Automatic Generation of Language Quizzes for ESL Learners Based on Video Transcripts

Aoi YOSHIDA, Nobuko KISHI, Tatsuya AOYAGI, Airi SONEHARA   Tsuda College, Tokyo, JAPAN

Introduction

Many authentic online videos in English

English as Second Language (ESL) learners want to utilizing those videos in their learning

To enhance their learning…

Online video **+** Teaching materials

Relevant to the video…
➢ Subtitles
➢ Explanation
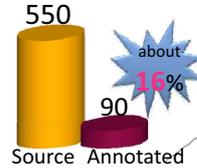➢ Quizzes   and so on…

Not enough!!

Since the development of them takes **too much cost and time**.

In case of COOLL …
COOLL is a learning system we are developing at Tsuda College. COOLL also doesn't have enough annotated materials.

http://cooll.tsuda.ac.jp

550

about **16**%

90

Source Annotated

To lower the cost of authoring language quizzes, we have designed the algorithm for generating quizzes based on video transcripts.

Our Algorithm

We have designed the algorithm for generating **multiple-choice quizzes** based on video transcripts.

In our multiple-choice quizzes, students are shown **1** English word and **4** choices of its translation in Japanese and asked to choose one correct translation.

Step 1:
Target word selection

Tokenized the video transcript

Stemmed
➤ Whether the words are in JACET8000
➤ Check the order of ranking in JACET8000
➤ Sorting in the increasing order
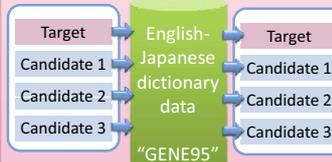➤ Selected the top of some words as target words

Step 2:
Candidate word selection

same part of speech

Search the closest to the target word in the JACET8000 ranking

Selected three candidate words

What is JACET8000?
Ranked list of 8000 English words compiled by the Japan Association of College English Teacher for Japanese learners of English. It is based on British National Corpus.

Level 1   Level 4   Level 8
high school   university

Step 3:
Translation

English word   Japanese translation

Target / Candidate 1 / Candidate 2 / Candidate 3 → English-Japanese dictionary data "GENE95" → Target / Candidate 1 / Candidate 2 / Candidate 3

If the same Japanese translation is selected for different English words, a different English word is selected by going back to **Step 2**.

Step 4:
Formatting

The target word, its translation, and the translation of the candidate words are all output.

Target word
Choose one correct translation.
1. translation of Target word
2. translation of Candidate 1
3. translation of Candidate 2
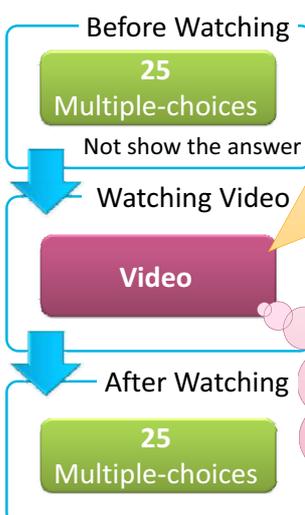4. translation of Candidate 3

Students solve the quizzes on the Moodle course site.

Experiment and Result

We have run the experiment to evaluate the relevance of automatically generated quizzes.

We selected from **8** videos and generated **25** multiple-choices for each video.

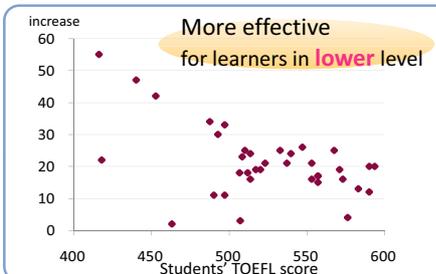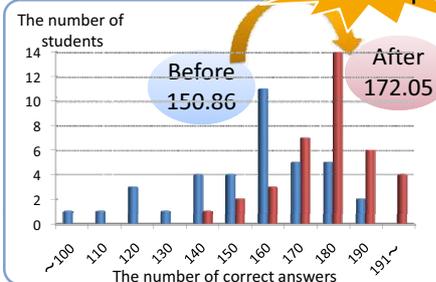**37** college students participated in this experiment.

Before Watching
25 Multiple-choices
Not show the answer

Watching Video
Video

After Watching
25 Multiple-choices

Experiment Conditions
➢ As many times as students wanted to within a time limit of **20** min
➢ allowed to use dictionaries and any web sites

Simulate **normal** online video viewing

4 videos   **4** videos
➢ approximately **4** min
➢ **500-700** words

The number of students

21.19 up!!

Before 150.86   After 172.05

14
12
10
8
6
4
2
~100 110 120 130 140 150 160 170 180 190 191~
The number of correct answers

increase
60
50
40
30
20
10
0
More effective for learners in **lower** level
400 450 500 550 600
Students' TOEFL score

Our simple algorithm has generated effective quizzes for authentic videos.

Using our algorithm, in case of COOLL…
550   550
90   →   220
about **2.4** times

Long-term Goal
Automatically generate **personal** and **high-quality** learning materials (quizzes)
➢ Language level
➢ Study record
➢ Want to learn   and so on…