Developing Web-Based English Reading-Aloud Practice App with Dictation Method Using Speech Recognition Technology

Conference Paper · October 2018

DOI: 10.1109/KCIC.2018.8628497

2 authors, including:

Aliv Faizal M
Electronics Engineering Polytechnic Institute of Surabaya
19 PUBLICATIONS **97** CITATIONS

SEE PROFILE

Developing Web-Based English Reading-Aloud Practice App with Dictation Method Using Speech Recognition Technology

Aliv Faizal Muhammad
Multimedia Broadcasting
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
aliv@pens.ac.id

Akhmad Alimudin
Multimedia Broadcasting
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
alioke@pens.ac.id

***Abstract*— Reading Aloud has been proven by numerous researches for its effectiveness to enhance not only reading aloud fluency but also pronunciation, vocabulary, comprehension, and many other aspects of language competences. Thus, reading aloud capability has become one of the important parts of Pearson Test of English (PTE) Academic. Providing reading aloud activities to English as Second Language (ESL) / English as Foreign Language (EFL) students can open opportunities for them to propel their English skills. Students need to see their achievement in practicing reading aloud. Their English teacher is mostly invited to check it by listening to the reading. This practice takes more effort and time in that they cannot flexibly arrange the time to get feedback from the teacher as a reviewer at any time. As PTE Academic used a computerized test to score reading aloud ability, it is necessary to develop an app for students to practice reading aloud that comes with grading and evaluating system so that students can monitor their own progress based on the scorecard. In response to that matter, we propose developing a web-based app for reading-aloud practice and grading to enhance reading aloud fluency. In this study, we focused on the usability. Thus, we examine it using Nielsen Heuristics testing. The result shows that the heuristic aspects related to feedback, error message, reversible actions, and document were mostly violated in which all reviewer rated 100% for serverity level 4. Thus, we need to better the these heuristic aspects before the app can be used by real users.**

Keywords—Reading aloud, Teaching English, Pronunciation, Speech Recognition, E-Learning

I. INTRODUCTION

Reading aloud has proven its effectiveness for building one's ability in building knowledge and acquiring information [1]. This is because reading aloud exposes one to complexity of linguistic and cognitive aspects in which one learns from rich vocabulary presented in the books [2]. Reading aloud has been proven effective to improve vocabulary mastery compared to silent reading [3]. Visual memory development in mind is influenced by reading aloud practice [4], [5]. Reading aloud enhances various aspects of language skills development. A study incorporated 139 universities in Korea revealed that reading aloud improves students' English reading and listening comprehension [6]. Regardless as being old fashioned method of learning language, this conventional practice does generate yielding benefits to language learning.

Reading aloud is also used to develop English pronunciation skills. This activity enhances ESL students' awareness of sounds, intonation and stress placement. It was proven that reading aloud is working instrument to improve English pronunciation ability [1]. It was analyzed that by reading aloud, one can cultivate his/her internal listening skills as well as sharpening one's ear for detecting authentic dialogue from a flowing narrative.

With the current trend of using technology for education, educators and researchers are incorporating technology with educational content. The use of technology for educational purposes are mainly for teaching, assessing, and evaluating [7]. Teaching pronunciation is one of the subjects to benefit the use of technology. Automated assessment of pronunciation has served practical way for classroom English teaching for many EFL/ESL teachers. Using ASR (Automatic Speech Recognition) technology helps teachers improve their students' English pronunciation[8].

Automated pronunciation assessment takes advantages of the development of natural language processing. Numorous groups of researchers, software engineers, and corporations develop their privite and shared API to be used for speech detection. One of speech recognition API used by a number of software and services using speech processing technology is Google Speech API [9].

In my setting in which English for specific purposes is served to our students in Politeknik Elektronika Negeri Surabaya (PENS), reading aloud activities may be able to provide learning opportunities in which students can enhance multiple aspects of English compentences including sentence structure, pronunciation, vocabulary, listening, grammar, comprehension, and etc. It is suggested to implement reading aloud for autonomous learning because this can reduce students' anxiety so that they can gain confidence in speaking [8].

Based on the review and recommendation discusssed earlier, in this research and development, we are establishing a web based app for reading aloud practice and grading that is served in daily basis so that students can monitor their progress..

II. LITERATURE

A. Reading-Aloud

Reading aloud refers to a comprehensive practice of pronunciation. The content for reading aloud includes the passage with particular material and settings. Every word has to be pronounced properly, and we ought to differentiate the meaning groups carefully, manage the pause according to the content. In the meantime, we ought to implement correct stress, as well as rhythm and intonation. Hence, reading aloud can be considered as a good practice of pronunciation. The students are from different places of the country, some

of them use accent that is strongly local. Reading aloud may help them effectively perfect their dialect [10].

Students who have perfect oral English ought to pronounce correctly and probably speak fluently. Many of the students study English with a focus on reading and writing skill. For some students with no confidence to enhance spoken English, reading aloud may serve them to develop fluency, and develop habit that strengthen good and natural pronunciation. Students can practice by reading with expression, changing the pitch that is high-low, tone that referst to gentle or rough, and adjusting volume from soft to loud of their voice to exhibit various characters or build a mood. Reading aloud can help students open their mouths, and improve their oral English palpably [10].

Reading is a multifaceted act that necessitates simultaneous synchronization across many tasks. To accomplish simultaneous coordination tasks transversely, immediate execution of component skills is mandatory. With immediate execution, reading fluency is attained so that performance can be speeded, superficially effortless, autonomous, and accomplished without ample consciousness or awareness [11].

It is not astonishing, therefore, that the most noticeable characteristic of skilled reading is the speed with which text is imitated into spoken language. The characteristic to which is termed as oral reading fluency, refers to the oral translation of text through accuracy and speed. Oral reading fluency may aid as an indicator of inclusive reading skill [11].

Researches propose that the most treasured feature of the read-aloud activity is that it provides children with experience with decontextualized language, demanding them to make sense of ideas that are related to something more than the here and now [12].

B. Related Research on Computer Assisted Pronunciation Teaching (CAPT)

In regard to teaching pronunciation, we can find varieties of different software, which cover all aspects of pronunciation skill. As an important parts of CALL pedagogy, CAPT had a great development other than other skills, and many new programs were devised which aimed to make the task of pronunciation teaching easier. These programs create a new form of context with a number of practices and opportunities for learners in a narrow and small space which is not bounded to the time and presence of instructor. Using computer in language teaching improve learner autonomy and increase their self-esteem. Many researchers and scholars in recent years tried to prove the efficiency of these programs in pronunciation teaching [13].

For example, the effect of MyET software for teaching pronunciation to college students in Taiwan was studied [14]. As My ET is an online software in this study students had access to the software both in and out of the classroom. The result of the study exposed that students developed their pronunciation and even though in some cases students preferred traditional classroom instruction, the tendency of most of the students toward using software in classroom instruction increased. Also Gorjian, et al in [15] in an attempt in helping Iranian university students to acquire prosodic features of English language such as stress and intonation by using Praat software found that the "learners that practiced stress and intonation through CALL approach were more successful than the students who were taught through traditional method" (p. 34). In addition, AbuSeileek [15] in a study for teaching stress to advance EFL learners of English by using Mouton Interactive Introduction to Phonetics and Phonology software found that the students' ability in perceiving and producing correct stress pattern in words, phrases, and sentences improved.

In another study Tanner & Landon [16] examined the effect of using Cued Pronunciation Readings (CPRs) software on intermediate ESL learners of English in US. The study aimed at learning suprasegmental features such as pausing, word stress, and sentence-final intonation. The learners in this study should work with the software in a self-directed form and the teacher had no interference in their learning. The result of the study showed the significant effects of this software on learners' perceptions of pausing and word stress and controlled production of stress.

The difference between the app developed in this research and the other apps that uses speech recognition technology is that while most of them focuses on teaching or assessing pronunciation lexically, this app is aimed at providing autonomous practice of reading aloud.

III. METHOD

The research method implemented in this research was Research and Development (R&D). It consisted of identifying the problems, exploring related literature, designing and developing the app, and testing the app. The details of the method are discussed in the followings:

A. Identifying the problems

In identifying the problems, we observed the students reading-aloud performance report during the last three years. We also conducted reflections about what we witnessed about the student's performance in practicing reading-aloud.

B. Exploring related literature

In finding studies related to this research, we observed books, journals, conference, websites, and apps to support the explanations. For research findings reports through journals and conferences, we focused on that of the last 10 years, while the theoretical basis we occasionally used some long existing theories that are still related to this research.

C. Designing and developing the app

Designing and developing the app for practicing English reading-aloud, we began with selecting content materials, creating the system structure, the mock-up, and preparing for testing the app.

1) Concepting system structure: We designed how the users interact with the app. The users start by choosing a topic from the selections. If the users decided to select a topic they would see the topics. Once the topic is selected, 10 sentences from the selected topic will show up and the user will see a sentence to read. There, the users input their voice by reading the sentence. If the sentence is correctly read by the user, another sentence will appear for the user to continue reading. If the user incorrectly read the sentence, the user will see the same sentence to correct it by reading it

again. This process continues untill all sentences are correctly read. This structure is displayed in **Figure 4**.

2) Selecting content materials: We chose materials from selected books covering some topics. The topics were organized into two levels. Level 1 consists of 5 topics with 10 100-character sentences including technology, health, business, law, and biology. Level 2 consists of 5 topics including computer, internet, heatlh, history, and arts with 10 150-characters sentences for each topic.

3) Designing mockup: We designed the mockup for web-based user interface. The homepage directly shows topic selections based on levels. The mockup of the homepage is shown in **Figure 1**. Once the user click on a topic, 10 random sentences will show up one by one. Under the sentence there is a microphone button for the user to touch when he/she is ready to start reading. This mockup is shown in **Figure 2**. After the user read the sentence, score card will show under the microphone icon. The percentage of the correctness shows in the left bottom, while the continue button is in the right bottom. This mockup is shown in **Figure 3**.

4) Testing: In testing the app, we adapted Nielsen usability heuristics testing [17] and [18]. This heuristic testing was used to examine educational apps for children developed based on Nielsen Usability Heuristics testing. We invited 5 English teachers to use and review the Reading-Aloud app and asked them to provide judgement based on the matrix as shown in TABLE I. The reviewers reviewed the app based on the guide as shown in TABLE II.

TABLE I. NIELSEN USABILITY HEURISTICS WITH THE EXPLANATIONS

No	Heuristic	Definition
1	Consistency	The users should not have to wonder whether different words, situations or actions mean the same thing. The interface should use accepted common standards
2	Visibility	The user should not have to wonder where they are in the system, what they can do next or what has happened after and action
3	Match	The image of the system perceived by the users should match the model the users have about the system
4	Minimalist	This involves judging whether any extraneous information is a distraction and a slow-down
5	Memory	Users should not have to memorize a lot of information to carry out tasks. Memory load reduces users' capacity to carry out main task
6	Feedback	The system should provide feedback about the user's actions
7	Flexibility and efficiency	Users should be allowed to use shortcuts or tailor frequent actions for their own needs
8	Error message	The system should alert the users to potential errors. The messages should be clear and precise
9	Prevent errors	The system has mechanisms in place to prevent errors from occurring
10	Closure	The completion of a task is clearly indicated
11	Reversible actions	The system allows the user to easily backtrack
12	Language	The language should be presented in a form, easily understandable by the intended user
13	Control	The user should be able to leave an unwanted state easily
14	Documentation	The user should be provided with easily accessible help and documentation

TABLE II. SEVERITY RATINGS OF HEURISTIC VIOLATIONS

Severity rating	Definition
0	I do not agree that this is a usability problem at all
1	Cosmetic problem only: need not be fixed unless extra time is available on project
2	Minor usability problem: fixing this should be given low priority
3	Major usability problem: important to fix, so should be given high priority
4	Usability catastrophe: imperative to fix this before product can be released

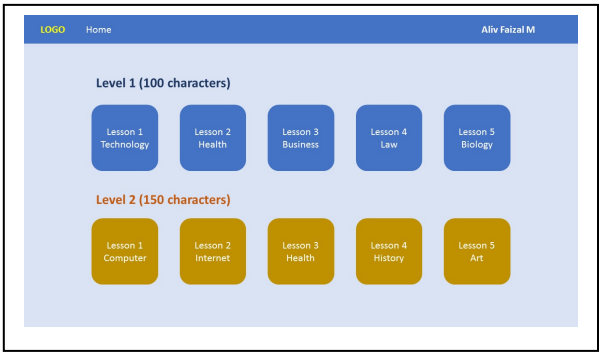

Figure 1. Mock Up of Topic Selections

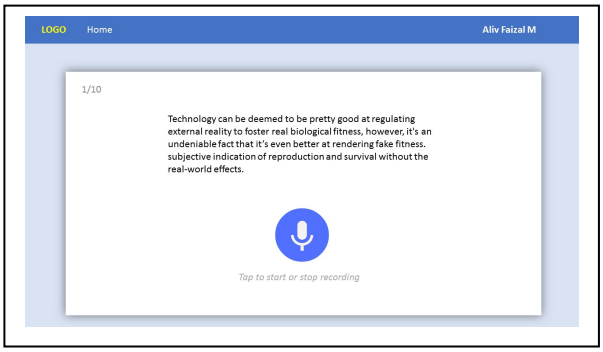

Figure 2. Mock Up of Sentence to read

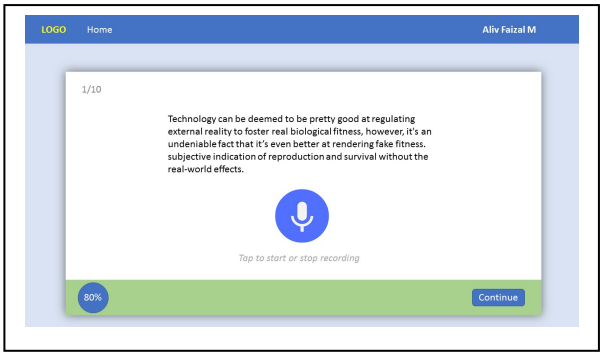

Figure 3. Mock Up score layout

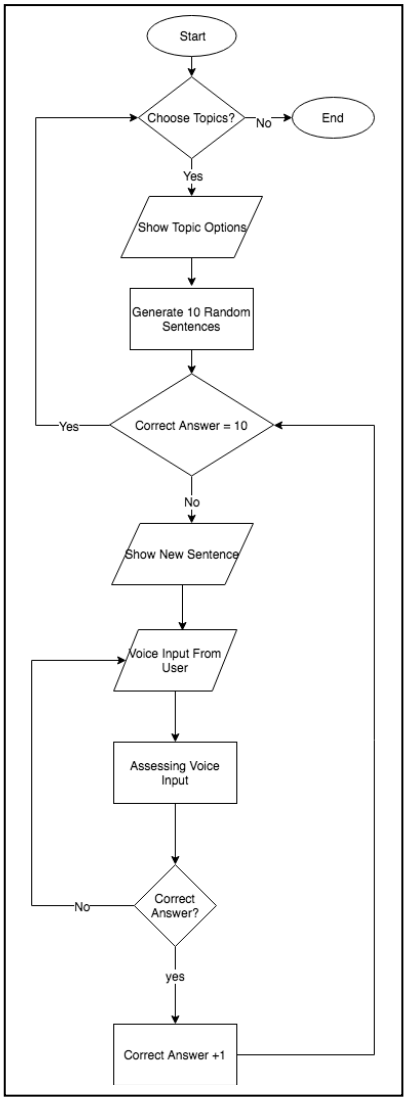

Figure 4. System Flowchart

IV. FINDING

Subsequent to the development of Reading-Aloud app that implements speech recognition technology, we tested it using adapted Nielsen Usability Heuristics testing to examine the web app. We presented the data of the test result based on the categories including visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimaslist design, help users recognize, diagnose, and recover from errors, and help and documentation.

A. Consistency, visibility, match, minimalist

The consistency aspect received 23% of rating 1, 24% of rating 2, and 53% of rating 3, while rating 0, 4, and 5 received 0%. From this result, we were suggested improving the consistency that we need to adjust the interface to be similar to common standards.

The Visibility data reported that severity score 2 received 67%, score 1 got 33%, the rest score got 0%. According to this data, the user still wonder where they are in the app system. We need to improve our design so that the user know what they can do and see what happened before and after the action.

The Match aspect received 77% for severity level 1 and 23% for severity level 2. This indicates that we can do minor revisions and improvement between what appears in the system and what the users may perceive.

The minimalist aspect received 100% for severity level 0. This means that our design does not suffer from being slow or any distraction.

The first four aspects of Heuristics are displayed in ***Figure 5***.

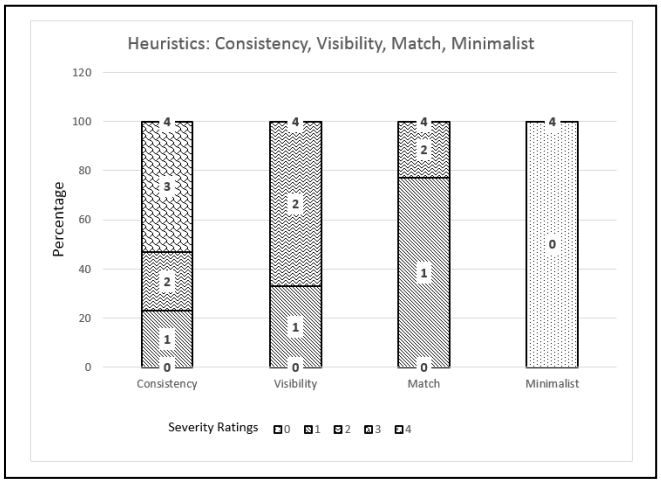

Figure 5. Ratings of Consistency, Visibility, Match, and Minimalist

B. Memory, Feedback, Flexibility

The heuristics aspects of memory, feedback, and flexibility are presented here. The memory aspect received 65% for severity level 1 and 35% for severity level 2. The other aspects received 0%. Based on this finding, we concluded that the users do not need to memorize much information in order to accomplish the task in this app. Severity level 1 resembles cosmetic problem in which we may revise the app as the project time is available.

The feedback aspect received 100% of severity level 4. This means that there is usability catastrophe in feedback aspect. Thus, it is imperative for us to fix this flaw before the app can be released.

The flexibility aspect got 82% for severity level 2 and 18% for severity level 1, while 0% for the rest severity levels. This indicates that our app suffers from flexibility. The flexibility has minor usability problem that need to be fixed but with low priority. These aspects can be seen in ***Figure 6***.

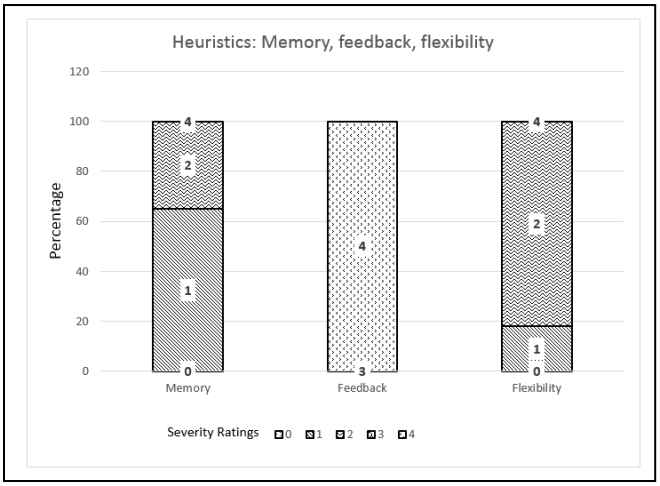

Figure 6. Ratings of Memory, Feedback, and Flexibility

C. Error Message, Prevent Errors, Closure, Reversible Actions

The aspect error message received 100% severity level 4, while the other aspects consequently received 0%. This indicates that our app does not have clear error message and it is imperative for us to fix this problem.

The error prevention received 72% of severity level 3 and 28% of severity level 3, while the other severity level got 0%. According to the judgment of the reviewer in which severity level 3 was mostly selected, the system mechanism of our app does not prevent errors from happening. Thus, it is considered as major usability problem. Fixing this part should receive high priority.

Our app received 55% of severity level 2 and 45% of severity level 1. This means that whenever the user has completed the task, they can not see and clear indication that the task is finished. This minor problem will be fixed with low priority.

According to the reviewers' judgment, the aspect of reversible action violate the heuristic usability aspect the most as all reviewers gave severity level 4 to this aspect. This is because the users cannot go back to the earlier sentence during the reading-aloud practice.

The result of these four aspects are is shown in ***Figure 7***.

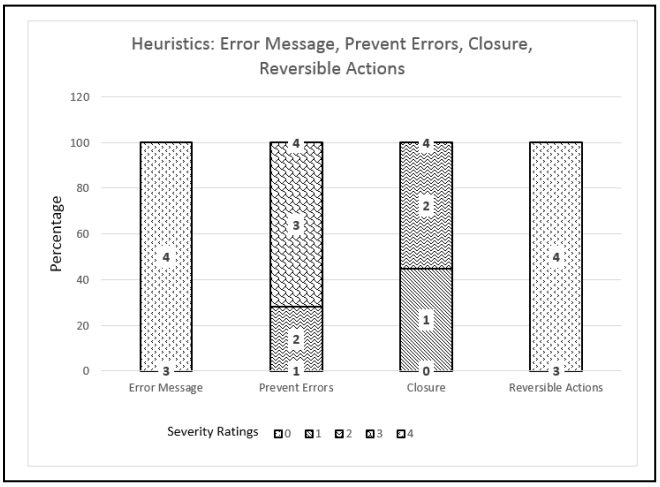

Figure 7. Ratings of Error Message, Prevent Errors, Closure, Reversible Actions

D. Language, Control, Document

All reviewers considered that the instructional language used in this app was presented in a form easily understood by the users, thus they all gave severity level 0.

The control aspect received 68% of severity level 1 and 32% severity level 2. This data means that the users were unable to abandon an unwanted state easily. The severity levels suggests cosmetic problem and minor usability problem in which fixing can be done with low priority by considering the availabe time during this project development.

This app has no documentation so that users can read to confirm things related to this app including how to use. Thus, the reviewers gave 100% for severity level 4. As this is categorized as usability catastrophe in which fixing is mandatory. The report of language, control, and document is displayed in ***Figure 8***.

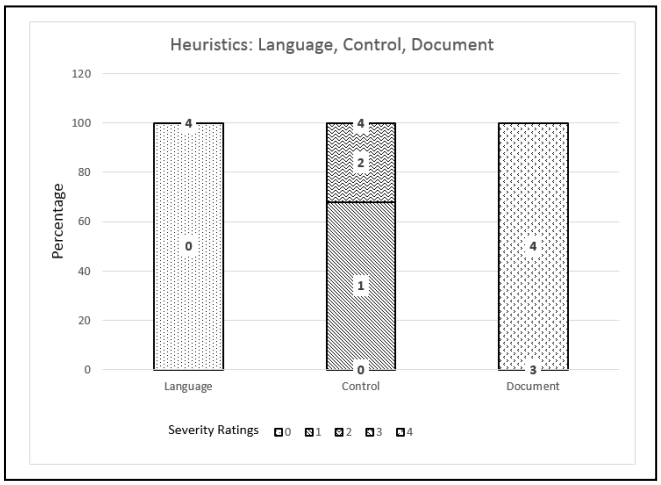

Figure 8. Ratings of Language, Control, and Document

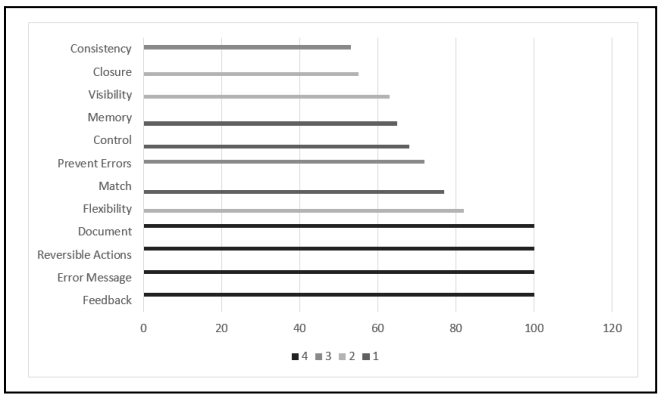

Figure 9. Summary of Heuristic Violation

Figure 9 shows the level of heuristic violations from the most severe to the lowest. The language and minimalist aspects were not shown in the chart as all reviewers rated 0 violations. The heuristic aspects feedback, error message, reversible actions, and documents violated the most. Thus, revisions are required.

V. CONCLUSION

According to the data presented in the finding, we concluded that some features of this app needs revisions and improvements before served to actual users. The fixings are based on Nielsen heuristics aspects including memory, feedback, error message, reversible actions, and documents as they were severely violated in which 100% reviewers rated severity level 4. In the future, as these revisions have been completed, we will test the effectiveness of this app by assigning two groups of classes into experimental research design to see the comparisons with another approach other than that of in this app.

ACKNOWLEDGMENT

We would like to acknowledge all corresponding parties to the development of this app and this research for their best support. We also thank to Politeknik Elektronika Negeri Surabaya (PENS) in providing financial support to this research.

REFERENCES

[1] Takita, Toshiaki, Kazuhisa Akimoto, and Junichi Hoshino. "Reading Aloud Training Game for Children with Auto Evaluation of Oral Reading Fluency." International Conference on Entertainment Computing. Springer, Cham, 2017.

[2] MASSARO, Dominic W. Reading aloud to children: Benefits and implications for acquiring literacy before schooling begins. American Journal of Psychology, 2017, 130.1: 63-72.

[3] GREENE BRABHAM, Edna; LYNCH-BROWN, Carol. Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. Journal of Educational Psychology, 2002, 94.3: 465.

[4] Smith, Sherri L., and M. Kathleen Pichora-Fuller. "Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss." Frontiers in psychology 6 (2015): 1394.

[5] Besner, Derek, and Evan F. Risko. "Thinking outside the box when reading aloud: Between (localist) module connection strength as a source of word frequency effects." Psychological review 123.5 (2016): 592.

[6] HEO, Sunyoung. The Effect of Reading Aloud on Listening Comprehension of EFL Learners. 교양교육연구, 2017, 11.3: 377-399.

[7] LOUKINA, Anastassia; DAVIS, Larry; XI, Xiaoming. AUTOMATED ASSESSMENT OF PRONUNCIATION IN SPONTANEOUS SPEECH. Assessment in Second Language Pronunciation, 2017, 153.

[8] Gibson, Sally. "Reading aloud: a useful learning tool?." ELT journal 62.1 (2008): 29-36.

[9] CHEN, Howard Hao-Jan; CHEN, Lisa Wen-Chun. Automated Speech Assessment. The TESOL Encyclopedia of English Language Teaching, 2018.

[10] HUANG, Liangguang. Reading aloud in the foreign language teaching. Asian Social Science, 2010, 6.4: 148.

[11] Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific studies of reading, 5(3), 239-256.

[12] Beck, I. L., & McKeown, M. G. (2001). Text talk: Capturing the benefits of read-aloud experiences for young children. The reading teacher, 55(1), 10-20.

[13] Witt, S., & Young, S. (1997). Computer-assisted pronunciation teaching based on automatic speech recognition. Language Teaching and Language Technology Groningen, The Netherlands.

[14] Tsai, P. H. (2006). Bridging pedagogy and technology: User evaluation of pronunciation oriented CALL software. Australasian Journal of Educational Technology, 22(3).

[15] Khoshsima, H., & Saed, A. Computer Assisted Pronunciation Teaching (CAPT) and pedagogy: Improving EFL learners' pronunciation using Clear...

[16] Tanner, M., & Landon, M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility.

[17] Alsumait, A., & Al-Osaimi, A. (2009, December). Usability heuristics evaluation for child e-learning applications. In Proceedings of the 11th international conference on information integration and web-based applications & services (pp. 425-430). ACM.

[18] Allen, M., Currie, L. M., Bakken, S., Patel, V. L., & Cimino, J. J. (2006). Heuristic evaluation of paper-based Web pages: A simplified inspection usability methodology. Journal of biomedical informatics, 39(4), 412-423.