So, H. J. et al. (Eds.) (2020). Proceedings of the 28th International Conference on Computers in Education. Australia: Asia-Pacific Society for Computers in Education

Development of Japanese Dictogloss Learning Support Environment for Pronunciation Learning of Japanese Speech

Satoru KOGUREa*, Hiroki HAKAMATAa, Yasuhiro NOGUCHIa, Tatsuhiro KONISHIa, Makoto KONDOa &, Yukihiro ITOHb
[a]*Faculty of Informatics, Shizuoka University, Japan*
[b]*Shizuoka University, Japan*
*kogure@inf.shizuoka.ac.jp

Abstract: In this paper, we describe the extension of the existing pronunciation evaluation framework of Japanese dictogloss support environment. To evaluate the pronunciation in the existing environment, it was necessary for the teacher to manually create a dictionary for recognition that include incorrect pronunciation. In addition, there were few variations in the dialogue part that indicated an error in the reproduction sentence of the collaborative learner agent and the sentence of pronunciation evaluation feedback. To solve these problems, we improved the system. We evaluated the system we created and obtained the results that suggest that the proposed method is effective.

Keywords: Dictogloss, pronunciation evaluation, language learning support, Japanese learning

1. Introduction

Dictogloss is an activity in language learning advocated by Wainryb (1988, 1990). An advantage of dictogloss is that it can teach speaking skills in addition to reading, writing, and listing. There are many practical reports on its application in the language-learning field (Snoder & Reynolds, 2019, Ajmal et al., 2020). However, in dictogloss, the learner needs a teacher and a collaborative learner. Therefore, it is difficult to use this activity when alone. To address this problem, we have developed a dictogloss system (Kondo et al., 2012; Tashiro et al., 2013, Kogure et al., 2015, 2016, 2017, 2018) that does not require teachers and collaborative learners. Our dictogloss environment has two agents: *a collaborative learner agent* (*CLA*). and *a teacher agent* (*TcA*). However, our dictogloss environment faces the following has three problem: (**P1**) the cost of generating a speech recognition dictionary for pronunciation evaluation is high. (*for teaching material creators*), (**P2**) there are limited interaction patterns that learners can experience. (*for learners*), and (**P3**) pronunciation evaluation feedback is binary (*for learners*). The purpose of our research is to solve these three problems, and in this report, we propose a method to solve them. We also implement a system that utilizes the method. In addition, we evaluate the system we created.

2. Dictogloss Learning Environment

Dictogloss is a learning activity advocated by Wajnryb (1988, 1990). It contains four activities:

(Act.1) The teacher reads a short text aloud (including some sentences).
(Act.2) Each learner listens to the text while taking notes.
(Act.3) Each learner reconstructs the original text while consulting with another learner.
(Act.4) The teacher reviews the reconstructed sentences in text and provides feedbacks to learners

There are many class practice reports that utilize dictogloss in actual language classes. Consider providing information and communication technology support for dictogloss learning activities. First,

in *Act. 1*, we can realize a system in which the learner can freely reconstruct by recording the teacher's speech. Next, in *Act. 2*, we can prepare an environment in which the learner writes down the words heard. However, in *Act. 3*, it is necessary to build an environment in which it is possible to interact with collaborative learners. Therefore, this step is difficult to realize. Also, *Act. 4* requires the same high level of difficulty in realization as in *Act. 3* because it requires analysis of the learner's utterance in *Act. 3*. We have devised a method that simplifies the interaction with collaborative learners in *Act. 3* and the method of analyzing the learner's input in *Act. 4*. We constructed a Japanese dictogloss learning support environment that supports from *Act. 1* to *Act. 4* according to the method (Kogure et al., 2018).

In *Act. 2*, the learner utters the sentence reconstructed by himself or herself. With this function, if the learner does not enter the reconstructed sentence correctly (or if it contains an unexpected error), the reconstructed sentence cannot be recognized correctly. Therefore, the learner can utter the reconstructed sentence only when the she or he inputs the reconstructed sentence within the range of the assumed error. In *Act. 3*, the learner compares the reconstructed sentences input by himself or herself and the collaborative learner. The learner specifies different morphemes and points out errors to the collaborative learner. At that time, the learner can ask "Was there WORD in Sn?". The system recognizes the pointed utterance by speech and then identifies the word for which the error was indicated the system generates a click event of the morpheme button corresponding to that word. In *Act. 4*, the system analyzes the learner's voice recognition results collected in *Act. 2* and *3* and evaluates pronunciation (Kogure et al., 2018).

3. Improvement of Pronunciation Evaluation

The teaching material creator (1) generates a speech recognition grammar and a dictionary with correct pronunciation, and (2) a pronunciation dictionary that includes erroneous pronunciations that are easily mistaken by a native speaker. They have to generate manually, and it especially takes time to complete (2) task for them. For example, it took an average of about 500 seconds for the author to create the incorrect pronunciation dictionary (for Korean native language) for 3 lessons out of 22 lessons prepared. We expect it to be more expensive to be created by someone who is not familiar. In addition, we judge that there is a possibility that the points to be corrected may be omitted or the quality of the created dictionary may be deteriorated by manual correction. In addition, since phonetic transcriptions that are expected to be incorrect differ depending on the mother tongue, it is necessary to manually correct each time the mother tongue changes. Therefore, when a mother tongue is specified, a framework that can automatically generate an erroneous pronunciation dictionary that considers pronunciation prone to error in the mother tongue is incorporated. The system only allows the learner to utter "Was there WORD in Sn?" and cannot use other utterances. Therefore, we increase the number of variations of dialogue that learners can experience. The actual pronunciation evaluation is performed on a word-by-word basis, but the confidence measure of pointing out feedback uses sentence confidence measure. The system does not look at word confidence measure. Therefore, we improved the usage of speech recognition. The system estimates the confidence measure for each word using the speech recognition results up to 5 best results.

4. Experimental Evaluation

For lessons 1–3, one of the authors manually created a speech recognition dictionary containing incorrect pronunciations. The generation time was 303 seconds for Lesson 1, 488 seconds for Lesson 2, and 741 seconds for Lesson 3. On the other hand, it took less than 1 second for all three lessons to automatically generate the dictionary. This result shows that **P1** was solved.

Regarding the part that points out the error in CLA in *Act.3*, we let four Japanese people use utterances that point out the error due to voice utterance in the preceding and the proposed system. After that, we acquired the goodness of dialogue by increasing the number of styles that can be spoken. We prepared two target lessons. Considering the order effect, we changed the order of the lessons used and the order of the system used by all four subjects. We asked the four subjects to evaluate whether each system was easy or difficult to utter to point out a mistake, using 1 to 5 points (1 meaning that previous system is very good, and 5 that the proposed system is very good). We obtained 2.25 points in the

preceding system and 3.75 points in the proposed system. This result suggests that increasing the variation of utterances that point out errors in *Act. 3* may improve the ease of pointing out by learners.

We made the same four subjects input the reconstructed sentence prepared here in *Act. 2*. Then, we read the reconstructed sentence to the subject with the pronunciation prepared here (both correct and incorrect). All four of them read with the pronunciation presented here (all the same). After that, we made them read the text of the indication of the correctness judgment of the feedback of *Act. 4*. In this experiment, the preceding system and the proposed system were used in order. Considering the order effect, two subjects used the preceding system first, and the other two used the proposed system first. After use, we asked the learners to evaluate whether the utterance pointed out was useful for language learning (role-playing a learner who is likely to make a specified pronunciation error) on a scale 1-5 (1 meaning that previous system is very good, and 5 that proposed system is very good). We obtained 3.50 points in the preceding system and 4.75 points in the proposed system. This result suggests that increasing the variation of utterances that point out errors in *Act. 3* may improve the ease of pointing out by learners, implying this result suggests that the proposed system could give a more appropriate feedback depending on the pronunciation of each word than the previous system.

5. Conclusion

In this research, we construct three improvements in the pronunciation evaluation framework of the existing Japanese dictogloss support environment. The first is the automatic generation of an erroneous pronunciation dictionary, the second is an increase in variability in utterances that point out mistakes to CLA, and the third is an increase in feedback variations for erroneous pronunciations. We evaluate each improvement and confirmed the usefulness of the system. It is necessary to conduct an evaluation experiment for international students who are learning Japanese. Also, we are currently planning to develop a dictogloss support environment for English.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 17K00483.

References

Ajmal, M., Iqbal, I., Mahmood, R., & Zafar, A. (2020). The effectiveness of using dictogloss technique to improve students' writing: a case study from University of Lahore. *Dilemas Contemporáneos: Educación, Política y Valores, 7*(2).

Kogure, S., Miyagishima, K., Noguchi, Y., Kondo, M., Konishi, T. & Itoh, Y. (2016). A teachable agent for the Japanese dictogloss learning support environment. *Proceedings of ICCE2016*, 88–90.

Kogure, S., Nakahara, T., Noguchi, Y., Konishi, T., Kondo, M. & Itoh, Y. (2018). Application of speech recognition in a Japanese dictogloss system, *Proceedings of ICCE2018*, 292–297.

Kogure, S., Okugawa, K., Noguchi, Y., Konishi, T., Kondo, M. & Itoh, Y. (2017). Improvement of the situational dialog function and development of learning materials for a Japanese dictogloss environment, *Proceedings of ICCE2017*, 104–106.

Kogure, S., Tashiro, A., Noguchi, Y., Kondo, M., Konishi, T. & Itoh, Y. (2015). An answer support environment based on grammar, context and situation for a dialogue to learner agent on Japanese dictogloss system. *Proceedings of ICCE2015*, 94–96.

Kondo, M., Sano, R., Tashiro, A., Noguchi, Y., Kogure, S., Konishi, T. & Itoh, Y. (2012). Development of a dictogloss system oriented for focus on form. *Proceedings of ICCE2012*, 1–8.

Snoder, P., & Reynolds, B. L. (2019). How dictogloss can facilitate collocation learning in ELT. *ELT Journal, 73*(1), 41–50.

Tashiro, A., Noguchi, Y., Kogure, S., Kondo, M., Konishi, T. & Itoh, Y. (2013). Evaluation of an improved dictogloss system oriented for focus on form. *Proceedings of ICCE2013*, 110–114.

Wajnryb, R. (1988). The dictogloss method of language teaching: A Text based communicative approach to grammar, *English teaching forum, 26*, 35–38.

Wajnryb, R. (1990). Grammar dictation. *Oxford: Oxford University Press*.