Computer Aided Reading and Pronunciation Practice System for Elementary Level: Development and Usability

Ricardo A. Catanghal Jr

***Abstract*—**This paper examines the initial study on teachers perception on usability of computer aided reading and pronunciation practice system and system development for elementary level in a rural area. The sample of this study consists of twelve (12) teacher, who teaches in an elementary level in the southern part of Antique, Philippines. The usability and user satisfaction questionnaire, which was adapted from Post-Study System Usability Questionnaire (PSSUQ), divided into seven categories namely design/layout, functionality, ease of use, learnability, satisfaction, outcome/future use and errors/system reliability was utilized. Prototyping model was used in the development of the system. The internal consistency of the subscale was tested using Cronbach's and the results are satisfactory.

***Keywords*—**Computer Aided Reading, automatic speech recognition, computer aided learning.

I. Introduction

One of the most basic skills taught in schools is word identification. Students who fail to develop strategies for word identification are at serious risk of academic failure because of the debilitating effects on other reading skills, such as fluency and comprehension [1]. Research has suggested that early word reading may be causally related with phonological processing abilities [1] [2]. When students do not have the ability to speak a word or its sounds aloud, teachers need to use specialized teaching strategies and adaptations to provide effective and efficient instruction. Reasons for these differences in reading abilities may be due to different background and reading experiences [3], ineffective reading strategies, and decreased phonological processing abilities [4].

The importance of beginning to train a child's pronunciation skills in a second language (L2) at an early age has long been known to researchers and educators. In the past, the main reason was to capitalise on the advantages that children have over adults in learning this skill. It was believed that, within the window of a critical period extending from approximately two years of age to puberty, children could learn an L2 with little if any effort, in contrast to adults [5].

Ricardo A. Catanghal Jr is a graduate student at the College of Information Technology and Computer Science , University of the Cordilleras, Baguio, Philippines and at the same time with the College of Computer Studies, University of Antique, Sibalom, Antique, Philippines (e-mail: catanghal.ric@gmail.com).

Golonka, Bowles, Frank, Richardson, and Freynik [6] reviewed over 350 studies of the effectiveness of technology use in foreign language (FL) learning and teaching, focusing on empirical studies that compared the use of technologies with more traditional methods or materials. Despite the abundance of available publications, evidence of efficacy was found to be limited. Technology was found to have made a strong impact on FL learning in two areas: the use of automatic speech recognition (ASR) for pronunciation training and the use of chat for increasing language production and language complexity. Moderate support was found for technology providing intelligent feedback and enhancing learners' noticing, focus on form, output and interaction, vocabulary learning, speaking proficiency, and affect and motivation.

Lyster [8] found that recasts, i.e. the correct repetition of a mispronounced utterance without any further explanation, as in most teacher-to-student interactions, might be sufficient to correct deviant pronunciation in the short term. Similarly, [9] hypothesize that detailed feedback might not be necessary for proficient learners, who are already familiar with the linguistic inventory (e.g. correct sounds of the target language) and only need to be directed to the correct alternative when they make a mistake. This study

II. Review Of Related Literature

A. Automatic Speech Recognition (ASR) Technology

This technology mainly uses the pattern recognition principle to recognize speech, and its theoretical framework is shown in Fig.1, this technology mainly consists of several units such as speech signal preprocessing, feature extraction and modeling.

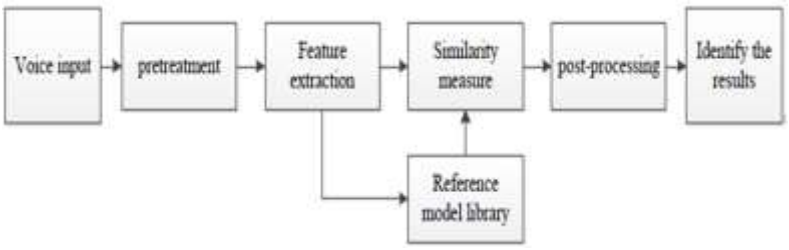

Fig. 1: Speech Recognition Principle

Automatic Speech Recognition offers many advantages for language learners. For example, this technology gives the teacher a chance to discover individual problems of the learners, which the learners can then practice independently. It offers a possibility to store student profiles in log-files, and both the students and the teachers can control the improvements and the problems recorded. Automatic Speech Recognition can also help students who are afraid from practicing speaking in public to improve their speaking skill.

The ideal automatic speech recognition based computer-assisted pronunciation training programs can be described as a sequence of five phases as Neri, Cucchiarini, and Strik [7] state:

1. Speech recognition: The automatic speech recognition engine translates the incoming speech signal into a sequence of words on the basis of internal phonetic and syntactic models. However, the main pedagogical advantage that automatic speech recognition-based computer-assisted pronunciation training can offer for training oral skills in the English as a foreign language is the provision of an evaluation of pronunciation quality.
2. Scoring: This phase makes it possible to provide a global evaluation of pronunciation quality in the form of a score. The automatic speech recognition system analyses the spoken utterance that has been previously known. The analysis can be done on the basis of a comparison between a student's utterance and a native's utterance. The benefit of automatic scoring for pronunciation training is that it gives the learner immediate information about the quality of his/her pronunciation.
3. Error detection: The system can locate the errors in the utterance and indicate to the learner where s/he makes mistakes. Referring to any problematic sound within a word can be useful to raise awareness in the learner of that problem and thus helpful for her/him to focus and practice more on that area.
4. Error diagnosis: The automatic speech recognition technology identifies the specific type of error that was made by the student and suggests how to improve it.
5. Feedback presentation: This phase presents the overall score as a graded bar, or as a number on a given scale. This phase is fundamental because the learner will only be able to benefit from all the information obtained by means of automatic speech recognition if this is presented in a meaningful way.

B. Technology Acceptance

The Technology Acceptance Model [11] relies on two factors explaining human behaviour: perceived usefulness and perceived ease of use. Perceived usefulness describes the user's point of view of enhancing his/her performance by using the system. Perceived ease of use is the degree of effort the user believes he or she will need for using a particular system. Another theoretical area relevant for this contribution is usability testing and system evaluation. According to ISO 9241-11 [12] usability is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction". User testing involves usability testing, pluralistic, informal, visual walkthroughs and contextual inquiry. Usability inspection comprises heuristic evaluation, cognitive walkthrough and GOMS (goals, operators, methods, and selection rules). But different contributions have a common definition of usability testing: persons performing a given task and evaluating the system.

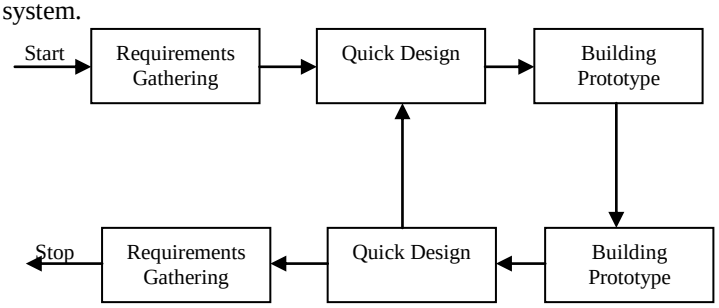

Fig. 2: Prototyping Model

The MS Speech Platform SDK 11was used as it includes both managed-code and native-code application programming interfaces (APIs) in this pilot research and most teachers and operating system are using Microsoft Windows. It is using the Speech Platform native-code API, you can acquire and monitor speech input, create speech recognition grammars that produce both literal and semantic recognition results, capture information from events generated by the speech recognition, and configure and manage speech recognition engines. The language used is 0x0409 or English –US, the implementation of lexicons is based on World Wide Web Consortium (W3C) Pronunciation Lexicon Specification (PLS) Version 1.0, which defines the structure and syntax for XML-based lexicon documents.

An integrated development environment (IDE) was used in the development , and implemented using the MS Visual C# language, which supported the .NET Framework 2.0 – 4.5.2.

C. Materials

The researcher uses a computer system and a laptop in which the computer aided reading and pronunciation system was installed, there are built in microphone and camera. The computer was part of the DCP project which is a 64-bit based operating system operating system.

D. Evaluation Instrument

Analyzing activities was done throughout the development process and phases. The informational data for analyzing purpose were gathered through reading, observation, interview and from the experts' advice. The analyzing process was not only limited in terms of physical appearance and the usage of the application but also includes the overall performance, mainly the usability and user satisfaction of the computer aided reading and pronunciation system. As for the usability and user satisfaction, test was conducted once the design and development processes were wholly completed. The usability and user satisfaction questionnaire, which was adapted from

Post-Study System Usability Questionnaire (PSSUQ) with 19 statements [10], was used to obtain feedback from the participants. The questionnaire slightly adapted to a nontechnical wording and extended by typical aspects relevant for recommendation systems (resulting in 21 statements in total), which were divided into seven categories namely design/layout, functionality, ease of use, learnability, satisfaction, outcome/future use and errors/system reliability. The questionnaire's item are 7-point scales with 1 for strongly agree, followed by next numbers in sequence with the end point 7 for strongly disagree.

III. Results and Discussion

Using the PSSUQ, the internal consistency of the subscales was found to be satisfactory (Table 1).

Subscale	Cronbach's alpha
Design/Layout	.776
Functionality	.925
Ease of Use	.743
Learnability	.877
Satisfaction	.842
Outcome/Future Use	.856
Errors/System reliability	.868

Table 1:

Design/Layout: Overall in design and layout aspects, the learners liked the interface of the application (M = 3.08, SD = 1.08), organization of the information presented was clear (M = 3.08, SD = 1.16), and the interface was pleasant to use (M = 2.83, SD = 1.11).

Functionality: The application has all the functions and capabilities that the learners expected (M = 3.67, SD = 1.07), the information retrieved from the application was effective in helping them to complete related tasks (M = 3.50, SD = 1.08), and the learners agree that all the features in the application functions well (M = 3.75, SD = 1.05).

Ease of use: The learner agree that the application was simple to use (M = 3.0, SD = 1.28), the information needed was easy to find (M = 2.75, SD = 1.05), the information provided was clear (M = 2.83, SD = 1.40), and overall, the application was easy to use (M = 2.92, SD = 1.16).

Learnability: The learners agree that it was easy to learn to use the application (M = 2.75, SD = 1.14)They found that, there was not too much information to read before able to use the application (M = 3.00, SD = 1.12).. They also agree that the information provided by the application was easy to understand (M = 3.17, SD = 1.11)

Satisfaction: In general, the learners felt comfortable using the application (M = 3.08, SD = 1.08). They also enjoyed exploring the application (M = 2.83, SD = .83), and overall they were satisfied with the application (M = 3.16, SD = 1.02).

Outcome/Future Use: The learners agree that they could become productive quickly using the application (M = 3.17 SD = 1.11). They were convinced that the application could improve their pronunciation skill (M = 3.25, SD = 1.21), and based on their current experience using the application, they would use it regularly (M = 3.08, SD = 1.24).

Errors/System reliability: The learners agree that whenever they made mistake using the application, they could recover easily and quickly (M = 3.50, SD = 1.24). They agree that the application gave error messages that clearly told them how to solve the problem (M = 3.83, SD = 1.03).

Overall outcome of the questionnaire clearly indicates the teachers' satisfaction of the computer aided reading and pronunciation system that was developed.

IV. Conclusion

Computer speech has significant contribution in education industry among various subject matters for the past years. As the researcher explored in this study the perceived usability of the computer aided reading and pronunciation is satisfactory among the teachers using the PPSUQ instrument. This result provide further confirmation of the positive influence of computer aided speech on perceived usability. The researcher would like to explore further other aspects in the future work.

References

[1] Adams, M. J. (1990). *Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.*

[2] J. Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33,* 468-479.

[3] Koppenhaver, D. A., & Yoder, D. E. (1993). Classroom literacy instruction for children wvith severe speech and physical impairments (SSPI): What is and what might be. *Topics in Language Disorders,* 13(2), 1-15.

[4] Vandervelden, M., & Siegel, L. (1999). Phonological processing and literacy in AAC users and students with motor speech impairments. *Augmentative and Alternative Communication, 15,* 191-211.

[5] Lennenberg, E.H. (1967). Biological Foundations of Language. New York: Wiley and Sons.

[6] Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. Computer Assisted Language Learning, 27(1), 70–105.

[7] Neri, A., Cucchiarini, C., & Strik, H. (2003). Automatic speech recognition for second language learning: How and why it actually works. Proceedings of 15th International Conference of Phonetic Sciences, 1157–2260.

[8] R. Lyster, "Negotiation of Form, Recasts, and Explicit Correction in relation to error types and learner repair in immersion classrooms", *Language Learning* 48, 183-218, 1998.

[9] H. Nicholas, P.M. Lightbown, and N. Spada, "Recasts as feedback to language learners", *Language Learning* 51, 719-758, 2001.

[10] Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. International Journal of Human Computer Interaction 7(1): 57-78.

[11] Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3): 319-340.

[12] ISO (1998). *ISO 9241 – 11. Usability Definitions - Guidance on Usability. Geneva, Switzerland. International Standards Organisation.* (http://www.usability.serco.com/trump/resources/standards.htm)