

A Contribution-Oriented Self-Directed Mobile Learning Ecology Approach to Improving EFL Students' Vocabulary Retention and Second Language Motivation

Zhuo Wang¹, Gwo-Jen Hwang², Zhaoyi Yin³ and Yongjun Ma^{1*}

¹Qingdao University, China // ²National Taiwan University of Science and Technology, China // ³University of Glaskow, United Kingdom // zhuowang@qdu.edu.cn // gjhwang.academic@gmail.com // piperyin@126.com // qdmayongjun@163.com

*Corresponding author

ABSTRACT: Vocabulary mastery is critical to English as Foreign Language students. Mobile technologies enable students to learn vocabulary without space and time limitations. However, existing mobile-assisted vocabulary learning research often employed teacher-directed activities that increased instructors' workload, undermined student motivation or targeted individual cognitive outcomes only. In this study, a Contribution-oriented Self-Directed Mobile Learning Ecology (CSDMLE) model is proposed for developing student-directed and motivational vocabulary learning activities in groups. Through a mixed-method design, we administered a survey and a vocabulary test to 55 freshmen students in a Chinese university, and conducted follow-up interviews. We found that students in the CSDMLE group outperformed those not in the group in the post-test vocabulary test, but there was no statistically significant difference between the two groups' delayed vocabulary test or L2 motivation. However, the treatment group displayed a highly favorable attitude toward the learning approach and a strong intention to use it continuously. The findings have implications for technology-supported vocabulary learning activities.

Keywords: EFL, Mobile learning, Vocabulary retention, L2 motivation

1. Introduction

Vocabulary acquisition is fundamental in learning a second or foreign language (L2) (Hwang & Wang, 2016; Tight, 2010). Learners who master vocabulary well are more likely to produce better language performance. However, it is often a long and tedious process (Chen et al., 2019); if learners are not motivated to learn or they do not know how to learn effectively, they might give up learning L2 vocabulary (Dörnyei & Csizér, 2002). Generally, learners can acquire vocabulary effectively under instructors' guidance, but class time is often limited, and language instructors cannot guide and monitor students out of class with traditional learning activities.

The advancements in computing technologies have dramatically changed the way we live as well as how languages are learned (Hung et al., 2018). Among others, the use of mobile technologies has grown rapidly worldwide (Sundberg & Cardoso, 2019), as evidenced by the rapidly increasing rate of device ownership, and the wider coverage of mobile-cellular networks in both developed and developing countries (Kaliisa et al., 2019; Huang et al., 2010). Due to its advantageous affordances such as connectivity, ubiquity and interactivity (Klopfer et al., 2012), many researchers have conducted studies on technology use during language learning processes, including vocabulary learning (Gürkan, 2019). Meta-analyses of mobile-assisted language learning studies in the last decade indicated that vocabulary outcomes were the most frequently researched variable (Hwang et al., 2019; Elaish, 2019).

Although such studies attest to the benefits of mobile-assisted vocabulary learning (MAVL), they also unintentionally conveyed two misleading messages: one, MAVL has to be teacher-directed or demands tremendous effort from instructors; two, whether more words were retained is the most important criterion for activity design or platform selection. In many of the reviewed studies, the content sent out to students was either originally created or appropriately tailored from existing resources, which would place heavy demands on the instructors, both cognitively and physically (e.g., Pirasteh & Mirzaeian, 2015). The workload of creating and frequently distributing course vocabulary content might deter many language instructors from incorporating mobile learning to its full potential. On the other hand, while we acknowledge that vocabulary memorization and retention is critical to language success, it should be neither examined solely, nor over-emphasized to such an extent that overshadows learner interest, ease of use, motivation or other affective and perceptive factors. When learners feel overwhelmed by using a technology, it is probably they will terminate its use once they have a choice. Thus, researchers need to balance between pursuing

cognitive outcomes and catering to students' attitudinal and emotional needs when they integrate MAVL approaches for long-term use.

In all, the above needs demand MAVL designs that not only alleviate instructor workload, but also examine cognitive and affective changes with equal attention. One of the most important affective factors in language learning is second language (L2) motivation. Distinguished from generic motivation, a term that is often loosely used to encompass various emotional aspects, L2 motivation refers specifically to one's motivation to acquire a second or foreign language (Dörnyei, 2005). It is measured with certain established instruments worldwide, such as AMTB (attitude and motivation test battery). Understanding whether and how MAVL could impact one's L2 motivation may provide insightful guidance for language instructors.

To address the above challenges in MAVL research, we propose a Contribution-oriented Self-Directed Mobile Learning Ecology (CSDMLE) model to comprehensively guide our design of effective MAVL experience. Specifically, the model was hypothesized to improve students' vocabulary retention better than traditional approaches through utilizing related pedagogical theories (i.e., theory of multimedia learning); to reduce instructor workload, we asked students to create and share vocabulary learning content themselves; to boost their L2 motivation, we set goals that stimulated their sense of contribution and responsibility. Accordingly, our research questions are listed as follows:

- Is there any significant difference in the post and delayed vocabulary retention performance of those who learned via the CSDMLE model and those who did not?
- Is there any significant difference in the L2 motivation of the participants using the CSDMLE model and those using conventional learning?
- How do students in the CSDMLE group perceive this learning approach?

2. Literature review

2.1. Vocabulary acquisition and retention

Vocabulary plays an indisputably vital role in students' L2 learning (Tight, 2010). Historically, language learners have expressed significant difficulty remembering vocabulary words (Chen & Chung, 2008), and retention is one of the most difficult learning problems to address, due to the unavoidable forgetting nature of human beings (Ebbihaus, 1913). Research on cognitive science has suggested that a list of principles be followed to enhance long-term memory of learning materials, such as using images or graphics to assist verbal learning (Driscoll, 2005). Regardless, traditional vocabulary instruction is often limited in terms of both class time invested and effective retention strategies employed. Students need to mainly rely on themselves for vocabulary learning, which could create problems and frustration for student learners and result in their loss of motivation (Dörnyei & Csizér, 2002).

Two prominent approaches were often used to elicit successful vocabulary learning and retention: multimodal presentation and spaced repetition (Kohnke et al., 2019). The former entails supporting word understanding and retrieval with multimedia, which is underpinned by the Cognitive Theory of Multimedia Learning (Mayer, 2009). It postulates that word knowledge is acquired through visual and verbal channels. When learners obtain word knowledge with multiple media forms, it stimulates both channels and strengthens one's memory retention. Spaced repetition refers to a programmed system with designated time intervals that provides a series of presentations or practices of vocabulary content (Kohnke et al., 2019). Being regularly and rhythmically exposed to word knowledge, learners can efficiently maximize their understanding and elongate their knowledge retention (Pellicer-Sánchez & Schmitt, 2010). Effective use of both the multimodal and spaced repetition approaches promises to yield satisfactory vocabulary learning outcomes.

2.2. MAVL research trends and limitations

The integration of mobile technologies and devices in vocabulary learning has gradually led to the field of MAVL. There has been a steadily increasing number of MAVL studies since the last decade. More recently, several meta-analyses have been conducted to synthesize MAVL research trends and gaps on different levels, including effect size, research settings, aspect of vocabulary knowledge, study duration, etc. For example, Lin and Lin (2019) found

that learners generally displayed a positive attitude toward MAVL. Mahdi (2018) concluded that receptive knowledge was exploited more frequently than productive knowledge. In this paper, we aimed to highlight certain limitations or gaps that warrant imminent attention for MAVL activity and research design.

First of all, in terms of goal-setting, most studies sought to improve individual outcomes, such as vocabulary retention and learning interest. For example, Alemi et al. (2012) conducted research upon 45 freshmen students and found statistically significant difference in treatment and control groups' delayed test mean score. Chen et al. (2019) found that primary students who learned vocabulary via their app-based self-regulated mechanism improved vocabulary retention and motivation significantly better than those in control group. While these are indeed important learning objectives, they do not emphasize sense of community, collaboration or socialization skills that are in rapid need today.

Secondly, regarding the content design and form of MAVL activities, there is a lack of studies employing student-centered approaches, such as Constructivism or self-directed learning. Most existing MAVL research reported activities that were still teacher-directed that undermined learner autonomy. This reflects a Behaviorist epistemology, placing students as passive knowledge recipients (Hu, 2013). For instance, undergraduate students in Pirasteh and Mirzaeian's (2015) study were reported to receive phrasal verb content prepared by course instructors through SMS every day for 25 days. This not only limited student output or productive skill development, but also created additional workload for instructors. It is observed that teachers who integrate mobile learning often need to commit more effort, such as digitizing the content to be placed in mobile devices, ensuring functionality, and solving emergent technical difficulties (Shih et al., 2010). Even with positive results attained, it remains questionable if instructors who went through the tedious process of creating and distributing content would persist such an endeavor in a longer term. Few studies epitomized student-generated content and self-directed learning. For example, 24 Iranian EFL students in Foomani and Hedayati's (2016) study took photos to demonstrate word usage and shared them on Padlet for peer discussion, but the study employed a pre-experimental design that was mainly descriptive. Botero et al. (2019) examined whether using Duolingo out of class could promote 118 university language students' self-directed learning, and found that students lacked sustained motivation in such learning and needed stronger sense of responsibility. Wong and Looi (2011) reported two case studies in which primary students took photos and created sentences for class discussion, and advocated to treat student-generated content as the "end."

Thirdly, L2 motivation, which is the most reliable predictor of language learners' long-term effort in L2 learning (Dörnyei, 2005) was rarely examined in MAVL studies. Although the term "motivation" was often mentioned in MAVL studies, it has been used more as an umbrella term for constructs like learning interest, intention to use, satisfaction, and may denote meanings that vary from study to study. For example, in Looi's et al. (2011) research, motivation was depicted as students' attitude toward and engagement in mobile learning, and relevant results were obtained based on classroom observation and a self-designed survey. More recently, in Loewen et al.'s (2019) study, eight participants' motivation level was inferred from their learning journal, in which they described their interest in and mood for learning Turkish via Duolingo. In terms of L2 motivation, AMTB developed by Gardner (1985) is a widely used instrument among language scholars worldwide. For example, Jain and Sidhu (2013) in Malaysia used AMTB to measure freshmen students' L2 motivation, and found that increasing anxiety would reduce their level of motivation, regardless of discipline, gender or language proficiency. Rahmany et al. (2013) used AMTB to determine the L2 motivation level of 60 Iranian EFL of different age groups and found that extensive reading did not elicit better L2 motivation.

Meanwhile, although university students were frequent participants in MAVL studies (e.g., Yuan, 2019; Hanson & Brown, 2020), there is a lack of research on pre-service teachers. Yet, how their perception of English, and the way they were taught English could potentially impact their future teaching philosophy and performance to great extent. Thus, investigating how pre-service teachers might benefit from innovative learning interventions could have a far-reaching significance.

Overall, the various limitations identified above demand a more comprehensive framework that is grounded upon solid pedagogical and instructional theories, and provides clear guidance for MAVL design in terms of content to be used, form it takes and a goal that is motivating and yet practical. The following model was designed as a response to this demand.

3. The contribution-oriented self-directed mobile learning ecology model for vocabulary learning

To meet language learners' both cognitive and affective demands, and begin to address the identified gaps, we developed the Contribution-oriented Self-Directed Mobile Learning Ecology model, specifically informed by three theoretical frameworks. In this model (see Figure 1), the triangle represents the three pedagogical aspects that were identified in our literature review as lacking improvement, namely the goal, content and form of the MAVL design. Secondly, the inner circle consists of three corresponding patterns that are deemed as problematic. Thirdly, the outer circle depicts our CSDMLE model with three key components, which are in direct contrast with the previous approach displayed in the inner circle. For example, while the majority of MAVL studies focused on producing individual outcomes such as increased vocabulary test scores, our model advocates collaborative gains in addition to individual growth. Finally, each component in the outer circle is supported by and grounded upon a particular theory as introduced further below.

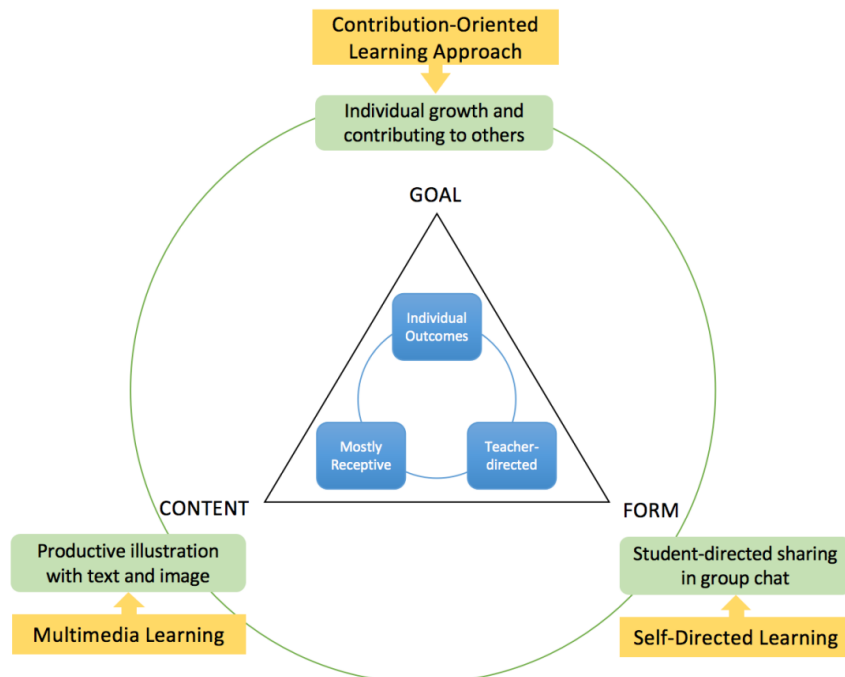


Figure 1. The contribution-oriented self-directed mobile learning ecology model

To begin with, the goal-setting was inspired by the Contribution-Oriented Learning Approach (COLA) proposed by Collis and Moonen (2001). COLA is a pedagogical theory that advocates the contributing role of individual learners in online environments. It characterizes the role of the instructor as facilitator and coordinator of activities, and that of students as learning resource creators and designers who should “contribute to make a difference” (Collis & Moonen, 2006). A distinct feature of COLA-informed activities is that students produce meaningful resources that can be practically used or reused by others for authentic purposes at a later time (Collis & Moonen, 2001). Such resources are in sharp contrast with traditional assignments that are often deemed as learning evidence and offer limited value beyond the individual students. The goal of making an actual contribution is believed to encourage students to take responsibility for their own learning, foster a sense of community as well as build a collaborative culture.

Next, the theory of Multimedia Learning, which guides an effective design of multimodal information presentation, was used to guide students' productive content design. Since its central premise is that using both verbal and visual channels is more effective than using either alone for promoting understanding and retention (Mayer, 2009), students were asked to create illustrations with text (contextualized sentence-making for a chosen word) and images (a corresponding picture that echoes the text) (see Figure 2).

The third theory is Self-directed learning (SDL) proposed by Knowles (1975). SDL is a “basic human competence--the ability to learn on one’s own” (Knowles, 1975, p. 17), and it has been well researched in the field of adult education. Self-direction is perceived as a significant component of achieving meaningful educational outcomes (Garrison, 1997). In order to explain what SDL encapsulates, Garrison (1997) proposed a comprehensive model consisting of three fundamental yet highly interconnected dimensions, including self-management, self-monitoring and motivation. In particular, self-management refers to learners’ active control during the learning process, but the control “must balance educational norms and standards with student choice and the responsibility for constructing personal meaning” (Garrison, 1997, p. 23). Self-monitoring encompasses the cognitive and meta-cognitive processes and refers to learners taking responsibility for active meaning making and critical reflection. Finally, motivational factors have a pervasive influence on learners’ goal-setting and subsequent task effort (Corno, 1989). More importantly, the entering motivation or the motivation to enter into a task plays a significant role in learners’ assessment of task value and attainability.



Figure 2. Two samples of student-created vocabulary illustrations

Congruently informed by all theories and components in the CSDMLE model, the activity was expected to proceed according to the following stages:

- (1) Entering motivation: the researchers describe the activity and allow students to determine whether they want to participate after assessing the task difficulty on their own.
- (2) Acquisition: students select a word from the required textbook glossary and study its meaning and usage.
- (3) Self-managing: students actively control their learning pace, the resources they want to consult, the applications (apps) they want to use, and the extent to which they conform with the task standards. Specifically, each student should produce an illustration that displays both the chosen word’s contextualized usage and an image that complements the text.
- (4) Contribution: students post their illustrations to a designated group chat in WeChat, the most popular social media app in Mainland China, so that students in the same group can view and learn about the shared resources.
- (5) Self-monitoring: Through viewing illustrations shared by other students in the group chat, learners actively compare and connect their own understanding and others’ presentations, and reflect on the quality of and strategies used for their last illustration.
- (6) Enhancing motivation: Feeling surprised or benefiting from others’ illustrations, students are motivated to continually improve their own and produce quality content for peers.

4. Methodology

4.1. Research design

According to Creswell (2009), a sequential explanatory mixed-method design refers to using quantitative data collection first and qualitative methods later that builds on the former. Such a strategy is appropriate when researchers intend to explain and interpret quantitative results by collecting and analyzing follow-up qualitative data (Creswell, 2009), and can be especially helpful when unexpected findings emerge from a quantitative study (Morse, 1991). The purpose of the present study was to determine the effect of the CSDMLE model on students’ vocabulary

retention and L2 motivation. Given the novelty of our model, we also anticipated that certain unexpected results might arise. Thus, follow-up qualitative data collection was added to help us interpret any perplexing findings.

4.2. Participants

Participants were two freshmen classes majoring in English at a Chinese normal university who were recruited via convenience sampling. Although the two groups were taught by different instructors of the same course, “Contemporary English,” their college entrance examination scores did not differ statistically significantly. A total of 28 students were in the experimental group (25 females and 3 males) and there were 27 in the control group (24 females and 3 males). The mean ages were 18.89 and 18.96, respectively.

4.3. Procedure

Both classes were first invited to complete the consent form and a pre-study survey online. Upon receipt of their responses, the researcher randomly assigned the participating classes to either the experimental group (EG) or the control group (CG). Each group then created a group chat on WeChat.

4.3.1. Pre-training

Before the study began, the researchers first joined both group chats on WeChat, and disseminated the study requirements via PowerPoint slides for the respective groups. Students in either group were then given one day to raise any questions or concerns about the study. For the CG, researchers explained the dates and form of upcoming vocabulary tests and surveys, as well as the use of the group chat for such purposes; for the EG, beside tests and surveys, the researchers also described the steps taken to create a quality illustration, demonstrated exemplary illustrations to help EG students visualize what was required, and elaborated on the posting schedule. The instructors for either group were not only asked to undertake the same instructional practice in class (selecting the same key words and phrases to instruct directly, and spend the same amount of time on vocabulary instruction), but also invited into the group chat so that they were aware of all activities and could respond promptly if unexpected problems arose.

4.3.2. Study participation

As shown in Figure 3, after completing a survey and a subsequent vocabulary test online, the CG continued to learn vocabulary in class and used their conventional approaches at will; the EG learned vocabulary in the same way as the CG when in class, but created and shared illustrations in the group chat in their spare time out of class. The EG’s schedule was as follows: posting their first illustration by 8pm on Tuesday, and their second by 8pm on Friday. The study lasted for 2 weeks, with two illustrations per student each week. Informed by Spaced Repetition, such a schedule spaced out the students’ illustrations across the week, and increased their times of exposure to vocabulary knowledge, which would highly probably lead to more effective and efficient vocabulary retention. It should be noted that while the instructor was present in the EG’s group chat, she had been politely asked to only intervene when an illustration contains incorrect information that was not timely revised by the author student him- or her- self. Immediately after the study ended, all students completed a survey and took the vocabulary test again. Two weeks later, all students took a delayed vocabulary test. Three weeks later, four EG students participated in the interview.

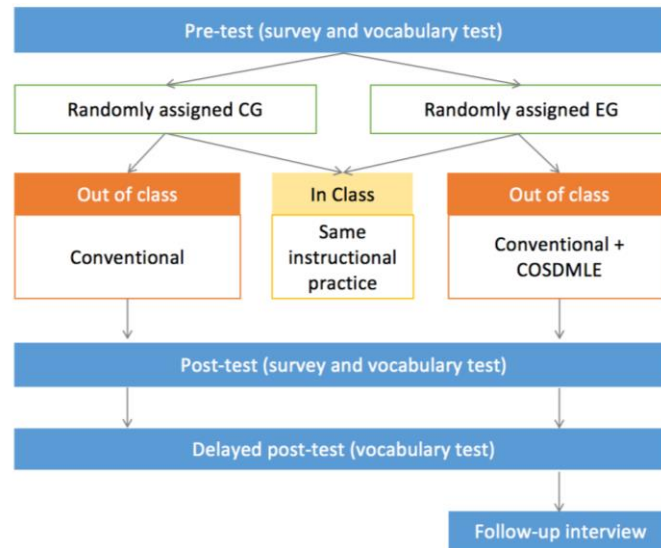


Figure 3. Study procedure

4.4. Instruments

The instruments used in the present study include a vocabulary test, a survey on L2 motivation, a questionnaire on learning experience and satisfaction, and an interview.

The vocabulary test took the form of active recall, asking students to provide the Chinese meaning for the given English lexical items. A total of 60 vocabulary items were selected by the English instructors as worth being included (considered new or difficult to memorize) in the target learning module. In random orders, these same items were tested in a pre-, post- and delayed-test fashion. Two raters scored the tests independently first, and then discussed those with differing opinions until they achieved agreement. The final analysis of data only included items that were illustrated by the students.

The L2 motivation survey was adapted from Gardner's (1985) AMTB, an instrument developed specifically to evaluate learners' L2 motivation. The Cronbach's α value is .90. In our study, three sub-scales were included: Attitude toward English Learning (ATEL), Motivational Intensity (MI), and Desire to Learn English (DTLE). ATEL evaluated respondents' general attitude toward English learning; MI assesses the intensity of a student's motivation to learn an L2 such as their effort in classroom assignments, future plans about their language study, etc. DTLE inquired about students' desire to learn an L2. A Chinese version of the survey, which has been validated by four English teachers through back-translation, was provided to the students.

The learning experience and satisfaction questionnaire has 18 items. It was developed by the researchers and then validated by two experts in the Instructional Technology field with over 10 years' experience. It asked about the EG's perceptions of their learning, usage of the illustrations and related behaviors. There were 13 multiple-choice questions, three checklists, one matrix, one ranking and one open-ended question.

Additionally, semi-structured interviews (through phone and instant messaging) were conducted with four EG participants to understand their perceptions and opinions based on the following questions:

- What do you consider the most beneficial features in this MAVL experience?
- What factors have hindered you from making the most of this learning?
- What did you do with the illustrations shared by others?
- What was the role of your English instructor?

5. Results

5.1. Vocabulary retention

The descriptive statistics for the three tests of EG and CG were presented in Table 1. We conducted an independent *t*-tests for pre-test, post-test and delayed test respectively (Table 2). Because of the voluntary nature of the study, some students chose not to participate in all three tests, and few missed the post-test or delayed test, during which they needed to address other priorities, such as course assignments or interest community meetings. Thus, the sample size varied in each test. The results indicated that there was statistically significant difference between the two groups' pre-test and post vocabulary test scores, but no statistically significance in the delayed test.

Specifically, in the pre-test, CG scored statistically significantly better than the EG ($t = -2.60, p < .05$). A possible explanation would be that some Chinese students had a habit of previewing or self-teaching learning content in upcoming modules in advance, so that they would understand better in class; therefore, these students would attain better scores even when being tested on content that was not taught yet. According to the results and line graph (Figure 4), EG caught up by the post-test and outperformed CG with statistical significance ($t = 2.42, p < .05$).

Table 1. Descriptive summary for vocabulary scores

	Group	<i>N</i>	Mean	Std. deviation	Std. error mean
Pre	EG	28	20.11	8.04	1.52
	CG	27	29.00	16.14	3.11
Post	EG	22	52.64	6.45	1.37
	CG	20	45.30	12.55	2.81
Delayed	EG	26	46.96	10.58	2.07
	CG	23	42.70	13.73	2.86

Table 2. Independent *t*-test results for vocabulary scores

	Levene's test for equality of variances		<i>t</i> -test for equality of means						
	<i>F</i>	<i>Sig.</i>	<i>t</i>	<i>df</i>	<i>Sig.</i>	Mean difference	Std. error difference	95% CI	
								Lower	Upper
Pre	13.45	0.00	-2.60	53.00	0.01	-8.89	3.42	-15.75	-2.03
Post	3.64	0.06	2.42	40.00	0.02	7.34	3.04	1.20	13.48
Delayed	0.42	0.52	1.23	47.00	0.23	4.27	3.48	-2.73	11.27

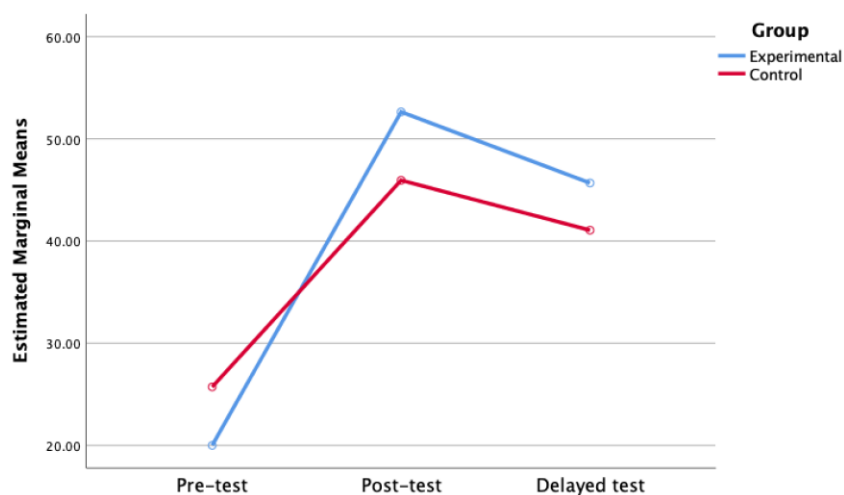


Figure 4. Vocabulary test mean score plot

5.2. L2 motivation

According to the two-way mixed ANOVA (see Table 3), there was no statistically significant interaction between treatment and time on motivation, $F(1, 51) = 0.51, p > .05$, partial $\eta^2 = 0.01$. In other words, the two groups were not statistically significantly different in terms of L2 motivation at pre- or post-test. Visually, it can be seen in Figure 5 that EG increased slightly more than the CG from pre-test to post-test, but the result needs to be interpreted with data from the survey and interview holistically.

Table 3. Two-way mixed ANOVA summary table for motivation

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Time	0.08	1	0.08	1.81	.18	0.03
Time * Group	0.02	1	0.02	0.51	.48	0.01
Error(Time)	2.28	51	.05			

Table 4. Correlation analysis for design factors predicting post L2 motivation

	Assistance of image	Sense of rapport	Sense of contribution	Dictionary use
Post-L2Motiv	0.526**	0.514*	0.579**	0.571*

Note. * $p < .05$; ** $p < .01$.

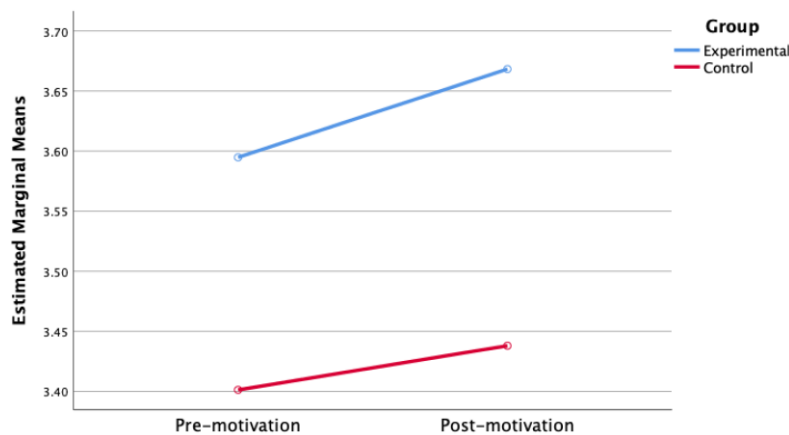


Figure 5. L2 motivation graph

A correlation analysis between students' perceived effective features of the treatment and the L2 motivation results was conducted to see what exact features predicted the EG's post L2 motivation, and the results are shown in Table 4. It can be seen that all the listed factors were statistically significantly correlated with post L2 motivation, with "sense of contribution" the strongest, followed by "dictionary use," "assistance of image" and "sense of rapport." Simply put, participants who acknowledged more of the image incorporation and propelled dictionary use, and who felt a stronger sense of rapport and contribution were more motivated to learn English subject as a whole.

5.3. Learning perceptions and satisfaction

When asked whether they wished to continue to learn this way, 92.3% of the EG participants responded positively. When asked to compare with traditional learning approaches on various dimensions by responding more, neutral or less, EG rated our approach as more satisfactory (80.8%), easier to use (80.8%), more memorable (69.2%), more flexible (65.4%), more interesting (65.4%), and more efficient (53.8%).

A correlation analysis further suggested that factors listed in Table 5 were statistically significantly correlated with each other. For example, students who considered our approach more flexible than traditional learning approach were very likely those who also rated high memorability ($r = .916, p < .01$). Additionally, among other factors, student who found MAVL more interesting were most likely to yield a higher level of satisfaction overall.

Table 5. Correlation analysis between perceived qualities

	Interestingness	Efficiency	Ease of use	Memorability	Flexibility
Efficiency	.636**				
Ease of use	.671**	.427*			
Memorability	.566**	.729**	.732**		
Flexibility	.660**	.778**	.671**	.916**	
Satisfaction	.671**	.597**	.505**	.520**	.465*

Note. * $p < .05$; ** $p < .01$.

5.4. Interview results

In order to gain an in-depth understanding of EG students' experiences, perceptions and motivation, we also conducted follow-up interviews. Though we attempted to recruit six EG participants with various level of commitment to MAVL, only four participants who consented to the interview were actually interviewed (one highly committed, three moderately). One was conducted over the phone, and the other via WeChat text messages.

Regarding the strengths of MAVL, all four participants mentioned that learning was more flexible and personalized this way. Three out of four stated that the illustrations created by their classmates were very helpful and of high quality. One student mentioned that such illustrations were "very down-to-earth, and conveyed a sense of proximity...that those standardized ones in the textbook or found online would never be able to achieve." Another student noted that MAVL helped build collegiality among classmates, because "whenever someone posts in the group chat, it was like saying, I'm studying vocabulary now or I'm with you." The frequent posted illustrations were also seen as automatic reminders to study or create illustrations, as the more often students reviewed the content, the better they would retain such knowledge.

In terms of ideal changes, one important note was that the retrieval of shared illustrations was a little cumbersome, and they wished to have someone store and organize them in a public folder every week. The number of required illustrations was mentioned as well, with some indicating more would be better and some preferring to create just one, because "the instructor may create a bigger group chat and invite students from other classes, so that we reap more but contribute less individually." It was also pointed out that sometimes students would illustrate the same word, and not bother creating another one, which could leave out some advanced or difficult words.

With regards to learning behaviors, all participants said that they would use a dictionary or discuss with friends when they were unsure about the usage of words illustrated by their classmates, especially when a word has multiple meanings or properties. It was emphasized by one student that normally she would seldom use a dictionary, because the textbook glossary was sufficient for understanding a word's meaning, but in order to be 100% sure about her illustrations, she had to use the dictionary more often, so that other students could learn correct knowledge. This was consistent with the survey data for Question 16, which asked students about their perceptions of their own illustrations: 73% of the respondents chose the option "I only sent out illustrations when I was 100% sure about the image and text accuracy." This indicated that students felt accountable for information they shared, and were thus more cautious of potential errors in vocabulary use than when they learned individually. Students also mentioned that they often guessed the meaning before using the dictionary if a word was unfamiliar, because "it was more fun...you can gain more confidence if you guessed right." One participant said that she did not learn about the illustrations as soon as she saw them in the group chat; rather she saved all illustrations to her photo album, and viewed them when she had longer chunks of time.

As for the role of the instructor, the interviewees agreed that the presence of instructors in the group chat was necessary, because "it means the instructor considers the task important" or "it makes students feel more secure because he or she can help when we need it." At the same time, they acknowledged that how much the instructor should be involved was a challenging decision to make: "...if they are involved too much, it would be overstepping; if it is too little, we may not treat it seriously." One student added that "the bottom line is, the instructor should be encouraging rather than judgmental. It's helpful to let us know when we did something wrong, but it could also be discouraging or devastating to sensitive classmates, because this is a public space."

6. Discussion and conclusions

Both vocabulary retention and L2 motivation are crucial for long-term language success. In this paper, we attempted to examine how and which MAVL features can enhance both, so that instructors and researchers can make more informed decisions when adopting and developing such activities.

First, we found that EG students, who attained a statistically significantly lower mean vocabulary score at pre-test than CG, outperformed in the immediate post-test with statistically significant difference; however, there was no statistically significant difference in the groups' delayed test. Such a finding is consistent with Zhang et al.'s (2011), in which two sophomore classes (one SMS group, one paper group) were compared, and statistically significant difference was only found in the two groups' post-test scores, but not the delayed test. This could be an indication of MAVL's apparent effectiveness in improving initial vocabulary acquisition, while traditional approaches like rote memorization can still make up for such disadvantage at a later time. This is especially true when tests use closed-ended questions that simply require students to recall, not to produce. For example, students who provide correct meanings for the same word may differ in their ability to use it accurately and meaningfully in a sentence. Such differences are unlikely to be captured via receptive tests, and thus may account partially for the inability to detect statistically significant results. Researchers also suggested that when there were few words to learn, the advantage of one approach (i.e., mobile learning) over the other might be too subtle to detect (Lu, 2008; Derakhshan & Kaivanpanah, 2011).

Secondly, in terms of motivation, there was no statistically significant difference either within group or between groups. However, results from the EG's post-test questionnaire and interviews suggested that a motivational increase might be yet to come. On the one hand, EG were satisfied with and thought high of this treatment experience. For example, 24 out of 26 students in the EG responded that they wished to continue to learn this way, and they attributed the highest score to the treatment's ease of use as contributing to MAVL effectiveness. This is consistent with Huang et al.'s (2007) study on 313 undergraduate and graduate students' use of mobile learning, where they found a statistically significant correlation between one's perceived ease of use and their intention to use it. Additionally, interviewees acknowledged the flexibility, sense of collegiality and quality of peer-generated illustrations which all inspired them to continuously learn this way. On the other hand, our correlation analysis indicated that certain MAVL features (i.e., sense of rapport and contribution, and dictionary use) could strongly predict students' post-test L2 motivation. Thus, it may be expected that if students continue to gain benefits from these features, their L2 motivation will eventually increase after a longer period of use.

Thirdly, informed by the CSDMLE model, we integrated features that targeted effective goal-setting, content design and form adoption. Results from the post-test questionnaire, interviews, and group chat behavior observation showed that all three aspects were relatively successful. For example, we aimed to promote both individual growth and collaborative outcomes through contribution-oriented learning, and it was indeed found that students' sense of contribution was most statistically significantly correlated with their L2 motivation; individuals attained higher scores from pre-test to post-test, and the few illustrations created by each student aggregated to a larger collection of high quality learning materials, which was a testament to their collaboration and contribution. This is consistent with Alghamdy's (2019) finding that students enjoyed sharing with others in the mobile language learning environment. In terms of content design, most students could meet the activity requirements and created illustrations that contained both sentence(s) and an image. Results showed that the use of image was correlated to their post-test L2 motivation, meaning image incorporation was an valuable feature in MAVL. Consistent with SDL, we expected students to follow the MAVL prerequisites and timely create and share illustrations during the study. According to the questionnaire, 94.5% of the EG respondents claimed that they met the requirements well. Simply put, students were able to autonomously persist in this activity with little instructor interference.

The interview results showed that students generally appreciated the benefits of the CSDMLE-informed MAVL design, including its flexibility and repeated encounters with vocabulary that were often reported in other MAVL studies (e.g., Liu, 2016). Congruent with COLA and ML, our participants deemed helping others and the use of both image and text as essential for deep learning. Also, consistent with the quantitative results, students appreciated the sense of rapport and making contribution. Moreover, they suggested that instructors encourage instead of judging in such activities. This not only echoed Knowles's (1985) SDL theory, emphasizing learners' active control of the learning process, but also partially supported Chien's et al. (2020) finding that teachers' criticism might harm EFL

students' performance and confidence. Additionally, it was pointed out that the difficulty of message retrieval and lack of coverage of important vocabulary needed improvement.

Overall, the CSDMLE model was effective in guiding student-directed collaborative MAVL design. The mixed results from vocabulary test, L2 motivation, questionnaire and interviews suggest that students' satisfaction with, and inclination to participate in, MAVL is impacted by multifaceted factors. However, vocabulary retention, which is often stressed by most language instructors and researchers, did not seem as much an important concern to students in this research. The questionnaire analysis showed that whether students were satisfied with MAVL was statistically significantly predicted by the perceived interestingness, efficiency, sense of rapport, ease of use, memorability and flexibility. Indeed, ease of use has always been identified as a critical indicator of users' intention to adopt a technology (Lee, Cheung & Chen, 2005), and engaging factors are also valued by MAVL students (Attewell & Webster, 2005). Therefore, instructors who aim to adopt MAVL should design learning experience that promotes these aspects.

Finally, there is still room for improvement. For example, more participants or a multiple-stage design could have increased the finding's generalizability; the study may also have been carried out for a longer duration so that students' L2 motivation change could be more observable. Moreover, owing to the voluntary nature of the present study, some EG students did not commit fully to this learning experience or take the tests seriously, which may have discounted their own and peers' test performances. Additionally, vocabulary performance may need to be measured in more innovative and diverse forms, so that students' progress can be accurately captured. It is also advised to use multiple instruments, including both vocabulary tests and those that evaluate their affective changes which are either conducive to or the result of students' cognitive growth. Lastly, the study may incorporate design elements that distinguish between high-, intermediate- and low-proficiency students, so that different groups can benefit the most.

Acknowledgement

We especially thank Dr. Sarah Lohnes Watulak and Dr. Scot McNary for providing valuable suggestions for this study.

References

- Alemi, M., Sarab, M. R. A., & Lari, Z. (2012). Successful learning of academic word list via MALL: Mobile assisted language learning. *International Education Studies*, 5(6), 99-109. doi:10.5539/ies.v5n6p99
- Alghamdy, R. Z. (2019). The Impact of mobile language learning (WhatsApp) on EFL context: Outcomes and perceptions. *International Journal of English Linguistics*, 9(2), 128-135. doi:10.5539/ijel.v9n2p128
- Attewell, J., & Webster, T. (2005). Engaging and supporting mobile learners. *Mlearn*, 2004, 15-19.
- Botero, G., Questier, F., & Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk? *Computer Assisted Language Learning*, 32(1-2), 71-97. doi:10.1080/09588221.2018.1485707
- Chen, C. M., Chen, L. C., & Yang, S. M. (2019). An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation. *Computer Assisted Language Learning*, 32(3), 237-260. doi:10.1080/09588221.2018.1485708
- 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. doi:10.1016/j.compedu.2007.06.011
- Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-speaking performance and learning perceptions. *Computers & Education*, 146, 103751. doi:10.1016/j.compedu.2019.103751
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53(4), 445-459.
- Collis, B., & Moonen, J. (2001). *Flexible learning in a digital world: Experiences and expectations*. London, UK: Kogan Page.

- Collis, B., & Moonen, J. (2006). The Contributing student: Learners as co-developers of learning resources for reuse in web environments. In D. Hung, & M. S. Khine, (Eds.), *Engaged Learning with Emerging Technologies* (pp. 49–67). Springer. doi:10.1007/1-4020-3669-8_3
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oak, CA: Sage Publications, Inc.
- Derakhshan, A., & Kaivanpanah, S. (2011). The Impact of text-messaging on EFL freshmen's vocabulary learning. *The Eurocall Review*, 19(19), 39–47.
- Driscoll, M. P. (2005). *Psychology of learning and instruction* (3rd ed.). Boston, MA: Pearson Education, Inc.
- Dörnyei, Z. (2005). *The Psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z., & Csizér, K. (2002). Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied Linguistics*, 23, 421–426.
- Ebbinghaus, H., (1913). *Memory*. New York, NY: Teachers College, Columbia University.
- Elaish, M. M., Shuib, L., Ghani, N. A., & Yadegaridehkordi, E. (2019). Mobile English Language Learning (MELL): A Literature review. *Educational Review*, 71(2), 257–276. doi:10.1080/00131911.2017.1382445
- Foomani, E. M., & Hedayati, M. (2016). A Seamless learning design for mobile assisted language learning: An Iranian context. *English Language Teaching*, 9(5), 206. doi:10.5539/elt.v9n5p206
- Gardner, R. C. (1985). *Social psychology and second language learning: The Role of attitudes and motivation*. London, UK: Edward Arnold.
- Gardner, R. C. (2006). The Socio-educational model of second language acquisition: A Research paradigm. *EUROSLA Yearbook*, 6(2006), 237–260. doi:10.1075/eurosla.6.14gar
- Garrison, D. R. (1997). Self-directed learning: Towards a comprehensive model. *Adult Education Quarterly*, 48, 18-33.
- Gürkan, S. (2019). Effect of annotation preferences of the EFL students' on their level of vocabulary recall and retention. *Journal of Educational Computing Research*, 57(6), 1436-1467. doi:10.1177/0735633118796843
- Hanson, A. E., & Brown, C. M. (2020). Enhancing L2 learning through a mobile assisted spaced-repetition tool: An Effective but bitter pill? *Computer Assisted Language Learning*, 33(1–2), 133–155. doi:10.1080/09588221.2018.1552975
- Hu, Z. (2013). Emerging vocabulary learning: From a perspective of activities facilitated by mobile devices. *English Language Teaching*, 6(5), 44–54. doi:10.5539/elt.v6n5p44
- Huang, J., Lin, Y., & Chuang, S. (2007). Elucidating user behavior of mobile learning: A Perspective of the extended technology acceptance model. *The Electronic Library*, 25(5), 585–598. doi:10.1108/02640470710829569
- Huang, A. F. M., Yang, S. J. H., Hwang, G. J., & Tsai, C. C. (2010). Situational language teaching in ubiquitous learning environments. *Knowledge Management & E-Learning*, 2(3), 312-328.
- Hung, H., T., Yang, J. C., Hwang, G. J., Chu, H. C., & Wang, C. C. (2018). A Scoping review of research on digital game-based language learning. *Computers & Education*, 126, 89-104.
- Hwang, G. J., Chu, H. C., Shih, J. L., Huang, S. H., & Tsai, C. C. (2010) A Decision-tree-oriented guidance mechanism for conducting nature science observation activities in a context-aware ubiquitous learning environment. *Educational Technology & Society*, 13(2), 53–64.
- Hwang, G. J., & Fu, Q. K. (2019). Trends in the research design and application of mobile language learning: A Review of 2007–2016 publications in selected SSCI journals. *Interactive Learning Environments*, 27(4), 567–581. doi:10.1080/10494820.2018.1486861
- Hwang, G. J., & Wu, P. H. (2014). Applications, impacts and trends of mobile technology-enhanced learning: A Review of 2008–2012 publications in selected SSCI journals. *International Journal of Mobile Learning and Organisation*, 8(2), 83–95. doi:10.1504/IJMLO.2014.062346
- Jain, Y., & Sidhu, G. K. (2013). Relationship between anxiety, attitude and motivation of tertiary students in learning English as a second language. *Procedia - Social and Behavioral Sciences*, 90(InCULT 2012), 114–123. doi:10.1016/j.sbspro.2013.07.072
- Kaliisa, R., Palmer, E., & Miller, J. (2019). Mobile learning in higher education: A comparative analysis of developed and developing country contexts. *British Journal of Educational Technology*, 50(2), 546–561.

- Klopper, E., Sheldon, J., Perry, J., & Chen, V. H. H. (2012). Ubiquitous games for learning (UbiqGames): Weatherlings, a worked example. *Journal of Computer Assisted Learning*, 28, 465–476.
- Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. New York, NY: Association Press.
- Kohnke, L., Zhang, R., & Zou, D. (2019). Using mobile vocabulary learning apps as aids to knowledge retention: Business vocabulary acquisition. *Journal of Asia TEFL*, 16(2), 683–690. doi:10.18823/asiatefl.2019.16.2.16.683
- Lee, M. K. O., Cheung, C. M. K., & Chen, C. Z. (2005). Acceptance of Internet-based learning medium: The Role of extrinsic and intrinsic motivation. *Information & Management*, 42(8), 1095–1104.
- Lin, J. J., & Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: A Systematic review and meta-analysis. *Computer Assisted Language Learning*, 32(8), 878–919. doi:10.1080/09588221.2018.1541359
- Liu, P. L. (2016). Mobile English vocabulary learning based on concept-mapping strategy. *Language Learning and Technology*, 20(3), 128–141.
- Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL*, 31, 293–311. doi:10.1017/S0958344019000065
- Looi, C. K., Zhang, B., Chen, W., Seow, P., Chia, G., Norris, C., & Soloway, E. (2011). 1:1 mobile inquiry learning experience for primary science students: A Study of learning effectiveness. *Journal of Computer Assisted Learning*, 27(3), 269–287. doi:10.1111/j.1365-2729.2010.00390.x
- Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning*, 24(6), 515–525.
- Mahdi, H. S. (2017). The Use of keyword video captioning on vocabulary learning through mobile-assisted language learning. *International Journal of English Linguistics*, 7(4), 1–7. doi:10.5539/ijel.v7n4p1
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed). New York, NY: Cambridge University Press.
- Motallebzadeh, K., & Ganjali, R. (2011). SMS: Tool for L2 vocabulary retention and reading comprehension ability. *Journal of Language Teaching and Research*, 2(5), 1111–1115. doi:10.4304/jltr.2.5.1111-1115
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language*, 22(1), 31–55.
- Pirasteh, P., & Mirzaeian, V. (2015). The Effect of short message service (SMS) on learning phrasal verbs by Iranian EFL learners. *Language in India*, 15(1), 144–161.
- Rahmany, R., Zarei, A. A., & Gilak, S. (2013). The Effect of extensive reading on Iranian EFL learners' motivation for speaking. *Journal of Language Teaching and Research*, 4(6), 1238–1246. doi:10.4304/jltr.4.6.1238-1246
- Shih, J.-L., Chuang, C.-W., & Hwang, G.-J. (2010). An Inquiry-based mobile learning approach to enhancing social science learning effectiveness. *Educational Technology & Society*, 13 (4), 50–62.
- Stockwell, G. (2007). Vocabulary on the move: Investigating an intelligent mobile phone-based vocabulary tutor. *Computer Assisted Language Learning*, 20(4), 365–383. doi:10.1080/09588220701745817
- Sundberg, R., & Cardoso, W. (2019). Learning French through music: The Development of the Bande à Part app. *Computer Assisted Language Learning*, 32(1–2), 49–70. doi:10.1080/09588221.2018.1472616
- Suwantarathip, O., & Orawiwatnakul, W. (2015). Using mobile-assisted exercises to support students' vocabulary skill development. *TOJET: The Turkish Online Journal of Educational Technology*, 14(1), 163–171.
- Tight, D. G. (2010). Perceptual learning style matching and L2 vocabulary acquisition. *Language Learning*, 60(4), 792–833.
- Wong, L.-H., & Looi, C.-K. (2010). Vocabulary learning by mobile-assisted authentic content creation and social meaning-making: Two case studies. *Journal of Computer Assisted Learning*, 26, 421–433.
- Yuan, Y. (2019). Empirical study on the mobile app-aided college English vocabulary teaching. *International Journal of Engineering and Technology*, 11(1), 68–74. doi:10.7763/ijet.2019.v11.1125
- Zhang, H., Song, W., & Burston, J. (2011). Reexamining the effectiveness of vocabulary learning via mobile phones. *Turkish Online Journal of Educational Technology-TOJET*, 10(3), 203–214.