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

Captioning Methods of Lecture Videos for Learning in English

Veri FERDIANSYAH & Seiichi NAKAGAWA*
Toyohashi University of Technology, Japan
*nakagawa@tut.jp

Abstract: This paper describes a study on the effect of using captions in lecture videos taught in English to assist non-native English speakers in learning. Lecture videos used in this paper were taken from MIT OpenCourseWare. We divided captions into 5 different types: (1) full text captions, (2) important sentence captions, (3) important phrase captions, (4) important keyword captions, and (5) no captions. Each caption was created manually and embedded in the videos. We prepared captions in two languages; English captions for foreign students and Japanese students, and Japanese captions for Japanese students. Participants for the experiment were chosen randomly from non-English native speaking countries and each participant must follow a set of tests. Experiment results showed that captions (2) or (3) helped participant the most in understanding the content of lectures.

Keywords: lecture speech, full text captions, important sentence captions, important phrase captions, important keyword captions, no captions, English captions, Japanese captions

1. Introduction

The widespread of Internet and the openness of education has made learning more accessible to everybody. Massachusetts Institute of Technology (MIT), with its MIT OpenCourseWare site, is the pioneer of the open courseware movement. Many other universities have followed the open courseware movement and opened their educational materials (e.g., course textbooks, presentations, videos, assessments, simulations, etc.) that are either (a) licensed under an open copyright license (e.g. Creative Commons) or (b) in the public domain so everybody can gain access to the resources at no costs. However, many of those educational materials, particularly lecture videos, are available without any captions or separate transcriptions. Therefore, for example, lectures in English, for students in non-English speaking countries such as Japan, will be very hard to understand the content of the lectures without getting any help from either captions or transcriptions.

The purpose of this study is to see whether captioning lecture videos in English will benefit students for learning in foreign language while investigating which type of caption will give more benefit to the learners. We divided captions into five different types and each caption was created manually by the authors and embedded in the videos using video editing software. We randomly chose participants for the experiment which are students from non-English speaking countries and each participant must follow a set of tests. During the tests, each participant will be shown 5 different videos with 5 different captions. This study showed that full text captions and important sentence/phrases captions helped participant the most in understanding the content of lectures. We also prepared the same five different types of captions in Japanese for Japanese students.

This paper is organized into the following five sections. In Section 2, previous works related to this study is explained. Methodologies are described in Section 3. Section 4 presented the results of our experiments and in Section 5, a conclusion is made and we explained our future work (Ferdiansyah and Nakagawa, 2013).

2. Related Work

Guillory et al. (1998) tested a keyword captioning method based on a hypothesis that keyword captions, while presenting learners with lesser material to read, didn't weaken their comprehension in

the spoken message. Zanon et al. (2006) distinguished three types of subtitling; (1) bimodal subtitling from English dialogues to English subtitles, (2) standard subtitling from English dialogues to subtitles in the learner's mother tongue, and (3) reversed subtitling from dialogues in the learner's mother tongue to English subtitles. Those subtitling types were then used by Zarei (2009) where the author investigated the effect of bimodal, standard, and reversed subtitling on L2 vocabulary recognition and recall. Winke et al. (2010) attempted to investigate the effects of captioning during video-based listening activities for second language learners.

Recently, Perez et al. (2014) also compared three types of captions; full captions, keyword only captions and no captions. They showed that the full caption outperformed the keyword only caption and no caption on the global comprehension questions, on the other hard, there was no difference on the detailed comprehension questions. Yang and Chang (2014) compared full captions, keyword only captions and annotated keyword captions for reduced forms such as the processes of contraction, elision, assimilation and reduction. Lwo and Lin (2016) compared four different types of captions; no captions, Chinese captions, English captions and Chinese plus English captions. They showed that the effect of different captions varied depending on student's proficiency levels.

Our purpose is to investigate how to present captions or translated captions by English and Japanese. Our final goal is to present the speech translated captions automatically based on this study. The following comparative experiments had been done in 2013, and then we extend the consideration and discussion.

3. Methodology

We used two methodologies for this study: (1) using English captions for targeting both foreign students and Japanese students and (2) using Japanese captions for targeting only Japanese students. The preparations and experiments conducted for both methodologies are the same and explained in the following sub-sections.

3.1 Video Preparations

In our experiment, we used videos from MIT OpenCourseWare lectures on computer science fields (Abelson, 2008). While screening and selecting the video materials that would be used, several criteria were maintained. The first and the most important criterion was that the video materials must have a transcription for our experiments, because manually transcribed videos needed are very impractical and time consuming. Thus, the pre-defined transcription is very important. The topics of the videos are also important criteria. Many topics are available but not all of them are common topics in other universities. It has to be a general topic, taught in every Computer Science department in any universities. With these criteria in mind, the lecture for video materials was selected.

From over 200 courses available in Electrical Engineering and Computer Science Department in MIT OpenCourseWare, one course was selected: Introduction to Computer Science and Programming (Fall 2008). From this course, we selected five different lectures to be used as our video materials. Once the video test materials were chosen, a shorter video about five minutes in duration was made from the original 60 minutes video for each lecture.

3.2 Captions

The chosen lecture video comes with a separate transcription. We utilized this transcription to create our captions. We created five different caption types manually: (1) full text captions, (2) important sentence captions, (3) important phrase captions, (4) important keyword captions, and (5) no captions.

To differ it from the full text captions, we limited the number of sentences in important sentence captions to be about 40-50% from the total number of sentences in the full text captions. Important phrase captions were created from the important sentence captions. From each sentence in the important sentence captions, we extracted 1 or 2 important phrases while maintaining the number of words in the important phrase caption is about 40-50% from number of words in the important sentence captions. As for important keyword captions, it was extracted from the important phrase captions. From each important phrase, we extracted about 1 or 2 important keywords. Using the

captions, we embedded the captions into the lecture videos. The final outcome was five videos with five different types of captions in English for each lecture.

For the Japanese captions, we also created the same five types of captions as the English version. Each caption type was a direct translation by a professional translator from its corresponding English caption.

3.3 Test set

We created a test set consisted of three different types of test to find out whether captioning video really helps learners to understand the content of the video and to measure which type of captions gives the most benefit to the learners: (1) pre-comprehension objective test, (2) comprehension objective test, and (3) caption helpfulness and understandability subjective test.

All questions and possible answers in all tests were originally written in English. For English captions experiment, in order to accommodate Japanese students who may have difficulties in reading English questions, we provided Japanese translation below the English questions but all possible answers were left un-translated in purpose. For Japanese captions experiment, all questions and possible answers are translated into Japanese language.

3.4 Scoring

Scoring method for the pre-comprehension and the comprehension test was different from the scoring method for caption helpfulness and understandability test. For the pre-comprehension and comprehension test, 10 points were given for each correct answer and 0 point for wrong answer. For caption helpfulness and understandability test, the score depends on participants' answer about each caption's level of helpfulness and understandability from one to five.

3.5 Participant Selection

Participant was divided into two different groups. The first group is international foreign students (FS) and the second group is Japanese students (JS). Several criteria must be met by each participant: (1) they have to be majored in Computer Science in Toyohashi University of Technology, (2) for international students, they have to come from non-English native speaking country, and (3) for Japanese students, they must have a TOEIC score not less than 400 (420 for TOEFL). A total of 21 subjects – 10 international students and 11 Japanese students participated in the testing sessions. Additionally, we recruited other 10 Japanese students as the third group who evaluated only five caption types of Japanese. The TOEIC score for some of them is less than 400 (420 for TOEFL). In order to make it comparable, we converted Japanese students' TOEIC score into TOEFL score by using the converting equation given by "TOEFL score = 0.348 × TOEIC score ＋ 296".

For references, we investigated the relationship between TOEFL score and correct word transcription rate (WTR) (Goto et al., 2016). The new 15 Japanese students were evaluated by transcribing while listening to English lectures. We allowed to listen them three times, but the transcribing time was limited to less than 150 seconds per utterance. The correct word transcription rate (WTR) was about 32% for students with average TOEFL score of 461 (TOEIC score of 475), 38% for 504 (TOEIC score of 600), and 42% for 556 (TOEIC score of 750). This shows that it is very difficult to listen or understand MIT English lectures for Japanese students. On the other hand, the automatic speech recognition rate by using our speech recognizer (Goto et al., 2016) which is better than students' transcription (the word correct rate is about 80%).

3.6 Procedure

All testing sessions were conducted in our Spoken Language Processing laboratory. All sessions were administered by the first author. Each testing session lasted approximately one hour.

The participants firstly took pre-comprehension test. There was no time limit for this test. They watched a series of five videos with the five different types of captions in randomized order. While watching the videos, the participants were not allowed to take notes, pause, rewind, and fast

forward the video. Each participant viewed all five videos consecutively without any break. After watching all videos, all participants were required to take comprehension test and caption's helpfulness and understandability test. There was also no time limit for both tests.

4. Results

Figure 1 presents the comparison among the pre-comprehension test score, comprehension test score, and TOEFL score for each participant who viewed the English captioned videos. The data showed that 19 participants had a higher score in their comprehension test if compared to the pre-comprehension test. Only 2 Japanese students had lower score in their comprehension test compared to the pre-comprehension test.

Figure 2 presents the comparison among the pre-comprehension test score, comprehension test score, and TOEIC score for each participant who viewed the Japanese captioned videos. The data showed that all participants had a higher score in their comprehension test if compared to the pre-comprehension test.

Figure 3 shows the average subjective evaluation score for caption helpfulness and understandability test. The result shows that the full text caption was the most helpful and the most understandable type of caption according to the participants, followed by the important sentence captions, important phrase captions, important keyword captions and the least helpful and the least understandable was no caption. There were no differences among the full text captions, the important sentence captions and the important phrase captions in English captions. However there was large difference between the full text captions or the important sentences and the important phrase captions in Japanese captions translated from English. We can conclude that the useful caption type depends on the learner's language ability. On the other hand, not all participants learned towards full text captions on their comments in the open question about what did thought about each type of caption.

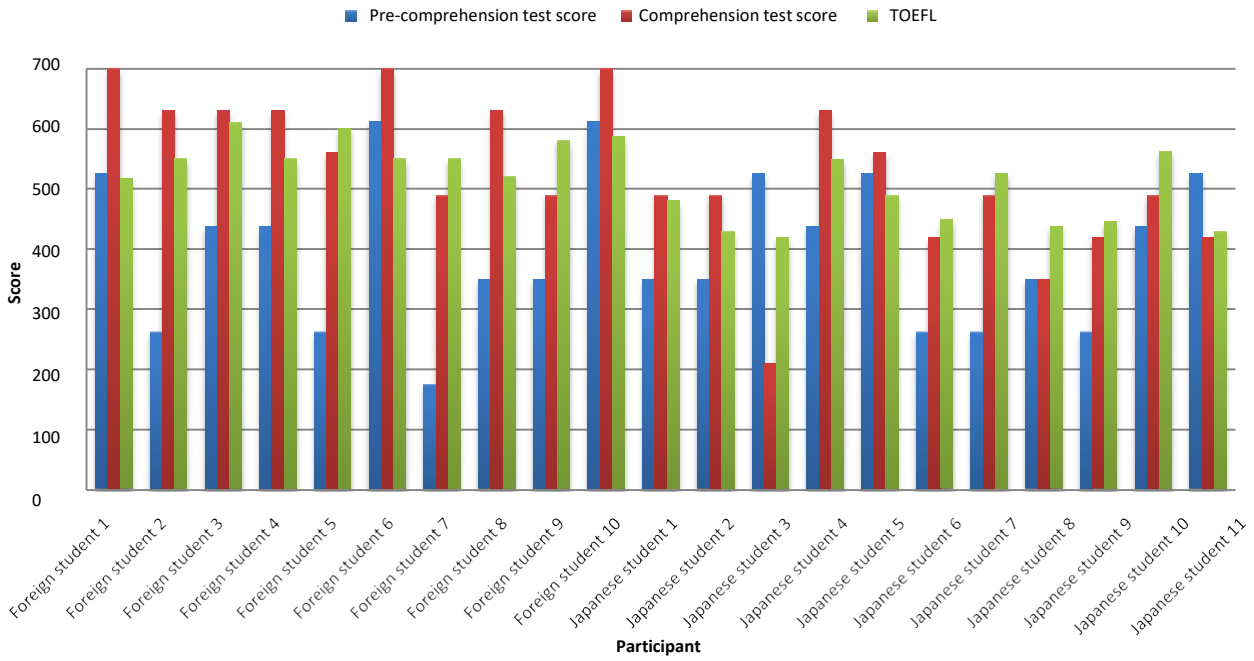

Figure 1. Pre-comprehension test vs. comprehension test vs. TOEFL score for English captioned videos

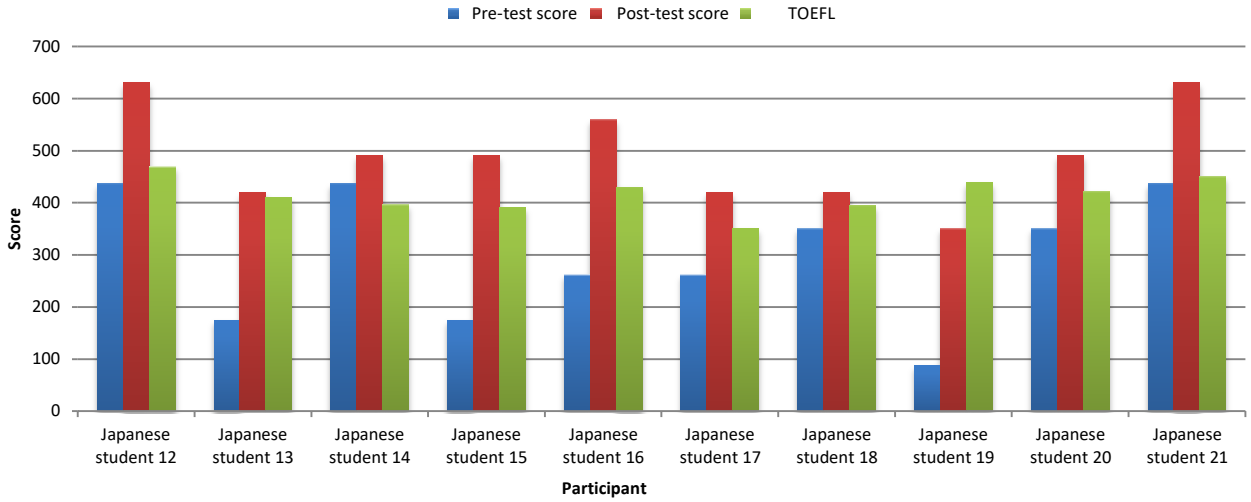

Figure 2. Pre/post-test and TOEFL scores for Japanese captioned videos

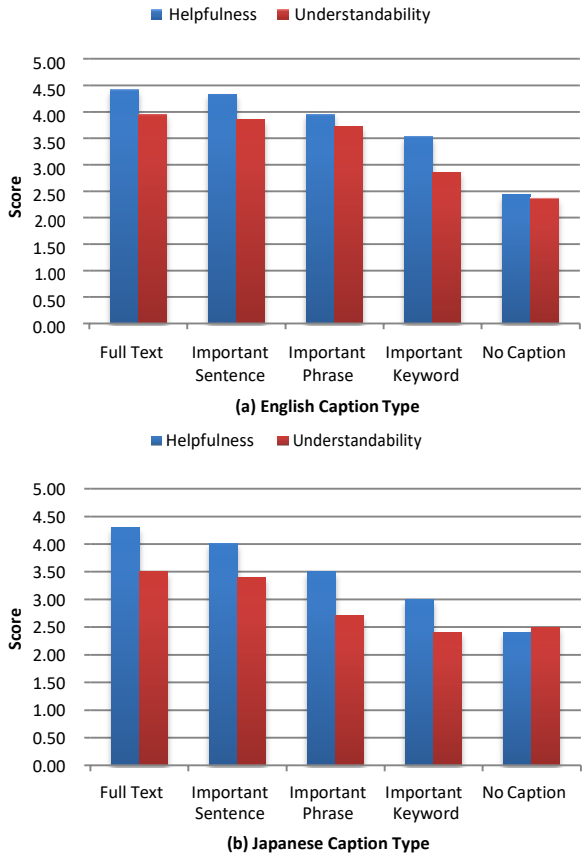

Figure 3. Test score for subjective evaluation

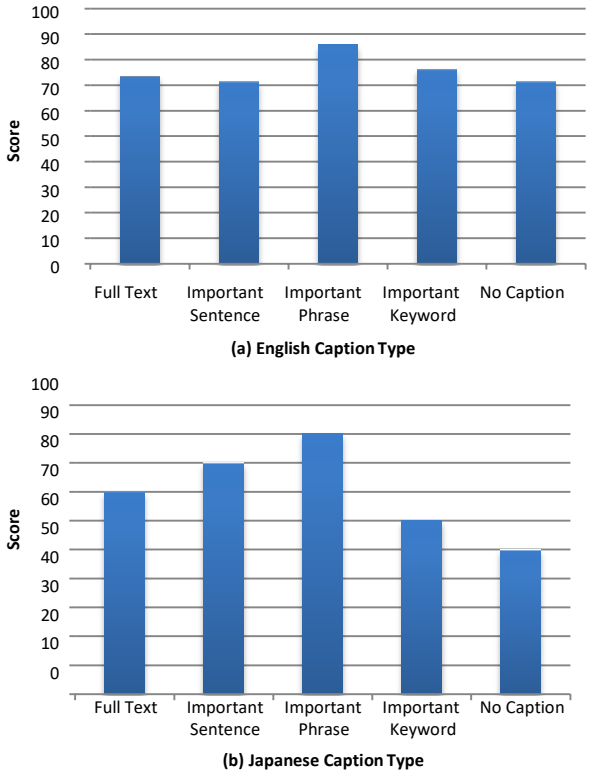

Figure 4. Captions Score for objective evaluation

Some participants' comments are: "Other captions than full text captions can be misleading." (foreign student having 550 TOEFL score); "Full text captions are the most helpful." (foreign student having 610 TOEFL score); "I prefer full text captions especially if the topic is hard so I can focus on the captions." (foreign student having 600 TOEFL score); "The speed of full text captions is too fast! I think the important sentence captions are more suitable for me." (Japanese student having 503 TOEFL score (595 TOEIC score); "Important sentence captions are easier to understand than full text captions." (Japanese student having 558 TOEFL score (690 TOEIC score); "Important phrase is good but still lacking information." (Japanese student having 487 TOEFL score (550 TOEIC score).

Figure 4 shows the average objective evaluation score for each type of caption. This score was calculated by grouping comprehension test questions based on caption type. Each caption has 2 questions in the comprehension test and we calculated the average of each caption type. The important phrase caption was the most useful. From Figures 3 and 4, we can expect that speech translation for only important sentences or important phrases is enough as understandable captions for Japanese students, especially, who have not good English ability.

5. Conclusion and Future Work

In this paper, we described a study of the effect of captioning lecture videos for learning in foreign language. We compared five different types of captions in English and Japanese for English lectures. We showed that captioning lecture videos benefited the students for learning in foreign language. From the caption helpfulness and understandability test results, we showed that subjectively, the full text captions and important sentence captions are the most helpful and understandable caption types but objectively by looking at the captions score, the important phrase captions was the most useful caption types.

We have been developing automatic Japanese captioning system for English lectures (Goto et al., 2016). Firstly, we create the ASR system for English lectures that we use to transcribe all lectures from MIT OpenCourseWare. Then we will create the captions, especially important sentence and important phrase captions by using the transcribed lecture while translating it into Japanese language.

Acknowledgment

This work was supported by JPS KAKENHI Great Number 2528002 (2013-2017).

References

Abelson, H. (2008). The Creation of opencourseware at MIT, Education and Technology, 13-2, pp.164-174.

Guillory, H.G. (1998). The effects of keyword captions to authentic French video on learner comprehension, Calico Journal 15-1/3, pp.89-108.

Goto, N. Yamamoto, K. & Nakagawa, S. (2016). Domain adaptation of a speech translation system for lectures by utilizing frequently appearing parallel phrases in-domain, Proc. APSIPA, 4pages

Lwo, L & Lim, M.C.T. (2012). The effects of captions in teenagers multimedia L2 learning, Recall, 24(2), 188-208.

Perez, M. M. et al. (2014). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension, ReCall, 26(1), 21-43

Ferdiansyah, V. and Nakagawa, S. (2013). Effect of captioning lecture videos for learning in foreign language, Vol.SLP-97, No.13, IPSJ

Winke, P. Gass, S. & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities, Language Learning & Technology 14-1, 65-86.

Zanon, N.T. (2006). Using subtitles to enhance foreign language learning, Porta Linguarum 6, pp.41-52.

Zarei, A. A (2009). The effect of bimodal, standard, and reversed subtitling on L2 vocabulary recognition and recall, Pazhuhesh-e Zabanha-ye Khareji 49, pp.65-85

Yang, J.C. & Chang, P. (2014). Captions and reduced forms instruction: The impact on EFL students listening comprehension, Recall, 26(1), 44-61