Vocabulary Learning Support System that Utilizes Suggestibility of Error by Image Generation

2110094 Kazuki Sugita

With the development of globalization, there is a need to use a common language to communicate with each other. English occupies the status of a common language in the world. English is spread worldwide in almost all fields such as science, engineering, technology, medicine, and education. Vocabulary learning has long been taught as a basis for L2 English language learning. To acquire vocabulary effectively, it should be learned through context.

However, learning through context requires the learner must learn through tons of reading. This method is effective for vocabulary learning, but for this reason, it is difficult for L2 learners. Another hand, a widely known method for learning English vocabulary is based on translations. Learners acquire English vocabulary through translation, a method that has been taught in school education for many years. However, such translation-based methods are not suitable. This is because, for example, native English speakers and native Japanese speakers have different ways of interpreting the same thing or situation. The same thing or situation is interpreted differently in each culture.

One of the methods of vocabulary learning that does not rely on these translated texts is to use visual clues. There is some research proposing a system utilizing image captioning, which generates captions from images. They regard the caption and images as the context of learning vocabulary.

In vocabulary learning in English by L2 learners, errors are perceived to occur even among learners at a high proficiency level. The similarity of the spell is one of the main difficulties of learning vocabulary, Japanese learners tend to figure out their meanings by looking at the words because Japanese words are ideograms. Another hand, English words are phonograms the meaning of the word cannot be understood by looking at the spell. The reason why spelling errors occur regardless of proficiency level has been attributed to the phonological awareness of native Japanese speakers.

Such causes of error in vocabulary learning can be the serious problem of fossilization. Fossilization makes errors a part of natural language and makes it difficult for L2 learners to improve their errors. Therefore, we need to consider the lack of suggestions for errors to be a serious problem for L2 learners.

Horiguchi et al. proposed an Error-Based Simulation focusing on the implication of error and its effectiveness for learning a physics task. In this study, the difference between the learner's prediction and the correct phenomenon is visualized by simulation and linked to error correction. This effectively helps the learner understand the cause of the error. We hypothesized that applying this visualization of learners' errors to vocabulary learning would prevent fossilization in English vocabulary learning.

In this research, we developed a vocabulary learning support system with image generation, LVIG(Learning Vocabulary with Image Generation) that automatically generates images corresponding to the learner's errors and shows these images to the learner to encourage reflection. We hypothesize that the difference between the image of an error and the image of a correct answer will be retained as an impressive memory.

Assuming that the reason for second language speakers' errors is spelling mistakes, we create an environment in which learners are more likely to cause errors. This environment allows learners to correct their mistakes.

We hypothesize that the difference between the image of an error and the image of a correct answer will be retained as an impressive memory. It is realized by the image generation model. To build such a system, we defined three tasks. The first is a Question Generation task, the task is for making learning material automatically from the MS-COCO dataset. To accomplish it, we define the vocabulary list COCO-word list by utilizing Natural Language Processing. We also made the module to generate a question automatically. The task is to extract the target word from the sentence from MS COCO and replace it with the blank. The module also utilizes Natural Language Processing. The second is an Answering Support task, which supports the learner in answering the question by extracting the candidate's word by similar spelling. It enables us to learn vocabulary easily through the candidate's words. The third is an Image Generation task. Utilizing the image generation model DALL-E2 which realizes generating images corresponds to learners' errors.

In addition, we conducted the experiment with 25 subjects. The results of the experiment were analyzed for 22 of the 25 subjects. The reason for this is that we excluded subjects whose data were missing due to a system error. The subjects in the experiment were graduate students with a high level of proficiency in English. The experimental setting consisted of two models. The first model includes image generation. When the learner makes an incorrect answer, the image generated corresponds to the sentence with an error. While the second model is a conventional model that does not include image generation. The vocabulary for the experiment consisted of a total of 100 high school to college-level questions. Eight parts of speech were used. In order to evaluate whether repetitive errors could be prevented, questions consisting of 20% error-only choices were included. The same subject tried on the two conditions and repeated them twice. Each condition is assigned 50 questions, 100 questions per term, and this is repeated twice for a total of 200 questions. Subjects completed a questionnaire when answering the questions. 1. The questions asked whether the images in the COCO dataset were associated with a sentence, 2. Whether the images generated during the wrong answer were associated with a sentence, 3. Whether the images generated during the wrong answer were impressive. In order to eliminate the effects of the order of the questions and the characteristics of the subjects, we adopted a counterbalancing technique. In this case, the learner solves the same problem twice.

To evaluate the learning effect, we measured whether the learner made the same mistake repeatedly when solving a problem in the second term as they did in the first term. To evaluate the learning effect, we compared the two models based on the number of vocabulary that the learners repeatedly got wrong. The Wilcoxon signed-rank test was used as the test method for the evaluation. The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test used either to test the location of a population based on a sample of data or to compare the locations of two populations using two matched samples. We also analyzed the effects of associativity and impressions of images on vocabulary learning, based on questionnaires obtained from subjects. To evaluate the model, we assessed whether the images represented the content of the sentences based on associativity collected from the subjects' questionnaires.

Our experiment result showed that the proposed method can prevent repetitive wrong

answers by comparing the proposed method with a model that includes image generation and a conventional method. We also found that the images in the existing dataset and the images generated by the image generation model are of high quality. On the other hand, due to its high accuracy, the proposed model has the potential to induce learners to learn vocabulary incorrectly when the learning content contains errors. Although these problems occurred because of the high reproducibility of the model, these experiments showed that our proposed method effectively prevents repetition errors and is effective in English vocabulary learning.

In conclusion, Recent technology contributes to the learning support system. A large amount of the data set also has a high potential for adopting education. Our proposed method utilizing MS COCO and sentence-to-image generation model has the possibility to enhance the learning vocabulary. In this study, we proposed a method consisting of three tasks to build a vocabulary learning support system that can suggest errors and learn from them by automatically generating images corresponding to the error. We described the possibility of using existing data resources to create word lists by difficulty level. We also described how these data resources could be used to generate questions with blanks for use in vocabulary learning automatically. We believe that the ability to reproduce concepts through the application of an image generation model has the potential to replace the experiential learning of our native language by enabling the conceptual acquisition of a second language. We aim to develop a vocabulary learning support system that can help L2 learners learn English vocabulary more effectively by using image generation to understand English concepts.