2022 International Conference on Advanced Learning Technologies (ICALT) | 978-1-6654-9519-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICALT55010.2022.00077

JaSenpai: Towards an Adaptive and Social Interactive E-Learning Platform for Japanese Language Learning

Mario E. Aburto-Gutierrez*, Gamar Azuaje*, Vipul Mishra*, Shaira Osmani, and Kazushi Ikeda
Graduate School of Science and Technology
Nara Institute of Science and Technology
Ikoma, Japan

***Abstract*—The Japanese language is an essential skill for many foreigners who plan to study or work in Japan, but it is very hard and time-consuming to learn. Given the current COVID-19 pandemic, the use of online and computer-assisted tools for Japanese language learning is indispensable. However, many of the currently available tools do not offer personalized content based on the user's performance and lack social interaction, which can lower the engagement level of the users. In this paper, we propose JaSenpai, a Japanese language E-learning platform that features an automatic generation of vocabulary exercises, a recommendation system based on previous answers, and a multiplayer game for social interaction. We believe these elements can provide a more engaging and effective learning experience.**

***Index Terms*—japanese language learning, interactive e-platform, personalized language learning, recommender system**

I. Introduction

The number of Japanese as a Foreign Language (JFL) learners has increased steadily for the past decade as Japan has become an attractive place for study and work [1]. Learning a foreign language in a classroom setting, with a teacher and other students, is effective. However, given the situation with the COVID-19 pandemic, many of the classes in traditional classroom settings had to be adapted into an online setting, not always giving the best results. Therefore, the development of E-learning tools to supplement the limitations of online classes is essential. In the case of JFL learning, it is crucial to develop effective tools that allow users to maintain a high level of engagement, given the difficulty and the large amount of time required to master the language.

Many applications (e.g. *Duolingo*, *Anki*, etc.) have proved to be helpful for learning Japanese as a foreign language. However, they have some limitations. First, they mostly rely on manually crafted exercises which are costly to generate and therefore limited. Additionally, they are designed in a single-user setting, even though, for example, Jabbari and Eslami [2] have shown using multiplayer games can enhance the learning experience in second language learning.

* Equal contribution. Corresponding author: Mario E. Aburto-Gutierrez (email: aburto_gutierrez.mario.ae2@is.naist.jp).

In this paper, we propose JaSenpai, a JFL E-learning platform that includes: 1) automatic generation of vocabulary exercises, 2) a recommendation system that provides personalized exercises based on the user's historical data, and 3) a multiplayer game that allows social interaction between users through exercise solving. We believe that these three elements combined can increase the engagement and effectiveness in the learning of the Japanese language.

II. Related work

There are multiple approaches for developing a Computer-Assisted Language Learning (CALL) system, but in our work, we focus mostly on employing gamification and recommendation systems, because they can provide an engaging and effective learning experience [3, 4]. In this section, we report on previous works that have used the aforementioned approaches in a language learning setting.

Gamification can be defined as the addition of game elements to a system to positively influence motivation, productivity and user behaviour [5]. In the context of JFL learning, Udjaja [3] developed a system to learn *hiragana* and *katakana* (Japanese syllabary characters) by including elements from games like role-playing, goals, and turn-based battles. They showed that including these elements not only made the learning more effective but also provided an engaging learning experience. However, their system is limited to a single-player, whereas our system proposes a multiplayer game, and they focus on *hiragana/katakana* while we focus on vocabulary learning.

In the context of employing recommendation systems for language learning, Takii et al. [4] proposed an e-book reader recommendation system for EFL (English as a Foreign Language) using a dataset of books and quizzes. Their system utilizes the users' past activity to recommend books/quizzes suitable to their level. Unlike our system, their target language is English, and they recommend books/quizzes based on the knowledge level, while we recommend exercises based on previous answers. Additionally, their system requires an existing dataset of quizzes, while ours can generate exercises automatically.

978-1-6654-9519-6/22/$31.00 ©2022 IEEE
DOI 10.1109/ICALT55010.2022.00077

III. System Description

Figure 1 shows the high-level architecture of our system. In the following subsections we describe each component in more detail.

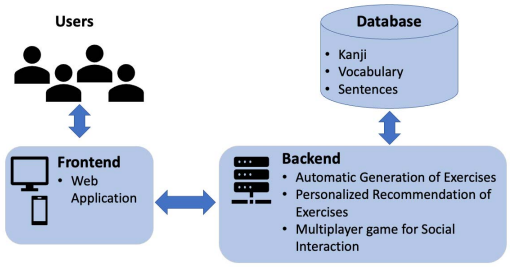

Fig. 1. High-level architecture of our system

A. Datasets

To generate exercises automatically, we used publicly available datasets for *kanji* (Chinese characters), vocabulary, and sentences. For *kanji*, we used the KANJIDIC dictionary [6], which includes the meaning, reading, JLPT (Japanese-Language Proficiency Test) level, among other information, for more than 6,000 standard Japanese *kanji*. For vocabulary, we used JMDICT [7], a Japanese-Multilingual online dictionary that includes readings, meanings in several languages, etc. For sentences, we use the corpus provided by Katsuta et al. [8], commonly known as the Yasashi Nihongo Corpus, for generating exercises. The corpus file comprises $50,000$ rows, each containing a Japanese sentence, a simplified version of the Japanese sentence, and its English translation.

B. Automatic generation of multiple choice vocabulary exercises

Our system generates vocabulary exercises that follow the JLPT format for *kanji reading*[1]. In these exercises, an underlined word is shown in a sentence, to provide context for its meaning and reading. The user has to select the correct reading among four different choices. (See Fig.2).

We filtered the vocabulary dataset for words containing two *kanji* characters to generate the questions. Then, for each word, we sampled a sentence from the Yasashi Nihongo Corpus. To generate the wrong choices, we combine alternative readings for both *kanji* in the word.

Table 1 shows the number of exercises for each JLPT level (N5 Beginner; N1 Advanced) that were generated from the datasets with the procedure described above.

C. Personalized recommendation of exercises

We recommend vocabulary exercises for a word depending on: 1) the number of times exercises for the word were shown ($count$), 2) the percentage of incorrect answers ($incorrect$), and 3) the time passed since an exercise for the word was last shown ($time$). We use weighted sampling [9] to sample the exercises for each user. The weight for each exercise is calculated according to the following formula:

$$w = \frac{1}{3}w_{count} + \frac{1}{3}w_{incorrect} + \frac{1}{3}w_{time}$$

where:

$$w_{count} = e^{-count}$$

$$w_{incorrect} = \begin{cases} 1 & \text{if } count = 0. \\ \frac{count - 0.99 * correct}{1.01 * count} & \text{otherwise.} \end{cases}$$

$$w_{time} = \begin{cases} 1 & \text{if } count = 0. \\ \min(1, time/(86400 * 7)) & \text{otherwise.} \end{cases}$$

The weights dynamically change to satisfy the following conditions. The probability of sampling a word will:

- Decrease if the word was already shown before.
- Increase if the word was answered incorrectly by the user.
- Increase as time passes since the word was last shown, the limit being 7 days.

TABLE I
Number of generated exercises by JLPT level

Level	Number of exercises
N5	90
N4	130
N3	285
N2	206
N1	281

D. Responsive web application and database

The web application is a Single-Page Application (SPA) comprised of: a login view, a home view, a practice view, a multiplayer-game view, and a profile view.

We provide vocabulary exercises for *kanji reading* in a multiple-choice format. The user is shown an underlined word written in *kanji* in the context of a sentence along with 4 different choices of readings and the user has to select the correct answer. We use the React JS framework to develop the frontend and the Flask framework to develop the backend API. All exercises are dynamically fetched to the frontend through backend API calls. The backend API retrieves user information and vocabulary exercises from a MySQL database, and uses the mechanism described in section III-C to sample the vocabulary exercises relevant to the user.

E. Multiplayer game for social interaction

The multiplayer game is designed to enhance the learning experience by enabling social interaction in a multiplayer environment. Users can join into different rooms where, together with other users, they answer vocabulary exercises.

Before joining a room, the user connects to the server and enters a lobby. Inside the lobby, they can create or join a room. A unique 6-digit hexadecimal code is generated when

[1]https://www.jlpt.jp/e/guideline/testsections.html

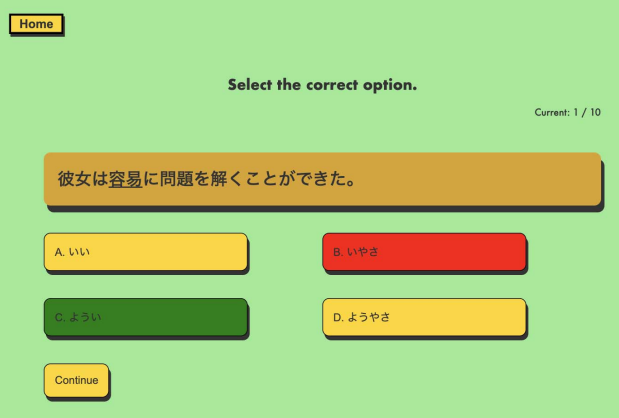

Fig. 2. Example of vocabulary exercise for a single user. The target word for practice is underlined in the sentence. In this case, the user selected a wrong choice (shown in red), and the system highlighted the correct answer (in green).

Fig. 3. Example of vocabulary exercise in the multiplayer game. The target word for practice is underlined in the sentence. In this case, Player1 selected the choice **c** by approaching the bottom-left desk, which was then highlighted in blue.

a user creates a room. The code can be used by other users to enter this room. The environment resembles a classroom (See Fig. 3), featuring a screen that shows an exercise and four desks that correspond to four different choices. The game has a GUI (Graphical User Interface) with the following elements: the scoreboard, and the timer.

Once the game starts, the backend API is used to fetch vocabulary exercises, which are then displayed on the screen. Users will have to answer before the timer runs out. The users choose their answer by approaching one of the four desks, which change to the color corresponding to the letter of the different choices displayed on the screen. When the timer reaches zero, the answer is highlighted, and the scoreboard is updated based on the answers of the users. A single game consists of ten exercises. The game was developed using the Unity engine, and the Photon package was used to enable asynchronous communication between the users.

IV. Conclusions and Future Work

We developed an E-learning platform for JFL learning that recommends personalized exercises while providing social interaction among its users. This system is our first step towards providing a more engaging and effective experience to learners of the Japanese language.

In the future, we intend to investigate the use of sophisticated methods like deep neural networks for automatically generating questions, borrowing from the recent advancements in deep question generation and language models. For the recommendation system, we plan to include more diverse content and recommend exercises based on the users' interests by using techniques such as contextual bandits. Additionally, the multiplayer game currently only allows users to play in a competitive mode, but we plan to add a cooperative mode to enable another form of social interaction between the users.

Finally, we plan to carry out an in-depth user study, tracking the learning progress of the users to validate the effectiveness of the generated questions, the recommendation system and the multiplayer game.

References

[1] JASSO, "Result of an Annual survey of International Students in Japan 2020," 2020. [Online]. Available: https://www.studyinjapan.go.jp/en/statistics/zaiseki/data/2020.html

[2] N. Jabbari and Z. R. Eslami, "Second language learning in the context of massively multiplayer online games: A scoping review," *ReCALL*, vol. 31, no. 1, pp. 92–113, 2019.

[3] Y. Udjaja, "Gamification assisted language learning for japanese language using expert point cloud recognizer," *International Journal of Computer Games Technology*, vol. 2018, 2018.

[4] K. Takii, B. Flanagan, and H. Ogata, "Efl vocabulary learning using a learning analytics-based e-book and recommender platform," in *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2021, pp. 254–256.

[5] I. Blohm and J. M. Leimeister, "Gamification," *Business & information systems engineering*, vol. 5, no. 4, pp. 275–278, 2013.

[6] J. Breen, "Multiple indexing in an electronic kanji dictionary," in *Proceedings of the Workshop on Enhancing and Using Electronic Dictionaries*, 2004, pp. 1–7.

[7] ——, "JMdict: a Japanese-multilingual dictionary," in *Proceedings of the Workshop on Multilingual Linguistic Resources*, 2004, pp. 65–72.

[8] A. Katsuta and K. Yamamoto, "Crowdsourced corpus of sentence simplification with core vocabulary," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.

[9] P. S. Efraimidis and P. G. Spirakis, "Weighted random sampling with a reservoir," *Information Processing Letters*, vol. 97, no. 5, pp. 181–185, 2006.