

# Effectiveness of Vision-Based Word Learning Using Head-Mounted Display

Naoki SEKIGUCHI, Keiji YASUDA, Naoki KAWAI & Hiroaki KIMURA

*KDDI R&D Laboratories. Inc.*

na-sekiguchi@kddilabs.jp

**Abstract:** We propose a word learning system using a lightweight head-mounted display (HMD). The system displays the relevant foreign word corresponding to the image recognized within the learner's field of view. We carried out an evaluation by comparing the number of memorized words under different conditions: i.e., where the words were related or unrelated to the learner's field of view using the HMD or a smartphone. The results of the experiments showed the effectiveness of the proposed system.

**Keywords:** Head-mounted Display, Image Recognition, Word Learning, Mobile Learning

## 1. Introduction

Many studies have examined the effectiveness of vocabulary learning methods using mobile devices, such as smartphones. Thornton and Houser (2005) sent three short mini-lessons via e-mail to university students each day, but only 10% of the subjects reported reading the messages each time. Lu (2008) sent two vocabulary lessons using the short message service every day and confirmed that there was a large deviation in study frequency, and it was clear that the lower the study frequency, the lower the vocabulary gain. According to these experiments, the problem was that the frequency of learner responses was low. If learners were enthusiastic about something other than learning or the learners' hands were not available, they ignored the message sent to their mobile devices. The possible cause for the infrequency of study was the small but troublesome action of taking the device out of their pockets or lack of interest. On the other hand, a context-aware learning system has been reported as an effective way of increasing learning opportunities. Li et al. (2012) developed the SCROLL system of recall learning quizzes via context i.e. location, time, and so on. Although the system contributed to an increase in learner responses, the consideration of gaining learner interest was not sufficient.

In this paper, we propose a vision-based word learning system that displays the foreign word related to the learner's field of view on a head-mounted display (HMD) to increase learning opportunities and learner interest. The latest lightweight HMD, such as Google Glass, can provide information to learners without any physical action. Hence, learners will not ignore the received information displayed on the lightweight HMD. In addition, the HMD can capture the learner's field of view as a still image using the built-in camera. Therefore, learners can acquire foreign words within the learner's direct field of view by using the HMD. Learning through experience and in real life contexts significantly facilitates the learning process (Brown, Collins, and Dugui 1989). Therefore, vocabulary acquisition is more effective for learners when the words are related to their activities, for example, when picking out a product from store shelves or looking at interesting signs. Little research has been conducted to survey the effect of learning using HMDs. Therefore, we evaluated the system by comparing the number of memorized words under different conditions. This study aims to answer the following research questions: (1) How much of an effect does the difference in visibility or comfort between HMD and smartphone have on learning and (2) is learning foreign words related to the learner's field of view a more effective method than showing words regardless of the context?

## 2. Word learning system under three different conditions

In order to explore the above research questions, we developed a word learning system as an Android application and prepared the following three different conditions.

- Condition A: Words only using a smartphone
- Condition B: Words only using an HMD
- Condition C: Words related to the learners' field of view using an HMD (proposed system)

Condition A uses a smartphone that displays the information as shown in Figure 1 (a), and conditions B and C use an HMD that displays the information as shown in Figure 1 (b), (c), respectively. In condition C, the proposed system captures the learner's field of view as a still image using a camera built into the HMD and immediately checks whether a part of the image matches the reference image in a prepared data set via image recognition technology based on binary local features (Uchida and Sakazawa 2013). The data set contains a number of reference images and relevant English-Japanese word pairs. A set of these processes is sequentially repeated about every second. If the image matches, the system selects the relevant word pair for display on the HMD in three seconds so that the learner can understand the information in the image. Thus, in condition C, the system displays the foreign words within the learner's field of view.

In conditions A and B, the learner learns the word regardless of the context. In the case of conditions A and B, the system provides English and Japanese word pairs for an eight-second interval. Then the system shows another word pair after two seconds. In the case of condition C, the learner can see the word on the HMD when looking at one of the 15 prepared reference pictures on the wall of the experiment room. For example, when the learner watches a picture of passports, the HMD displays the word "immigration." And then, learners memorize the word in each condition. Through the experiments, our research question (1) will be addressed by comparing the number of memorized words in conditions A and B, and research question (2) will be addressed by comparing the number of memorized words in conditions B and C.

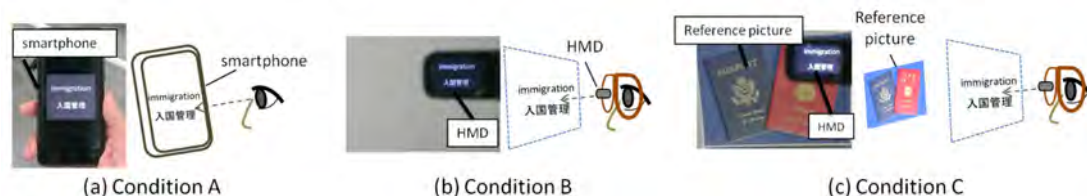


Figure 1. The view that a learner sees and schematics under each conditions

### 3. Experiments and result

The subjects were issued instructions on the procedures for the experiments. Before starting the experiment, all subjects became accustomed to wearing the HMD to accustom them to seeing images through it as naturally as possible. We divided subjects into two groups. One group conducted experiments A, B, and C in that order (first using a smartphone). The other group conducted experiments B, C, and A in that order (first using the HMD). The word test was conducted before and after each experiment, and the number of memorized words was enumerated. Other experimental conditions were as follows.

- Number of subjects: Eight Japanese college students
- Data set of word translation pair: 45 (15 words/each condition of A to C) of high difficulty i.e. low frequency, English nouns were selected from JACET 8000. Data sets used in each condition varied with the subjects.
- HMD: Vuzix M100 Smart Glasses

In order to evaluate the effectiveness of the system, the subjects took the same word translation tests before and after each experiment. We consider the increased numbers of correctly answered words as learning effects. We performed a Wilcoxon signed-rank test to compare the learning effects of the three conditions. The results of the test are shown in table 1. Here, the average number of memorized words under each condition and the significant difference between three pairs of conditions are shown. The results indicated no significant difference between conditions A and B. However, condition B had the potential of being more effective for learning because the information sent to the smartphone was often ignored whereas the information sent to the HMD was seldom ignored. The results of the test between conditions B and C showed a marginally significant difference at a significance level of 5% in a one-sided test. These results indicated that the proposed system was more effective for learning words than just showing the word regardless of the context. The reason was that the system made it easier to memorize words by associating foreign words with an image within the learner's actual field of view. However, we did not confirm a significant difference between conditions A and C. Some subjects did

not concentrate on learning because the HMD was uncomfortable, even though they had some practice using it before the experiments.

After the experiments, we conducted a survey in the form of a questionnaire, e.g., Q1: Is the HMD easy to wear? and Q2: Is the HMD a useful device for learning foreign words? The answers were collected and scored using a five-point scale. The results of the questionnaire are shown in Figure 2. The responses to Q1 indicated that only two subjects considered the HMD comfortable to wear; the others regarded the HMD as uncomfortable. In fact, some subjects took the time to acquire the ability to properly watch the display. On the other hand, the result of Q2 indicated that half of the subjects regarded the HMD as a useful device. The two subjects who answered, "The HMD is not useful for learning foreign words" also remarked as follows.

- "This HMD is not comfortable."
- "I was distracted from learning because I was unfamiliar with the HMD."

Considering these points, the proposed system was apparently a more effective learning system once the user became more accustomed to the HMD.

Table 1 : Resultt of Wilcoxon Signed Rank Test

Condition	M	SD	Amount of statistics	
			compared with A	compared with B
A	6.5	2.5	N/A	** -
B	6.25	3.07	**	N/A
C	8	2.55	**	*

\* Significance level of 5 %

\*\* No significant difference

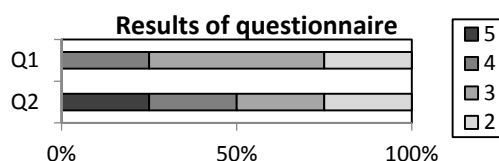


Figure 2. Results of questionnaire

#### 4. Conclusions

In this paper, we proposed a word learning system using a lightweight HMD that displayed a foreign word within the learner's field of view. After comparing the number of memorized words among three different conditions, we concluded that the proposed system was more effective than simply showing the word regardless of the context. We intend to carry out experiments to evaluate the learning opportunities using the system as a future research task.

#### Acknowledgements

We would like to thank Yukimi Imaoka, assistant director of the Media Education Center, Kanda Institute of Foreign Language; Dai Ujihara, Graduate School Administration Office, Digital Hollywood University; and the students of these institutions for the support we received in conducting our experiments.

#### References

- Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *JCAL*, 21(3), 217-228.
- Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *JCAL*, 24(6), 515-525.
- Li, M., Ogata, H., Hou, B., Uosaki, N., & Yano, Y. (2012). Personalization in context-aware ubiquitous learning-log system. *WMUTE, 2012 IEEE Seventh International Conference on* (pp. 41-48). IEEE.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, 18(1), 32-42.
- Uchida, Y., & Sakazawa, S. (2013). Image Retrieval with Fisher Vectors of Binary Features. In *Pattern Recognition (ACPR), 2013 2nd IAPR Asian Conference on* (pp. 23-28). IEEE.
- JACET Basic Words Revision Committee. (2003). JACET List of 8000 Words (JACET 8000).