Development and Evaluation of ASR-based Speaking System to Support English Proverbs Learning

Yi Hsuan Wang, Shelley Shwu Ching Young, Chi-Ting Wei
Institute of Information Systems and Applications, National Tsing Hua University, Taiwan
annywang12345@hotmail.com

Abstract: In this study, the researchers from different background cooperatively completed the interdisciplinary research. The aim of this study was to explore a CALL system by combining the speech recognition (ASR) technology with English learning in the hope that the state-of-the-art technology could provide learners with more opportunities of bi-directional language learning in both of formal and informal English learning. The ASR-based CALL system was constructed which provided learners with the opportunities to practice English speaking with immediate diagnosis of their utterance and three levels of pedagogical feedback was embodied to assist learners enhancing English speaking. The quasi-experimental design was adopted in this study and a total of 32 Taiwanese students participated in the experiment. The results revealed that using the ASR-based CALL system in learning had positive effects on learners' speaking performance especially and students had positive attitude toward using the ASR-based system for language learning. Moreover, the three-level corrective feedback of the ASR-based CALL system could help students to improve language faults.

Keywords: CALL, English Learning, Automatic Speech Recognition, Feedback Design

Introduction

The advancement of computer assisted language learning (CALL) facilitates learning and teaching. The automatic speech recognition (ASR) based CALL system benefits learners by providing them with integrated learning stimulation and opportunities of enhancing learners' English speaking (Chiu, Liou, & Yeh, 2007; Chen, 2011). The ASR technology provides a flexible learning environment where learners obtain immediate evaluation of their English speaking and they can practice at any time that suits him. However, many issues in the feedback design of ASR-based CALL system requires further research (Chen, 2011). Related research has shown that in the speaking-practiced CALL system, learners tend to produce more accurate utterances when they are provided with corrective feedback, instead of the opportunity of to speak (Lyster & Ranta, 1997). Nevertheless, little study evaluated the feedback design of the ASR-based CALL system because the requirement to develop the ASR technology is technically demanding and challenging which needs researchers from different domain knowledge, both from technical expertise and technology learning theory background, to work together. Understanding the above research background, we attempted to construct and evaluate a CALL in which the ASR web-service and pedagogical corrective feedback were integrated to provide learners with a flexible learning environment for English speaking. The researchers from different backgrounds of computer science and learning technology worked together with the English educators to bring out the interdisciplinary study. The results will inform our following research project

on how to well design and implement an ASR-based CALL system to address the actual needs of learners in Taiwan.

1. Literature Review

1.1 Computer assisted language learning

Over the last few decades years, there has been a dramatic increase in the variety of research on CALL including the mobile assisted language learning (Shield & Kukulska, 2008), tangible companions for learning conversation (Young, Wang & Jang, 2010) and the web-based voice recognition system for acquiring second language learning (Chiu, Liou, & Yeh, 2007). An intelligent Computer Assisted Language Learning (ICALL) involved the application of state-of-the-art computing technology such as automatic speech recognition technology (ASR). The timely evaluation based on learners' speaking performance forms an individual learning environment which provides each learner with step by step learning opportunities The ASR-based pronunciation learning system attracts more and more interests from researchers and English instructors and the future of applying ASR into CALL system in language learning is promising (Lu & Jaw, 2010).

1.2 The importance of feasibility feedback in language learning

The importance of providing learners with corrective feedback while using CALL system has been recognized from several studies (Neri et al., 2006; Chen, 2011). Learning feedback or immediate reward of learners' performance is necessary for learners to improve their ability (Hawkins, 1987). It could help learners learn effectively while providing them with corrective feedback rather than only giving them learning input (Lyster & Ranta, 1997). Corrective feedback could be presented in the implicit form or explicit form. According to the study results from Lyster and Ranta (1997), it revealed that teachers used implicit feedback, recast, in the language class the most to correct learners' language errors. Students pointed out that the explicit and segmentation feedback was more efficient than implicit ones for learners to acquire their language faults (Bigelow et al., 2006). While considering students' responses to various types of feedback, it is also important to consider the pedagogical purpose at the same time. An integrated feedback should be designed for the target learners and an explicit feedback may come after the implicit one. After reviewing the related literatures, we approached the research goal from the view point of associating ASR techniques with web-based learning concept and aimed to develop a CALL system which integrating multiple levels of corrective feedback to facilitate learning.

2. The ASR-based CALL System

The ASR-based CALL consisted of four modules, Expert module, Instruction module, Student module and ASR module, and each module had its specific role in language learning. The Expert module and Instruction module played the role of an English tutor that provided students with appropriate learning materials and guideline depending on the learner's speaking proficiency. The ASR module acted as an on-line learning partner and speaking evaluator that listens to each learner's English utterance and provided language feedback. The Student module acted as a recorder which recorded the learner's operation of the system and generated a language portfolio.

2.1 Corrective feedback design and presentation

Under this framework, both implicit and explicit feedbacks were provided and were organized into three levels. At the first level, which focused on providing implicit feedback, shows the learner's pronunciation score and the audio waveform. At the second level, which aimed to provide explicit feedback, contained a comment, an emoticon (a smiley face or crying face depending on the score), list of words that were pronounced accurately and inaccurately and an audio toolbar for replay of the learner's utterance. At the third level, demonstration of the accurate utterance, in both full sentence and single-word form were available. The learners could also play the sentence in normal and slow speed.

3. Methodology

To understand the learning effectiveness of using this system, we adopted the quasi-experiment. The control group practiced English speaking using single-level-feedback system. The design of single-level-feedback system was referred to the previous CALL study discussed in the literatures in which only the waveform diagram were presented as feedback to evaluate learners' speaking. On the contrary, the experimental group was given the three-level-feedback system which integrating the implicit and explicit elements into feedback presentation. A total of 32 seventh graders in the middle Taiwan participated in this study. The experimental group contained 16 students and the control group contained the other 16 students. The study used comparative test data and empirical experiments to report on the performance of learning English in the ASR-based CALL system with different levels of learning feedback. The learning topic chosen for this study was based on our previous research results (Wang & Young, 2012) that indicated English proverbs were one of the most desired learning contents for Taiwanese learners. Learners were required to take a pre-test and a post-test each contained an oral evaluation and several multiple-choice questions.

3.1 Research questions

The research questions of this study are as follows:
1. Can students achieve better English speaking after using the ASR-based CALL system?
2. Can the three-level feedback ASR-based CALL system promote learning effectiveness and motivation more effectively than the one-level feedback system?

4. Data Analysis

Following the research questions, the researchers analyzed and reported the study results based on quantitative data collected from the pre-test and post-test and the qualitative results from questionnaires and system login records. The Independent Samples T-test and Paired Samples T-test were conducted using the SPSS 12software package.

4.1 Learning effectiveness of experiment group (E.G.) and control group (C.G.)

Learning performance of the E.G. and C.G. are reported below. The independent sample t-test showed that there was no significant difference in the scores of pre-tests and post-tests for E.G and C.G. However, there were significant differences between the pre-test and post-test speaking scores for the E.G. and C.G. Base on the audio recordings of the

participants, it was found students tended to pronounce the vocabulary correctly after practicing speaking with the ASR-based CALL system. Besides, their speaking fluency improved toward the end of our data collection period. For example, the proverb "Don't put off till tomorrow what should be done today", during the first few weeks, students tended to experience difficulty of pronouncing full sentences. They only could pronounce the first three words accurately but pronounce the last part inaccurately. However, after several weeks of practice, their speaking was more complete and they spoke with much more confidence.

4.2 Evaluation of the corrective Feedback

The ASR-based CALL system provided both explicit and implicit feedback to learners. Three levels of corrective feedbacks were provided for students in the E.G. and only one level of corrective feedback for the C.G. From the pre-test and post-tests data, the result indicated that both the E.G and CG's speaking fluency was enhanced after eight weeks of practice. Considering the possible influence of students' learning achievement, we further analyzed their speaking fluency according to their achievement level. It was found that the means of the speaking post-test scores were higher than the speaking pretest scores in E.G.. Furthermore, there were significant differences between pre-test and post-test scores from the results of Paired Sample T-test for the low-achievement learners in E.G. (t= -4.79, P= 0.01). It showed that the low-achievement learners in E.G. had improved significantly with the English speaking but not for the learners in C.G.. On the other hand, from the results of questionnaire, it showed there were only 40% of the students reflected they could read the information from the first-level of feedback. More than half of the students were confused and had problems understanding the audio waveform diagram. Furthermore, 53% of the students suggested that they did not know how to adjust their pronunciation just by using the first-level feedback information. Approximately 60% of students in the C.G. suggested the need for more details from the system.

5.   **Discussion and Conclusion**

The study results showed that the ASR-based system efficiently improved Taiwanese students'' English speaking. Research analyses indicated that students' English speaking fluency and pronunciation were enhanced significantly. It was observed that some students experienced difficulty pronouncing English proverbs at the beginning but they were able to produce full sentences after several weeks. From the recorded system login data, it is found that learners became much more active in speaking English. Besides, the corrective language feedback in the proposed system was organized into three levels, containing both implicit and explicit forms. The research results in this study accorded with the earlier ones (Lyster & Ranta, 1997; Giuliani, Mich & Nardon, 2003; Chiu, Liou & Yeh, 2007). Students in the C.G. pointed out they had difficulty detecting their pronunciation errors from the level-one implicit feedback. The audio waveform was not helpful in providing useful information to the learners and in-depth analysis was needed and expected. On the other hand, the system provided students in E.G. with all three levels of learning support, feedback in implicit format firstly and then in explicit format. It indicated that corrective feedback in explicit format of immediate audio replay (recast) with the textual description and model pronunciation benefited learners the most especially for the low-achievement ones. The pre-test and post-test speaking scores for the low-achievement learners in the experiment group were improved through the integrated information of feedback. The textual feedback interpreted the waveform of students' utterance more specific and the

audio replay recast students' speaking utterance immediately so that they could read and hear the evaluation of the system accordingly to achieve better language learning.

5.1 Future work

It can be concluded from the current study that the ASR-based system is promising and the cooperation of the interdisciplinary study could provide tremendous value to language learning. In the future, the researchers will continue to improve the ASR-based CALL system and develop more interactive activity types such as interactive storytelling or speaking cloze test. Besides, we tend to extend the use of the system and evaluate the system for learners at different ages for addressing the actual needs of learners in Taiwan. Other research findings yielded from this series of ASR technology in language learning will be shared in the near future.

Acknowledgements

References

[1]   Bigelow, M., Delmas, R., Hansen, K., and Tarone, E. (2006). Literacy and the Processing of oral recasts in SLA. TESOL Quarterly, 40 (4), 665-685.

[2]   Chen, H., N. (2011). Developing and evaluating an oral skills training site supported by automatic speech recognition technology.ReCALL Journal, 23(1), 59-78.

[3]   Chiu, T. L., Liou, H. C., & Yeh, Y. (2007). A study of Web-based oral activities enhanced by automatic speech recognition for EFL college learning. Computer Assisted Language Learning, 20, 209-233.

[4]   Hawkins, E. (1987). Modern Languages in the Curriculum, 2nd Edition. Cambridge, Cambridge University Press.

[5]   Lu, J.h. & Jaw, G.Y. (2010). Applying MyET Oral Practice Database to Upgrade College Students' English Speaking Abilit. Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, USA, 1111-1115.

[6]   Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. Studies in Second Language Acquisition, 19, 37-66.

[7]   Neri, A., Cucchiarini, C. & Strik, H. (2002). Feedback in computer-assisted pronunciation training: Technology push or demand pull? Proceedings of ICSLP 2002, Denver, Center for Spoken Language Research, UK, 1209-1212.

[8]   Young, S. S.-C., Wang, Y. S., & Jang, J. S. R. (2010). Exploring Perceptions of Integrating Tangible Learning Companions in Learning English Conversation. British Journal of Educational Technology, 41(5), 78–83

[9]   Wang, Y. S., Jang & Young, S. S.-C. (2012). Exploring young and adult learners' perceptions of the corrective feedback in the ASR-based CALL system. British Journal of Educational Technology, 43(3), 77-80.