Personalizing Augmented Flashcards Towards Long-Term Vocabulary Learning

YUICHIRO IWASHITA, Osaka Metropolitan University, Japan
ANDREW VARGO, Osaka Metropolitan University, Japan
MOTOI IWATA, Osaka Metropolitan University, Japan
KOICHI KISE, Osaka Metropolitan University, Japan

Flashcards are one of the most popular tools for learning vocabulary for second-language learners. While flashcard mediated learning is efficient, it may not induce motivation for continued use. Some researchers have proposed augmented flashcards that provide multimedia contexts to motivate people to study. However, the augmented flashcards also have a problem that it takes time to learn each target word. Understanding this tradeoff, we introduce a system that users can learn vocabulary with both standard and augmented flashcards. In addition, our system recommends the best learning strategy to users adaptively, and realizes the long-term vocabulary learning. In this paper, we describe the system and present the results of the preliminary data analysis towards the long-term vocabulary learning.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**; • **Applied computing → E-learning**.

Additional Key Words and Phrases: Vocabulary Learning, Multimedia, Intelligence Augmentation, Adaptive Learning

ACM Reference Format:
Yuichiro Iwashita, Andrew Vargo, Motoi Iwata, and Koichi Kise. 2024. Personalizing Augmented Flashcards Towards Long-Term Vocabulary Learning. In *The Augmented Humans International Conference (AHs 2024), April 4–6, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3652920.3653045

1 INTRODUCTION

Flashcards have been one of the most popular methods of learning vocabulary, especially in the context of mobile environments [5]. Since flashcards usually contain only the target word and its definition, it is efficient to study English words repeatedly in a short period (we define them as standard flashcards). However, it may be monotonous to study English words through flashcards, and reduce your motivation to learn English. Therefore, while learning English vocabulary with flashcards is efficient, it may not be suitable to continue learning English for a long term.

In order to continue learning for a long term, it is effective to motivate people to study by providing materials that interest them [4, 10]. Some researchers propose augmented flashcards that provide multimedia contexts such as example sentences and videos [8, 11]. This allows you to study vocabulary in an immersive and motivating way. However, the augmented flashcards may reduce the efficiency because it takes time to learn each target word. In addition, it is known that some people are more effective with standard flashcards while others are more effective with augmented flashcards.

Understanding the tradeoffs, we consider using both standard and augmented flashcards in a certain combination, and realize the long-term vocabulary learning. However, it is difficult for people to determine when to use which type of flashcards. Therefore, we introduce a system that adaptively recommends learning strategies to users.

In this paper, we introduce a system that personalizes augmented flashcards towards long-term vocabulary learning. We collected data from 14 university students by asking them to learn English words, and built a prediction model that can predict the optimal learning strategy. The results of the data analysis showed that the prediction model can predict the optimal learning strategy with high accuracy. The main contributions in this paper are as follows:

- We introduce a system that personalizes augmented flashcards towards long-term vocabulary learning.
- We built a prediction model that can predict the optimal learning strategy by finding the appropriate features.

In the next section, we describe the background of this research. Then, in section 3, we present the system and describe the data collection and the preliminary data analysis. Finally, in section 4, we conclude this paper and discuss future work.

2 BACKGROUND

Researchers have discussed how motivation affects learning languages [3, 7, 9]. Walkington investigated the effects of learning with interest-based learning materials [9]. In the experiment, students were divided into two groups, one group studied the material they were interested in and the other group studied the material they were not interested in. The study showed that the group of students who studied the material they were interested in solved the problems faster and more accurately than the group of students who studied the material they were not interested in.

Yamaguchi et al. and Vargo et al. proposed multimedia flashcards augmented by articles and videos of interests [8, 11]. The system collects English words that learners do not know the meaning of from media content that they are interested in and provides learners with a context-sensitive way to learn English words. They compared the effects of learning English using only context-added flashcards with those of learning English using only standard flashcards. The results showed that some learners showed more benefit from the use of context-added flashcards and some did not.

3 METHODOLOGY

3.1 System

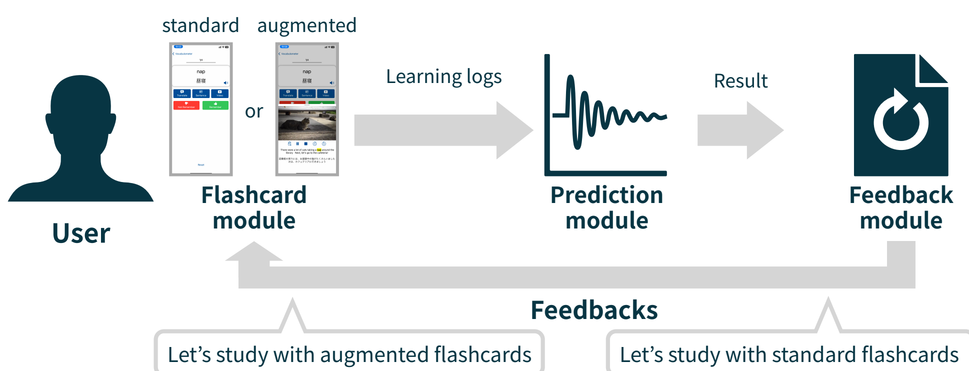

Fig. 1. Overview of our system. Users can choose standard or augmented flashcards. The prediction module predicts which type of flashcard should be used next. The feedback module recommends the user the best option.

We introduce a system that personalizes augmented flashcards towards long-term vocabulary learning. This system is an extension of a system called Vocabulometer [1, 6, 8, 11]. As shown in Figure 1, the system provides a user with two types of flashcards, standard flashcards and augmented flashcards. Then, the system predicts which to use next based on the user's learning logs. Finally, the system recommends the user to use the flashcard predicted by the prediction model.

3.1.1 Flashcard Module. This module provides the user with two types of flashcards, standard flashcards and augmented flashcards, and the user can freely choose which type of flashcard to use for each target word. The standard flashcard contains only the target English word and its translation. The augmented flashcard contains the target English word, its translation, and the example sentence or video. While the user is learning English, the module collects the learning log data of the user per target word and passes it to the prediction module. The data contains linguistic information such as the target word and its general frequency of appearance, and the user's behavioral information such as data obtained from the smartphone motion sensors and the time spent on learning the target word.

3.1.2 Prediction Module. The prediction model predicts the best learning strategy for the learner based on the learning strategies of the learner who learns efficiently and motivationally.

The model is trained in the following steps. First, the model receives the learning log data of the learner, and generates the features and label for each card. The list of features is shown in Table 1. To account for time series, the features include labels for the last three time points in addition to the learning logs described above. The label consists of two classes, which the user chooses the standard flashcard or the augmented flashcard as the next card. Then, an over-sampling method called SMOTE [2] is applied to the data to balance the number of samples for each class. Since there are many features in the data, the model selects features using the Sequential Backward Selection method. Finally, the model is trained using the Random Forest.

The prediction is performed in the following steps. First, the model receives the learning log data of the learner, and generates the features for each card, and the features are selected based on the combination of the features selected in the training phase. Then, the model predicts the best learning strategy for the learner using the Random Forest.

Table 1. A list of features the model receives for each card.

No.	Features
1-3	Linguistic information (e.g., frequency of appearance)
4-5	The position of the target word in the learning sequence
6	User subjective difficulty of the media associated with the card
7-9	Labels for the last three time points
10	Time spent on learning the target word
11-46	Data obtained from the smartphone motion sensors
47-49	Timing of the learning (month, day, hour)

3.1.3 Recommendation Module. The recommendation model recommends the best learning strategy based on the prediction of the prediction model described in section 3.1.2. It compares the selection of the user and the prediction of the model. If the user selects the same learning strategy as the prediction, the model does not recommend anything, but if the user selects a different learning strategy from the prediction, the model recommends the learning strategy predicted by the model. This aims to move users' learning strategies closer to optimal learning strategies.

3.2 Data Collection

In order to build the prediction module described in section 3.1.2, we have collected data from 14 Japanese university students in-the-wild setting. Two of them are female and the others are male, and the average age was 21.5 years old (SD = 1.53). Their English proficiency ranged widely, with a maximum TOEIC Listening & Reading test score of 815, a minimum score of 225, and an average score of 616.

During the data collection, the participants were asked to freely use the two types of flashcards described in section 3.1.1, and we collected the learning log data. 10 participants used their own iPhone, and the others who did not have an iPhone used the iPhone provided by us. The data collection lasted for 4 weeks, and for each week, participants were asked to use the system for the first 5 days and take a vocabulary test on the 6th day. The last day of the week was set as the rest day, and the participants were asked not to use the system on that day.

3.3 Preliminary Data Analysis

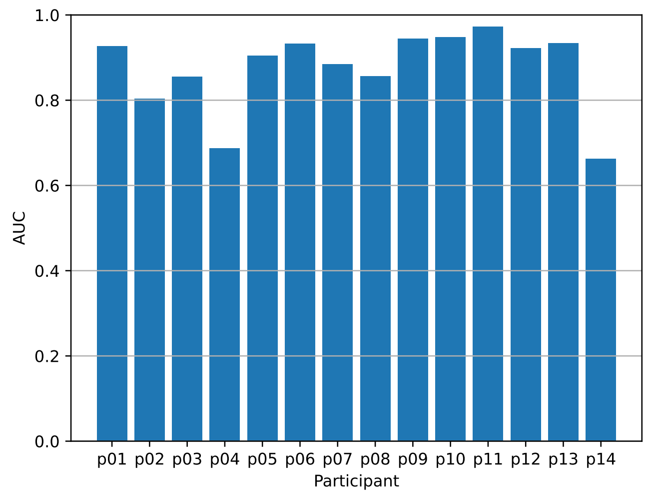

Fig. 2. The AUC values of the prediction module for each participant.

By using the data collected in section 3.2, we evaluated the prediction module described in section 3.1.2. The evaluation was conducted by leave-one-participant-out cross-validation, a cross-validation method in which the data of one participant is used as the test data and the data of the other participants is used as the training data.

Figure 2 shows the results of the evaluation. It shows the area under the ROC curve (AUC) value for each participant. According to the results, the AUC was greater than 0.8 for 11 of the 14 participants, and the AUC was greater than 0.9 for 8 of these 11 participants. The average AUC of the proposed method was 0.87 with a standard deviation of 0.09. Therefore, it is suggested that the prediction module can predict the best learning strategy accurately.

4 CONCLUSION

In this paper, we introduced a system that provides feedback to learners who are less effective in learning English based on the English learning methods of those who are effective in learning English, we predicted the next English learning method that learners would choose based on their behavior, and investigated the best index for the prediction. The experimental results suggest that using features generated from English vocabulary and smartphone sensor data in addition to English learning time can predict English learning with higher accuracy than using only English learning time as a feature. The results also suggest that the prediction accuracy is high for a large number of learners.

ACKNOWLEDGMENTS

This work was supported in part by grants from JST Trilateral AI Research (Grant No. JPMJCR20G3), JSPS Fund for the Promotion of Joint International Research (Fostering Joint International Research (B)) (Grant No. 23KK0188), Grand Challenge of Initiative for Life Design Innovation, MEXT Society 5.0 Realization Research Support Project, Osaka University.

REFERENCES

[1] Olivier Augereau, Clément Jacquet, Koichi Kise, and Nicholas Journet. 2018. Vocabulometer: A Web Platform for Document and Reader Mutual Analysis. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*. IEEE Computer Society, Los Alamitos, CA, USA, 109–114. https://doi.org/10.1109/DAS.2018.59

[2] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 16 (June 2002), 321–357. https://doi.org/10.1613/jair.953

[3] Zoltán Dörnyei. 1998. Motivation in Second and Foreign Language Learning. *Language Teaching* 31, 3 (July 1998), 117–135. https://doi.org/10.1017/S026144480001315X

[4] Ulrike Hammerschall. 2019. A Gamification Framework for Long-Term Engagement in Education Based on Self Determination Theory and the Transtheoretical Model of Change. In *2019 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, Dubai, United Arab Emirates, 95–101. https://doi.org/10.1109/EDUCON.2019.8725251

[5] Marissa K. Hartwig and John Dunlosky. 2012. Study Strategies of College Students: Are Self-Testing and Scheduling Related to Achievement? *Psychonomic Bulletin & Review* 19, 1 (Feb. 2012), 126–134. https://doi.org/10.3758/s13423-011-0181-y

[6] Clément Jacquet, Olivier Augereau, Nicholas Journet, and Koichi Kise. 2018. Vocabulometer, a Web Platform for Ubiquitous Language Learning. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp '18)*. Association for Computing Machinery, New York, NY, USA, 361–364. https://doi.org/10.1145/3267305.3267614

[7] Rebecca Oxford and Jill Shearin. 1994. Language Learning Motivation: Expanding the Theoretical Framework. *The Modern Language Journal* 78, 1 (1994), 12–28. https://doi.org/10.2307/329249 jstor:329249

[8] Andrew Vargo, Kohei Yamaguchi, Motoi Iwata, and Koichi Kise. 2024. A Context-Based Multimedia Vocabulary Learning System for Mobile Users. *Informatics* 11, 1 (March 2024), 1. https://doi.org/10.3390/informatics11010001

[9] Candace A. Walkington. 2013. Using Adaptive Learning Technologies to Personalize Instruction to Student Interests: The Impact of Relevant Contexts on Performance and Learning Outcomes. *Journal of Educational Psychology* 105, 4 (2013), 932–945. https://doi.org/10.1037/a0031882

[10] Keith Weber. 2003. The Relationship of Interest to Internal and External Motivation. *Communication Research Reports* 20, 4 (Sept. 2003), 376–383. https://doi.org/10.1080/08824090309388837

[11] Kohei Yamaguchi, Motoi Iwata, Andrew Vargo, and Koichi Kise. 2020. Mobile Vocabulometer: A Context-Based Learning Mobile Application to Enhance English Vocabulary Acquisition. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*. ACM, Virtual Event Mexico, 156–159. https://doi.org/10.1145/3410530.3414406

Received 22 February 2024; accepted 28 February 2024