Comparison of unknown word estimation performance between Japanese and French people when reading English comics aloud

Group 3   Hayato Seki

1. Introduction

Globalization makes English learning crucial in information gathering and communication increasingly. Reading aloud is one of the English learning methods and enable us to improve reading, listening, speaking and writing skills. However, if learners find words they do not know the meanings of (hereinafter called "unknown words") when reading aloud, they are concerned about unknown words and read aloud more inefficiently. Therefore, we propose a system which automatically estimates unknown words when learners read aloud. Then, learners can read aloud without their concern for unknown words and can also review unknown words later. The proposed system uses comics, and learners can enjoy because of its entertainment and remember unknown words more easily with its story and images.

This paper proposes a method estimating unknown words by learners' speech and eye gaze when learners read comics aloud. In this situation, eye gaze cannot be divided into words, but can be divided into speech balloons. Therefore, In our research, we estimate speech balloons including unknown words using Takaike et al.'s method [1] and extract unknown words from the estimated speech balloons. In order to consider accuracy improvement methods, we used Takaike et al.'s method for French people which have different cultural and linguistic background from Japanese people and analyzed the difference between Japanese and French people. It should be noted that this research has been approved by the Osaka Metropolitan University Graduate School of Engineering Ethics Committee.

2. Experiment

We conducted an data measurement experiment recording speech and eye gaze information for French people, and estimated speech balloons by using Takaike et al.'s method [1] for the obtained data in the experiment. Then, we compared the estimation using French data with that using Japanese data.

2.1. Data Measurement

We asked 13 French students to read English comics aloud using a PC and recorded their speech and eye gaze information. The 13 participants read pages specified in 3 English comics aloud and recorded unknown words every time they finished reading one English comics aloud. The measurement time is around 60 minutes including experiment explanation.

2.2. Evaluation

We performed user-independent leave-one-user-out cross validation and evaluated based on AUPR, which is the AUC of the PR curve. In this paper, the estimation using only text information was used as the baseline.

2.3. Result

Figure 1 shows AUPR averages of the estimations for Japanese data, for French data and for both data, which are expressed as JP, FR and JP+FR, respectively. As
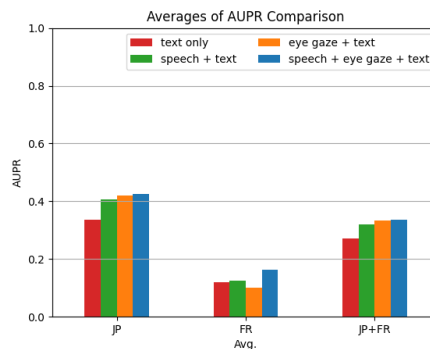

Figure 1: AUPR averages of estimation using Japanese and French data

shown in Fig. 1, the estimation using speech, eye gaze, and text information exceeded the baseline in all estimation. Also, since the AUPR average of FR is lower than that of JP and JP+FR, it is possible that Takaike et al.'s method is more effective for Japanese people than for French people.

In addition, we checked the selected percentage of 9 speech features regarding reading time and confidence for IBM Watson Speech to Text, and 12 eye gaze features, in each estimation for Japanese and French data. In this paper, the selected percentage of feature is calculated by dividing the number of times the feature is selected by the total number of Japanese or French participants. In terms of speech features regarding reading time and confidence for IBM Watson Speech to Text, the selected percentage in FR is higher than that in JP for 7 out of 9 features. Confidence for IBM Watson Speech to Text is the accuracy of pronunciation. Therefore, it is possible that speech features regarding reading time and the accuracy of pronunciation are effective for French people. In terms of eye gaze features, the rate of JP is higher than the rate of FR for 10 out of 12 features, so it is possible that eye gaze features are effective for Japanese people.

3. Conclusion

This paper compared unknown word estimation performance for French participants with that for Japanese participants, and confirmed that the estimation performed better for Japanese ones than for French ones. In addition, This paper confirmed that speech features regarding reading time and the accuracy of pronunciation are effective for French participants and that eye gaze features are effective for Japanese participants.

References

[1] T. Takaike, M. Iwata, and K. Kise. Estimation of unknown words using speech and eye gaze when reading aloud comics. In *Pattern Recognition, Computer Vision, and Image Processing. ICPR 2022 International Workshops and Challenges*, pp. 91–106, 2023.