Developing a Mobile-Assisted Software Application to Observe University Students’ Vocabulary Growth Through Extensive Reading

Wei-Wei Shen¹, Zeng-Wei Hong²,³, Chun-Po Huang³, Jim-Min Lin¹

¹ Department of Foreign Languages and Literature, Feng Chia University, Taiwan
² Department of Computer Science, Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, Malaysia
³ Department of Information Engineering and Computer Science, Feng Chia University, Taiwan
wwshen@fcu.edu.tw, hungcw@utar.edu.my, z2301010@gmail.com, jimmy@fcu.edu.tw*

Abstract

Based on a consensus of what makes good training of extensive reading, this paper aims to demonstrate how a mobile-assisted extensive reading (MAER) tool including a recommendation service and an online assessment was designed to support a reading process of the printed articles in a setting where English is learned as a foreign language. It was intended that learners could still use a paper format to read the topics closely matched their interests and vocabulary levels. This paper also shows how this tool was implemented in a class as a case study. 35 English majors studying in a freshmen class were observed in a ten-week period to use the system. Up to 3 hours guided by MAER, they were able to read up to 24 articles on average. Moreover, their vocabulary recognition level was found to have a significant improvement according to a paired-samples t-test. The implications and suggestions for using or developing MAER are then offered in consideration of the findings.

Keywords: Mobile-assisted reading, Extensive reading program, K-level vocabulary size, Vocabulary competence

1 Introduction

Scholars have claimed that extensive reading can improve language learners’ vocabulary competence [1-4]. On the other hand, the growth of vocabulary ability fostered by extensive reading can then help reading speed or fluency because vocabulary is an important factor in enhancing reading [5]. Recently, with the fast development of technology in the research of extensive reading and vocabulary, scholars have substantiated that the two factors can result in positive language learning effects [6-12].

Pedagogically, there are principles to implement extensive reading stated in the frequently cited article of Day and Bamford [13]. Important principles of extensive reading are that “a variety of reading material on a wide range of topics must be available,” “learners choose what they want to read” and “reading is individual and silent, and teachers orient and guide their students.” To achieve the above principles relating to the control of reading material, it is simpler to combine the perspectives of extensive reading and vocabulary research; the former can be based on comprehensible input (i+1) theory in extensive reading proposed by Krashen [2] and the latter can be the index of the 95% or 98% known words in a reading text [10]. The idea of combining these two major research trends is that a learner has a current vocabulary and reading ability but the unknown words in a reading text given to each learner should be carefully controlled, otherwise learners may lose interest in reading [11].

Different from traditional extensive reading programs, some studies have considered using online materials. Information technology can support personal software features such as an online reading forum, online testing, online training and so on [10]. Moreover, the Internet has quite a lot of free and authentic reading resources, such as e-books, online news and digital articles on web-based forums [6].

In addition, mobile computing or mobile applications [14-17] also provide opportunity for educational purpose. Regarding the methods of implementing online extensive reading, the popular trend is called mobile-assisted language learning (MALL). For example, Lin [9] conducted an experiment to assess the effects of mobile- and website-based extensive reading. Using e-books, the result showed that the mobile group outperformed the PC group in terms of reading achievement and appreciation. Another similar research [12] aimed at improving reading fluency by mobile-based extensive reading.

Despite the fact that the present studies of MALL have shown the benefits to language education [18-20],

*Corresponding Author: Zeng-Wei Hong; E-mail: hungcw@utar.edu.my
DOI: 10.3966/160792642020052103006
it is rare to find more existing research directly on the development of the extensive reading system. Moreover, although digital reading materials or e-books are widely used in online extensive reading, students still read printed materials [21-25]. Conventionally, printed articles and books can be major materials used in EFL reading classes because teachers normally use printed course books in a class for teaching and testing purposes.

Therefore, the purpose of this paper aims to describe the design and implementation of the mobile-assisted extensive reading tool, named MAER, to help learners studying English as a foreign language (EFL) to read printed articles in an extensive reading program. MAER can help them to look for articles according to their interests. It can assess reading ability and recommend articles based on vocabulary proficiency level. Moreover, learners can sit for online quizzes anytime to measure and recognize their current proficiency level. Every printed article contains a unique QR code. Each user can get access to online quizzes by scanning the QR code on the printed article. After each online quiz, the system gets a user’s responses, and then it will “know” if a particular reading is or is not difficult for this specific user. If it is difficult, there is a possibility that the same article can be repeatedly displayed on the recommendation list again. That is, we adopted a hybrid approach to controlling students’ vocabulary levels in an extensive reading program so that lower levels of EFL learners may feel “the reading material is easy” [13].

2 Design of MAER

2.1 Reading Materials

The 96 articles used in our system were from a graded book series, Active Skill for Reading, designed by Anderson [26]. Every article contains 300-600 words. Three volumes were selected from this series ranging from the beginning to intermediate levels. There were four reasons to use these articles. First, Neil Anderson is an internationally famous expert on lexical and reading research domains who allowed us to use his book content for our research. Second, Neil Anderson had already designed vocabulary comprehension quizzes for each article which can be helpful to maintain good validity for the online vocabulary quizzes. Third, his book was used as a unified reading textbook in the English program for first-year students at the university where this study was undertaken and the majority of the participants did not possess a good vocabulary size necessary to read for pleasure. Fourth, as mentioned, MAER is a hybrid design integrating IT technology into printed materials because of a concern about students’ different reading preferences and an insufficient quantity of vocabulary. Thus, the readings in MAER consist of different genres and very fundamental but important vocabulary.

2.1 Reading Interest

We referred to an approach proposed by Hsu et al. [27] to design an expert system to recommend articles that meet readers’ preferences. In our system, the articles cover thirteen topics, such as Family Life, School Life, Medicine, Arts and so on. The first step of the methodology is to construct two repertory grids: Articles and Reader Preference. For the article repertory grid, it is a matrix whose rows represent the thirteen topics and columns represent the articles. For every entry in the article repertory grid, \( gA_{ij} \), represents the relevance of the \( j \)th article to the \( i \)th topic. The rating of a reading topic is based on a 5-scale scheme. The value of an entry is from 1 to 5 represents the degree of relevance, where “5” means that this article is highly relevant to the topic and “1” is somewhat irrelevant. In this way, the \( j \)th column of the repertory grid represents the thirteen relevance ratings of the \( j \)th article.

We invited two English teachers to determine the relevance ratings of the 96 articles. The Spearman Correlation Coefficient was .83 with \( p < 0.01 \), indicating a high correlation between the two teachers. For the reader preference repertory grid, the rows represent the thirteen topics and the columns represent readers’ preferences towards those topics. Thus, for every entry in this grid, \( gS_{ij} \), represents the preference of the \( i \)th reader to the \( j \)th topic. The preference for a topic is rated from 1 to 5, where “5” means this reader prefers to read the topic and “1” means the reader dislikes the topic of this article. Hsu et al. [27] used equation (1) shown below to compare the preferences of readers with the traits of each article. In equation (1), \( N \) is the number of topics. In our case, \( N \) is 13. \( MaxScore \) represents the maximum rating in the repertory grids. Here \( MaxScore \) is 5. \( A_i \) is the \( i \)th article, \( S_j \) is the \( j \)th reader, and \( |gA_{ik} - gS_{jk}| \) is the distance between the \( i \)th article and the \( j \)th reader based on the \( k \)th topic in the repertory grid. On this basis, the tool can recommend the articles with the highest fitness value for the reader.

\[
Fitness(A_i, S_j) = 100 \times \frac{\sum_{k=1}^{N} |gA_{ik} - gS_{jk}|}{MaxScore - 1} \times \frac{100}{N} \tag{1}
\]

The following is an example; the values for topic relevance for the articles \( A_i \) and the reading interests of the student \( S_j \) are shown in Table 1. Column \( A_1 \) is \{1, 1, 1, 1, 2, 2, 5, 1, 1, 1, 1, 1\}. A student, \( S_2 \), registered the following values for reading interest in the thirteen topics, \{2, 3, 3, 2, 2, 1, 4, 3, 3, 3, 2, 1\} (see Table 1). The fitness value of the article \( A_2 \) and the preference of the reader \( S_2 \) can be determined by the following calculation:
Fitness \( (A_i, S_j) \)
\[
= 100 - \frac{1 + 2 + 2 + 1 + 0 + 1 + 1 + 2 + 2 + 2 + 1 + 0}{4} \times \frac{100}{13}
\]
\[
= 69.231
\]

Table 1. An example of the fitness of an article and a student’s reading interest

<table>
<thead>
<tr>
<th>Article ( A_i )</th>
<th>Student ( S_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family life</td>
<td>1</td>
</tr>
<tr>
<td>School life</td>
<td>2</td>
</tr>
<tr>
<td>Jobs</td>
<td>1</td>
</tr>
<tr>
<td>Food</td>
<td>1</td>
</tr>
<tr>
<td>Entertainment</td>
<td>2</td>
</tr>
<tr>
<td>Biology or environment</td>
<td>2</td>
</tr>
<tr>
<td>Medicine</td>
<td>5</td>
</tr>
<tr>
<td>Astronomy or geography</td>
<td>1</td>
</tr>
<tr>
<td>Art</td>
<td>1</td>
</tr>
<tr>
<td>Fashion</td>
<td>1</td>
</tr>
<tr>
<td>Information or science</td>
<td>1</td>
</tr>
<tr>
<td>History</td>
<td>1</td>
</tr>
<tr>
<td>Literature</td>
<td>2</td>
</tr>
</tbody>
</table>

The fitness values for a student’s reading interest in all the articles can be determined by equation (1). Finally, fitness values for all the articles are sorted into descending order. Overall, our scaffolding tool recommends a maximum of 10 highest-ranking articles in a list after each survey of reading interest for each student is done. Users can select any articles they want to read from the list.

### 2.2 Grading Text Difficulty

In extensive reading, selecting the articles that match the readability level of a learner is a very important issue. After filtering the articles based on readers’ preferences, our next step is to recommend articles appropriate to their level. Thus designing a scoring scheme to measure the recognition difficulty of an article is necessary. We used two scales, Reading Ease and Vocabulary Size, to represent the reading difficulty of an article.

Flesch proposed the measurement shown in equation (2) to test the difficulty of an article [28]. The equation uses the average sentence length (i.e. ASL) and the average number of syllables in a word (i.e. ASW) to calculate the reading ease of an article. The score from this equation is from 0 to 100, and a higher score indicates that the text is easier for the reader.

\[
\text{FRE} = 206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW}) \quad (2)
\]

In our system, equation (2) is adjusted to the one shown in equation (3). The reading difficulty, \( D_{\text{FRE}} \), calculated by equation (3) is from 0 to 1, and the higher the score of an article \( A \) is, the more difficulty a reader may encounter.

\[
D_{\text{FRE}}(A_i) = 1 - \frac{\text{FRE}(A_i)}{100} \quad (3)
\]

Another scale to determine the difficulty of an article is vocabulary size. The readability of an article depends on a person’s word recognition ability. If readers can recognize more words in an article, they are likely to feel that the article is easier. The minimum vocabulary size of an article in our research varies from the first 1000- (i.e., K1) to the fifth 1000-word family (i.e., K5). For this scale in our research, a method was needed to calculate the number of words for each 1000-word family in an article. Fortunately, Paul Nation along with Beglar, Cobb, and others designed an online tool called VocabProfilers that is able to satisfy our requirements [29-30]. We can simply input an article and the tool displays the distributions of word amounts for each 1000-word family. The difficulty of vocabulary size, \( D_k \) for the article \( A_i \) can be calculated using equation (4), where \( K_i \) represents the number of words in the \( i \)th 1000-word family.

\[
D_k(A_i) = \frac{1}{2} K_1(A_i) + \frac{2}{5} K_2(A_i) + \frac{3}{5} K_3(A_i) + \frac{4}{5} K_4(A_i) + K_5(A_i) \quad (4)
\]

Finally, the average for the two scales, Reading Ease and Vocabulary Size, can be calculated using equation (5) which represents the reading difficulty of an article. The difficulty is rated from 0 to 1, and a higher value can mean a more difficult article.

\[
D_k(A_i) = \frac{D_{\text{FRE}}(A_i) + D_k(A_i)}{2} \quad (5)
\]

### 2.3 Recommendation Strategy

Our MEAR scaffolding tool can link to Nation’s online Vocabulary Size Testing tool to measure each reader’s vocabulary size. This online test covers 14 K-level vocabulary and each K-level contains 10 questions. Thus the whole test has 140 questions in total. The format of this online test is a multi-choice question and a user has to choose the correct meaning of the keyword. After submitting answers, the tool immediately shows vocabulary size. To pass each level, users should have at least 8 correct answers.

Moreover, MAER recommends articles in a list that match users’ vocabulary level. If their vocabulary size grows, the tool will recommend articles in the next 1000-word family. However, if their vocabulary size does not grow, they can still request a word comprehension test for the article they are reading. Then the tool can extract an article that corresponds with the current vocabulary size and matches their comprehension ability. The maximum score in a word comprehension test is 100. Equation (6) determines the reading ability \( A(S_j) \) of the student \( S_j \). The reading ability is scored from 0 to 1.
\[ A(S_j) = \frac{\text{SCORE}}{100} \]  

Equation (7) determines the most appropriate article \( a_{fit} \) that matches a user’s reading ability, where \( A \) is the total collection of 96 articles and \( A' \) is the articles that have been recommended. In this equation, \( A(S_j) \) means readability, which can be obtained from the calculation of \( A - A' \) to represent the articles that have not yet been recommended because these texts may be beyond a user’s current ability. \( D(a_j) \) is article difficulty. The smaller gap \( |A(S_j) - D(a_j)| \) there is between the two indexes, the more appropriate the texts are for a user. These fitness values are placed in ascending order. Our tool recommends no more than ten articles.

\[
a_{fit} = \arg \min_{a, A - A'} \text{Fitness}(A(S_j), D(a_j)) = \arg \min_{a, A - A'} |(A(S_j), D(a_j))|
\]

\[ \text{(7)} \]

### 2.4 The Process of Operating MAER

When new users sign up for this tool, they complete a reading preference survey via the function. They can modify their registered preferences anytime. In addition, users are requested to let MAER know their result from the initial test of vocabulary size and then our students can manually type their level in MAER to get the appropriate and suitable vocabulary level of reading materials.

Based on the user’s preference and vocabulary size, the articles recommended by the scaffolding tool are listed for users. Then they can look for the corresponding printed article according to the index numbers and titles (see Figure 1).

\[ \text{Figure 1. Recommended articles} \]

Regarding every printed article, there are two QR codes for vocabulary learning and testing (see Figure 2). For extensive reading, looking up words in a dictionary may interrupt reading, so before reading the article, users can browse the important words in this article to scan the code on the top of the book page if they want. After reading an article, they can scan another code on the bottom to link up with the vocabulary size and comprehension test (see Figure 3).

\[ \text{Figure 2. QR codes on the printed article} \]

\[ \text{Figure 3. Words comprehension test} \]

### 3 Methods of Evaluating MAER

The MAER was implemented on students who were studying the English language as their major in a university class in Taiwan. These students just completed their high school education and enrolled at this university with the Common European Framework Levels generally fell on the ranges of A2 to B1.

#### 3.1 Research Questions

- How many articles could students read?
- What was the growth of vocabulary after students completed the reading period?
- Could the articles recommended by MAER match the proficiency level and interest of students?
3.2 Subjects and Procedures

This experiment was conducted in an EFL literacy program. 35 university students who were first-year students majoring in English in central Taiwan enrolled in this experiment. The participants aged 19-20 years (8 males and 27 females). All of them were willing to follow the guidance of MAER for 10 weeks. In the first week, they were trained to use the designed tool and received the 96 articles. They then did the K-level vocabulary pre-test and recorded their initial ability. The result showed that three students were for K-level 2, twenty were for K-level 3, ten were for K-level 4 and two were for K-level 5. Afterward, they read the articles in their free time. Finally, in the tenth week, the subjects completed the vocabulary competency test again and gave their feedback on MAER to the teacher.

3.3 Results and Discussion

Regarding the first question, how many articles could students read?, the results show that the average number of the articles that were read was 24.08. The average time using the tool was 198.03 minutes. Thus, one participant averagely spent around 3 hours to complete 24 articles within 10 weeks. This autonomous reading quantity was optimistic because comparing to the past teaching experience in a unified reading curriculum, this teacher was only required to complete no more than 16 articles in a semester course (about 14 weeks long).

As for the second research question, what was the growth of vocabulary after students completed the reading period?, Table 2 shows the results of the 35 participants' vocabulary size growth as a group analyzed by a paired samples t-test in SPSS. The 2-tailed test value was reported because we did not presuppose that there was only one positive direction of the learning outcome. In the pre-test stage, their average vocabulary recognition K-level was 3.31, meaning 3310 words, but in the end, the K-level increased to 3.74 (3740 words). In the pretest, the standard deviation was .718 (standard error mean = .121), but in the posttest, the standard deviation was .817 (standard error mean = .138). This may mean that some students’ progress may be more obvious than others. In other words, individual differences in vocabulary growth became extreme. Some may progress a lot, but others may regress.

Table 2 also shows that the t-test difference is significant (t=2.380, p<.05), which indicates that after 10 weeks of the experiment, students’ growth of vocabulary size as a whole could be found to have significant improvement. Thus, the results suggest that the students on average could increase their vocabulary by 430 words and this achievement was obvious and significant in terms of vocabulary size increase within the monitored time span.

Moreover, the improvement of individual K-levels ranging from K2 to K5 is presented further in Table 3. Fourteen out of thirty-five students progressed in this class, which is 40%, indicating that these students might be able to do reading alone without learning from the class teacher in the course. When examining each level closely in this table, we see that all three participants in K2 reached higher levels. In the K3 group, seven of the twenty subjects increased their vocabulary size; four of the twenty subjects progressed to K4 and three of them reached K5. Moreover, four out of the ten students originally at the K4 level proceeded to K5. However, neither of the two students originally at the K5 level managed to progress to K6. It is worthwhile indicating the fact that no one in the 10 weeks could progress to K6 actually; this could mean that K-level 5 was the upper bound in this experiment or 10 weeks would not be sufficient time to increase vocabulary size up to K6. Future studies can prolong the learning time to see if the results will be different.

Overall, the result indicates our tool, MAER, can result in a more significant effect on the students whose initial K-levels are 2, 3 and 4. It is also interesting to note that the good design of the contents of the textbooks, with 96 readings selected by the scholar, Neil Anderson, was indeed quantitatively sufficient for the beginning and intermediate levels of the EFL learners studying at this university.

Finally, our answer to the last research question, could the articles recommended by MAER match the proficiency level and interest of students?, can be supported by the fact that all of the 96 reading topics were chosen to read; one of the most popular articles with the keyword about ‘Food’ was read up to 31 times, but the least popular one was ‘Shakespeare’ which was only read three times. Moreover, our students did not easily fail the vocabulary comprehension test for each reading according to the record in MAER. Calculating the overall 1020 readings covered by 35 students, only 106 readings were found to have been returned to the students because they did not pass the vocabulary comprehension test. The failure rate is only about 10%, indicating that the system can reach 90% precision in recommending suitable reading texts.

Table 2. The results of the two tests obtained from the paired samples statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>3.31</td>
<td>35</td>
<td>.718</td>
<td>-2.380</td>
<td>34</td>
<td>.023*</td>
</tr>
<tr>
<td>Posttest</td>
<td>3.74</td>
<td>35</td>
<td>.817</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.
Table 3. Students’ vocabulary levels before and after the use of MEAR

<table>
<thead>
<tr>
<th>K-level group (pretest)</th>
<th>No. of upgraded subjects in the post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K3</td>
</tr>
<tr>
<td>K2</td>
<td>3</td>
</tr>
<tr>
<td>K3</td>
<td>20</td>
</tr>
<tr>
<td>K4</td>
<td>10</td>
</tr>
<tr>
<td>K5</td>
<td>2</td>
</tr>
</tbody>
</table>

4 Conclusions

We have designed a personal mobile application, MAER, which facilitates extensive reading by taking students’ interests and current vocabulary abilities into account. This system allowed students to read paper format after they got the system recommendations for any of the pre-selected articles.

The result of the experiment showed that the learners’ vocabulary size had significant growth. The relationship between students’ extensive reading training and vocabulary growth could be easily monitored and recorded with the help of the MAER system. Thus, it could help teachers to give feedback easily. Moreover, the positive learning outcome may encourage students to read more autonomously if they could be informed of their reading achievement.

Although our results shed some light on extensive reading using the system, further improvements were still discovered. Firstly, current system operates on the Android system without any problem. We can employ more platforms such as Apple. Finally, although we did not claim that printed reading materials are economical advantageous, quickly installing other authentic materials or more advanced learning materials into the MAER system to guide extensive reading may be worthwhile.

Acknowledgements

This study was supported by the Ministry of Science and Technology (MOST) under Grants: MOST 107-2511-H-035-004-MY2 and MOST 108-2511-H-035-002-MY2.

References


**Biographies**

**Wei-Wei Shen** is an associate professor of Feng Chia University and was a chair of the Department of Foreign Languages and Literature at the university and an advisory member & teacher trainer of the English Language Education Board at Taichung City Council in Taiwan. She received her MA degree at Essex University and Ph.D. at the University of Leicester in the UK. Her research interests include teaching English as a second/foreign language, vocabulary teaching and learning, cultural keywords, writing, and computer-assisted language teaching.

**Zeng-Wei Hong** is an associate professor at the Department of Information Engineering and Computer Science, Feng Chia University, Taiwan. His current research interests include e-learning and software engineering.

**Chun-Po Huang** received his M.S. degree from the Department of Information Engineering and Computer Science, Feng Chia University, Taiwan. His current research interests include the mobile application development and e-learning on language education.

**Jim-Min Lin** was an associate professor at the Department of Information Engineering and Computer Science, Feng Chia University, Taichung City, Taiwan. Since August 2005, he has been a Full Professor in the same department and served as the chairman of the department from August 2009 to July 2012. From August 2012 to July 2017, he served as the Director of the Main Library. He also served as the Secretary-General of the Computer Society of the Republic of China (CSROC) from November 2008 and October 2011. His research interests include Operating Systems, Testable Design, Software Integration/Reuse, Embedded Systems, Robotic technology and applications, and Software Agent Technology.