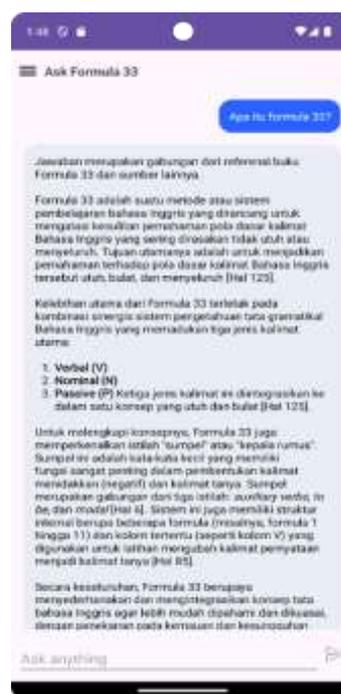


Figure 5. Formula 33 chatbot features

This implementation allows users to study the material, ask questions about sentence structure, or understand the sentence formation rules contextually from the Formula 33 book. The results of the test showed that the chatbot was able to provide relevant responses in accordance with the Formula 33 material without causing problems.

The use of RAG-based chatbots also demonstrated a positive contribution, where users frequently asked questions regarding tenses and sentence structure. This indicates that the chatbot serves not merely as a supplementary feature but as a key element in supporting the self-directed learning process.

The expert website is developed using Bootstrap and JavaScript with a direct connection to the same Firebase database as the mobile app. Its main features include a system dashboard, question input, question bank management, formula input, formula management, and user value statistics. This website can only be accessed by users with special credentials (experts).



Figure 6. Formula management features

The results of the implementation show that each feature can function properly, such as adding, editing, and deleting question data and formulas used in the mobile application. The dashboard also displays statistical data on user scores in the form of tables and graphs.

Evaluation

Evaluation of the system is carried out to ensure that the Formula 33 application that has been developed and delivered to the user is running according to the user's needs and specifications. The evaluation process is carried out using two approaches, namely Black Box Testing to test system functionality and System Usability Scale (SUS) to measure system usability from the user side.

Black box testing was carried out on all features in the Formula 33 application, starting from user registration, login, quizzes, practice questions, and learning formulas. Meanwhile, on the expert website, testing is also carried out on the login features, question input, formula input, question bank management, formula management, and user value statistics. Black box testing focuses on the input output flow without interacting directly with the system's source code. Every user interaction with the system is done as a test scenario. This scenario is used to ensure that every user interaction with the system can proceed smoothly. The following in table 1, table 2, and table 3 are the results of testing on some of the main features of the application and website of the Formula 33 system.

Table 1. Black Box Testing Result for Quiz Features

Scenario	Test Result
Users can see the information page and instructions for working on the quiz before starting to work on the quiz questions.	As Expected
Users can start a quiz.	As Expected
Users can choose answers for sentence type points.	As Expected
Users can choose answers for Aspect points.	As Expected
The user can choose the answer for the Time point.	As Expected
Users can choose the answer to the Formula points.	As Expected
Users can choose the answer to the Translation point.	As Expected
Users see information on the question items that have been done.	As Expected
The user successfully answers one question and proceeds to the next question.	As Expected
The user successfully answered all the questions in the quiz.	As Expected
Users can see the information page and instructions for working on the quiz before starting to work on the quiz questions.	As Expected

Table 2. Black Box Testing Result from Formula Management Features

Scenario	Test Result
Experts can see a list of all formulas.	As Expected
Formula information is displayed complete.	As Expected
Experts can edit the content of the formula.	As Expected
Experts can remove formulas	As Expected
Experts can perform a formula search with the search bar.	As Expected

The results of the tests carried out show that every feature of the application runs according to the specifications. The input-output process for each feature runs properly, navigation between pages runs

consistently, and the integration between applications and websites with the database also runs well. These results show that the system has met every planned functional requirement.

Testing is also carried out on the user experience in using the application to measure the usability of the application from the user's perspective. Evaluation to assess the usability of this application was carried out using the System Usability System (SUS) method. A total of 32 respondents, consisting of English students and the public, were involved in trying each feature of the Formula 33 application. After that, respondents were asked to fill out a questionnaire containing 10 both negative and positive statements related to the use of the application and user experience. The test scoring results showed an average score of 73.3, which was included in the grade of B, with the adjective rating "Good" and the level of acceptability at the level of "Acceptable". This indicates that the Formula 33 app has a high level of usability, the content and learning flow in the app are easy to understand, and it provides good user learning experience.

Based on the results of the evaluation, it can be concluded that the developed system has functioned well in technical and functional terms and has a high usability value for users. The black box test showed that all features on the system were running properly and there were no functional errors. In the SUS test, it shows positive results from users with high average scores. From these results, the Formula 33 application can be said to be suitable for use as an English learning medium with the Formula 33 method.

Interpretation and Discussion

The results demonstrate strong functional performance and high system usability. The integration of the RAG-based chatbot significantly enhances learner autonomy by providing contextual grammar explanations that remain consistent with the Formula 33 framework. Unlike conventional MALL applications that rely on general-purpose responses, this system offers more structured and controlled guidance, ensuring that learners receive accurate explanations grounded in validated material. These improvements are further supported by the redesigned interface, which allows users to navigate learning features more easily and contributes to a smoother overall learning experience. The system's ability to manage content dynamically through the expert website also strengthens its adaptability, enabling continuous updates without modifying the application itself. Overall, the findings indicate that the combined use of Formula 33, MALL, and RAG chatbot technology creates effective digital learning environment.

Strengths and Implications

The strengths of the developed system lie in its ability to integrate Formula 33, MALL, and a RAG-based chatbot into a unified learning environment. This

combination offers a novel approach that delivers structured and consistent grammar guidance, as the chatbot retrieves explanations directly from a validated source. Such consistency is important in avoiding misinformation and ensuring that learners receive reliable feedback aligned with the Formula 33 rules. The presence of an expert website also strengthens the system, allowing dynamic management of learning materials, question banks, and user progress without requiring changes to the mobile application. This flexibility supports continuous content updates by experts. In addition, the improved interface design enhances usability and supports a more intuitive learning experience, enabling users to navigate learning features with ease. These strengths imply that the system can serve as an effective digital tool for grammar learning and has the potential to support broader technology-enhanced language learning initiatives.

Limitations and Opportunities

Several limitations were identified in this study. The chatbot is limited to the Formula 33 content embedded in the system, which ensures consistency but restricts its ability to provide explanations beyond the available material. The application is also currently available only on Android, reducing accessibility for users on other platforms. In addition, the number of respondents involved in the usability evaluation is still limited, so broader testing is needed to strengthen the findings. Another limitation is that long-term learning outcomes were not measured, meaning the study has not yet examined whether the system effectively improves grammar proficiency over time.

Future development may include expanding the chatbot's knowledge base beyond Formula 33, integrating voice-based interaction for speaking support, and developing versions for iOS and web platforms to increase accessibility. Longitudinal studies are also needed to evaluate long-term improvements in learners' understanding and the sustained effectiveness of the Formula 33 learning system.

Conclusion

Based on the results of functional tests, the Formula 33 system has met all functional needs. Functional testing was carried out with black box testing on applications and websites, showing results that the entire system workflow in input/output was running according to the specifications and that no errors or bugs were found that could hinder the functionality of the application. In terms of application usability, the test was carried out using SUS (System Usability Scale) with 32 respondents who tried the application directly. From the results of this test, an average score of 73.3 was

obtained, which was included in grade B, with the adjective rating "Good", and the acceptability level was at the level of "Acceptable". The results of this functional and usability evaluation show that the learning content, usage flow, and application interface are easy to use and can provide a good learning experience. The developed model successfully integrates various technological and pedagogical elements into a comprehensive English learning ecosystem. Through the incorporation of gamification features, the application enhances user engagement and motivation while making the learning process more enjoyable and goal-oriented. Experts can easily input and update materials and question banks through an integrated expert system, ensuring that content remains accurate, relevant, and adaptive to learners' needs. Simultaneously, users benefit from progress-tracking features that allow them to monitor their independent learning achievements in real time, thus fostering autonomy and self-regulated learning. A key innovation of this model is the inclusion of a document-based chatbot that supports contextualized learning by referencing the Formula 33 materials within interactive discussions. This feature enables users to study grammar structures and examples that are directly related to the learning documents, promoting deeper comprehension and practical application of the rules. Altogether, these combined features demonstrate that the Formula 33-based model is not merely a digital adaptation of traditional learning but a transformative platform that bridges expert knowledge, learner interactivity, and intelligent content delivery in a unified and scalable system.

Acknowledgments

The authors would like to thank the Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemendikti Saintek) for funding support through the Publication Assistance Program in Reputable Journals, and Universitas Islam Indonesia for supporting this interdisciplinary collaboration.

Author Contributions

Conceptualization, S.M., A.H. and N.S.; methodology, S.M. and N.S.; software, G.I.N.; validation, N.S. and A.H.; formal analysis, S.M.; investigation, N.S. and A.H.; data curation, G.I.N.; application concept, S.M.; application review, N.S.; writing original draft, S.M., G.I.N. and A.H.; writing review and editing, A.H. and S.M. All authors have read and agreed to the final version of the manuscript.

Funding

This research was funded by the Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemendikti Saintek) through the Publication Assistance Program in Reputable Journals.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Abidin, S. R. Z., Noor, S. F. M., & Ashaari, N. S. (2017). Guidelines of brain-based learning through serious game for slow reader students. *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*, 1–6. IEEE. <https://doi.org/10.1109/ICEEI.2017.8312461>
- Ahmadi, M. R. (2018). The Use of Technology in English Language Learning: A Literature Review. *International Journal of Research in English Education*, 3(2), 115–125. <https://doi.org/10.29252/ijree.3.2.115>
- Badia, A. (2015). Research trends in technology-enhanced learning/Tendencias de la investigación en el aprendizaje favorecido por la tecnología. *Infancia y Aprendizaje*, 38(2), 253–278. <https://doi.org/10.1080/02103702.2015.1016744>
- Chaib, S. O., Joti, I., & Khouliji, S. (2023). Gamification in mobile language learning applications: A case study examining a mobile learning tool in a Moroccan university context. In *E-learning and smart engineering systems (ELSES 2023)* (pp. 177–185). Atlantis Press. https://doi.org/10.2991/978-94-6463-360-3_19
- Crompton, H., & Burke, D. (2018). The use of mobile learning in higher education: A systematic review. *Computers & Education*, 123(2), 53–64. <https://doi.org/10.1016/j.compedu.2018.04.007>
- Danyaro, K. U., Abdullahi, S., Abdallah, A. S., & Chiroma, H. (2025). Hallucinations in Large Language Models for Education: Challenges and Mitigation. *International Journal of Teaching, Learning and Education*, 4(6), 13–19. <https://doi.org/10.22161/ijtle.4.6.2>
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... Wang, H. (2024). Retrieval-Augmented Generation for Large Language Models: A Survey. *Proceedings - 2024 Conference on AI, Science, Engineering, and Technology, AIxSET 2024*, 166–169. Retrieved from <http://arxiv.org/abs/2312.10997>
- Grieve, N., Braaten, K., MacPherson, M., Liu, S., & Jung, M. E. (2025). Involving End Users in the Development and Usability Testing of a Smartphone App Designed for Individuals With Prediabetes: Mixed-Methods Focus Group Study. *JMIR Formative Research*, 9(1), e59386. <https://doi.org/10.2196/59386>
- Husmildin, R. S. P. P., Rosyid, H. A., & Elmunsyah, H. (2020, September). Developing gamified mobile learning to increase student motivation to learn WAN technology for vocational high school

- students. In *2020 4th International Conference on Vocational Education and Training (ICOVET)* (pp. 218-222). IEEE. <https://doi.org/10.1109/ICOVET50258.2020.9229957>
- Husni, M., Sumardjoko, B., & Darsinah, D. (2021). Computer Game-Based Learning Media as a Strategy to Increase Students' Motivation Learning. *At-Taqaddum*, 13(1), 85-102. <https://doi.org/10.21580/at.v13i1.8337>
- Isal, R. Y. K., Santoso, H. B., & Novandi, E. R. (2021). Development and Evaluation of a Mobile-Learning Application Based on the Felder-Silverman Learning Styles Model. *International Journal of Emerging Technologies in Learning (IJET)*, 16(15), 107. <https://doi.org/10.3991/ijet.v16i15.24165>
- Jabbar, S. A. (2024). Mobile-Assisted Language Learning (MALL) in English Education: Trends and Developments. *International Journal of English Literature and Social Sciences*, 9(6), 113-116. <https://doi.org/10.22161/ijels.96.20>
- Kortum, P., & Sorber, M. (2015). Measuring the Usability of Mobile Applications for Phones and Tablets. *International Journal of Human-Computer Interaction*, 31(8), 518-529. <https://doi.org/10.1080/10447318.2015.1064658>
- Lang, G., & Gürpınar, T. (2025). AI-Powered Learning Support: A Study of Retrieval-Augmented Generation (RAG) Chatbot Effectiveness in an Online Course. *Information Systems Education Journal*, 23(2). <https://doi.org/10.62273/ZKLLK5988>
- Maulidiyah, A., & Garnida, S. C. (2021). The Up-Ranking Of The Source Language Structure In The Translation Of Brown's Principles Of Language Learning And Teaching, Fifth Edition. *Parafrese: Jurnal Kajian Kebahasaan & Kesastraan*, 21(1). <https://doi.org/10.30996/parafrese.v21i1.5399>
- Mohammed, M. (2020). The Effectiveness of Instructional Technology on English Language Learning in EFL Classroom. *International Journal of Linguistics*, 12(1), 13. <https://doi.org/10.5296/ijl.v12i1.16205>
- Ozer, O., & Kiliç, F. (2018). The Effect of Mobile-Assisted Language Learning Environment on EFL Students' Academic Achievement, Cognitive Load and Acceptance of Mobile Learning Tools. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(7), 2915-2928. <https://doi.org/10.29333/ejmste/90992>
- Refat, N., Kassim, H., Rahman, M. A., & Razali, R. (2020). Measuring student motivation on the use of a mobile assisted grammar learning tool. *Plos One*, 15(8), e0236862. <https://doi.org/10.1371/journal.pone.0236862>
- Sabron, M. Z. M., Hashim, R., Abdullah, A., & Shamsudin, N. (2020). Humanizing Technology: Smartphone usage in the learning environment. *Environment-Behaviour Proceedings Journal*, 5(13), 107. <https://doi.org/10.21834/e-bpj.v5i13.1958>
- Sadiq, N. (2015). Measurement Of The Students' Attitude Toward A Whole-Languageapproach-Based Lecture Using Score Questionnaires. *Jurnal Cakrawala Pendidikan*, 2(2). <https://doi.org/10.21831/cp.v2i2.4825>
- Sadiq, N. (2016). Teaching English To Non-English Students' Department At Islamic University Of Indonesia Through Formula 33: Students' Reflections. *Journal of English and Education*, 6(2), 84-102. <https://doi.org/10.20885/jee.vol6.iss2.art8>
- Shadiey, R., Liu, T., & Hwang, W. Y. (2020). Review of research on mobile-assisted language learning in familiar, authentic environments. *British Journal of Educational Technology*, 51(3), 709-720. <https://doi.org/10.1111/bjet.12839>
- Swacha, J., & Gracel, M. (2025). Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications. *Applied Sciences*, 15(8), 4234. <https://doi.org/10.3390/app15084234>
- Wiboolyasarın, W., Wiboolyasarın, K., Tiranant, P., Jinowat, N., & Boonyakitanton, P. (2025). AI-driven chatbots in second language education: A systematic review of their efficacy and pedagogical implications. *Ampersand*, 14, 100224. <https://doi.org/10.1016/j.amper.2025.100224>
- Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., & Wijekumar, K. (Kay). (2019). The effectiveness of educational technology applications on adult English language learners' writing quality: a meta-analysis. *Computer Assisted Language Learning*, 32(1-2), 132-162. <https://doi.org/10.1080/09588221.2018.1501069>