Ogata, H. et al. (Eds.) (2015). Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education

A Boosting Algorithm to Implement a Mobile APP-Based Language Learning System

Wen-Lin, CHEN [a]*, Yu-Chieh WU [b] & Jie Chi YANG [a]
[a] *Graduate Institute of Network Learning Technology, National Central University, Taiwan*
[b] *Department of Communications and Management, Ming Chuan University, Taiwan*
*lynnchen1993@hotmail.com.tw

Abstract: In this study, we proposed a mobile APP-based system which integrates story-based animations and simulation quizzes for English learning. Users could interact with both components of the app. The story-based animation component provides the function of word/phrase reading which displays short animation clips. For the simulation quiz component, a number of simulated quizzes are presented. The app automatically displays the scores obtained upon completion of the quiz, and regenerates another quiz based on the student's performance in the preceding quiz and the proposed boosting algorithm. To evaluate the effectiveness of the system, we conducted an experiment which compares the proposed system with traditional class-based learning. The results of the experiment show that the proposed mobile learning system yields superior scores to the control group. The results suggest the APP-based learning is helpful for students to improve their English skills. Statistical analyses demonstrate that the proposed APP-based system significantly outperforms the control group. Finally, the satisfaction survey also showed that users are highly satisfied with the mobile APP-based language learning system.

Keywords: Game-based learning, language learning, mobile learning, ubiquitous learning,

1. Introduction

Over the past few years, several prior studies have shown to the study of multimedia learning (Mayer & Moreno, 2003; Leacock & Nesbit, 2007). Multimedia learning is defined as learning from text, pictures and other media. Previous studies (Mayer & Chandler, 2001; Mayer, 2003) reported that the use of multiple media sources such as video and audio provides opportunities for more interactions and visual perceptions than a single media could do. For example, Yang, Huang, Tsai, Chung, and Wu (2009) reported that video-based learning showed better learning outcomes than traditional learning. For the mobile-based language learning, Chen and Chung (2008) presented an English vocabulary learning system based on memory learning cycle. In their system, administrators can control and assign vocabulary individually according to learners' outcome. Sandberg, Maris and de Geus (2011) also observed that mobile-based English learning improved the traditional learning significantly.

Recently, APPs running on mobile devices play an important role in learning (Falloon, 2013; Wu, 2015). However, few past studies focus on user's quiz history and quiz level for mobile learning. The quiz should not only capture the user's learning experience but also provide opportunities to learn from past errors. Even the mobile or computer assisted learning system presents variant learning styles, it remains an open question with regard to adaptively generating quizzes that specifically suit the need for individual learners.

In this study, we proposed a language learning APP-based system on mobile devices. It comprises two main modules, i.e., animation lectures and quizzes. Learners can read and learn words/phrases via displaying animations in the animation lecture. It also provides different scenarios for different words and phrases. The quiz module supports the simulation test which selects questions from the question bank based on an error analysis on the students' previous performance. The error analysis reports incorrectly answered questions and also input to the proposed boosting algorithm to form the next round quiz. More specifically, the system auto-updates the quiz and selects not only new questions from the question bank but also the questions which were incorrectly answered by the user. The system achieves the following purposes: 1) improve learning effectiveness while reducing learning costs; 2) generate quizzes automatically based on the boosting algorithm. In addition, this system is highly flexible and extensible where new lectures could be easily added into the question bank.

2. System overview

Figure 1 shows the proposed language learning APP system. As shown in Figure 1, the system consists of two modules. The first module, Reading Module, focuses on reading exercise. Learners can learn new or review learned words or phrases in this module. It displays picture and animation interactively for every word or phrase, with the meaning of the word being explained with the animation. Users can choose to stop learning at any time. More specifically, we prepare still pictures for vocabulary words and animation clips for English phrases.

For the second module, Quiz Module, the algorithm is used to determine which questions to be included in the next quiz on the basis of the user's previous performance on the quiz. During the quiz, users can directly access the evaluated score and know which part is wrong. After testing, all questions that are not correctly answered are stored in a database that is link to a personal profile and the boosting algorithm. Instead of randomly selecting questions from the question bank, the proposed algorithm is designed to select new questions based on a review of errors made in the previous tests, thus preempts the chances of repeating the same questions. Once the question is incorrectly answered by the user, the algorithm gives it the weight and selects it as a candidate for next round question. To share the common errors with others, we also add a shared database to re-weight questions. It takes both personal and public incorrect answers into consideration for the next round quiz.

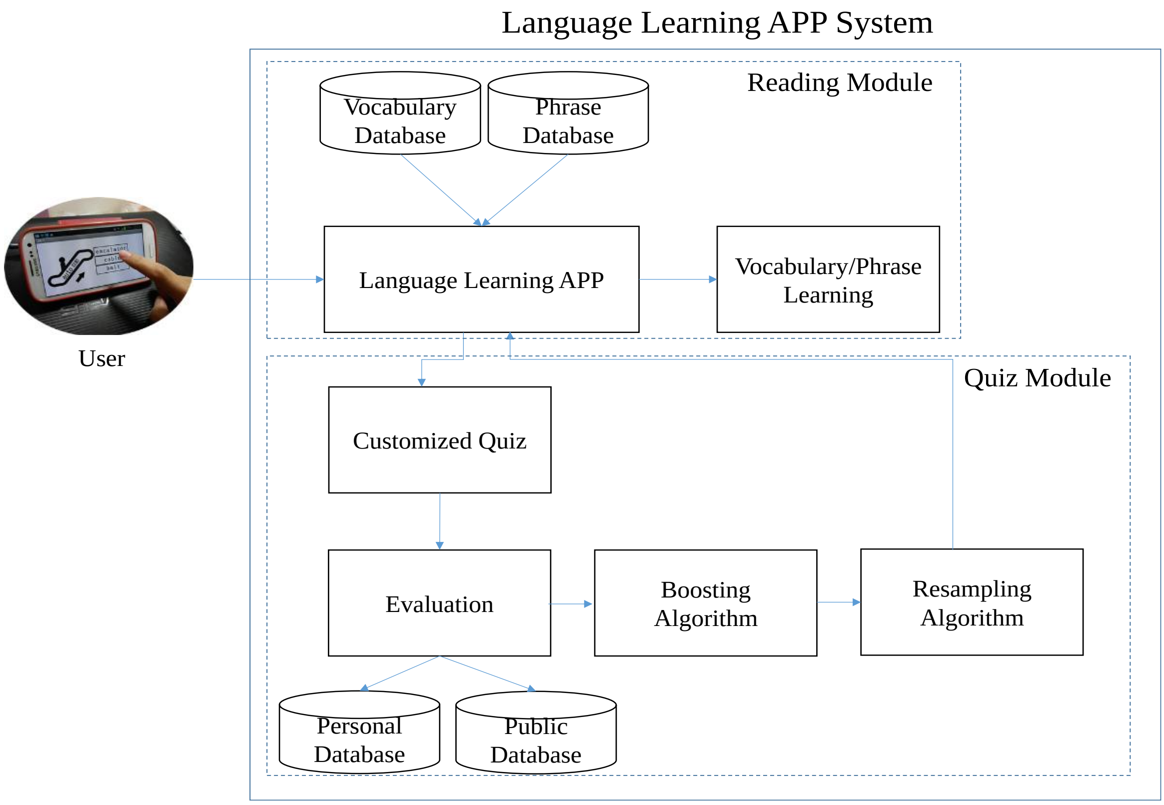

Figure 1. System Architecture.

2.1 Boosting algorithm

The questions were collected which were not correct answered by users from previous tests. To ensure that the errors could be learned from users, the error-based boosting method is applied. Here, we define the training question set is $\bar{Q}$ and it contains the question items, $\bar{Q}=\{q_1,q_2, \ldots, q_N\}$. Initially, the weight vector of $\bar{Q}$ is initialized randomly. The weight of each item in $\bar{Q}$ forms the probability to generate the question. For example, the questions with high weight have much higher possibility to form the quiz. In this paper, we define the weight function as follows.

$$W(q_i) = \begin{cases} 2^{err_i}, & \text{if } err_i > 1 \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

$$\text{where } err_i = \sum_j Ans_i - User_{i,j} \quad (2)$$

At the beginning, we trim down the errors that is rare occurs. Equation (1) represents the weight of each question q_i. The term err_i is the indicator which is the frequency of incorrect answered by user. The number j means the trial number. The incorrect answer is mainly judged by the ground-truth data. Ans_i is the true answer of the q_i, while $User_{i,j}$ means the user answer of q_i in the j-th trial. The overall boosting algorithm is shown in Figure 2.

As shown in Figure 2, the boosting algorithm stores the incorrect answered questions and updates the weight vector of Q (line 4~5). The weight vector is updated by equation (2). Next, the resampling algorithm (next sub-section) is used to reform the next round quiz.

<u>Boosting Algorithm</u>

1. Randomly initialize question set Q_0. (or using prior knowledge)
2. j:=0; // the j-th trial
3. for each trial, do{
4. Do the test and get the error question items Q_i';
5. Reweighting the question $\bar{Q}$={q_1,q_2, …, q_N};
6. Resampling the question set Q_{j+1} for next test according to QW;
7. j++;
8. }

Figure 2. The proposed boosting algorithm.

2.2 *Resampling algorithm*

In the boosting algorithm, the updated weight for each question item can be used to resample for next round. The question generation process can be viewed as a way of unfair sampling process. Questions with high error rate should have higher priority than low error questions to be selected. In this paper, we define three error types according to its error frequency, namely, HF (high frequency), MF (middle frequency), and LF (low frequency). Figure 3 lists the proposed re-sampling algorithm.

$$\begin{cases} q_i \in HF & \text{if } W(q_i) > \theta_1 \\ q_i \in MF & \text{if } \theta_1 \geq W(q_i) > \theta_2 \\ q_i \in LF & \text{otherwise} \end{cases} \quad (3)$$

<u>Sampling Algorithm</u>

1. <u>Define</u>
2. *HF*: high-frequent error question, *MF*: mid-frequent error question, *LF*: low-frequent error question;
3. <u>Algorithm</u>
4. Clean the target question set;
5. Classifying questions according to the weight using equation (3)
6. While (target question size is not reached) {
7. Select one question type *QF* from {*HF*, *MF*, *LF*} based on the frequency
8. Randomly select question from *QF*;
9. Add the question to the target question set;
10. }

Figure 3. The resampling algorithm.

The algorithm initializes the question set as empty set. Line 5 in Figure 3 is to classify each question item according to its weight (equation (1)). To simplify this process, we use two thresholds to classify the HF, MF, and LF. Equation (3) represents the criterion. The two thresholds, θ_1=4 and θ_2=2 were empirically determined. The class of question q_i is mainly judged by the two thresholds. After the question type classification, next is to sample questions from the three type (line 6~10). The process consists of two steps. First, the algorithm samples one question type according to the overall frequency (line 7). Second, one question is randomly selected to form the target question set (line 8-9). The algorithm iteratively runs until the number of sampled questions meets the target size.

3. The experiment

3.1 Methods

To evaluate the effects of the proposed APP-based language learning system with the proposed boosting algorithm, we conducted an experiment on two groups, experimental group (use the APP) and control group (traditional textbook learning). Both pre-test and post-test are applied to the two groups. The procedure of the experiment is listed below.

1) Apply pre-test for both groups,
2) Traditional lecture-based learning is applied to the control group and the APP-based learning is applied to the experimental group (see Figure 4 for screenshots of the system),
3) Apply post-test for both groups,

Each group was assigned to run its own learning strategy. In mobile group, we installed the APP to their mobile phones, while in the control group, the traditional lectures were assigned. The experiment was applied within one week. After the defined period, we evaluated both groups. Then, the evaluated scores are kept to analyze the results. Besides, the satisfaction questionnaire is taken for the participants after they finished the whole experiment. It was based on how many times, how often that the user used and degree of satisfaction to analyze if these factors affect the results.

Figure 4. English Learning application content.
Users can study by reading the writing question bank and the picture question bank. Besides, users can also wacth videos which have explanations of words and usages.

3.2 Results

The statistical result is shown in Table 1 and 2. We used t-test to compare the result of the experiment group and the control group. The pre-test result shows that there is no significant difference between the two groups, at a significance level of p=.390. After making sure the two groups are equally proficient in their English levels, we begin to analyze the pre- and post-test scores of the control group. The mean score of the pre-test is 52 points and 71.5 points for the post-test. The scores improved by an average of 19.5 points. The p-value of t-test is 0.001 ($p<0.01$). That means the use of traditional method to learn English is still positive for students. Next, we analyzed the pre- and post-test scores of the experimental group. The t-test result shows significant differences. The mean score is 58 points for pre-test and 89.5 for the post-test. Overall improvement of the score is 31.5 points. The p-value of t-test is 0.000 ($p<0.001$). It means using the APP to learn English is useful to learners. Next, we compared the post-test results between the experimental and the control group. The p-value of t-test is 0.005 ($p<0.01$), indicating that there is significant difference between the experimental group and the control group. Based on the result of the post-test, the use of the proposed APP system is shown to have achieved outcomes than the traditional approach exercised by the control group.

Table 1: Results of paired samples t-test between pre-test and post-test.

		Pre-test		Post-test				
		Mean	*SD*	*Mean*	*SD*			*p*
Group	N					*t*-value	df	
Experiment	10	58.00	17.51	89.50	9.56	-6.851***	9	.000
Control	10	52.00	12.52	71.50	15.10	-4.993**	9	.001

*** $p < .001$, ** $p < .01$

Table 2: Results of independent samples t-test between experiment and control groups.

Group	N	*Mean*	*SD*	*t*-value	df	*p*
Experiment	10	89.50	9.56	-3.185**	1	.005
Control	10	71.50	15.10			

** $p < .01$

3.3 Satisfaction

The degree of students' satisfaction was shown in Table 2. As shown in Table 3, we can find that the students who improved most are those who used the APP most. On the contrary, students who improved less are those who used the APP less with shortest time. Therefore, it indicates that this proposed APP is useful for students' English learning. In addition, in this experiment, the scores of students are all improved and almost all students thought this APP is satisfied.

Table 3: Degree of satisfaction.

Pre-test scores	Post-test scores	Frequency of use	Time of use	Degree of satisfaction (1~5)
60	95	1~5	< 10 mins	4
85	95	1~5	30~60 mins	4
60	90	1~5	10~30 mins	4
40	90	6~10	< 10 mins	4
25	65	1~5	< 10 mins	4
70	90	1~5	30~60 mins	4
65	95	1~5	10~30 mins	5
75	85	1~5	< 10 mins	4
50	100	11~15	< 10 mins	3
50	90	6~10	10~30 mins	5

4. Conclusion

This study presents an APP-based system for English learning. The system consists of two modules, story-based animation and simulation quiz. The former provides the vocabulary learning from animations while the later supports the quiz test based on the proposed boosting algorithm. To evaluate the system, we compared with traditional class-based learning. The experimental results show that the proposed system yielded better post-test scores than the traditional learning group. The significant test results also reveal that the proposed system yielded statistical significance. The degree of students' satisfaction also showed that users are highly satisfied with the mobile APP-based learning system.

In the future, we plan to integrate additional language learning materials to the system and expand the question bank for quizzes in the cloud and social media.

Acknowledgements

The authors would like to thank all the subjects who participated in the study.

References

Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education, 51*(2), 624-645.

Falloon, G. (2013). Young students using iPads: App design and content influences on their learning pathways. *Computers & Education*, *68*, 505-521.

Leacock, T. L., & Nesbit, J. C. (2007). A Framework for Evaluating the Quality of Multimedia Learning Resources. *Educational Technology and Society, 10(2), 44-59.*

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13*(2), 125-139.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390–397.

Mayer, R. E., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 43-52.

Sandberg, J., Maris, M., & de Geus, K. (2011). Mobile English learning: An evidence-based study with fifth graders. *Computers & Education*, *57*(1), 1334-1347.

Wu, Q. (2015). Designing a smartphone app to teach English (L2) vocabulary. *Computers & Education*, *85*, 170-179.

Yang, J. C., Huang, Y. T., Tsai, C. C., Chung, C. I., &Wu, Y. C. (2009). An Automatic Multimedia Content Summarization System for Video Recommendation. *Educational Technology and Society*, *12*(1), 49-61.