

TABLE OF CONTENT

Anita Christina Anthony Pre-Service ESL Teachers' Mindsets and their Approaches towards Assessments https://doi.org/10.24191/mvahvg19	1–21
Maryam Muidh Alsulami, Sharifah Shahnaz Syed Husain, Aini Faridah Azizul Hassan Saudi EFL Learners' Attitudes Towards Arabic-English Code-Switching https://doi.org/10.24191/yjvd8a60	22–37
Julazzmie Kambutong, Nur Farha Shaafi Content Validity of a Survey on Knowledge, Skills, and Readiness for Generative AI https://doi.org/10.24191/xnvcmw23	38–57
Aini Syahira Jamaluddin, Nabila Huda Nasir, Nur Azyyati Mohamad Sa'edin, Ummu Habibah Mohd Sakri @ Shukri Can Artificial Intelligence Replace Teachers? Perspectives from Pre-Service Teachers in Elementary Education https://doi.org/10.24191/vjk6sz19	58–73
Xia Yutong, Talaibek Musaev Grammar of Obligation: Syntactic Patterns in Malaysian Reddit AIED Discourse https://doi.org/10.24191/rs3g1v95	74–97
Mohd Hafiz Mat Adam, Geethanjali Narayanan, Juriani Jamaludin, Rafidah Abd Karim Fostering Independent Expository Essay Writing Development Using Song-Assisted Learning (SAL) to Avert AI Over-Reliance https://doi.org/10.24191/nhh8d038	98–112
Farhana Atiqah Ahmad Radzuan, Marina Mohd Arif Teacher Perspectives on ChatGPT Integrated Vocabulary Instruction in a Flipped ESL Classroom https://doi.org/10.24191/9deysg59	113–135
Damia Syafiqah Mohd Salleh, Norlina Mohd Sabri, Fazlin Marini Hussain SpeakEasy Nihongo: Mobile Application for Japanese Beginner Learners https://doi.org/10.24191/1k11n645	136–152



SpeakEasy Nihongo: Mobile Application for Japanese Beginner Learners

Damia Syafiqah Mohd Salleh

damiasyafiqah0@gmail.com

Faculty of Computer and Mathematical Sciences

Universiti Teknologi MARA Cawangan Terengganu, Malaysia

Norlina Mohd Sabri*

norli097@uitm.edu.my

Faculty of Computer and Mathematical Sciences

Universiti Teknologi MARA Cawangan Terengganu, Malaysia

Fazlin Marini Hussain

fazlinmarini@uitm.edu.my

Faculty of Computer and Mathematical Sciences

Universiti Teknologi MARA Cawangan Terengganu, Malaysia

Corresponding author*

Received: 15 January 2026

Accepted: 29 January 2026

Published: 25 May 2026

CITE THIS ARTICLE:

Damia, S. M. S., Norlina, M. S., & Fazlin, M. H. (2026). SpeakEasy Nihongo: Mobile application for Japanese beginner learners. *Journal of Creative Practices in Language Learning and Teaching*, 14(1), 136–152. <https://doi.org/10.24191/1k11n645>

ABSTRACT

Studying a new language may bring in a lot of surprises and new opportunities, yet it also entails a number of difficulties. To achieve fluency, most learners do not only need to memorise new vocabularies and learn grammar. Words are not everything but sounds and rhythms are a part of the language that needs to be mastered to communicate. Nonetheless, most learners have problems in enhancing their language abilities particularly in the aspects that extend to reading and writing. In this project, the proposed development involves a mobile application to learn Japanese language named SpeakEasy Nihongo. The primary goal of the application is to assist the novice learners to learn and enhance their pronunciation skills in Japanese with the help of interactive quizzes and lessons with the help of the mobile platform. This app stands out of the crowd of other applications of language because it heavily focuses on the feedback and



engagement received through the combination of Artificial Intelligence, speech recognition, and gamification features. The app enables the users to train Hiragana and Katakana, as well as number and other simple words. The app utilizes the Speech API of the Microsoft Azure to grade pronunciation of users and give real-time review thus allowing the learners to diagnose the errors and rectify them easily. This study adopted development-based approach that included requirement analysis, interface design, system development with Flutter and user testing. The interface and app features were measured with the help of a User Experience Questionnaire (UEQ). The findings indicated positive feedback on all six UEQ scales, particularly on such areas as Perspicuity and Stimulation. In accordance with these results, the suggested areas of improvement might be the inclusion of additional lessons in the future and the development of more chatbot features and enhanced voice input to more natural learning.

Keywords: Japanese pronunciation, mobile application, speech recognition, Microsoft Azure, gamification

INTRODUCTION

One of the most common challenges in language learning is pronunciation. Unlike grammar or vocabulary, pronunciation is often harder to practice, and learners may not get immediate feedback on how well they are doing. Many traditional language programs focus on vocabulary and grammar, leaving pronunciation training as an afterthought (Pennington, 2021). However, pronunciation is essential for clear and effective communication. In languages like Japanese, where even a slight change in sound can alter a word's meaning, accurate pronunciation becomes even more important. Learning Japanese as a second language can be challenging, particularly for beginners who are not familiar with its unique sounds and writing systems (Chaib et al., 2023). Unlike English, Japanese includes phonemes like “tsu” and “fu” which are difficult to replicate without prior exposure or proper guidance (Ueyama, 2020). Most learners begin with vocabulary memorization and grammar drills, while pronunciation is either introduced too late or practiced without proper feedback (Rethinasamy et al., 2021; Amalia, 2020). As a result, beginners often develop poor pronunciation habits that are hard to fix later in their learning journey (Amalia, 2020).

Another common issue in traditional language classes and standard apps is the lack of real-time correction (Qin, 2024). Learners are expected to repeat words without knowing whether their pronunciation is correct or not (Olson & Offerman, 2021). Without feedback, they lose confidence and become hesitant to speak. Moreover, traditional classroom settings can cause anxiety, especially when learners are afraid of being judged by others during oral practice (Qin, 2024). A solution that allows users to self-correct in a private, supportive environment could significantly improve their experience and progress (Brata & Brata, 2020).

As the mobile technology emerges, opportunities to learners to practice pronunciation more effectively increase. It is now possible to get instant feedback through voice recognition using mobile apps and learners can view and hear their own mistakes in real-time. The conventional language learning methods can leave students with a challenge to re-assess themselves in terms of their pronunciation abilities or make them spend hours before they receive feedback. With the



incorporation of technology in the learning curve, students are now able to improve their pronunciation skills on a consistent basis and locate themselves anywhere in the world. Mobile-Assisted Language Learning (MALL) has proved to be one of the most productive aids in most language learning mentoring, particularly in lessons that target pronunciation elimination difficulties. By adopting this approach, the platform can offer a balanced learning experience that builds speaking and listening skills while boosting learners' confidence. It can also support educators in creating engaging lessons that cater to diverse learning styles (Hoxhaj, 2024).

MALL uses mobile communication devices in learning, which encourages mobility and fast learning. The MALL techniques are getting more acceptance in the context of language teaching today since it is aimed at improving the language knowledge especially speaking and pronunciation. Amalia (2020) states that students taught by MALL methods on pronunciation exercises achieve higher results in comparison to students taught more traditional processes. Learners in MALL can set the rate of their learning process, get access to material at their own time and practice pronunciation at a time that fits them. This flexibility becomes quite important in perfecting the articulation abilities of the learners especially to those handling the non-native students. The creation of SpeakEasy Nihongo was based on this requirement, whereby the beginner learners are guided through the Japanese sounds and an opportunity in using audio, repetition, and scoring is offered to them. In a new language, even the smallest disparity in the pronunciation attainable can alter the whole meaning of the word. It is therefore necessary at early stages to pronounce words correctly. To address these issues, a mobile application that has voice feedback appears to be an advantageous solution that allows correcting oneself and remaining motivated.

The objectives of this project are:

1. To design and develop a mobile app that supports Japanese pronunciation learning for beginners using voice recognition technology.
2. To evaluate the usability of the mobile application.

Literature Review

The following review explores the algorithm and technology applied in the app, along with related studies that support its development and functionality.

Algorithm

The project uses the Fisher-Yates Shuffle algorithm to randomize quiz questions, ensuring varied and non-repetitive sequences during each practice session (Kirana et al., 2021). This algorithm works by swapping each item in a list with another randomly chosen item that appears later, creating a fair and unpredictable order. This feature plays a significant role in enhancing the learning process, as it prevents users from memorizing question orders and encourages them to focus on the content of each quiz.

Fisher-Yates Shuffle increases the app's replay value by ensuring users do not see the same sequence of questions twice. This makes each quiz feel fresh and encourages learners to practice

more. It also supports gamified learning by adding unpredictability and variety, which helps keep learners engaged (Klock et al., 2020; Smiderle et al., 2020). Using a lightweight algorithm like Fisher-Yates is particularly suitable for mobile applications where performance and memory efficiency are important (Kirana et al., 2021).

Technology for Pronunciation – Azure Speech API

The project utilizes Microsoft Azure’s Speech-to-Text API, particularly the Pronunciation Assessment feature, to evaluate spoken Japanese input (Ferraro et al., 2023). This API analyzes speech and returns word-level accuracy, fluency, and completeness scores, which are important for assessing beginner pronunciation skills (Olson & Offerman, 2021). Azure works smoothly on different devices and lets developers easily include pronunciation scoring in mobile apps. Many learning platforms choose Azure because of its user-friendly tools and clear documentation (Ferraro et al., 2023).

Azure provides more specific feedback compared to other speech recognition tools like Google’s or Apple’s. It includes phoneme-level accuracy and segmental correctness, which makes it easier for beginners to understand where they need to improve (Pennington, 2021). Because Azure gives instant feedback, users can quickly fix mistakes. This saves time and helps learners avoid repeating incorrect pronunciation.

Similar Works

Table 1 below shows several applications for pronunciation learning. The table shows the applications, objectives and the key features for the apps.

Table 1. Table of Similar Work

Application	Objective	Key Features / Strengths
Mondly (Nushi et al., 2024)	Evaluate suitability for EFL learners	Follows Reinders & Pegrum framework; Potential EFL learning benefits
Falou (Saragih et al., 2024)	Improve students’ English speaking skills	Boosts grammar and confidence; user-friendly interface
Duolingo (Anderies et al., 2023)	Analyze user experience using UEQ	Gamified UX; widely used globally
Rosetta Stone (Nst et al., 2023)	Support speaking imitation for future educators	Increased engagement and motivation; improved speaking skills
Miryoku (Wibawa et al., 2023)	Support Japanese learning via speech recognition & gamification	Uses Android Speech API; beginner-friendly gamified design
ELSA Speak (Kholis, 2021)	Improve pronunciation with ASR	Real-time speech feedback; supports accurate pronunciation
FluentU (Altynbekova & Zhussupova, 2020)	Develop public speaking through mobile learning	Video-based learning; contextual vocabulary exposure

ELSA Speak, for example, is a well-known English language app that uses speech recognition to provide real-time pronunciation feedback (Kholis, 2021). It allows users to see where their



pronunciation falls short and gives suggestions to improve. However, its focus is on English, and it lacks gamification elements that could keep users engaged over time.

Duolingo is another popular language learning app that includes Japanese as one of its languages (Anderies et al., 2023). It uses gamification extensively, with features like streaks, leaderboards, and badges. While it is engaging, its pronunciation practice is limited. Users are often asked to speak into the mic, but there's no scoring system that gives precise feedback on how accurate the pronunciation was. It's more focused on vocabulary and grammar than speaking skills.

Another app called *Miryoku* was developed specifically for Japanese beginners. It uses Android's speech recognition API to help users practice basic Japanese words and receive audio feedback (Wibawa et al., 2023). However, the app has a simpler interface, and the feedback system is not detailed enough to help learners improve pronunciation at a deeper level. Also, the gamification elements are quite limited.

SpeakEasy Nihongo is a mix of these applications, but the best part is that it incorporates their strengths. It scores using the Pronunciation Assessment API of Microsoft Azure and provides feedback, but the interface is user-friendly and simple to use. To enhance the motivation of users, the app also has gamified rewards, Fisher-Yates randomized quizzes, and content related to the lesson. Its attention to the correct pronunciation along with its interesting learning aspects make it more appropriate among the beginners who want to see the actual improvement in speaking the Japanese language.

METHODOLOGY

Based on the modified Waterfall Model exemplified in Figure 1, this project was created and divided into five major steps: analysis, design, development, testing and documentation.

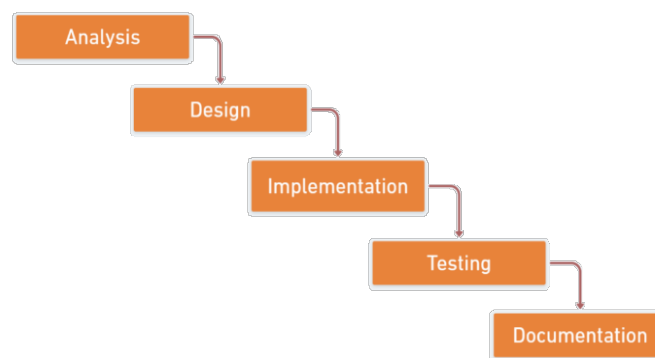


Figure 1. Adapted Waterfall Model

Software development begins with the analysis phase during which project goals and issues are stated, and options are sought. In this project's case, the analysis was geared towards the challenges in learning Japanese such as pronunciation feedback as well as user engagement. A literature review was conducted to try and address these issues while looking at the existing



research relating to motivation and retention. Beginner voice recognition education and gamification of the Japanese language were looked into, to create more engagement. During this phase, problems with current pronunciation learning apps were identified, and requirements for beginner-friendly pronunciation practice were gathered. Azure Speech Services was chosen after testing several APIs for voice recognition.

In the design phase, the user interface was designed using Figma, and each app screen was mapped out to cover lessons, quizzes, and pronunciation scoring pages. Figure 2 below shows interfaces of Figma design in the form of wireframe to clearly visualize idea of navigation in within the application for implementation stage. In this phase, the materials for the application are designed to enhance pronunciation practice and engagement for beginner learners. The materials for the app include audio files that consist of:

- a. **Characters:** 46 Hiragana and 46 Katakana characters, each with an audio file for accurate pronunciation.
- b. **Words/Phrases:** Approximately 1000-1500 common words and phrases for beginners, with audio files to help learners listen to the correct pronunciation before attempting to speak.

The primary purpose of these materials is to allow learners to listen first, ensuring they hear the correct pronunciation of each character, word, and phrase. Listening to the audio, it is possible to get familiar with the sounds and then learners can practice their pronunciation. In this way, the retention and understanding are improved because learners will be able to better identify the sounds with the corresponding characters and meanings.



Figure 2. Wireframe

The architecture was also described as presented in Figure 3 which indicated how the interaction of components like audio playback, recording, scoring, and storage would be. Figure 3 compares the way the app functions in the background and the flow of the data. There are three primary components of the system, namely, Learning Module, Pronunciation Module, and Gamification Module. It all begins with the User Interface (UI) where the user can do all things such as listening to lessons, practicing vowel sounds, and quiz. All the three modules relate to the UI and assist to facilitate the work of users in the process of learning.

The application has been built on a layered architecture and consists of Presentation Layer, Business Layer, and Data Layer. It also engages cloud services of Google Cloud and Azure. All these layers have their respective functions which ensure that all functions are in place such as efficient interaction, quick processing, and safe data management within the application.

It is the Presentation Layer that is seen and used. It includes such items as quiz questions and pronunciation result and gamification such as badge, progress bar. This layer is also in charge of the voice input whereby the user can speak into the microphone, and the app will record the sound to be checked and provide feedback at a later time.

Business Layer performs the primary functions of the application, and it consists of three modules. The Learning Module taps into learning audio content that is stored on Google Cloud and plays it to users. The Pronunciation Module records voice data by using the device microphone and transfers the audio to the Azure Speech Service API that processes the audio and gives voice recognition as feedback on pronunciation. The Gamification Module takes care of such features as daily quizzes, streak tracking, reminders, push notifications, and badges to keep the users motivated and consistent.

The Data Layer will deal with all user data. Simple data, e.g. quiz scores, gamification status are stored locally through SharedPreferences. At the same time, Service Tools will be charged with the responsibility of aligning data with cloud services like Google Cloud and Azure, so that audio files and pronunciation comments are never out of date.

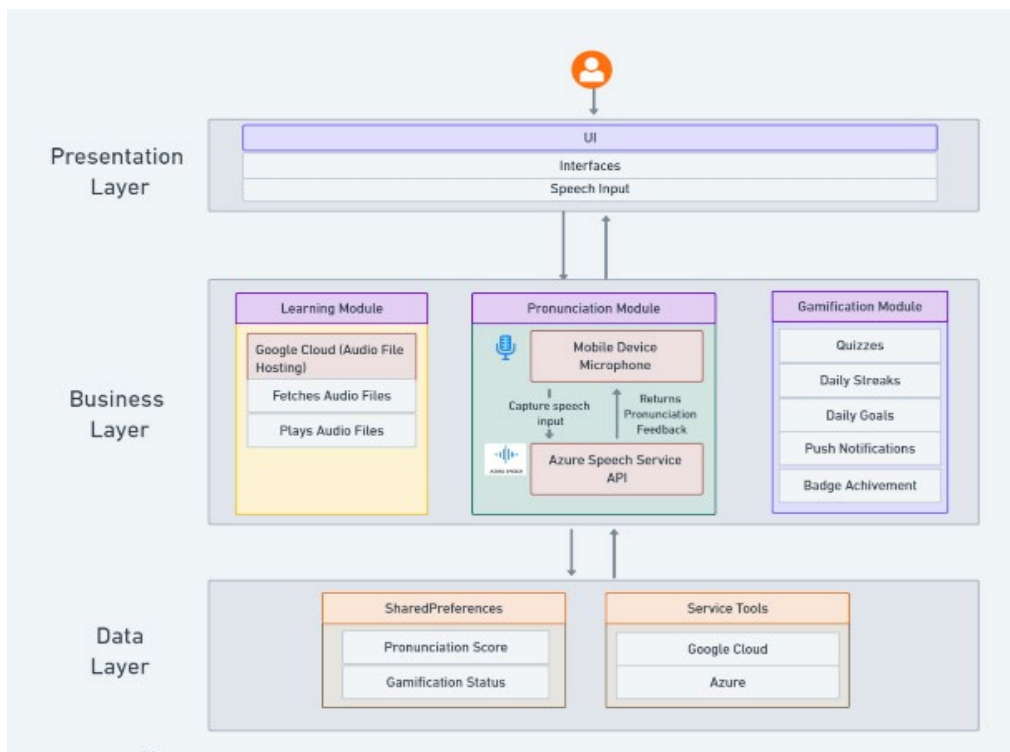


Figure 3. System Architecture

Figure 4 depicts the flow of the use case of the SpeakEasy Nihongo application that demonstrates how the user interacts with the application through the core functions, such as tracking progress, learning vocabulary, and practicing pronunciation. To enhance the user experience on the app, the user needs to create a profile first. With the profile set up, the user can view different pre-defined challenges that are assigned to assess their current level and offer tasks that are appropriate. Later, the user is allowed to click on “Learn Alphabet” as a basic feature of the language interface. As for improvement, the user can track their progress and critique their level of understanding. Alongside, there is the available option of Practice Pronunciation that provides the user with the active practice of speaking and improves their articulations. The user is then evaluated to determine how well they performed, which is an important step to take in order to avoid making the same mistakes again. The sequence of these



activities leads the user to a complete cycle of learning starting from practice to receiving feedback that is aimed at self-improvement.

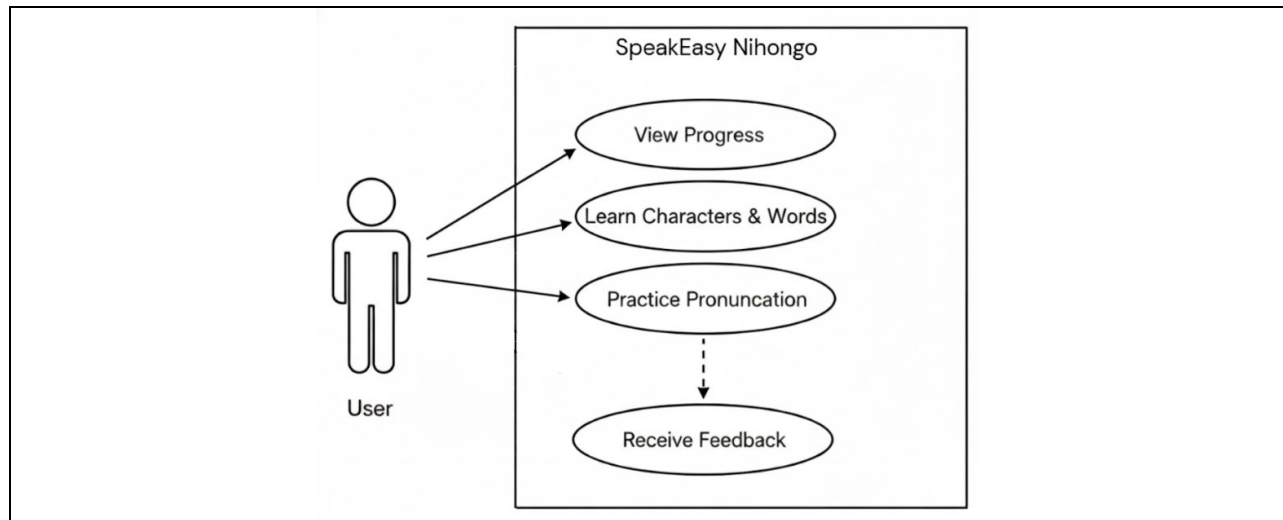


Figure 4. Use case Diagram

The development stage implied creating the app with Flutter and connecting it with gamification and Microsoft Azure. In this phase, the plans and designs made in the earlier stages of the project are converted to a usable product. The main issue is the writing, debugging and making the application functioning. This project had a mobile application developed in the implementation stage which combines voice recognition with game-based learning, as well as API integration. There were several limitations faces during development process. One significant constraint was the fully reliance on REST API for pronunciation assessment. This shows that the main function can only be accessed in the presence of internet connection. The way the use of Microsoft Azure server receive input through internet stability shows how this can affect learners that lives in rural area or those who prefer offline access. Another limitation is that the app does not include feature that is most likely to exist in language gamification app which is competitive leaderboard. Human nature of wanting to win and compete can help with longer user retention. It can also make learning more exciting and less of repetition and routine.

The testing phase was oriented to functional testing, user assessment, and performance measurement. The evaluation stage is critical during the development of software systems since it guarantees the application will work correctly, reliably, and as intended by end users. In this phase, the primary focus is on assuring the project expected the software to accomplish the desired goals. For this project, evaluation will help determine both the functionality and the Interaction of end users with the mobile app integrated with voice recognition technology and gamified learning. In this phase, the intended users have to download the mobile app and test all the functionalities provided. At the end, users are required to answer the questionnaire given for the usability evaluation. Participants in the evaluation were consisting of 32 high school students which were selected by the Japanese language teachers at SMK Putrajaya Presint 11(1). Another 10 participants were college students and adult learners. Specifically, accuracy of the pronunciation scoring mechanism was tested by comparing system produced scores with

reference pronunciations. The scoring system gives a fair estimate of the pronunciation accuracy, which justifies its usage as a valuable feedback instrument for beginner learners. The findings from the evaluation would reveal the effectiveness of SpeakEasy Nihongo as a mobile-assisted language learning platform, and its educational importance. Educationally, the application is able to solve critical problems encountered by beginner learners because it offers flexible, self-paced and interactive learning opportunities. These results indicate that pedagogically informed design incorporated into mobile applications can greatly contribute to the early-stage language learning experience. Regarding the effectiveness, future assessment could be conducted to evaluate whether the application has shown improvements in vocabulary acquisition through learner engagement.

Finally, there was documentation of the summary of all the diagrams, user feedback, technical configuration and evaluation results. Moreover, documentation was also used to specify the project scope, project objectives and technical requirements such that developers could easily consult them. Documentation was used to maintain clear records in order to minimize risks of miscommunication and to minimize the negative effects of unnecessary changes which contributed to more efficient development process overall.

RESULTS AND DISCUSSION

This section discusses the results obtained from the development and evaluation of the application. The final application that is presented in Figure 5 included the final overall application interfaces inside the application. The app has 6 major user interfaces that are Home, Quiz, Learning, Pronunciation Assessment, Achievements and Detailed Learning Pages.

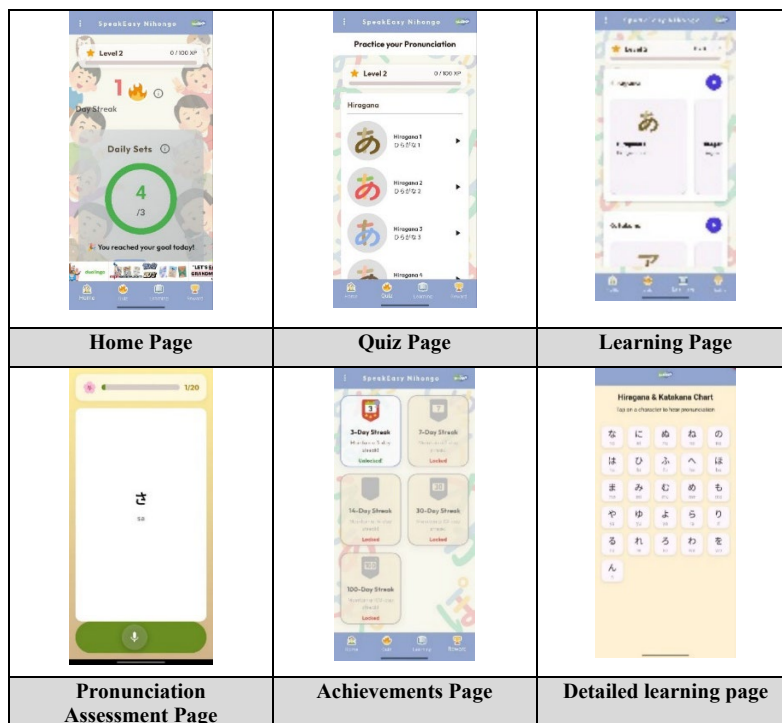


Figure 5. Final App Interface

Figure 6 presents clear pronunciation screens, a record button to record user speech and a results page, which displays pronunciation scores. The scoring option in Microsoft Azure described scoring which comprises accuracy, fluency and pronunciation. This gave useful feedback which allowed the users to learn about their errors. Each score has the following purposes:

1. **Accuracy:** Measures how closely each word or sound in the user's speech matches native pronunciation.
2. **Fluency:** Measures the smoothness, pacing, and rhythm of the user's speech.
3. **Pronunciation:** A combined score showing how natural and native-like overall speech sounds.

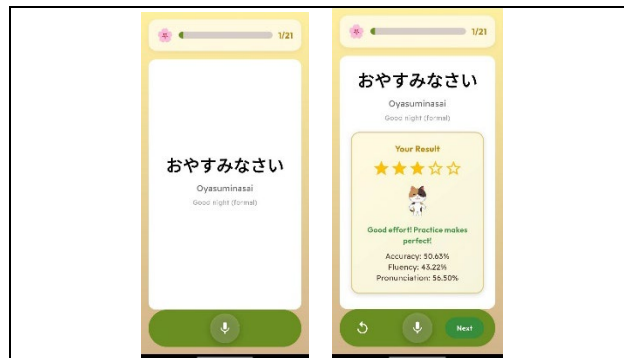


Figure 6. Pronunciation Page

In the application, The Fisher-Yates Shuffle algorithm has verified to work properly in the quiz module, to deliver randomized sequences. The algorithm allows users to see different questions each time of opening to keep the experience fresh.

To test the usability of SpeakEasy Nihongo, the User Experience Questionnaire (UEQ) was given to 42 respondents who were all high school students of SMK Putrajaya Presint 11(1), college students and adult learners with different backgrounds in learning Japanese. With the participants representing different backgrounds, the data gives a general understanding on how the users who belong to various demographics were about the app.

The UEQ consists of 26 items, each containing two opposite terms (“annoying – enjoyable”), rated on a scale from -3 to +3. Table 2 below presents the item-level scores, which mostly range from 1.9 to 2.4, indicating that users generally had a positive experience with the app. Although the details are given in individual items, the interpretation with references to the six UEQ scales is given in the next section.

Table 2. UEQ Table Result Summary

Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	2.1	0.9	1.0	42	annoying	enjoyable	Attractiveness
2	2.4	1.3	1.1	42	not understandable	understandable	Perspiciuity
3	2.3	1.1	1.0	42	creative	dull	Novelty
4	2.2	1.5	1.2	42	easy to learn	difficult to	Perspiciuity



						learn	
5	2.1	1.0	1.0	42	valuable	inferior	Stimulation
6	2.1	1.6	1.3	42	boring	exciting	Stimulation
7	2.1	1.7	1.3	42	not interesting	interesting	Stimulation
8	1.9	1.4	1.2	42	unpredictable	predictable	Dependability
9	2.0	1.0	1.0	42	fast	slow	Efficiency
10	1.9	1.4	1.2	42	inventive	conventional	Novelty
11	2.2	1.4	1.2	42	obstructive	supportive	Dependability
12	2.2	1.4	1.2	42	good	bad	Attractiveness
13	2.3	1.0	1.0	42	complicated	easy	Perspicuity
14	2.2	0.8	0.9	42	unlikable	pleasing	Attractiveness
15	2.0	1.4	1.2	42	usual	leading edge	Novelty
16	2.2	1.0	1.0	42	unpleasant	pleasant	Attractiveness
17	-2.0	1.5	1.2	42	secure	not secure	Dependability
18	2.2	1.2	1.1	42	motivating	demotivating	Stimulation
19	2.1	1.2	1.1	42	meets expectations	does not meet expectations	Dependability
20	2.1	1.0	1.0	42	inefficient	efficient	Efficiency
21	2.4	0.9	1.0	42	clear	confusing	Perspicuity
22	2.0	1.1	1.1	42	impractical	practical	Efficiency
23	2.1	1.1	1.1	42	organized	cluttered	Efficiency
24	2.2	1.5	1.2	42	attractive	unattractive	Attractiveness
25	2.4	1.2	1.1	42	friendly	unfriendly	Attractiveness
26	1.9	1.6	1.3	42	conservative	innovative	Novelty

For a clearer interpretation of these results, the 26 items can be grouped into six UEQ scales:

1. **Attractiveness** – the pleasantness of the overall appearance of the app.
2. **Perspicuity** - the ease and clarity of the app.
3. **Usability** – ease with which users navigate the app features.
4. **Efficiency** – how well the app helps to deliver the lesson and give feedback.
5. **Stimulation** – how entertaining or inspiring the app is when using it.
6. **Novelty** – the difference or uniqueness of the app in relation to other language apps.

Grouping the 26 items into the six UEQ scales allows for a clearer interpretation of the results. Table 3 below presents the mean scores for each scale. As per the standard UEQ rules, a mean value of above +1.00 is considered to be on a positive side, between -1.00 to +1.00 is thought to be on a neutral side and below -1.00 is regarded as an impression. In the positive range, the more user satisfaction is expressed, the better, whereas the less the user satisfaction is expressed, the more problematic or weak areas.

Table 3. Mean scores for UEQ scales

UEQ Scales	Mean
Attractiveness	2.214
Perspicuity	2.327
Efficiency	2.054
Dependability	1.060
Stimulation	2.125
Novelty	2.012

The scores were high in Perspicuity which means that the app was simple to use and comprehend by the users. The respondents also indicated that the interface was user friendly and the audio playback, voice recording and feedback were user-friendly. This is in tandem with the design objectives of the app which is to assist novices to learn in a non-threatening environment. Stimulation and Attractiveness also got a good score which indicates that the app was able to keep the users occupied with its colorful UI and its gamified features. The system of feedback was positively received by learners, who had an opportunity to monitor the performance in pronunciation and enhance it every time. This indicates the motivational effect of incorporating speech scoring, visual and interactive characteristics.

However, Dependability scores lowest, suggesting lower satisfaction in that area. This could be due to the app’s context, where security and critical error handling are not top priorities. The novelty of the app was perceived as innovative and new, and the users liked its originality and design as opposed to more traditional learning systems. The outcome also indicate that the users could smoothly continue with their work without time or energy wastage, which justifies the high efficiency score.

The result of UEQ Scale mean score as discussed can be clearly visualized with a bar graph as in Figure 7. The green area indicates positive mean result. The yellow indicates moderate mean score while red indicates low and negative score.

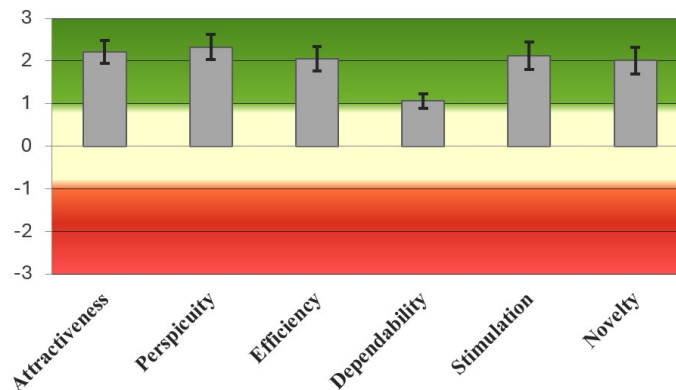


Figure 7. UEQ Scale Graph by Mean



The UEQ findings in general reveal that SpeakEasy Nihongo provides a convenient and encouraging platform to learners studying Japanese pronunciation who are starting with basics. The upper score presented in usability dimensions implies that the app can fulfill the expectations of the target users and especially in making the pronunciation practice felt convenient and enjoyable. These evaluation responses have shown that there are areas of improvements that can be made, including making voice input more accurate and the ability to give more personalized advice which will be looked into in further updates.

The SpeakEasy Nihongo has a pedagogical effect as it is designed to assist learners to acquire Japanese language at the beginning level by providing structured and interactive learning. The application uses some of the principles of microlearning where content is offered in small and manageable units to enable learners to progressively develop their vocabulary and simple sentence structures. Also, repetitive exercises and instant feedback systems reinforce memory and enhance the accuracy of pronunciation. This method conforms to known theories of language learning in general, especially the behaviorist reinforcement and constructivist learning, in which students are actively involved in what they learn and learn by doing.

The results of this research are not only supportive but also complimentary to the current studies in applications of pronunciation learning and Mobile-Assisted Language Learning (MALL) settings. As with other studies, repetitive audio exercises can help to increase the accuracy of pronunciation and vocabulary levels. Nevertheless, in contrast to most of the currently available apps, SpeakEasy Nihongo incorporates interactive feedback and structured microlearning, which seem to contribute to better engagement and understanding of the concept among learners. Also, although the flexibility and accessibility are pronounced as the main benefits of MALL studies, this research shows that the inclusion of pedagogically guided material may also enhance the effectiveness of learning. Moreover, the portability of the application is also in line with the theory of MALL, which emphasizes flexibility, accessibility, and autonomy of learners. Combined, these theoretical views describe the usefulness of the application in facilitating the acquisition of the Japanese language at the beginner level. Although these contributions have been made, the limited sample size and limited duration of evaluation in the study limits it, implying that future studies need to examine the long-term learning outcomes and extend the research to more users.

CONCLUSION AND RECOMMENDATION

The development of the application validates the fact that real-time pronunciation scoring can be easily incorporated into mobile applications among learners of Japanese who are of a novice level. Using the Speech API of Azure, the app is useful in providing a structured lesson and audio guidance with immediate and data-driven feedback. This is what will change a passive learning process into a corrective practice environment. The benefits of developing an app for learning Japanese intended for beginners that pays special attention to pronunciation are its ability to help specific users like schoolchildren, university students, and working professionals intending to learn the language. Such app development considers sustainable development goals, which for instance in this case SDG 4 is relevant. SDG 4 contributes to Quality Education, which emphasizes relevant and effective learning in order to build essential skills. Pronunciation is



crucial in communication and as such many language learning applications tend to disregard speaking exercises for students. Effective app development goes a long way in enabling people of all ages, genders and nationalities to learn and speak the Japanese language with fluency and pronounce the relevant words correctly. At the same time, the additional gamification elements engage and motivate users to practice consistently, making learning both accessible and enjoyable. This approach ensures that language education is not only high in quality but also more adaptable and inclusive, providing learners with tools tailored to diverse learning needs such as school students seeking foundational skills and university students pursuing language as part of their studies. The further development of the application can be achieved by focusing on lexical enrichment, with more categories of vocabulary being included to reinforce the language background of the learner. Additionally, the use of AI chatbot is suggested to imitate the way of natural conversation. This would allow users to transition from isolated word practice to applying their pronunciation skills within a dynamic, real-world conversational context. In terms of assessments, the future evaluation would be the assessment of the efficiency or summative assessment for the mobile application. This is to evaluate beginner learners' improvement especially in the vocabulary recognition and pronunciation accuracy.

REFERENCES

- Altynbekova, G., & Zhussupova, R. (2020). Mobile application Fluentu for public speaking skills development. *SHS Web of Conferences*, 88, 1–6. <https://doi.org/10.1051/shsconf/20208802008>
- Amalia, I. (2020). The application of Mobile Assisted Language Learning (MALL) in teaching pronunciation. *International Journal of Language Education and Culture Review*, 6(2), 194–203.
- Anderies, A., Agustina, C., Lipiena, T., Raaziqi, A., & Gunawan, A. A. (2023). User experience analysis of Duolingo using user experience questionnaire. *Engineering, Mathematics and Computer Science Journal*, 5(3), 155–159. <https://doi.org/10.21512/emacsjournal.v5i3.9227>
- Brata, K. C., & Brata, A. H. (2020). User experience improvement of Japanese language mobile learning application through mental model and A/B testing. *International Journal of Electrical and Computer Engineering*, 10(3), 2659–2667. <https://doi.org/10.11591/ijece.v10i3.pp2659-2667>
- 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 *Proceedings of the E-Learning and Smart Engineering Systems* (pp. 177–185). https://doi.org/10.2991/978-94-6463-360-3_19
- Ferraro, A., Galli, A., La Gatta, V., & Postiglione, M. (2023). Benchmarking open source and paid services for speech to text: An analysis of quality and input variety. *Frontiers in Big Data*, 6, 1–8. <https://doi.org/10.3389/fdata.2023.1210559>
- Hoxhaj, S. (2024). Interactive language learning apps: Enhancing student motivation and participation in EFL activities at the University. *Academic Journal of Business, Administration, Law and Social Sciences*, 10(2), 21–32. <https://doi.org/10.2478/ajbals-2024-0012>



- Kholis, A. (2021). Elsa Speak app: Automatic speech recognition (ASR) for supplementing English pronunciation skills. *Pedagogy: Journal of English Language Teaching*, 9(1), 1–14. <https://doi.org/10.32332/joelt.v9i1.2723>
- Kirana, C., Wijaya, B., & Holil, A. (2021). Implementation of the fisher-yates shuffle algorithm in exam-problem randomization on M-learning applications. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, 7(2), 47–51. <https://doi.org/10.23917/khif.v7i2.11761>
- Klock, A. C., Gasparini, I., Pimenta, M. S., & Hamari, J. (2020). Tailored gamification: A review of literature. *International Journal of Human-Computer Studies*, 144, Article 102495. <https://doi.org/10.1016/j.ijhcs.2020.102495>
- Nst, A. H., Daulay, S. H., & Dewi, U. (2023). Using Rosetta Stone application as a mobile assisted language learning (MALL) in EFL: Documentary study. *Scope: Journal of English Language Teaching*, 8(1), 8–13. <https://doi.org/10.30998/scope.v8i1.17653>
- Nushi, M., Fattahi, N., & Ebn-Abbasi, F. (2024). An evaluative review of Mondly: A mobile language learning application. *The EuroCALL Review*, 30(2), 69–85. <https://doi.org/10.4995/eurocall.2023.15763>
- Olson, D. J., & Offerman, H. M. (2021). Maximizing the effect of visual feedback for pronunciation instruction: A comparative analysis of three approaches. *Journal of Second Language Pronunciation*, 7(1), 89–115. <https://doi.org/10.1075/jslp.20005.ols>
- Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *RELC Journal*, 52(1), 3–21. <https://doi.org/10.1177/00336882211002283>
- Qin, J. (2024). Research on foreign language anxiety in Japanese acquisition. *Journal of Education, Humanities and Social Sciences*, 38, 70–75. <https://doi.org/10.54097/gz3ap916>
- Rethinasamy, S., Pae, R., & Ramanair, J. (2021). Malaysian undergraduates' beliefs and motivation for learning Japanese as a foreign language. *Turkish Online Journal of Qualitative Inquiry*, 12(6), 1519–1529.
- Saragih, L., Ramadhan, A., & Daulay, E. (2024). The effect of Falou application in improving students' speaking skill. *Morfologi: Jurnal Ilmu Pendidikan, Bahasa, Sastra dan Budaya*, 2(5), 229–237. <https://doi.org/10.61132/morfologi.v2i5.955>
- Smiderle, R., Rigo, S. J., Marques, L. B., Peçanha de Miranda Coelho, J. A., & Jaques, P. A. (2020). The impact of gamification on students' learning, engagement and behavior based on their personality traits. *Smart Learning Environments*, 7, 1–11. <https://doi.org/10.1186/s40561-019-0098-x>
- Ueyama, M. (2020). New approach to teaching Japanese pronunciation in the digital era: Challenges and practices. In G. Pappalardo & P. Heinrich (Eds.), *European Approaches to Japanese Language and Linguistics*. <https://doi.org/10.30687/978-88-6969-428-8/010>
- Wibawa, R., Lokacarya, A., Kurniawan, F., & Udjaja, Y. (2023). Japanese language learning game “Miryoku” using Android-based speech recognizer API. *Procedia Computer Science*, 216, 547–556. <https://doi.org/10.1016/j.procs.2022.12.168>



Declaration of Generative AI and AI-assisted Technologies in the Writing Process

This manuscript was developed with the assistance of StealthWriter for the purpose of correcting sentences in English. The tool supported the authors in reviewing and refining the content, and the authors assume full responsibility for the final version submitted for publication.

Conflict of Interest

The authors have no conflicts of interest to declare.

Acknowledgement

The authors wish to express appreciations to Faculty of Computer and Mathematical Sciences of Universiti Teknologi MARA Cawangan Terengganu for the continuous support given towards development of research projects.

Authors' Contributions

Damia Syafiqah Mohd Salleh developed the mobile application and drafted manuscript; Norlina Mohd Sabri supervised the development and edited the manuscript; Fazlin Marini Hussain reviewed the final draft. Finally, all authors approved the final version of the manuscript.