

Personalized Word-Learning based on Technique Feature Analysis and Learning Analytics

Di Zou¹ and Haoran Xie^{2*}

¹Department of English Language Education, The Education University of Hong Kong, Hong Kong //

²Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong
// dizoudaisy@gmail.com // hrxie2@gmail.com

*Corresponding author

ABSTRACT

Many studies have highlighted the importance of personalized learning, and most current e-learning systems are able to personalize materials, activities, etc., based on individualized learner-factors. However, none of the extant word-learning systems provides a personalized learning experience that is guided by a comprehensive word learning theory. In this study, we develop such a system based on Nation and Webb's checklist for technique feature analysis - a thorough set of factors that promote effective word learning. This system recommends personalized word learning tasks based on the technique feature analysis scores of different tasks and user models. To examine the effectiveness of the proposed system, we conducted an experiment among 105 English learners, grouped them into three teams randomly, and asked them to learn forty target words through three approaches: a non-personalized approach, a personalized approach guided by a partial version of the technique feature analysis, and a personalized approach guided by the full list of the technique feature analysis. Significant differences were observed among the effectiveness of the three approaches in promoting word learning, with the personalized approach guided by the complete checklist leading to the best learning performance. It is therefore suggested that e-learning systems should be designed based on comprehensive learning theories.

Keywords

Personalized learning, Vocabulary acquisition, Learning analytics, Technique feature analysis, User model

Introduction and background

With the rapid development of information technology and pervasive use of digital devices in recent years, language education is steadily moving toward utilizing more technology-based tools and approaches. Computer- and Mobile-Assisted Language Learning (CALL, MALL) have gained increasing popularity, and a large number of language-enhancement systems have emerged. For example, Chen and Hsu (2008) developed a personalized intelligent mobile-learning system which recommends appropriate English news articles to learners based on their reading abilities as evaluated by a fuzzy item response theory proposed in the research. Liu (2009) built a ubiquitous learning system - the Handheld English Language Learning Organization (HELLO) - which integrates augmented reality and sensors for supporting learners' development of listening and speaking skills. Similarly, Kwon et al. (2010) presented a personalized computer-assisted language learning system based on learners' cognitive abilities with respect to their language proficiency levels; this system employed a strategy of retrieval learning, a method of learning memory cycle, and a method of repeated learning for improving learning effectiveness. Focusing on learning environments, Wu, Sung, Huang, Yang and Yang (2011) developed a situated and reading-based English learning system that integrated a reading guidance mechanism into the development of an e-learning environment. Moreover, Hsieh, Wang, Su and Lee (2012) designed a fuzzy logic-based personalized learning system to support adaptive English learning. Hsu, Hwang and Chang (2013) developed a personalized recommendation-based mobile learning approach to improving reading skills of language learners. Additionally, Xie et al. (2017) proposed a profile-based approach to discovering learning paths for group users to improve the learning effectiveness and efficiency of a whole group. A common feature of all aforementioned systems and approaches is that they provide a personalized learning experience based on the needs, prior knowledge, preferences and/or learning styles of individual learners, and thus are more effective than non-personalized systems.

With close connections to grammatical knowledge and other language skills, word knowledge is widely acknowledged by linguists and teachers as the foundation of language acquisition (Lightbown & Spada, 2006). Language learners also attribute great importance to word knowledge and are particularly interested in effective approaches to word learning, because they believe that word knowledge is central to their communicative competence (Schmitt, 2000). However, many learners feel frustrated about vocabulary learning because they often forget words that they have previously learned. They also do not know what words should be learned first and how to learn them effectively. The probability of knowing words through implicit learning is very small, whereas it is difficult to engage in explicit learning for a long period of time (Nation, 2001), a large number of

learners therefore regard word learning as a time-consuming, yet not necessarily rewarding, activity. Identifying effective word learning methods or techniques is therefore paramount.

Personalized learning systems for vocabulary learning

Given the importance of word knowledge for language acquisition and students' preference for greater use of digital technological tools in education (Zou & Lambert, 2017), numerous e-learning systems are specifically designed for supporting effective vocabulary learning. Most of them provide a personalized learning experience for the users, as learning vocabulary by using desired and relevant language resources is essential for effective acquisition (Zou et al., 2017). A few representative studies are reviewed as follows.

Barker (2007) proposed a personalized approach that enables language learners to make their own decisions about the costs and benefits of learning new words through analyzing these words via consideration of both word- and learner-specific factors when they encounter the words. Jung and Graf (2008) developed a word association game to facilitate personalized vocabulary learning in a web-based system; this system can effectively increase the motivation of learners and cater to their individual needs. Moreover, Chen and Chung's (2008) personalized mobile English vocabulary learning system, which is based on item response theory and learning memory cycle, can appropriately recommend vocabulary for learning according to individual learners' word knowledge and memory cycle. Chen and Li (2010) also designed a personalized context-aware ubiquitous vocabulary learning system based on learner locations as detected by wireless positioning techniques, learning time, individual English vocabulary abilities, and leisure time. This system enables learners to adapt their learning content to effectively support English vocabulary learning in a school environment. Furthermore, Huang, Huang, Huang and Lin (2012) developed an easy-to-use ubiquitous English vocabulary learning system to assist students in experiencing a systematic vocabulary learning process. Employing fuzzy inference mechanisms, memory cycle updates, learner preferences and analytic hierarchy processes, Hsieh et al. (2012) proposed a personalized English article recommendation system which selects appropriate articles for learners by using accumulated learner profiles. This system can effectively improve learners' English proficiency levels in an extensive reading environment via helping them comprehend new words quickly and review words that they knew implicitly. Similarly, the mobile learning system of Hsu et al. (2013) includes a reading material recommendation mechanism that suggests articles to learners based on their preferences and knowledge levels. It also involves a reading annotation module that enables students to take notes of English vocabulary translations for the reading content in individual or shared annotation mode. Sandberg, Maris and Hoogendoorn (2014) compared the learning performance of two groups of learners who participated in a mobile learning application, and noted that gaming contexts and intelligent adaptation have additional value for mobile vocabulary learning. Zou, Xie, Li, Wang and Chen (2014) presented a personalized word learning task recommendation system based on Laufer and Hulstijn's (2001) involvement load hypothesis, and found that this system which personalizes learners' learning experience via load-based learner profiles promotes effectively word learning. Additionally, Wang and Shih (2015) examined the effects of self-paced use of smart phones as mobile learning tools on English vocabulary learning, and found that the group with mobile learning scored significantly higher than the control group. Likewise, Huang, Yang, Chiang and Su (2016) investigated the effects of a situated mobile learning approach on students' English learning motivation and performance by developing a five-step vocabulary learning strategy and a mobile learning tool in a situational vocabulary learning environment. The research findings showed that the proposed strategy and tool are effective in increasing students' learning motivation and performance. With similar purposes, but a different theoretical framework, Xie, Zou, Lau, Wang and Wong (2016) developed an e-learning system for recommending vocabulary learning tasks based on topic-based profiles obtained from social media platforms and load-based profiles measured by the involvement load hypothesis. The experiment results demonstrated that the proposed system not only improves learning effectiveness, but also increases learning enjoyment.

However, most of the aforementioned personalized vocabulary learning systems involve just one or two factors that are conducive to effective word learning, and thus are limited in other respects. For example, Chen and Chung's (2008) system recommends appropriate English vocabulary for learning according to different learners' prior word knowledge and memory cycle, and thus it is effective in adjusting learning modes of various learners to promote their learning performances and interests. Nevertheless, this system promotes little development of productive word knowledge, as no generative use is involved in the learning experience. Similarly, although the system designed by Huang et al. (2016) takes into account facilitative factors for word learning, such as high motivation, retrieval of words and meaningful contexts, it is limited in terms of the spacing between retrievals, linking of form and meaning, and generative use of the words. This is probably because the design of these systems is not guided by a comprehensive checklist that covers all important factors that are essential for

effective vocabulary learning. Therefore, it is necessary to develop a vocabulary learning system under the umbrella of a comprehensive set of word learning techniques. Nation and Webb's (2011) checklist for technique feature analysis (hereafter, TFA) is selected as the theoretical framework because it constitutes an elaborate set of criteria which provides a reliable guide for predicting, evaluating, and explaining the effectiveness of diverse word-focused tasks.

The checklist for technique feature analysis

Nation and Webb's (2011) checklist for technique feature analysis, which operationalizes cognitive notions, such as depth of processing and richness of encoding, involves five main components: motivation, noticing, retrieval, generation, and retention. Each of these five main components further includes three to five questions, covering various factors that are effective in promoting word learning. There are altogether 18 questions in the checklist, and point values are used to evaluate different word learning techniques (see Table 1).

Table 1. The checklist for technique feature analysis (adapted from Nation & Webb, 2011, p. 7)

Component	Criteria	Scores	
Motivation	Is there a clear vocabulary learning goal?	0	1
	Does the activity motivate learning?	0	1
	Do the learners select the words?	0	1
Noticing	Does the activity focus attention on the target words?	0	1
	Does the activity raise awareness of new vocabulary learning?	0	1
	Does the activity involve negotiation?	0	1
Retrieval	Does the activity involve retrieval of the word?	0	1
	Is it productive retrieval?	0	1
	Is it recall?	0	1
	Are there multiple retrievals of each word?	0	1
Generation	Is there spacing between retrievals?	0	1
	Does the activity involve generative use?	0	1
	Is it productive?	0	1
Retention	Is there a marked change that involves the use of other words?	0	1
	Does the activity ensure successful linking of form and meaning?	0	1
	Does the activity involve instantiation?	0	1
	Does the activity involve imaging?	0	1
	Does the activity avoid interference?	0	1
Maximum score			18

Specifically, the component "motivation" concerns whether an activity has a clear word learning goal, whether it motivates learners, and whether the learners select the words. The component "noticing" questions whether an activity induces attention on the target words, whether awareness of new word learning is raised, and whether negotiation is involved. The component "retrieval" consists of receptive retrieval, productive retrieval, recall, multiple retrievals, and spacing between retrievals. The component "generation" comprises generative use, productive generation, and marked changes that involve the use of other words. The component "retention" mainly refers to whether an activity ensures linking of form and meaning, and whether it involves instantiation, imaging, and avoids interference (Nation & Webb, 2011). Scores of the questions are measured in a binary manner: one will be given if a learning activity involves a certain criterion, and zero will be given if not.

Taking the task reading comprehension and performing cloze-exercises with textual annotations as an example, its total TFA score is 7. As demonstrated in Table 2, because this task requires the learners to compare different words and fill them in the blanks where the contexts are suitable for the words, it has a clear word learning goal with focused attention on the target words. Moreover, as the students are aware of the learning of the target words while matching them with the appropriate contexts, receptive generative use of these words is induced, and the students are motivated. The generative use of the target words induced here is not productive since the contexts are given, and the students do not generate original contexts themselves. However, as the students need to fill in the blanks with the target words, the activity ensures successful linking of form and meaning. Lastly, since the words are normally not members of the same lexical set, the activity avoids interference. The total FTA score of cloze-exercises is therefore 7. Compared to cloze-exercises, writing original sentences using target words has two more TFA scores because learner-created original contexts for the target words are generated, and hence productive generative use of target words and marked changes that involve the use of other words are induced. Except for this point, these two tasks are similar in other aspects of the TFA criteria. They both induce

clear learning goals, raise awareness of word learning, draw attention to the target words, ensure linking of form and meaning, motivate learning, and avoid interference (see Table 2).

Table 2. The technique feature analysis scores of two common word learning tasks

Component	Criteria	Scores of two tasks	
		Cloze-exercises	Sentence-writing
Motivation	Is there a clear vocabulary learning goal?	1	1
	Does the activity motivate learning?	1	1
	Do the learners select the words?	0	0
Noticing	Does the activity focus attention on the target words?	1	1
	Does the activity raise awareness of new vocabulary learning?	1	1
	Does the activity involve negotiation?	0	0
Retrieval	Does the activity involve retrieval of the word?	0	0
	Is it productive retrieval?	0	0
	Is it recall?	0	0
	Are there multiple retrievals of each word?	0	0
	Is there spacing between retrievals?	0	0
Generation	Does the activity involve generative use?	1	1
	Is it productive?	0	1
	Is there a marked change that involves the use of other words?	0	1
Retention	Does the activity ensure successful linking of form and meaning?	1	1
	Does the activity involve instantiation?	0	0
	Does the activity involve imaging?	0	0
	Does the activity avoid interference?	1	1
Total score		7	9

According to Nation and Webb (2011), tasks with higher TFA scores promote better word learning than tasks with lower scores. Hu and Nassaji's (2016) study also provided empirical evidence for the reliability of TFA in evaluating the effectiveness of diverse word learning tasks. This article also noted that TFA has good explanatory power in predicting word learning gains. In the present study, we develop a personalized vocabulary learning system under the umbrella of the checklist. The detailed contributions of this article are listed as follows.

- A user model is developed with the checklist for technique feature analysis as its theoretical framework, and thus a comprehensive set of factors that facilitates effective word learning is covered.
- Personalized vocabulary learning processes are generated by the proposed system based on the user model to assist learners to select appropriate learning tasks.
- Real participants are invited to use the proposed system, and the effectiveness of this personalized task recommendation system is verified by the learning performance of the participants.
- Implications of the research findings are elaborated from the perspectives of how technology-enhanced word learning tasks and personalized e-learning systems should be designed.

The remaining sections of this article are organized as follows. The methodology section describes the development of the user model and the generation of personalized vocabulary learning. The experiment and results sections explain the settings, processes, and results of the experiment. The implications and conclusion of the study and future research directions are discussed in the last section.

Methodology

In this section, we will introduce the method of the study, focusing on the development of the user model and the generation of personalized vocabulary learning.

User modelling based on the checklist for technique feature analysis

As mentioned previously, the theoretical foundation of the user model in the proposed vocabulary learning system is the checklist for technique feature analysis. To measure the TFA score of a word, the TFA scores of all tasks that include this word as a target word will be counted. For example, if learner A has completed two tasks

(i.e., task *A* and task *B*) that focus attention on the word *B*, the TFA score of word *B* for learner *A* is the sum of the TFA scores of task *A* and task *B*, as shown in Figure 1. Formally, a user model is defined as a matrix of TFA scores of each word with respect to the 18 TFA criteria as follows:

$$L_i = (s_{mn}^i) \in \mathfrak{R}^{m \times n} \quad (1)$$

where L_i is the user model for learner i ; s_{mn}^i is an entry in m -th row and n -th column; m is the size of the collection of target words in the system; and n is the total number of criteria in the TFA (i.e., $n = 18$).

To calculate the TFA score of each entry, the scores of all tasks that induce learning of the target word are considered. Specifically, the overall TFA score of entry s_{xy}^i is the total TFA scores of all tasks that include the target word x for criteria y as follows:

$$s_{xy}^i = \sum_{t \in T_x^i} s_y(t) \quad (2)$$

where T_x^i is the set of all tasks, which include word x as a target word, studied by learner i ; and $s_y(t)$ denotes the binary scoring function, as shown in Table 1. A larger value of s_{xy}^i implies that learner i has a higher load of learning on the target word x in terms of criteria y .

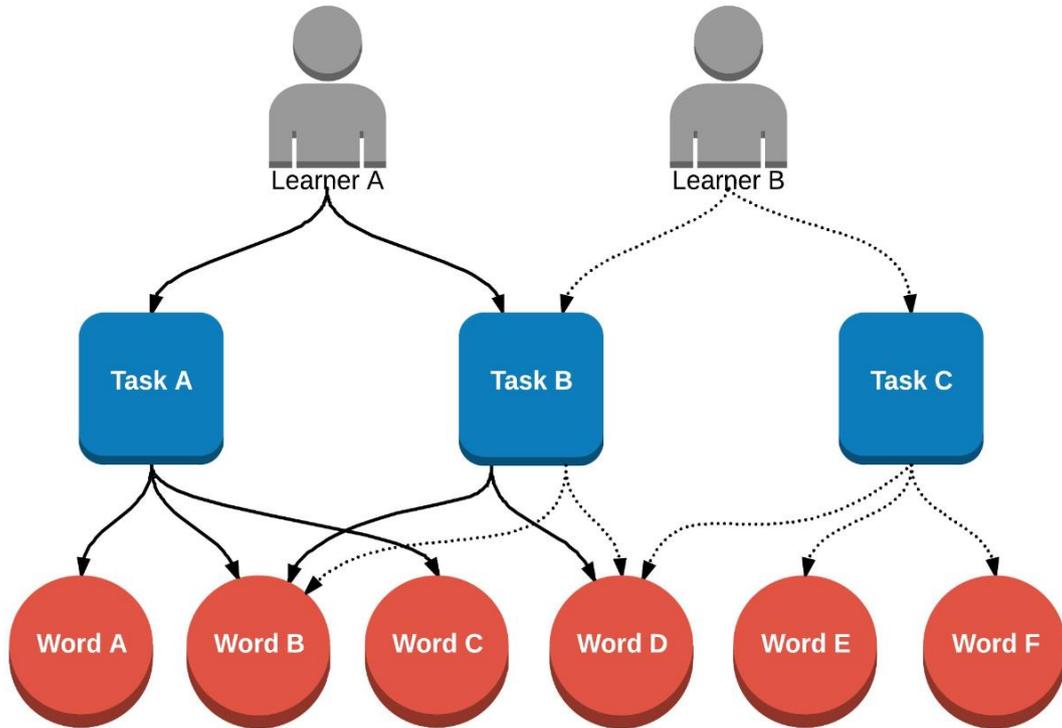


Figure 1. The relationships among tasks, target words, and learners

The generation of personalized vocabulary learning

Personalization is a learning technique which has been widely adopted by various disciplines, including natural science (Hwang, Kuo, Yin, & Chuang, 2010; Hwang, Sung, Hung, Huang, & Tsai, 2012), mathematics (Chen & Liu, 2007), and management (Xie et al., 2017). The user model based on the TFA, as explained above, enables the system to track the TFA scores of all target words in each criterion during a user's learning process. To better exploit such a user model, it is important to track the learning processes of every learner and then generate personalized learning to improve his or her learning effectiveness. The generation of a personalized learning process in this learning system is achieved via two approaches: TFA utility and task diversity.

TFA utility

The TFA checklist is utilized from two aspects by the learning system when recommending a task to a learner. Firstly, it is used to evaluate the effectiveness of several candidate tasks, and one is selected that can best promote the learning of target words. The selection of the task follows the principle that the new task induces components of the TFA criteria that are different from, but related to, the previous tasks. This aims to help users learn various aspects of knowledge of the target words via performing different tasks, and practice learning them in different ways. For example, if Task 1 focuses on the component “retention” and involves imaging and linking of form and meaning, Task 2 may then emphasize the component “generation” by asking learners to create original contexts for the target words. Moreover, Task 3 may impose retrieval of the previously learned target words as guided by the component “retrieval.”

Therefore, the first part of the TFA utility is formally defined as follows:

$$util_1(t) = \sum_{y \in C} s_y(t) \quad (3)$$

where C is the set of all TFA criteria. Based on these principles, the system recommends a task with maximal utility for learner i .

The second aspect of the TFA utility enables the system to recommend a task that can assist a learner to increase the TFA scores of entries with lower scores in the user model. This is achieved by recommending tasks that emphasize the TFA components with lower scores as induced by previous learning tasks. In this way, the TFA scores in the user model can be “fully utilized” after the participants complete the current learning task. Formally, the second part of the TFA utility is defined as follows:

$$util_2(t, L_i) = \sum_{y=1}^n \sum_{x=1}^m \frac{\Delta s_{xy}^i}{1 + s_{xy}^i} \quad (4)$$

where s_{xy}^i is an entry of the TFA score in the user model; Δs_{xy}^i is the change of the TFA score of this entry after the completion of task t ; and $1 + s_{xy}^i$ is to avoid a zero score for the denominator. The core principle here is to find a learning task that can increase the scores of more entries with low TFA scores in the user model. The priority is to fill entries with low TFA scores. The overall utility score is consolidated by using the following aggregation method:

$$util(t, L_i) = e^{util_1(t)} \times e^{util_2(t, L_i)} \quad (5)$$

where $util_1(\cdot)$ and $util_2(\cdot)$ are the two functions as defined in the equations (3) and (4), respectively. The proposed learning system basically prioritizes tasks with larger overall utility and recommends them to users.

Task diversity

Task diversity refers to the degree to which a task is different from previous tasks. For two tasks A and B which have a similar TFA utility, task A can better motivate a learner if it is more dissimilar from previous learning tasks than B . The principle here, which is suggested by Nation (2001) and partially supported by a study (Xie et al., 2016), is that tasks with greater diversity lead to better word retention than tasks of similar types. Formally, the diversity of a task t to the set of previously learned tasks T^i is defined as follows:

$$div(t, T^i) = 1 - \frac{1}{|T^i|} \sum_{t_* \in T^i} \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{t}\| \times \|\mathbf{t}_*\|} \quad (6)$$

where \mathbf{t} is a vector of TFA scores with five components (i.e., $\mathbf{t} \in \mathfrak{R}^5$) for the task t ; \mathbf{t}_* is a previous task in the set T^i ; and $|T^i|$ is the total number of previous learning tasks. The overall diversity function $div(\cdot)$ measures the dissimilarity of a task t to the set of previous learning tasks T^i by calculating the inverse of the average cosine similarity.

Table 3. An example of calculating task diversity in different granularities

Component	Criteria	t_a	t_b	t_c
Noticing	Does the activity involve generative use?	0	0	1
	Is it productive?	0	0	1
	Is there a marked change that involves the use of other words?	0	0	1
Retention	Does the activity ensure successful linking of form and meaning?	1	1	0
	Does the activity involve instantiation?	1	0	0
	Does the activity involve imaging?	1	1	0
	Does the activity avoid interference?	0	1	0

Note that we adopt a vector of the five main components of the TFA rather than all 18 criteria when calculating the degree of diversity, because the main components have better granularity than the sub-criteria in terms of deciding how similar two learning tasks are. Suppose that there are three tasks t_a , t_b and t_c , as shown in Table 3. For simplicity, only seven criteria from two main components are displayed in this example. If the criteria are used as the vector dimension for measuring diversity, the diversity between t_a and t_b is the same as that between t_a and t_c . However, this is unreasonable, as t_c compared to t_b is more different from t_a . The TFA scores of t_a and t_b are mainly in the component “retention”, while the TFA scores of t_c are mainly in the category “noticing”. Thus, the degree of diversity is more reasonable when measured from the main component level (i.e., $div(t_a, t_c)$ is larger than $div(t_a, t_b)$). Therefore, in our model, the diversity among different tasks is measured from the perspective of the main components, rather than the sub-criteria of the TFA.

Moreover, as each component of the TFA has different scales of scores (e.g., scores of “noticing” range from 0 to 3 and scores of “retention” range from 0 to 4), the score of each main component of a task $c(t)$ is normalized in the model by using the following method:

$$c_j(t) = \frac{\sum_{s_i \in c_j} s_i(t)}{|c_j|} \quad (7)$$

where c_j is a component which includes several criteria; $\sum_{s_i \in c_j} s_i(t)$ is to aggregate the TFA scores under a component; and $|c_j|$ is the number of criteria under a component.

Learning process generation

The learning process is generated in an interactive manner. As shown in Figure 2, the detailed steps of the whole interactive process are as follows.

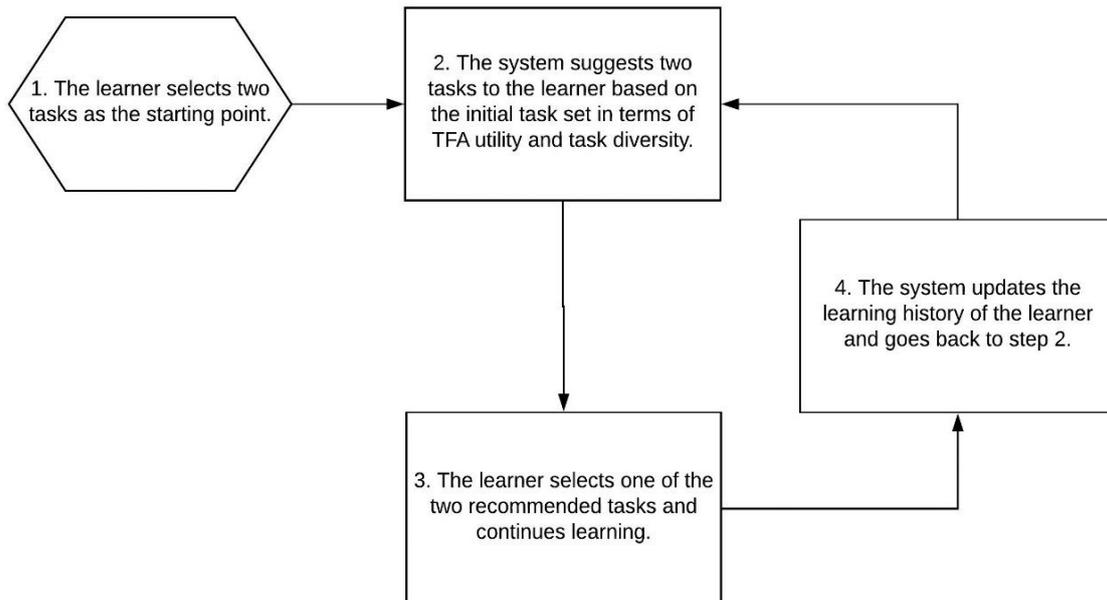


Figure 2. The interactive process of the system

- Step 1: The learner selects two tasks as the starting point;
- Step 2: The system suggests two tasks to the learner based on the initial task set in terms of TFA utility and task diversity;
- Step 3: The learner selects one of the two recommended tasks and continues learning;
- Step 4: The system updates the learning history of the learner and goes back to step 2.

An illustrative example of the overall generic framework

To better illustrate the overall generic framework, the learning process of an example user is shown in Figure 3. The learning logs record all relevant information about the tasks that the user has completed, including the task id, time, task types, target words, and the TFA scores of the tasks. Taking *Task 7* as an example, it was completed by the learner on April 23, 2017 at 16:15; it is cloze-exercises, and the target words include *renege*, *trait*, etc. The total TFA score of this task is 7, and the sub-scores of the 18 TFA criteria are individual (1, 1, 0, 1...) as listed in brackets. The user model is a matrix as defined in Equation (1). Each entry of the matrix denotes the history record of a participant’s learning of a target word in terms of one of the 18 TFA criteria. For example, entry 4 of the first column of the first row denotes that the first target word has been learned four times by this user through performing four tasks with clear vocabulary learning goals. According to the learning logs and user models, the system suggests learning tasks based on the TFA utility and task diversity as respectively defined in Equation (5) and Equation (6). Normally, the suggested learning tasks: (1) are of different types as compared to the tasks that have been completed by the user recently; and (2) tend to focus more on the target words that the user has not encountered through performing other tasks. In each cycle, when the user has selected and completed a task, the system updates the user’s learning logs and user models automatically, and based on the latest data, the system further makes recommendations of new learning tasks. As the system continues to iterate, the user’s learning experience throughout the entire learning processes is personalized with respect to his or her learning logs and user models, and the recommended tasks are guided by TFA utility and task diversity.

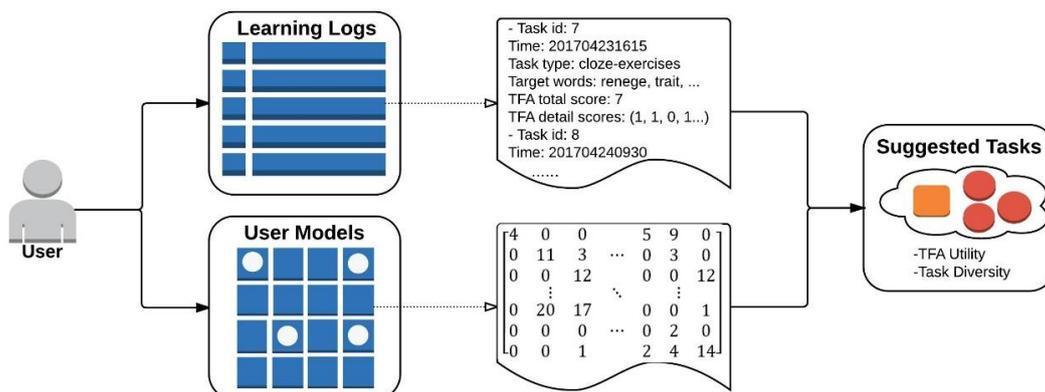


Figure 3. An illustrated example of the overall generic framework

Experiments

Subjects

To verify the effectiveness of the proposed system, real subjects were invited to participate in the experiments. A total of 105 students from universities in Hong Kong and mainland China participated in the study. Their ages ranged from 18 to 28, and a wide variety of majors were covered, including business studies, engineering, humanities, social sciences, biology, medicine, and physical sciences (see Table 4).

These participants were randomly assigned into three groups. Each group adopted a different learning approach to using the word-learning task recommendation system. Group A adopted a non-personalized learning approach, while Group B and Group C adopted personalized learning approaches. The main difference between the approaches adopted by Group B and Group C is that the generation of personalized learning processes for the participants of Group B was guided by an earlier version of the TFA, where only three main components: noticing, retrieval, and generation are included; whereas, the generation of personalized learning processes for the participants of Group C was guided by the complete checklist of TFA, as shown in Table 1. The earlier version of the TFA does not quantify features concerning elaboration, while the current checklist of TFA covers

such features and increases the number of elaboration parameters (Nation, 2001; Hu & Nassaji, 2016). The information related to the three groups is summarized in Table 5.

Table 4. Information about participants

Attributes	Attribute values	Counts
Age	18-21	79
	22-25	25
	25-28	1
Programmes	Business studies	18
	Engineering	29
	Humanities, social sciences	22
	Biology and medicine	12
	Physical sciences	25
Gender	Male	49
	Female	56
Region	Hong Kong	54
	Mainland	51

Table 5. Allocation of participants to groups

Groups	Learning strategy	Counts
A	Non-personalized learning (self-paced learning)	35
B	Personalized learning using the earlier version of the TFA with three components	35
C	Personalized learning using the current version of the TFA with five components	35

Experimental procedures

To verify the effectiveness of the proposed system in promoting vocabulary learning, the following experimental procedures with two stages, as shown in Figure 4, were developed and conducted.

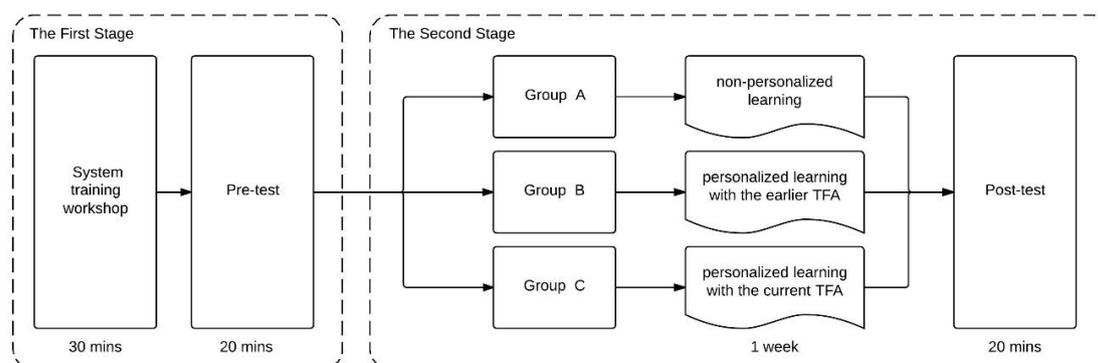


Figure 4. The experimental procedures

The first stage

The first stage involves two steps. Firstly, the subjects were invited to participate in an online training workshop on how the system can be used. After this, the participants' prior knowledge of 60 candidate target words were evaluated. Based on the test results, 40 words that were unknown to almost every participant were selected as the target words of this study, and thus the participants' pre-knowledge of the 40 target words was almost zero. All participants were randomly grouped into three teams.

The second stage

The second stage also includes two steps: (1) vocabulary learning via the system; and (2) post-testing. The participants were asked to learn the 40 target words through performing different tasks as recommended by the system in a week. The system includes a bank of 20 types of tasks, and the participants were asked to complete

at least one task per day. The three groups of participants applied three approaches to word learning: a non-personalized approach for Group A; a personalized approach guided by a partial version of TFA for Group B; and a personalized approach guided by the full list of TFA for Group C. Specifically, the participants of Group A decided what tasks to do by themselves, as no recommendations were given by the system. Tasks were recommended to the participants of Group B based on the earlier version of the TFA. The participants of Group C were recommended with tasks based on the current version of the TFA. After one week of learning, all participants were tested to evaluate their knowledge of the target words by utilizing a modified vocabulary knowledge scale. This assessment tool has been used in Zou (2017), and the same grading criteria used by Zou (2016) were employed to mark the post-test.

Learning logs

There were 1057 learning logs in total, recording the tasks that were completed by different users during their learning processes. Each user completed a total of 10.07 tasks on average, and 1.43 tasks per day. The maximum number of tasks that were completed by a single user was 20, and the minimum was seven. This is probably because the participants were required to complete one task per day. The distribution of the numbers of learning logs is presented in Figure 5. The vertical axis denotes the number of users, and the horizontal axis denotes the number of logs. 80% of users (i.e., 84 users) completed 7 to 12 tasks within the learning period. There was no significant difference among the numbers of tasks completed by the three groups of participants.

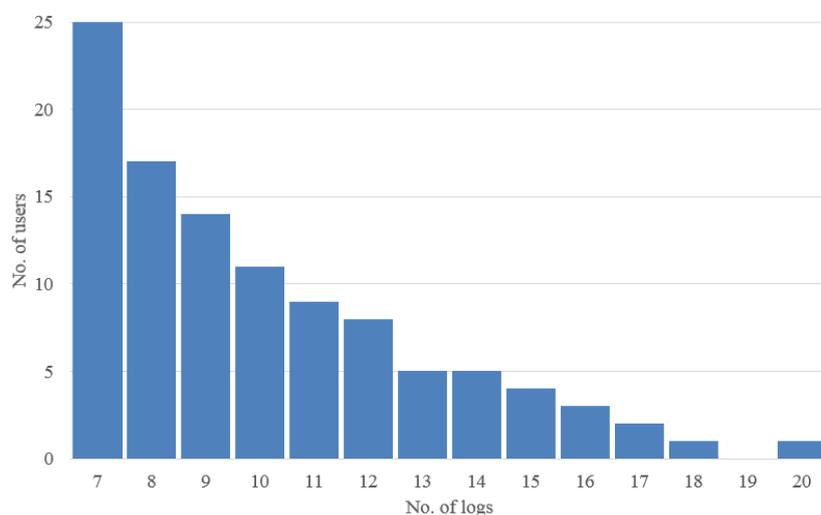


Figure 5. The distribution of the numbers of learning logs

Results

As shown in Table 6, the proposed learning system is very effective in promoting the learning of the target words, given that the participants' prior knowledge of these words was almost zero. The learning performance of the participants of Group C (a mean score of 74.74) was better than that of the participants in Group B (a mean score of 68.28), and the mean score of the participants in Group A (a score of 61.60) was the lowest among the three groups.

Table 6. Group performance of post-test

	<i>N</i>	Mean	<i>SD</i>
Group A	35	61.60	9.20
Group B	35	68.28	8.47
Group C	35	74.74	8.68

Table 7. Results of the one-way ANOVA test of participants' scores

	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2
Between groups	3023.16	2	1511.58	19.54	0.00	0.28
Within groups	7888.22	102	77.33			
Total			10911.39			

To further examine whether any significant differences existed among the three groups, a one-way ANOVA test was applied, the results of which, as demonstrated in Table 7, indicated statistical significance ($F(2, 102) = 19.54, p < .001, \eta^2 = .28$).

Discussion and conclusion

The results of the research provide empirical supporting evidence for the effectiveness of the proposed e-learning system which recommends word learning tasks based on the TFA scores of different tasks and user models, as the group of participants who learned the target words through a personalized approach guided by the full list of TFA criteria had the best learning performance among the three groups. The effectiveness of this personalized recommendation system results, to a large extent, from the fact that it provides a personalized learning experience and suggests tasks based on TFA utility and task diversity according to users' learning logs and user models. The results also indicated that e-learning systems should be designed based on comprehensive learning theories, because the group of participants who learned the target words through an approach guided by a partial list of TFA criteria had less effective learning performance than the participants whose learning approach was guided by the full list.

Moreover, it is suggested that effective word learning requires encountering or processing information of target words in different circumstances while performing a wide range of word learning tasks. In other words, a combination of several tasks with different TFA scores is more conducive to word learning than simple repetition of similar tasks. Because tasks with different focuses on the five main components of the TFA checklist promote the learning of different aspects of knowledge of the target words, a combination of different focuses entails a higher probability of building up networks of the target words. Learners are advised to perform tasks with lower TFA scores first, and then move to tasks with higher scores for the learning of certain target words. Additionally, the results further support the argument that a personalized vocabulary learning system is more conducive to word learning than a non-personalized system.

In sum, this study developed a personalized vocabulary learning system under the umbrella of the checklist for TFA, and the experiment results demonstrated that the proposed system is very effective in promoting word learning. The major contribution of the study is to evidence that a comprehensive theoretical framework is essential for the optimal design of learning systems. This research also implies that language education should move toward employing more personalized learning systems, considering that they are very conducive to language learning.

Future studies are suggested to focus on: (1) how to exploit user models to facilitate the learning of various aspects of word knowledge; (2) how to better generate vocabulary learning tasks by using the checklist of the technique feature analysis; and (3) how to integrate the prior knowledge and learning styles of participants in the recommendation process so as to improve the personalized learning experience.

Acknowledgements

This study was fully supported by the Start-Up Research Grant (RG 54/2017-2018R) and the Internal Research Grant (RG 63/17-18R) of The Education University of Hong Kong, and the Dean's Reserve-funded Learning and Teaching Project (75.8ACD), The Hong Kong Polytechnic University.

References

- Barker, D. (2007). A Personalized approach to analyzing "cost" and "benefit" in vocabulary selection. *System*, 35(4), 523-533.
- 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.
- Chen, C. M., & Hsu, S. H. (2008). Personalized intelligent mobile learning system for supporting effective English learning. *Educational Technology & Society*, 11(3), 153-180.
- Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, 18(4), 341-364.

- Chen, C. J., & Liu, P. L. (2007). Personalized computer-assisted mathematics problem-solving program and its impact on Taiwanese students. *The Journal of Computers in Mathematics and Science Teaching*, 26(2), 105-121.
- Huang, Y. M., Huang, Y. M., Huang, S. H., & Lin, Y. T. (2012). A Ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education*, 58(1), 273-282.
- Huang, C. S., Yang, S. J., Chiang, T. H., & Su, A. Y. (2016). Effects of situated mobile learning approach on learning motivation and performance of EFL students. *Journal of Educational Technology & Society*, 19(1), 263-276.
- Hsieh, T. C., Wang, T. I., Su, C. Y., & Lee, M. C. (2012). A Fuzzy logic-based personalized learning system for supporting adaptive English learning. *Journal of Educational Technology & Society*, 15(1), 273-288.
- Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A Personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education*, 63, 327-336.
- Hu, H. C. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement load hypothesis versus technique feature analysis. *System*, 56, 28-39.
- Hwang, G. J., Kuo, F. R., Yin, P. Y., & Chuang, K. H. (2010). A Heuristic algorithm for planning personalized learning paths for context-aware ubiquitous learning. *Computers & Education*, 54(2), 404-415.
- Hwang, G. J., Sung, H. Y., Hung, C. M., Huang, I., & Tsai, C. C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educational Technology Research and Development*, 60(4), 623-638.
- Jung, J. Y., & Graf, S. (2008). An Approach for personalized web-based vocabulary learning through word association games. In *International symposium on applications and the Internet, SAINT 2008* (pp. 325–328). doi:10.1109/SAINT.2008.63
- Kwon, D. Y., Lim, H. S., Lee, W., Kim, H. C., Jung, S., Suh, T., & Nam, K. (2010). A Personalized English vocabulary learning system based on cognitive abilities related to foreign language proficiency. *Transactions on Internet and Information Systems*, 4(4), 595-617.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The Construct of task-induced involvement. *Applied linguistics*, 22(1), 1-26.
- Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford, UK: Oxford University Press.
- Liu, T. Y. (2009). A Context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning*, 25(6), 515-527.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.
- Sandberg, J., Maris, M., & Hoogendoorn, P. (2014). The Added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning. *Computers & Education*, 76, 119-130.
- Schmitt, N. (2000). *Vocabulary in language teaching*. New York, NY: Cambridge University Press.
- Wang, Y. H., & Shih, S. K. H. (2015). Mobile-assisted language learning: Effects on EFL vocabulary learning. *International Journal of Mobile Communications*, 13(4), 358-375.
- Wu, T. T., Sung, T. W., Huang, Y. M., Yang, C. S., & Yang, J. T. (2011). Ubiquitous English learning system with dynamic personalized guidance of learning portfolio. *Journal of Educational Technology & Society*, 14(4), 164-180.
- Xie, H., Zou, D., Lau, R. Y., Wang, F. L., & Wong, T. L. (2016). Generating incidental word-learning tasks via topic-based and load-based profiles. *IEEE multimedia*, 23(1), 60-70.
- Xie, H., Zou, D., Wang, F. L., Wong, T. L., Rao, Y., & Wang, S. H. (2017). Discover learning path for group users: A Profile-based approach. *Neurocomputing*, 254, 59-70.
- Zou, D. (2016). Comparing dictionary-induced vocabulary learning and inferencing in the context of reading. *Lexikos*, 26(1), 372-390.
- Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research*, 21(1), 54-75.
- Zou, D., & Lambert, J. (2017). Feedback methods for student voice in the digital age. *British Journal of Educational Technology*, 48(5), 1081-1091.
- Zou, D., Xie, H., Li, Q., Wang, F. L., & Chen, W. (2014). The Load-based learner profile for incidental word learning task generation. *Advances in Web-Based Learning, LNCS*, 8613, 190-200.
- Zou, D., Xie, H., Rao, Y., Wong, T. L., Wang, F. L., & Wu, Q. (2017). A Comparative study on various vocabulary knowledge scales for predicting vocabulary pre-knowledge. *International Journal of Distance Education Technologies*, 15(1), 69-81.