Chen, W. et al. (Eds.) (2017). Proceedings of the 25th International Conference on Computers in Education. New Zealand: Asia-Pacific Society for Computers in Education

Designing a Personalized Language-Learning Environment for an Automatic Quiz Generation System Utilizing Online Video Content

Ai NAKAJIIMA[a*] & Yuichi ONO[b]
[a]*School of Engineering, University of Tokyo, Japan*
[b]*Center for Education of Global Communication (CEGLOC), University of Tsukuba, Japan*

Abstract: In this paper, we describe our work-in-progress research on the construction of the foreign language–learning support system "YouTutors," which is intended to enhance personalized learning to support a learner's vocabulary acquisition by using online educational video platforms such as YouTube and TED Talks. Our research so far has demonstrated that the proposed system, which automatically generates listening and vocabulary quizzes from online video content, has been successful in raising learners' motivation for self-study. However, we have not sufficiently incorporated data about learning behavior and learners' proficiency levels. This work-in-progress paper describes the improved approach of how the system incorporates these data to generate quizzes to maximize the potential for providing a personalized learning environment for an individual learner.

Keywords: Personalized Learning, Automatic Quiz Generation, Online Video Content, Open Educational Resources, Creative Commons, Web Mining

1. Introduction

In this research, we propose a method of re-utilizing online videos for foreign language learning. Our earlier papers have dealt with the development and evaluation of an automatic quiz generator (AQG) based on open educational resources (OER) on the Web for Japanese English as foreign language (EFL) learners. The research started with TEDNote (Nakajima & Tomimatsu, 2012) as a first prototype and continued with YouTutors (Nakajima & Ono, 2014) as an application for practical use in university's English courses. After some modifications, Ono, Nakajima and Ishihara (2017) demonstrate that YouTutors raises learners' motivation to look for new materials among TED Talks videos based on their interests.

1.2 Overview of the Application

The current version of YouTutors has been upgraded to a learning management system (LMS) that contains the following functions, as shown in Figure 1: (i) a Web mining system, which constructs educational material out of scraped data from streaming video services such as TED.com or YouTube.com, (ii) an AQG system, which generates English listening exercises in the forms of a typing game, a word-ordering game and a fill-in-the-blank game, (iii) a game system, with which users play the listening and vocabulary exercise with streaming videos, and (iv) a portfolio system, which visualizes the result and the user's learning behaviors in each exercise.

To create a new exercise, user needs to copy and paste a URL of a specific video into the entry form of YouTutors and then the system analyzes the subtitle data of the video. The system creates each stage by splitting the video into each short video clips for each sentence of subtitle. Then, the system generates a new quiz in the chosen game mode format by converting subtitle data of each stage. Once the game starts, the students interactively take the listening quiz of a stage, which corresponds with the short video clips loaded in the center of the interface, as shown in Figure 2. Learner is allowed to use several helpful functions such as "Repeat," which plays the video clip repeatedly until the he or she catches the answer, and after answering all quizzes of a stage, he or she can challenge the next stage.

During the game, the application monitors the use of the following functions as students' learning behaviors: (i) words typed incorrectly, (ii) words replayed by the "Repeat" function, (iii) words cheated with the "Hint" function, (iv) words checked with the "Dictionary" function, (v) and words

stored in “Bookmarks.” All the marks of learning behaviors on each exercise are mapped in the “Portfolio.”

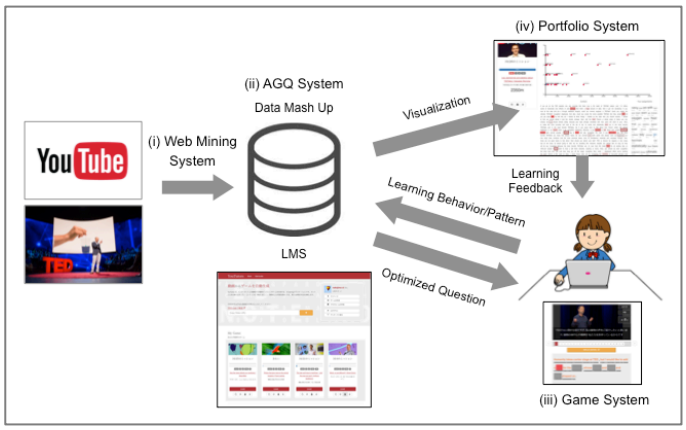

Figure 1 System Design of YouTutors

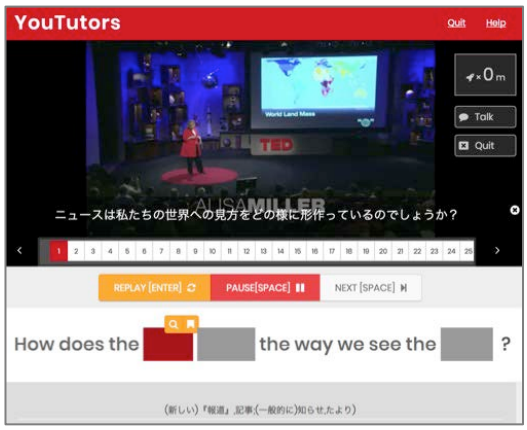

Figure 2 Interface of Game System (Fill-in-the-blank Game)

2. Challenges

Although we were successful in demonstrating the usefulness and merits of employing the system in actual English courses as suggested in the papers above, Ono, Nakajima and Ishihara (2017) also mention some required challenges, which are based on interview data collected from students who participated in a user evaluation test of one-week continuous study with YouTutors. Most students showed a high interest in continuing to use YouTutors for their self-study, but they said it was arduous to complete each game every day. Some students strongly felt that authentic English content from TED Talks is difficult to follow, depending on their choice. To encourage students’ continued learning, the application should have the system to suggest appropriate video content among its huge collection and the AGQ system should be designed to generate the optimal level of questions with moderate learning stress according to a student’s level and needs. In this work-in-progress, we describe about our approach for the latter challenge especially.

Ono and Nakajima (2017) argue that the following three points need to be carefully considered when we transform content taken from the Web into educational material for an EFL classroom: (i) authenticity, (ii) diversity and (iii) multimodality. In particular, points (ii) and (iii) are directly relevant to the issue of how to minimize the gap between students’ expected target words and the actual question items generated by the application. From a learning analytics point of view, it has been recommended that students’ learning behaviors should be reflected visually and taken into account to produce more optimal questions for a student, depending on the student’s learning patterns (Ono, Nakajima & Ishihara, 2017).

3. Our Approach

With the challenges described in the last section in mind, this paper proposes and constructs an advanced version of the AQG, integrating the following components to maximally optimize the difficulty of the question items provided by the system by incorporating students’ learning patterns and natural language processing (NLP) technology in order to create smoother and more comfortable learning. Using a fill-in-the-blank quiz as an example, in order to determine the proper position of blanks and number of them in each quiz, we need to estimate the level of each student’s stress and proficiency. Throughout the investigation, based on user interviews and the learning behavior observed in their portfolios, we found that we can create a hypothesis to estimate their level and status using a combination of learning habits.

For example, in order to predict user proficiency level, we focused on the relationship between the use of the “Dictionary” and “Bookmark” functions after the incorrect answer (Table 1.) We utilized the JACET8000 index to evaluate each vocabulary level. Based on the hypothesis as illustrated in Table 1, the system monitors the JACET index of each vocabulary and calculate an average index in each

category to estimate user current proficiency level. To optimize occurrence probability of blank for each vocabulary in the next sentence, system applies a decision-tree algorithm with updated thresholds.

Table 1. Hypothesis for setting a blank probability of each vocabulary based on users' learning behaviors

Behavior			**Hypothesis**	**Quiz Design**	
Answer	**Dictionary**	**Bookmark**	**Student Condition**	**Level Design**	**Occurrence probability of blank**
Mistake	Yes	No	The student did not know the meaning, but he thinks he does not need to memorize it yet.	Non-targeted Level	Eliminate the occurrence probability of blank
Mistake / Correct	Yes	Yes	The student was not sure of a meaning and wants to memorize it.	Challenge Level	Decrease the occurrence probability of blank
Mistake / Correct	No	Yes	The student knows the meaning, but he / she felt it was necessary to bookmark the word to strengthen his / her memory.	Targeted Level	Increase the occurrence probability of blank
Mistake	No	No	The student knows the meaning of the word, and he does not need to memorize it.	Current Skill Level	Give the highest occurrence probability of blank
Correct	No	No	The student knows the meaning of the word, and he does not need to memorize it.	Non-targeted Level	Eliminate the occurrence probability of blank

4. System Development

We developed the application interface (front end) with JavaScript as same as the previous version but the PHP back end was replaced with Python in order to introduce NLP and data mining technology. As an initial AQG setting, the occurrence probability of blank was influenced by the grammatical role of words, which was analyzed by the Natural Language Toolkit (NLTK). Vocabularies, which have closer indexes to 1500 (the High school elementary level in JACET8000), are given the highest occurrence probability of blank.

5. Conclusion

This paper describes our updated version of YouTutors to provide a personalized learning environment based on learning behavior. Our pilot evaluation shows that the system works fine. The next task is to carry out an experiment to validate its optimality as a personalized learning support system in an actual classroom.

References

Nakajima, A., & Ono, Y. (2015). The prospect of open online e-learning system based on the Free Culture movement: Development of the YouTutors as an auto-assignment generator by utilizing Creative Commons contents online. *Proceedings of 4th International Conference on Learning Technologies and Learning Environments (LTLE2015)* (pp. 97-402). IIAI-AAI.

Ono, Y., & Nakajima, A. (2015). Automatic quiz generator and use of open educational web videos for English as general academic purpose. In H. Ogata et al. (Eds.), *Proceedings of the 23rd International Conference on Computers in Education* (pp. 559-568). Asia-Pacific Society for Computers in Education

Ono, Y., Nakajima, A., & Ishihara, M. (2017) Motivational effects of a game-based automatic quiz generator using online educational resources for Japanese EFL learners. In P. Resta & S. Smith (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2017* (pp. 189-196). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Ono, Y., & Nakajima, A. (in press) Discussion Course Model Using Online Educational Resources to Enhance EFL Learners' Motivation and Critical Thinking. In Chen, W. et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education*. New Zealand: Asia-Pacific Society for Computers in Education.