Shih, JL. et al. (Eds.) (2023). Proceedings of the 31st International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Toward Game-Based Learning of Japanese Writing for Elementary School Students

Kazumasa OMURA[a*], Kei KUBO[b], Frederic BERGERON[c] & Sadao KUROHASHI[a,d]
[a]*Graduate School of Informatics, Kyoto University, Japan*
[b]*Center for Japanese Language, Waseda University, Japan*
[c]*SystemI Co., Ltd., Japan* [d]*National Institute of Informatics, Japan*
*omura@nlp.ist.i.kyoto-u.ac.jp

Abstract: It is a long-standing problem that many elementary school students in Japan have an aversion to writing compositions. To address the problem, we designed an AI educational game for elementary school students to study Japanese writing utilizing existing language resources. In the game, players construct simple and complex sentences by connecting given word cards with particle marks. The constructed sentences are automatically scored using large-scale language resources, allowing players to receive on-the-spot feedback, such as how to improve the use of a case marker. We also developed smartphone and web applications of the game and conducted a user study to assess it. The results of the user study demonstrated that our application can be used as a good introduction to studying Japanese writing.

Keywords: Game-based learning, Japanese writing, Language resources

1. Introduction

Since we use language as a means of cognition and communication, language education plays an essential role in our lives. In language education, written language production is also indispensable for learning how to express our thoughts.

However, in Japanese writing education, it is a long-standing problem that many elementary school students have an aversion to writing compositions (National Institute for Educational Policy Research, 2008; Ritsumeikan University Library, 2017). One of the possible reasons is that they learn Japanese writing only through open-ended writing assignments such as book reports and essays on topics related to daily life. Such assignments rarely motivate the students to go further than merely complete them, and they often struggle to come up with what to write due to the open-ended format. In addition, they receive little feedback on their writing, leading to a vicious cycle where they become increasingly aware of being poor at writing without learning how to improve.

In order to ameliorate the current situation, Japanese education is in need of an educational material that offers a more engaging experience, i.e., allows students to construct sentences with fun and receive feedback on their writing. However, there are two major challenges to achieving this: how to reduce an aversion to constructing sentences and how to evaluate writing automatically.

Game-based learning (Vandercruysse et al., 2012), which aims to make the learning process more fun with a game, is a promising solution to the first challenge. There has been no AI educational game for studying Japanese writing to the best of our knowledge; thus, it is worth attempting to develop such a game and investigate its effectiveness.

Regarding the second challenge, we focus on existing language resources. In natural language processing, large-scale language resources have been built so far to teach linguistic and world knowledge to computers (Fellbaum, 1998; Kawahara & Kurohashi, 2006; Speer et al., 2017), some of which can be utilized for human learning as well as machine learning. By exploiting them, it is now possible to automatically generate and score simple writing practice questions, and the aforementioned educational material has become more feasible.

Against the above background, we design an AI educational game for elementary school students to study Japanese writing utilizing existing language resources. Hereafter, we call it “Kotoba-musubi”. (“Kotoba” and “musubi” are the Japanese words that mean “words” and “connecting”, respectively.) Section 2 describes the details.

We also develop smartphone and web applications of Kotoba-musubi and conduct a user study to investigate its effectiveness. Section 3 presents the positive results showing the effectiveness in reducing an aversion to writing compositions. We have released the game for educational purposes, the link to which is provided in Section 4 with a conclusion.

2. Kotoba-musubi

2.1 Game Design

Kotoba-musubi is a word-based game, where the player constructs simple sentences by connecting content words with case markers and combining the simple sentences with discourse markers into complex sentences with contingent relations (cf. Figure 1). In order to construct sentences, players arrange given 12 *word cards* and nine *particle marks* according to the empty rectangle and oval frames, respectively.

Figure 1. Screenshot of Kotoba-musubi with gloss (left) and its scoring results (right). “NOM” and “ACC” in the left figure stand for “nominative” and “accusative” cases, respectively.

Word Card: The 12 word cards break down into six noun cards, five verb/adjective cards, and one wildcard that allows players to enter a word freely. The role of each card is distinguishable by its color or shape; thus, players can learn how to construct sentences even if they do not understand the concept of parts of speech. In addition, ruby characters are written above Chinese characters to facilitate reading. Regarding verb/adjective cards, the predicates can be conjugated in the present, past, negative, or past-negative form, which makes it possible to create more diverse sentences.

Particle Mark: The nine particle marks break down into five of major case particles (nominative, accusative, dative, and instrumental cases, and “with” or “and”) and four of discourse connectives representing contingent (causal and conditional) relations. As with the word cards, the role of each mark is distinguishable by its color or shape. Each mark is arrow-shaped, its direction indicating the dependency between words.

The score of each sentence is automatically computed by pattern matching with language resources, allowing players to receive instant feedback. For instance, if no examples of a simple sentence are found in the Japanese case frames (Kawahara & Kurohashi, 2006) due to incorrect usage of a case marker, our system gives feedback suggesting a more appropriate case marker (cf. Figure 1). The player’s objective is to achieve a higher score by constructing basic simple sentences and complex sentences with contingent relations.

2.2 Method for Building Word Card Sets

In order for the game to be more fun, it is preferable that players can construct several simple and complex sentences from a given word card set. To guarantee this, we focus on *core events* (Omura et al., 2020; Omura & Kurohashi, 2022).

Core events are defined as the high-frequency predicate-argument structures acquired from case frames (Kawahara & Kurohashi, 2006); that is, each core event consists of one predicate and a few of its frequently co-occurring arguments and case markers. We adopt core event pairs with contingent relations such as "雨が降る→長靴を履く (it rains→wear rain boots)", which can be extracted at scale from the Kyoto University Commonsense Inference dataset (cf. https://nlp.ist.i.kyoto-u.ac.jp/EN/?KUCI).

Our method is to automatically generate word card sets from core event pairs. It roughly consists of the following three steps (cf. Figure 2).

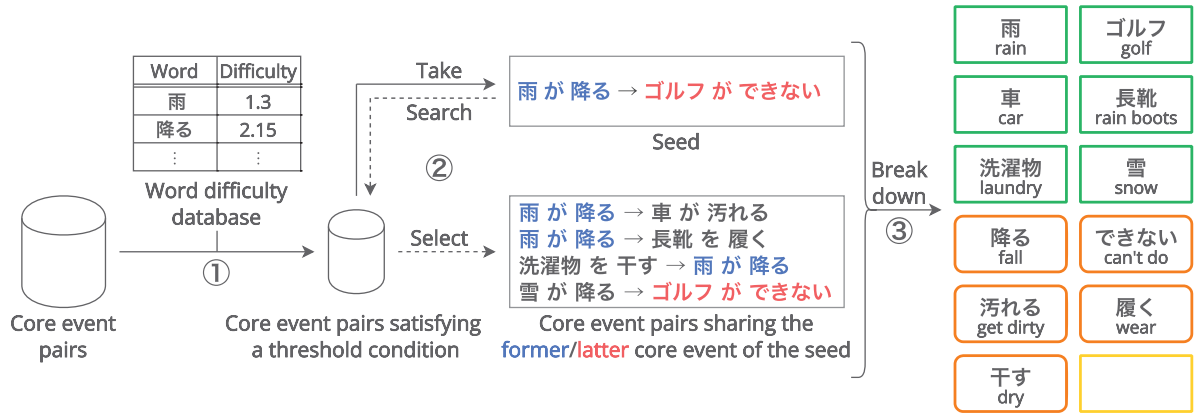

Figure 2. Overview of the method for building word card sets. Green rectangles and orange rounded rectangles correspond to noun and verb/adjective cards, respectively.

Step 1: Filter Core Event Pairs by Word Difficulty. First, we adjust vocabulary using the Japanese word difficulty database (Mizutani et al., 2019), considering the target users are elementary school students. This database contains 26k words of basic Japanese vocabulary labeled with their average acquisition time using crowdsourcing. Acquisition time is regarded as word difficulty and expressed as a number from 1 to 5 (cf. Table 1).

Table 1. *Correspondence between numbers and acquisition times in the Japanese word difficulty database (Mizutani et al., 2019). ES and JHS stand for elementary school and junior high school, respectively.*

Number	1	2	3	4	5
Acquisition time	Before ES	Early ES	Late ES	After JHS	Never seen or heard

In this study, we set the following threshold conditions.

- **Easy** (for early elementary school students): Maximum word difficulty of words in a core event pair does not exceed 2.0.
- **Medium** (for middle elementary school students): Average word difficulty of words in a core event pair exceeds 1.5, and maximum word difficulty does not exceed 2.5.
- **Hard** (for late elementary school students): Average word difficulty of words in a core event pair exceeds 2.0.

We exclude core event pairs not satisfying a threshold condition. We set a strict upper bound but allow easy words to get mixed in harder word card sets to smooth the change in difficulty.

Step 2: Select Core Event Pairs. Second, we select five core event pairs from the ones satisfying a threshold condition. Specifically, we take one core event pair (hereafter, **seed**) and randomly select the other four pairs that share the former or latter core event of the seed. If we fail to get five pairs, including a seed, we skip generating a word card set.

Step 3: Generate a Word Card Set. Lastly, we generate a word card set by breaking down the five pairs selected in the previous step into predicates and arguments. If we obtain six nouns and five or six verbs/adjectives after de-duplication, we regard the words as a word

card set. In case that six verbs/adjectives are obtained, we replace the least frequent verb/adjective with a wildcard.

2.3 Careful Examination of Core Event Pairs and Building Word Card Sets

We must be extremely careful not to give inappropriate questions as an educational application. Accordingly, we request two linguistic experts to manually classify the core event pairs used for word card sets into the following categories.

- **Valid:** A core event pair has contingent relation.
- **Invalid:** A core event pair has no contingent relation.
- **Inappropriate:** A core event pair contains educationally inappropriate expressions.

We use the "valid" core event pairs for generating word card sets.

Prior to the examination, we excluded core event pairs containing minor case particles or words unregistered in the Japanese word difficulty database. As a result, 9k core event pairs were left and examined. We also automatically assigned reading to each word and conjugated forms to each predicate using the Japanese morphological analyzer, Juman++ (Morita et al., 2015; Tolmachev et al., 2018). We asked the experts to correct auto-assigned readings in addition to the examination.

As a result, we obtained 4,362 valid, 3,560 invalid, and 1,120 inappropriate core event pairs; then, we generated word card sets from the verified core event pairs according to the method described in Section 2.2. Finally, we obtained 99 easy, 594 medium, and 284 hard word card sets, achieving a size that is sufficient for solo play.

2.4 Automatic Scoring

When players tap a scoring button, the information on how cards and marks are placed is sent to a back-end server and automatically scored. The automatic scoring is performed in the following three steps.

Step 1: Recognize Simple and Complex Sentences. A simple sentence is recognized as a sequence of a verb card at the end and noun cards in the rest connected by case particle marks, and a complex sentence as two simple sentences connected by a discourse connective mark. We consider all possible combinations of simple sentences.

Step 2: Score Each Sentence Automatically. Our policy is to prioritize simple sentences whose predicate and argument frequently co-occur (i.e., idiomatic) or whose length is longer. The score of each simple sentence is determined based on the number of examples in the Japanese case frames (Kawahara & Kurohashi, 2006). Specifically, regarding the case frame cf where examples are found, we compute a score of each argument and case pair (a, c) with the following function S and sum it up.

$$S(a, c, cf) = 0.5 + 0.5 \times \min\left(1, \frac{f_{a,c,cf}}{f_{cf}} \times 5\right)$$

where $f_{a,c,cf}$ is the frequency of argument a in case c of case frame cf, and f_{cf} is the frequency of case frame cf.

Each score is converted into a symbol for further interpretability. Specifically, a simple sentence is graded as "☆☆" if the score is 1 or more, as "☆" if the score is between 0 and 1, and as "?" if the score is 0. Regarding complex sentences, they are graded as "☆☆☆" if they contain one of the verified core event pairs; otherwise, as "?". When examining the containment relation, we take into account the polarity of negation of a predicate.

Step 3: Generate Feedback based on the Results. The following feedback is generated from the sentences graded as "?" depending on the reasons behind them.

- There is a more appropriate case marker that allows for the sentence to achieve a score.
- The sentence is grammatically incorrect (e.g., the sentence starts with a verb/adjective, a verb/adjective depends on a noun, and so on.)
- The sentence is not found in language resources.

Regarding the third point, we expect to collect unknown contingent relations through an error reporting function and thus improve our evaluation system.

3. User Study

We developed iOS, Android, and web applications of Kotoba-musubi and conducted a user study to assess engagement.

3.1 Settings

We recruited 80 pairs of elementary school students and their parents across Japan and had the children play the game for an hour in total over two days. 80 students consist of 10 boys and 10 girls each from third to sixth grade in elementary school. In order to let them play freely, we neither set their quota nor specify the difficulty of questions they tackled. After playing the game, they answered the questionnaire described in Table 2.

Table 2. *Main items and options of the questionnaire answered by the participating children. Regarding the second and subsequent items, the option with a larger number is preferable.*

Do you like writing compositions?				
5. Like	4. Like a little	3. Neither	2. Dislike a little	1. Dislike
Did you enjoy playing Kotoba-musubi?				
5. Enjoyed	4. Enjoyed a little	3. Neither	2. Didn't enjoyed a little	1. Didn't enjoy
Do you want to continue to play Kotoba-musubi?				
5. Yes	4. Yes, a little	3. Neither	2. No, a little	1. No
Which do you think Kotoba-musubi is: "game" or "study"?				
5. Game	4. A litte game	3. Neither	2. A little study	1. Study
Do you think it was worth playing Kotoba-musubi?				
2. Yes	1. No			

3.2 Results and Discussion

Figure 3 shows the aggregate results of the questionnaire. We can see that elementary school students tend to dislike writing compositions, as mentioned in Section 1. Despite this disadvantageous situation, 70% of the participating children enjoyed Kotoba-musubi, and 90% answered it was worth playing the game. In addition, 70% expressed their will to continue to play the game, which suggests that it is a good introduction to studying Japanese writing.

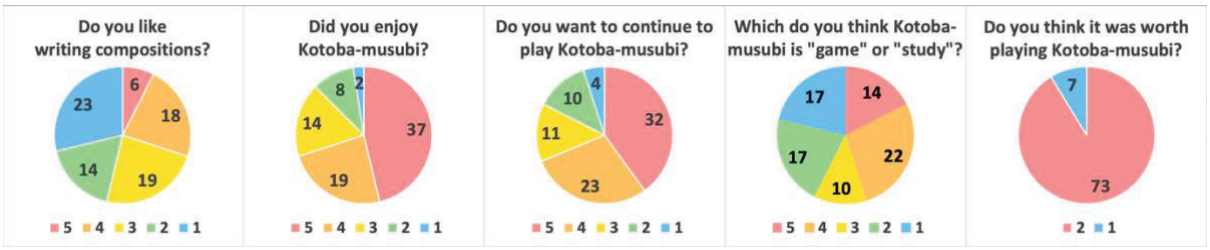

Figure 3. Aggregate results of the questionnaire described in Table 2. The numbers in the figure represent those of children who chose the option.

We also investigated the number of children who disliked writing compositions but enjoyed the game. We found that 18 of 37 children who disliked writing compositions enjoyed the game. The result supports effectiveness in reducing an aversion to writing compositions.

The excerpted feedback comments from the participating children are listed below.

- + It was fun and refreshing to create sentences using words I didn't usually use.
- + I'd honestly like to see the game introduced to school tablets.
- ± It will be more fun if we can compete with our friends for higher scores.
- − I couldn't fully understand how to play the game.
- − It was hard to get sentences scored, which made me frustrated.

While many children noted they enjoyed themselves, there were also several comments that it took time to understand how to play, which raises the need for improving a tutorial.

The user study also revealed some current issues of Kotoba-musubi. For instance, there is room for improvement in reducing the feeling of studying while playing the game. As one of the feedback comments suggests, we need to enhance the enjoyment of our application by introducing game elements such as a match game system. Another issue is the quality of automatic scoring. One of the possible remedies is to incorporate neural language models into automatic scoring for more flexible evaluation. We will explore how to utilize neural language models considering the computational cost.

4. Conclusion

We proposed an AI educational game for elementary school students to study Japanese writing, which fully utilizes existing language resources such as case frames, a word difficulty database, and a commonsense inference dataset. We also developed smartphone and web applications of the game and carried out a user study involving 80 pairs of elementary school students and their parents to assess it. The results of the user study demonstrated the effectiveness in reducing an aversion to writing compositions. We expect elementary school students to develop their vocabulary and reasoning skills in a ludic manner with the game. The link to the game and the supplementary material of this paper are available at https://nlp.ist.i.kyoto-u.ac.jp/EN/?Kotobamusubi.

We will address the remaining issues referring to the feedback comments and further investigate long-term educational effects such as vocabulary development. We also consider collecting unknown contingent relations through error reports.

Acknowledgements

This work was supported by the Japan Kanji Aptitude Testing Foundation. The first author was also supported by Information/AI/Data Science Doctoral Fellowship of Kyoto University and Grant-in-Aid for JSPS Fellows #22J15958.

References

National Institute for Educational Policy Research (2008). Survey on a particular subject (Japanese) results. https://www.nier.go.jp/kaihatsu/tokutei/04002010000004000.pdf (in Japanese)

Ritsumeikan University Library (2017). Detail of Reference Example. National Diet Library, Japan. https://crd.ndl.go.jp/reference/detail?page=ref_view&id=1000209543 (in Japanese)

Vandercruysse, S., Vandewaetere, M., & Clarebout, G. (2012). Game-Based Learning: A Review on the Effectiveness of Educational Games. Handbook of Research on Serious Games as Educational, Business and Research Tools. vol. 1. 628-647.

Fellbaum, C. (1998). WordNet: An Electronic Lexical Database. MIT Press.

Kawahara, D., & Kurohashi, S. (2006). Case Frame Compilation from the Web using High-Performance Computing. Proceedings of LREC 2006.

Speer, R., Chin, J., & Havasi, C. (2017). ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. Proceedings of AAAI 2017.

Morita, H., Kawahara, D., & Kurohashi, S. (2015). Morphological Analysis for Unsegmented Languages using Recurrent Neural Network Language Model. Proceedings of EMNLP 2015.

Tolmachev, A., Kawahara, D., & Kurohashi, S. Juman++: A Morphological Analysis Toolkit for Scriptio Continua. Proceedings of EMNLP 2018: System Demonstrations.

Omura, K., Kawahara, D., & Kurohashi, S. (2020). A Method for Building a Commonsense Inference Dataset based on Basic Events. Proceedings of EMNLP 2020.

Omura, K., & Kurohashi, S. (2022). Improving Commonsense Contingent Reasoning by Pseudo-data and Its Application to the Related Tasks. Proceedings of COLING 2022.

Mizutani, Y., Kawahara, D., & Kurohashi, S. (2019). Construction of a Word Database based on Educational Level Specific Words using Crowdsourcing Questionnaires. Proceedings of NLP2019. https://www.anlp.jp/proceedings/annual_meeting/2019/pdf_dir/C6-4.pdf (in Japanese)