INVESTIGATING EFL LEARNERS' PERSPECTIVES ON VOCABULARY LEARNING EXPERIENCES THROUGH SMARTPHONE APPLICATIONS

by **Saman Ebadi** and **Saba Bashiri**

Razi University, P. O. Box 67149-67346, Baghe Abrisham, Kermanshah, Iran

samanebadi @ gmail.com, bashiri.s1368@yahoo.com

Abstract

This study investigated EFL learners' perspectives about their vocabulary learning experiences via a smartphone application. An online demographic questionnaire was used for recruiting 50 EFL learners from a language teaching channel in Telegram messenger required to use a smartphone application called Vocabulary Flashcards 2016 for a month. After finishing the sampling procedure, the participants were asked to take part in Dialang online diagnostic test to specify their vocabulary level proficiency based on CEFR (Common European Framework of Reference). The quantitative and qualitative data were collected utilizing evaluation questionnaires and semi-structured interviews respectively. The evaluation questionnaire adapted from Chapelle's (2001) evaluation criterion was used to evaluate the application from the users' perspectives. This study investigated the effects of learners' proficiency level and gender differences on using the application, and their perspectives on the negative and positive aspects of the application were also uncovered. The findings showed that the users held positive attitudes towards the application because it influenced their learning positively and provided them with both form and meaning-focused instruction, but they were dissatisfied with the app's levels and authenticity. Results of independent t-test and ANOVA respectively showed that gender and vocabulary proficiency level did not make significant difference on participants' app usage patterns. The findings of this study highlighted the users' localized needs which could be used as guidelines for customized vocabulary apps' development purposes. The study's implications for learners, teachers, and app developers are discussed in detail.

Keywords: MALL; smartphones apps; vocabulary learning; evaluation criteria

1. Introduction

Using technology has become one of inseparable aspects of life in the 21st century. Almost everybody can feel and appreciate its penetration into all aspects of life. Information and Communication Technology (ICT) by introducing smart devices enabled people to have access to knowledge and information with no spatial and temporal constraints (Sampson, Isaias, Ifenthaler & Spector, 2013). Probably the most important impetus for utilizing

technology in the process of language learning is its ability to emancipate all stakeholders from time and space limitations (Burston, 2011) and solve the time boundary problems between instructors and their students (Salleh & Binti, 2010).

Roughly speaking, this learning which is aided by technology and especially by computers is called Computer Assisted Language Learning (CALL). As Savchenkova (2003) states, "Starting in the early 60s… CALL has become a common practice of language teaching and learning" (p.1). Two positive aspects of CALL are providing learners with authentic learning materials (Martiz, 2015) and widening the potential of language learning by increasing its effectiveness and decreasing its tedium (Savchenkova, 2003). Being considered as an almost new branch of CALL, Mobile Assisted Language Learning (MALL) came into vogue with the advent of "Portable Digital Assistant (PDA) and i-Pod" (Burston, 2011.p. 57). MALL is the process of learning a language by the aid of a mobile learning device which is defined as "a handheld, portable computing instrument with Internet or some other network access, which allows for mediated activity for information access and learning in multiple contexts" (Walters, 2012, p.16).

Enhancing language learning opportunities needs special attention to the aspects which form the basis of language. One of these aspects is vocabulary acquisition. Vocabulary, as a key component of any language, has been paid considerable attention with the aim of finding techniques that foster its acquisition (Vahedi, Ghonsooly & Pishghadam, 2016). It should be noted that the role of this component has undergone changes in L2 instruction through time, which has resulted in different approaches towards its role in L2 learning (Celce-Murcia, Brinton & Snow, 2014). According to Leal Alves and de Oliveira (2014), the difficulties faced by EFL learners in vocabulary acquisition are caused by several variables. Furthermore, they believe that these variables "are somehow dependent on factors such as socioeconomic, ideological and cultural conditions beyond their own teaching/learning and the intellectual characteristics of learners" (p.51). Sanchez and Manchon (2007) asserted that there has always been concentration on the best pedagogical way in developing learners' vocabulary or lexicon.

The vast body of literature in the area of technology-aided vocabulary learning and teaching shows some important trends that need to be summarized here. Montero Perez, Peters, Clarebout and Desmet (2015) investigated the impact of video captioning on both incidental vocabulary acquisition and video comprehension. Wang, Teng and Chen (2015) showed the effectiveness of using iPad App in English vocabulary acquisition compared to traditional methods, while mre recently Vahedi et al. (2016) investigated the effect of

different gloss types on L2 vocabulary learning. The unconvincing issue about studies of this kind is that they only seek for technology's influence on participants' performance.

The advent of iPad in 2010 resulted in developing freely available computer programs specialized for mobile devices use which are called applications (apps) (Deng & Trainin, 2015). To the best of the authors' knowledge, a significant point of concern about vocabulary apps studies so far is that their main focus has been almost exclusively on evaluating the apps and tapping into participants' attitudes. It is hoped that this study's findings will pave the way for designing new apps which are tailored and customized to specific audiences' needs within countries, within races, or even within genders. However, it should also be noted that localization and customization of instructional tools, especially digital ones, does not seem an easy practice and needs time, fund, and patience investment.

2. Literature review

In this section studies in the areas of vocabulary learning, Computer Assisted Vocabulary Learning (CAVL), Mobile Assisted Vocabulary Learning, and vocabulary apps will be presented. Gui (2015) tried to find a positive correlation between EFL learners' vocabulary size and their proficiency level in a non-English context. This correlation was explored using such tests as College English Test Band-4 (CET4), College English Test Band-6 (CET6) and Vocabulary Size Test (VST). 96 male and female Chinese English learners were selected as participants. Results of the study showed that vocabulary size was highly influential in predicting learners' listening, reading, and overall proficiency. However, the researcher contented that in cases of rote vocabulary learning and lack of in-depth knowledge, participants' improvement in vocabulary size did not necessarily result in their overall EFL proficiency.

In CAVL area Stockwell (2011) compared two types of online vocabulary learning in one of which vocabulary items were provided in form of online materials selected by teachers (teacher-centered) and in the other learners themselves were responsible for compiling and entering lexis into a designed online system. Data for this purpose were collected through administering vocabulary pre- and post-tests along with an attitudinal questionnaire to elicit learners' perceptions of both systems. 55 first-year law students studying an obligatory English course were divided into "teacher-centered" (28), and "learner-centered" (27) classes. The pre-test results showed that LC and TC class members were not so different in their achievement (TC: 61.3% and LC: 61.4%). However, this difference was very sharp in post-test results as in this stage TC class members achieved 93.2% while their counterparts

achieved 85.9%. Totally speaking both tests results proved that TC class compared to LC class made greater improvement, which was because LC class spent more time on the activities. Both groups evaluated the activities as useful but LC class members held more positive attitudes towards their activities. They expressed the view that vocabulary data input was very interesting and useful.

In classroom settings smartphones and tablets started to attract attention since the beginning of the 21st century which was a consequence of introducing iPod touch and iPhone, and finally iPad (Leis, Tohei & Cooke, 2015). An investigation in mobile assisted vocabulary learning exercise was pursued by Suwantarathip and Orawiwataakul (2015). 80 EFL university learners were put into two sections of 40 students each. One section was provided with paper based vocabulary exercises and the other one was sent SMS-based exercises for 7 weeks. Before starting the study the students were examined with a pre-test to realize their proficiency level which was then covered to be at the same level. In the parallel pretest the score mean of experimental group (33.25) was higher than that of the control group (29.70). This data revealed SMS-based exercises advantage over paper-based activities in developing vocabulary knowledge. To gain evidence of participants' attitudes after the experiment section, they were given an attitudinal questionnaire to express their opinions about SMS-based vocabulary improvement. The responses revealed their overall satisfaction with the activity, their acceptance of mobile phones as learning aids, and mobile phones potential to remove spatial limitations.

Effectiveness of *WhatsApp* educational mobile application was studied in a 4-week long project by Basal et al. (2016). A pre-test and post-tests were employed to compare 50 first year English students in two equal groups. The app provided a corpus of 40 figurative idioms. Data was collected through a 40-item achievement test. Before starting the experimental phase of the study, the 40 idioms were presented to both groups' members in a pre-test to check groups' differences. Pre-test results indicated no significant differences in their knowledge. After that the control group was given learning material and activities in printed form (paper-based) including idioms' meaning, their usage example, and fill in blanks. The experimental group was provided with MMS via *WhatsApp* on their smartphones which included the idioms, their meanings, their pictorial representation, and three sample sentences followed by a test to be answered and sent back after two hours. Although both groups' immediate post-tests means improved significantly compared to their pre-test, the experimental group's improvement in the targeted skill was much greater. This implies that both traditional and technology-based instructional tools indeed resulted in better post-test

performance when compared within each group, but between groups comparison showed the app's advantage over the traditional method.

3. Methodology

3.1. Research questions

The following questions guided this study:

1) What is EFL learners' perspective on vocabulary learning experiences through using the vocabulary application called *Vocabulary Flashcards 2016* at different levels of vocabulary proficiency according to CEFR?
2) Do gender and proficiency differences influence learners' app usage patterns during vocabulary learning experiences using the app?
3) What are learners' perspectives on the advantages and disadvantages of their learning experience with the app? What are their suggestions for making the app more efficient?

3.2. Research methodology and data collection instruments

Following a mixed-methods approach to research, this study combines qualitative and quantitative methods. Based on Dörnyei (2007) the mixed-method type used in this study was "questionnaire survey with follow-up interview or retrospection" (p.170). For the quantitative part the data was collected through questionnaires while semi-structured interviews were utilized to collect qualitative data.

The first instrument used in this study was the demographic questionnaire. This online English questionnaire retrieved from https://www.surveymonkey.com/r/SCDCJMH (Appendix A) was distributed prior to starting the project. It was designed in *Google Docs* service and shared to the channels' members by providing them with the link and a brief introduction to the study's design and purpose.

The other instrument was an online diagnostic test called *Dialang* available at https://dialangweb.lancaster.ac.uk/ sent to the participants via a *Gmail* group called 'vocabulary team'. This team included those who completed the demographic test and provided the researchers with their emails for further contact. Designed to determine language learners' proficiency level in 14 European languages, *Dialang* use for purposes other than diagnosing like granting certificates or employment purposes is rejected by its developers. The test in each language section is divided into three parts after which the examinees are assigned into different proficiency levels. Learners' proficiency in different language

components (grammar, vocabulary) and skills (listening, reading, writing, and speaking) is diagnosed through scores related to the *Common European Framework of Reference for Languages* (CEFR). Based on these scores the examinee is assigned into one of CEFR level from A1 (the least proficient) to C2 (the most proficient). Participants were requested to complete the vocabulary test and then inform the researcher of their levels via email. After that they were divided into three groups based on the test results. Group A included A1 and A2 levels (elementary), group B included B1 and B2 levels (intermediate), and group C included C1 and C2 levels (advanced).

Vocabulary Flashcards 2016 as the targeted vocabulary learning application was another instrument to be installed on participants' smartphones and used for a month. In this app high-frequency English words (in total 1,200 words) are divided into three groups based on their difficulty level (easy, medium, and hard). Each entry includes pronunciation, meaning, contextualization in a sentence, semantic relation (synonym, antonym), and a memory trick for better memorization. For each level quizzes are designed and after their completion the user is informed of her/his right and wrong responses. While taking the exams immediate feedback is provided after each question. The final feedback specifies all answers as right or wrong, and in the latter case the correct option is provided again. The users' control over the app includes selecting the words to be ordered alphabetically or randomly, and opting to be shown either all words from the selected level or alternatives such as seen words, new words, and learnt words. Words can be bookmarked for easy access by tapping on a like symbol. By tapping each level the studied words can be recognized from new ones. To be sure that the participants studied the words of their level, the researcher requested them to send screenshots on app's pages in which the words that were not studied were marked as new (Appendix B).

The main instruments for collecting data were a questionnaire and an interview, the purpose of which was evaluating the application and tapping into its users' attitudes and perceptions towards it after one month of usage. This questionnaire was designed in a 3-point Likert-scale format in Persian (Appendix C). The first section included instructions, the second part included providing personal information and some questions about manner and amount of using the app. In the last part the items were in the form of statements followed by three options ('yes', 'somewhat', 'not at all') to be selected by respondents. The total number of items other than those which were about personal information and proficiency level were 17. These statements were designed based on evaluation criteria proposed by Chapelle (2001) and adapted from Jamieson, Chapelle and Priess (2005), including language learning

potential, meaning focus, learner fit, authenticity, positive impact, and practicality. The justification for using a three-point Likert scale was the relative similarity of these two studies in nature. The mentioned criteria were used intact but the items were changed to be suitable for current study purposes.

To assure content validity of the questionnaire the researcher compared it to the similar questionnaire used by Jamieson et al. (2005), which resulted in ensuring the content validity on part of the researcher by realizing that the responses to the items were not affected by any other factors. A common way of measuring questionnaires' reliability is using Cronbach Alpha coefficient. To measure reliability, each response was assigned a scale and then put into the Cronbach's Alpha formula in *SPSS* software version 21. The reported coefficient index was .81 which is considered reliable for this questionnaire.

Questionnaire piloting was conducted to get rid of any ambiguities and pitfalls and evaluating its appearance, clarity, and answering time (Dörnyei, 2003). It was administered to a sample of 5 persons not included in the study and based on their verbal opinions and their answering time modifications were applied. These modifications included changing some ambiguous and loaded words. The final versions of the questionnaires were administered in two ways. The researcher prepared some hardcopies to administer them to those participants who were known and nearby. For those participants who were not available, a file containing the questionnaire was sent via their *Telegram* accounts.

Interview questions (8 items) were extracted from the questionnaire items indirectly but they were not identical. Although the questionnaires were distributed to all the participants, only 2 representatives (one male and one female) from each group (A, B, C) were interviewed voluntary (in total 6 persons). Questionnaire piloting was done by asking the items from three nearby participants who did not volunteer to take part in the main interview. Accordingly ambiguities were removed or modified. The interviews were conducted in Persian to hinder any misapprehension between the interviewer and the interviewees. *Skype* call service was the channel for conducting interviews with two participants, while others were available for personal performing. Each interview was audio recorded for further analysis. Finally, each interview was transcribed, translated into English, and then analyzed using thematic analysis.

3.3. Participants

English teaching public *Telegram* messenger channel called @drebaditoefl was targeted as a recruitment source for participants' selection. The justification for this selection was that the

participants had joined the channel in order to improve their English proficiency and this channel provided them with rich material including useful vocabulary so they seemed eager to improve their vocabulary learning. The members were required via an online demographic questionnaire yielding data about their age, gender, education, field of study and email. Out of the 55 submitted questionnaires, all the respondents were selected as the final participants in the study. These participants were both male and female, their age ranged from 20 to 45 years, and all of them had university degrees in different fields of education. From among them, 5 did not answer the emails and were excluded from the study. Filling in the demographic questionnaire did not cause any obligations to continue participating in the study. After one month of using the app, the researcher sent them the online evaluation questionnaire.

The following table represents the demographic information of the questionnaire participants.

Table 1. Demographic information of the questionnaire participants

Gender	**N**	**Age Range**		**Vocabulary Proficiency Level**	**N**	**First Language**	**N**	**Academic degree**	**N**
Male	19	Max	44	A	10	Persian	50	BA	24
Female	31	Min	20	B	22			MA	25
		Mean	28	C	18			Ph.D.	1

The interview participants were not different form the questionnaire participants, that is; two participants were selected to be interviewed from among each proficiency level.

3.4. Data analyses

The first two research questions were answered using data obtained from the evaluation questionnaire. *Statistical Package for the Social Science* (SPSS) version 21 (descriptive and inferential statistics) was used to process the quantitative data. The questionnaire items were analyzed in this way to see how many of the participants had selected each option for each question and then the percentage was calculated. T-test and ANOVA were inferential statistical tools used for making inferences from the selected options to questionnaire's items. For answering the last question the interviews were transcribed in Persian and then translated into English. After that they were analyzed using open-thematic coding method. Through this

method, common patterns are looked for in the transcribed data to render what is called theme (Seidman, 2006), so commonalities in thematic terms are put into a single category.

4. Results and findings

4.1. Quantitative results

As regards the first research question, only those questions which were developed based on the evaluation criteria were taken into consideration to tap into participants' perception of vocabulary learning using the app. These widely-known criteria proposed by Chapelle (2001) were used as evaluation criteria in studies like Jamieson et al. (2005) for evaluating a CALL product called *Longman English Online* (LEO). Following this approach and by references to Chapelle (2001), Jamieson at al. (2005), Hubbard (2006), and Leakey (2011) each criterion will be elaborated on and participants' attitudes in this regard will be explained.

4.1.1. Evaluation criteria

1. language learning potential

This criterion was described as the degree of opportunity the product presents for users to focus on form in a useful manner (Leakey, 2011). As Figure 1 shows, participants' attitude towards this criterion in reference to the used app is rather favorable as more than two thirds of them selected either the first or the second option. In line with this finding, Bensalem (2018) found that EFL learners who used WhatsApp enjoyed more vocabulary learning compared to those who did not use it.

2. meaning focus

While the previous criterion emphasized focusing on form, this one is more in favor of focus on meaning. This means that both form and meaning should be taken into consideration in instruction. This criterion states that when the learner is learning a language via CALL or MALL products, his/her attention should be directed toward the meaning of that language (Leakey, 2011). This criterion is assumed to be considered in the app used in this study as the majority of participants agreed on it. This might be due to the fact that every word was contextualized in sample sentences and its antonyms and synonyms were also provided, which led learners to pay attention to semantic relations.

3. learner fit

Learner fit criterion, as its name suggests, proposes materials to be fine-tuned to learners' characteristics. Leakey (2011) contends that an equal amount of opportunity for engagement with language should be provided for learners based on their characteristics. Chapelle (2001) argues that "learner fit refers not only to appropriate difficulty but also to appropriate instructional strategy relative to individual differences" (p.158). A quick look at Figure 1 reveals that the app was not successful in fulfilling this criterion. More than two thirds of the participants selected the third option, which was the least positive one.

4. authenticity

According to Hubbard (2006), when in instructional CALL/MALL material the learning activity corresponds to real activities out of classroom and CALL practice, that piece of material is assumed to enjoy a high degree of authenticity. Like the previous criteria this one was also negatively evaluated by the participants (72% disagreed). This implies that activities did not resemble real life activities.

5. positive impact

The impact of the CALL activity on app's users is evaluated through this criterion. In this study the participants were asked whether this product had any influence on their desire to improve their vocabulary ability and also whether it led to their search for similar apps. This criterion was the most positively evaluated with regard to the app, as 87% of learners authenticated this by selecting the first option.

6. practicality

Practicality concerns the sufficiency of resources that support using the CALL/MALL product. Furthermore, this criterion refers to the degree of learner control over the time and place of use. Based on the responses presented in Figure 1 (73% selected 'yes, very much'), it is inferred that practicality of the app was considerably high because using the app needed no specialized skill on the part of users, and did not need network for operation.

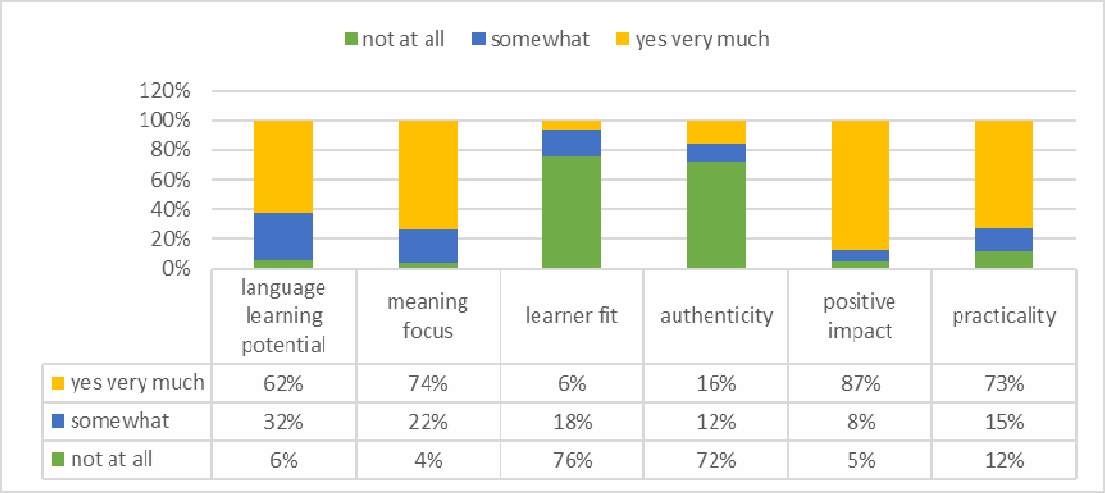


Figure 1. Learners' attitudes based on the evaluation criteria

4.1.2. Gender and proficiency influences

The second research question was aimed to determine the influence of gender and proficiency level on app usage patterns during vocabulary learning experiences using the app. The results of t-test and ANOVA are represented in Tables 2 and 3 respectively.

Table 2. Group statistics

gender		N	Mean	Std. Deviation	Std. Error Mean
New	male	19	1.3308	.29375	.06739
	female	31	1.3548	.24431	.04388

Table 3. T-test statistics

	Levene's test for Equality of Variances	t-test for Equality of Means							
	F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	1.542	.220	-.312	48	.756	-.02401	.07690	-.17863	.13061
Equal Variances not assumed			-.299	32.945	.767	-.02401	.08042	-.18763	.13961

The independent samples t-test comparing the two genders in their app usage patterns shows that there is no significance difference between the genders as the p>.05 (.22). This finding is

in contrast to a study about gender differences in accepting CALL programs for EFL learning by Lai and Kuo (2007), who used a different program and found that male learners preferred to spend more time on this kinds of programs. As regards the use of CALL in a classroom setting Awad and Alkaraki (2013) found that gender and proficiency level were not a determining factor in shaping participants' attitudes, which is in line with the results of the current study. However, they conducted their research in self-directed vocabulary learning, rather than teacher-guided.

As for the result of the t-test the p value for ANOVA was also more than .05 (p>.05) so the proficiency level did not result in significant difference between groups.

Table 4. Descriptive statistics

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
A	10	1.3857	.33705	.10659	1.1446	1.6268	1.00	2.00
B	22	1.3182	.22013	.04693	1.2206	1.4158	1.00	1.71
C	18	1.3571	.27392	.06456	1.2209	1.4934	1.00	1.86
Total	50	1.3457	.26150	.03698	1.2714	1.4200	1.00	2.00

Table 5. ANOVA statistics

	Sum of Scores	df	Mean Score	F	Sig.
Between Groups	0.35	2	.018	.248	.781
Within Groups	3.316	47	.71		
Total	3.351	49			

This is in line with Kawauchi (2008), who targeted proficiency differences in CALL-based vocabulary learning experiences. The two proficiency levels in his study perceived the program called *PowerWords* as favorable. In another study by Amer (2014) those who reported the highest TOEFL score exhibited the greatest usage of a MALL product called *Idiomobile* which is somehow in contrast to this study finding. Soleimani and Morshedian (2013) concluded that more proficient participants showed more tendencies to take benefit from technology-supported instruction. In line with Maleki et al. (2015), most learners did not have technical problems with implementing technology-based vocabulary instruction as previously addressed by the practicality criterion. This was also declared as a positive aspect in the interviews.

4.2. Qualitative results

Semi-structured interviews with 6 participants were conducted in Persian in order to answer the final question. The interviewees volunteered to participate in this phase of study. Two of the interviews were conducted via *Skype* and the others were performed in interviewees' presence one by one. On average each interview lasted about 15 minutes. All interviews were audiotaped, transcribed for further analysis and coded with open-thematic coding. During the process follow-up questions were asked in order to gain more insight into participants' responses (Lai et al., 2016). Interview participants' background information and their pseudonyms are displayed in Table 6. Based on the research question the transcribed data was assigned into themes for further organization. Accordingly three themes emerged from the data:

1) participants' reasons for using the app which somehow implied its positive aspects,
2) any shortcomings in the app perceived by participants that revealed app's negatives aspects,
3) participants' suggestions for modifying the app.

The interview questions, the coding system and illustrative segments from interviews are presented in Table 7. The 8 guiding questions were extracted from questionnaire's items indirectly. The aim was to gain further insight into app users' recommendations for alleviating its shortcomings.

1) *What was something specific that you enjoyed about this vocabulary learning application?*
2) *What were some specific concerns or difficulties that you had during using this application?*
3) *What were your typical approaches to studying and the average effort you put into each lesson?*
4) *Were different parts designed in accordance with your expectations?*
5) *Were the words in different sections taught in a good way?*
6) *What is your overall evaluation of this app?*
7) *Is there any way to redevelop the app into a more efficient version?*
8) *How different learners' needs can be satisfied by this app?*

Table 6. Interview participants' background information

Background Pseudonyms	Gender	Proficiency level	Major	Amount of app use per week
Sara	Female	A	Philosophy	1 hour
Farhad	Male	A	Arabic Literature	1.5 hour
Simin	Female	B	Statistics	2 hours
Dana	Male	B	IT	4 hours
Hoda	Female	C	TEFL	4 hours
Reza	Male	C	English literature	7 hours

Table 7. Themes, codes, and segments extracted from interviews

Themes	Codes	Segments
Positive aspects	a) systematic	• Being in flashcard form (3) • Leveling the words (2) • Recognizing seen words, new words, and learnt words (2) • Showing words in sentences (3)
	b) up to date	• Developed in 2015 and updated in 2016 (6)
	c) easy to use anytime anywhere	• Smartphones are portable (5) • Does not need network for operating (2)
	d) included exams	• Exams were followed by feedback (3)
	e) vivid explanations	• Explaining words in fluent English (2) • Providing synonyms/antonyms (3) • Providing memory trick (1)
Negative aspects	a) levels	• Fuzzy boundaries between levels (4) • Easy level was not suitable for basic learners (3)
	b) tests	• Were only of one kind (4) • Did not provide comprehensive feedback that would lead to improvement (3)
Recommendations for making the app more efficient	a) changing way of presenting material b) including some other elements	• Putting words in form of paragraphs (3) • Putting words in form of dialogues (3) • Putting related words in a lesson format and giving instructions for study (2) • Adding photos (6) • Adding phonetic symbols (6)
	c)implementing variety in test	• Including fill in blank tests (2) • Including open-ended questions (2) • Including bilingual translation (1) • Providing feedback in form of hints and tips (4) • Adding game quizzes(1)

4.2.1. Positive points

Almost all interviewed participants held positive views about their vocabulary learning experiences using the app. These points included the following aspects which will be followed by related interview segments.

1. Systematic presentation of material, providing different information for each entry, leveling words, and allowing for organization of studied and new words

Simin was a student bored with keeping up with paper-based vocabulary learning:

> It was very interesting for me to find a vocabulary app which is designed in form of flashcards. I used to write down new words on one side of white papers and on the other side I wrote its equivalent in Persian. Because it was hard to keep big papers this small flashcards did not allow me to write monolingual explanation and this bilingual method of vocabulary learning caused little, if any, improvement. (Simin)

The traditional method of writing a long list of vocabulary along with students' first language equivalents was regarded as an inefficient way which did not lead to deep knowledge of words. This participant had somehow negative attitudes towards vocabulary learning, which changed after using this app. This attitude shift from negative to positive after using a certain product was also found by Tabatabaei (2012). Similarly, in another study by Shafeii Ebrahimi (2016) the participants of interviews declared that they preferred using those kinds of language learning materials which are integrated with technology instead of old-fashioned printed ones. This can be due to learners' unsuccessful experiences in target skills development while using traditional methods:

> In traditional paper-based vocabulary learning when I was preparing myself for Konkor exam, I had no idea of how to use words in sentences because I was taught and practiced with a long list of words along with their Persian equivalence either on blackboard or in book to be memorized. This app by contextualizing the words in sentences allows us a deep practice. (Dana)

2. Being systematic in word categorization

> In past when I resumed my vocabulary practices, I had to search all papers to find those words that I had not learnt or studied and this took me a lot of time. I faced this problem no more as they were recognized by different labels as seen, new, learnt. (Hoda)

Other positive points declared by these participants were app's being up to date, its easy way of using, including quiz section, and vivid explanations provided for each entry. The title of app was a determining factor in attracting EFL learners who responded the researcher's mail to take part in the study as represented in the following extract:

> When I received an email containing the link of a vocabulary app titled *Vocabulary Flashcards 2016*, I really got excited to download and use it as soon as possible. Its name implied its being updated in current year. I found it in Google Play store with some difficulties. When I saw the reviews of other users I got more interested in it…..I think one of my main reasons for using this app was its operation on my smartphone which could be used anytime anywhere. (Hoda)

Another reason expressed by participants was their dissatisfaction with previously used apps which could not be used as easily as this one:

> I had downloaded a lot of apps before this one but most of them were not possible to be used at all times and in all places. This shortcoming was alleviated here as no network was needed for working with app's different parts except for pronunciation. Besides these some of previously download apps in my mobile phone either did not contain any examination or did not divide words into levels. (Farhad)

Feedback provision after taking the quizzes was of interest for a level A participant:

> The quizzes were very interesting for me because at the end I was informed of my wrong and right answers. Even the meanings that I selected wrongly were assigned to related word. I think providing feedback just after a test is very effective in memorizing the material. (Sara)

The app was useful for the study's teacher participant in order to respond to his learners' needs:

> I am studying and teaching TEFL and so I should constantly be in contact with this language otherwise I will face many problems in my career. As my students are in advanced level and ask for new words' English translation or equivalence, this app was a good help for me. I introduced them this app as the explanations were really vivid and contained synonyms and antonyms for each word. It would definitely be said that it can function as a 2 in 1 dictionary. (Reza)

4.2.2. Negative points

Besides mentioning a fair number of positive aspects about the app, negative aspects had also been discovered by them. These negative points were targeted to the quiz part and problems with levels.

1. Unsuitable division of words in each specified level

In answering first research question, it was realized that learner fit criterion was the most negatively evaluated one which was mostly unfavorable for least proficient participants:

> I know a limited range of vocabulary, when I was introduced this app it seemed to be a good opportunity for improving my proficiency. Before that I had not struggled in this regard because I believed I do not need it. When I clicked the option learn words and I was given options about the levels labeled as easy, medium, and hard my motivation was doubled. I was expecting basic and more frequent words for easy level but to my surprise they were not so

> easy and I really lost my motivation to use it very much. Even if I wanted to spend more time on it I could not, because sometimes practicing a single word took me so much time that I got bored to continue. (Sara)

Even intermediate participants did not assume the words to be at appropriate defined levels:

> At early days of downloading this app I set a plan to start from the easy level until I progress to the hard one. But when I took a quick look at words I came to realize that there could not be put a boundary between the three defined levels. In my opinion all of them should be divided into two levels and be labeled as intermediate and advanced. (Dana)

The participant teacher, though satisfied with words' explanations, did not like the exams:

"An app through which one can both learn words and test himself is a strong point for attracting users. One of my reasons for using the app was this capability. But the tests were very monotonous. I mean the leaner was only given a decontextualized word followed by 4 options one of which had to be selected. In this way a mere memorization could result in correct answers for all questions and could give learner the impression that she/her had made a great improvement." (Reza)

This was also expressed by Hoda as she thought just knowing meaning cannot result in vocabulary improvement:

> I liked the way of teaching words in this app but the way of testing was not desirable. I think just knowing the meaning of a word cannot lead to its mastering and sticking in one's memory. (Hoda)

Although the provided feedback at the end of each test was interesting for participants like Sara, for others this feedback was not considered as so helpful.

> Always after taking part in a test I look forward to receiving my score and my teachers' feedback in form of her comments especially in writing tasks. It was a strong point that this app contained such a system to inform me of my wrong and right answers. But in my opinions the provided feedback was not so helpful to lead to a significant improvement. (Dana)

4.2.3. Suggestions for improving the app

Finally, the participants' suggestions for improving the efficiency of app were elicited. As learners are thought to benefit the most from app development projects (Lindaman & Nolan, 2015), their suggestions can be useful in this way.

1. Adding more elements in words' presentation and more test types

2. Increasing the amount of contextualization of words

> As I mentioned previously the explanations provided for each entry were very helpful but many other ways can be implemented in order to further this efficiency. In my opinion if words are first put in a sentence and then in a paragraph it would be very helpful. In this way more focus on meaning is achieved. (Simin)

Reza, the teacher participant, was more in favor of an authentic way of contextualizing words:

> I often ask my students to form a group of three persons and use new words in form of a dialogue then perform it for others. I think using this method, I mean putting words in dialogues, can be a good idea. (Reza)

A level B participant suggested a more organized manner of word instruction in the app:

> I studied most of this app's words and I came across words which were related to each other in semantic terms or appearance. I think it is a good idea to put these words in weekly or daily lessons. For those words that are similar in appearance instructions and hints can be provided in order not to mix them or their meanings. (Dana)

Alghamdi (2016) stated that many learners are multimodal in their learning style which means that diverse modalities in combination are conducive to their learning. In the following extract Sara indirectly pointed to different learning styles:

> Some learners have a good visual memory and including visual material to instruction can be very helpful for them. (Sara)

Sara's point was also expressed by the teacher participant:

> In learning eyes can function as very helpful tools. For example some of my students are very good at dictation and when I ask them the reason they say they memorize the form of words while writing them. Because English orthography system contains some exceptions and in all cases sound-symbol correspondence cannot be found I think adding pictures is a very helpful idea. (Reza)

3. Adding words' pronunciation

Although the app was useable anytime anywhere, pronunciation was not paid much attention in it. Below Farhad's suggestion can be seen in this regard.

> As I mentioned in response to previous questions, this app was very easy to use regardless of time and place. As some of the words were totally new for me and I had never seen before, I had no idea about their pronunciation and I even could not guess about it. The problem in this regard was that only if my phone had any network connection the pronunciation was able to be reached. One way to solve this problem is adding phonetic symbols. (Farhad)

4. Adding challenging quizzes (suggested by more proficient learners)

> I really like quizzes on the conditions that my abilities are challenged. I think this was not taken into consideration in this app. All the quizzes were in one form. I think filling in blank questions followed by four or more words can be another form to be added here. (Hoda)

Inclusion of learners' mother tongue was favored by level A participant Sara:

> For EFL learners like me who have a very basic knowledge of vocabulary knowing the Persian meaning of words is of prime importance. If in learning words and exam sections Persian meaning can be included I will continue using it with more interest. (Sara)

The teacher participant suggested a new method of quiz in mobile apps which needs an advanced operating system:

> I think if the app can be designed very smart it can also include open-ended exams in which the user is required to provide more than two one word or even a sentence as the response. (Reza)

Maybe this last suggestion can be considered as a dynamic assessment approach to be implemented in the app's quiz section. This consideration can be justified because the proposer is a graduate of TEFL:

> Providing feedback in form of some mediation while taking the quiz is a helpful idea to improve the exam section. For example if the learner chooses the wrong answer, the right answer will not be revealed on the spot. I mean that some guidance be provided to the learner to make more guesses. This can also be done by adding games which are more exciting. (Hoda)

As these extracts from interview data showed, almost all the participants held some positive views about this app's different parts. Besides possessing remarkable advantages, also some limitations inherent in the app were mentioned by interviewees. Dissatisfaction with the levels difficulty range was said to increase learners' fear of language and also decrease their efficient functioning in language teaching (Lai et al., 2016). According to Javdani et al. (2011), facing difficulty in this situation results in perceiving the tool to be unhelpful for independent and autonomous learning.

Different expectations were reported by participants at different proficiency levels. While intermediate-level students were in favor of mixing Persian and English for instruction and assessment, other interviewees preferred more challenging strategies to enhance their vocabulary proficiency via apps. It has been revealed by language teachers that diverse media provision by CALL and MALL aids learners' to acquire more language (Lindaman & Nolan, 2015), which was suggested by the study's participants in the form of adding pictures. Another widely suggested idea in this regard was showing pronunciation by using phonetic symbols. This is in line with Maleki et al.'s (2015) study, in which more than two thirds of participants agreed (52.5%) or strongly agreed (17.5%) that technology-supported vocabulary learning can be of more interest and usefulness if pronunciation of words is provided. The preference of users for adding pictures, changing way of presenting the material, and inclusion of more exam types were new suggestions not encountered in the literature.

5. Final conclusions and implications for the future

This study was an attempt to address EFL learners' needs to improve vocabulary learning with the aid of technology. To address the issue three research questions were developed to be answered in a mixed-methods approach design. The first two questions were answered using

data collected via questionnaires and analyzed in quantitative terms, while the last research question was answered via semi-structured interviews conducted with two volunteers (one male and one female) from three different proficiency levels.

Quantitative results of the study showed that all participants, regardless of their age, gender, and proficiency level, held positive perspectives about *Vocabulary Flashcards 2016*. The most positively viewed criterion regarding the app was its positive impact (87%) while the most negatively viewed one was learner fit (6%). Results of t-test and ANOVA showed no significant difference as regards gender and proficiency level in terms of participants' preferences and app usage patterns. Most interviewees shared similar positive and negative viewpoints and also suggested similar ideas except for adding game quizzes, which was only proposed by a level C female.

The findings can be useful for EFL learners in all proficiency levels, instructors, and also app developers not only in vocabulary instruction but also when teaching all other language skills and components. Instructors learn how to change the class atmosphere to be learner-centered by asking learners' opinions about materials and material development while app developers should operationalize all these considerations. The evaluation conducted in this study intended to highlight the users' needs which could serve as guidelines for customized vocabulary apps development purpose.

Major limitations of this study were a short span of time devoted to using the app and a limited number of participants. Developing new apps or modifying existing ones is not possible unless longitudinal and more comprehensive research is conducted. The limited number of participants' evaluation of an app cannot be considered as the final judgment to its rejection or acceptance, therefore, more ideas from a diverse number of learners and teachers with different experiences need to be taken into consideration in a longer-term project.

References

Amer, M. (2014). Language learners' usage of a mobile learning application for learning idioms and collocations. *CALICO Journal*, *31*(3), 285-303. doi: 10.11139/cj.31.3.285-302

Awad, A. K. A., & Alkaraki, S. M. S. (2013). Attitudes of EFL students towards using computers in learning English. *ESP-World*, *13*(37), 1-20. Retrieved July 21, 2018 from http://www.esp-world.info/Articles_37/Sulaiman_Attitudes_of_EFL_students.pdf

Basal, A., Yilmaz, S., Tanriverdi, A., & Sari, L. (2016). Effectiveness of mobile applications in vocabulary teaching. *Contemporary Educational Technology*, *7*(1), 47-59. Retrieved July 21, 2018 from http://dx.doi.org/10.1016/j.system.2004.12.008

Bensalem, E. (2018). The impact of WhatsApp on EFL students' vocabulary learning. *Arab World English Journal*, *9*(1). doi: https://dx.doi.org/10.24093/awej/vol9no1.2

Burston, J. (2011). Realizing the potential of mobile phone technology for language learning. *IALLT Journal of Language Learning Technologies*, *41*(2), 56-71. Retrieved July 21, 2018 from https://ialltjournal.org/index.php/ialltjournal/article/download/115/106

Chapelle, C. (2001). *Computer Applications in Second Language Acquisition.* Cambridge: Cambridge University Press.

Celce-Murcia, M., Brinton, D. M., & Snow, M. A. (Eds.). (2014). Teaching and learning vocabulary for second language learners. *Teaching English as a second or foreign language,* (4th ed.). (pp. 288-302). Boston: National Geographic Learning.

Deng, Q., & Trainin, G. (2015). Learning vocabulary with Apps: From theory to practice. *The Nebraska Educator: A Student-led Journal. Paper 29.* Retrieved July 21, 2018 from http://digitalcommons.unl.edu/nebeducator/29

Dörnyei, Z. (2003). *Questionnaires in Second Language Research: Construction, Administration, and Processing*. Mahwah: Lawrence Erlbaum Associates.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.

Gui, L. (2015). Predictability of vocabulary size on learners' EFL proficiency: Taking VST, CET4 and CET6 as instruments. *Studies in Literature and Language*, *10*(3), 18-23. doi: 10.3968/6679

Hubbard, P. (2006). Evaluating CALL software. In L. Ducate & N. Arnold (eds), *Calling on CALL: from Theory and Research to New Directions in Foreign Language Teaching*. San Marcos: CALICO. Retrieved July 21, 2018 from http://web.stanford.edu/~efs/calleval.pdf

Jamieson, J., Chapelle, C. A., & Preiss, S. (2005). CALL evaluation by developers, a teacher, and students. *CALICO Journal*, *23*(1), 93-138. doi:10.1558/cj.v23i1.93-138

Javdani, F., Mahboudi, H. R., Ghafoori, N., & Abdullah, A. B. (2011). EFL students' attitudes towards self-access language learning centers (SALC): The case of Iranian ESP students. *The Journal of Applied Linguistics*, *4*(2), 64-96. Retrieved July 21, 2018 from http://www.sid.ir/en/VEWSSID/J_pdf/1009720110204.pdf

Kawauchi, C. (2008). Proficiency differences in CALL based vocabulary learning: The effectiveness of using "PowerWords". *FLEAT*, *5*, 55-65. Retrieved July 21, 2018 from http://fleat5.byu.edu/_files/71Kawauchi.pdf

Lai, C. C., & Kuo, M. M. (2007). Gender difference in CALL programs for English as a Second Language Acquisition. *Online Submission*. Retrieved July 21, 2018 from http://files.eric.ed.gov/fulltext/ED496190.pdf

Lai, C., Yeung, Y., & Hu, J. (2016). University student and teacher perceptions of teacher roles in promoting autonomous language learning with technology outside the classroom. *Computer Assisted Language Learning*, *29*(4), 703-723. doi: 10.1080/09588221.2015.1016441

Leakey, J. (2011). *Evaluating Computer-Assisted Language Learning. An Integrated Approach to Effectiveness Research in CALL*. Bern: Peter Lang.

Leal Alvez, S. M., & de Oliveria. B. F. (2014). Vocabulary acquisition in second language. *International Journal of Humanities and Social Science*, *4*(13). Retrieved July 21, 2018 from http://www.ijhssnet.com/journals/Vol_4_No_13_November_2014/8.pdf

Leis, A., Tohei, A., & Cooke, S. D. (2015). Smartphone Assisted Language Learning and autonomy. *International Journal of Computer-assisted Language Learning and Teaching (IJCALLT)*, *5*(3), 75-88. doi: 10.4018/IJCALLT.2015070105

Lindaman, D., & Nolan, D. (2015). Mobile-Assisted Language Learning: Application development projects within reach for language teachers. *IALLT Journal of Language Learning Technologies*, *45*(1), 1-22. doi:10.1111/bjet.12224

Maleki, H., Ghasemi, A. A., & Moharami, M. (2015). CAVL: Does it develop learner's attitude? *Turkish Online Journal of Distance Education*, *16*(2), 122-136. http://dx.doi.org/10.17718/tojde.18103

Martiz, G. (2015). *Qualitative Case Study on Cell Phone Appropriation for Language Learning Purposes in a Dominican Context* (Doctoral dissertation). Logan: Utah State University. Retrieved July 21, 2018 from http://digitalcommons.usu.edu/etd/4414/

Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal*, *99*(2), 308-328. doi: 10.1111/modl.12215

Salleh, A. M., & Binti, M. (2010). *A Study of Students' Perception of Mobile Learning in Probability Lessons* (Master's thesis). Available from ProQuest Dissertation and Theses Database. (UMI No. 1506902)

Gourova, E., Asenova, A., Dulev, P. (2013). Integrated platform for mobile learning, In D. Sampson, P. Isaias, J. M. Spector, D. Ifenthaler (eds.), *Ubiquitous and Mobile Learning in the Digital Age* (pp. 67-92). New York: Springer Science+Business Media.

Sanchez, A., & Manchon, R. M. (2007). Research on second language vocabulary acquisition and learning: An introduction. *International Journal of English Studies*, *7*(2), vii-xvi.

Savchenkova, V. B. (2003). *Evaluation Criteria for Language Teaching Software: A Delphi Study of Professional Perspectives* (Doctoral dissertation). Knoxville: The University of Tennessee. Retrieved from: https://www.editlib.org/p/128412/

Seidman, I. (2006). *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences* (3rd. ed.). New York: Teachers College Press.

Shafiei Ebrahimi, Sh. (2016). Effect of digital reading on comprehension of English prose texts in EFL/ESL contexts. *International Journal of English Language and Literature Studies*, *5*(2). doi:10.18488/journal.23/2016.5.2/23.2.111.117

Soleimani, H., & Morshedian, M. (2013). The relationship between attitude to Computer Assisted Language Learning and language proficiency: Evidence from mobile text message. *International Journal of Basic Sciences & Applied Research*, *2*(5), 523-530.

Stockwell, G. (2011). Online approaches to learning vocabulary: Teacher-centered or learner-centered? *International Journal of Computer-assisted Language Learning and Teaching*, *1*(1), 33-44. doi: 10.4018/ijcallt.2011010103

Suwantarathip, O., & Orawiwatnakul, W. (2015). Using mobile-assisted exercises to support students' vocabulary skill development. *Turkish Online Journal of Educational Technology-TOJET*, *14*(1), 163-171. Retrieved July 21, 2018 from https://www.learntechlib.org/p/160575/

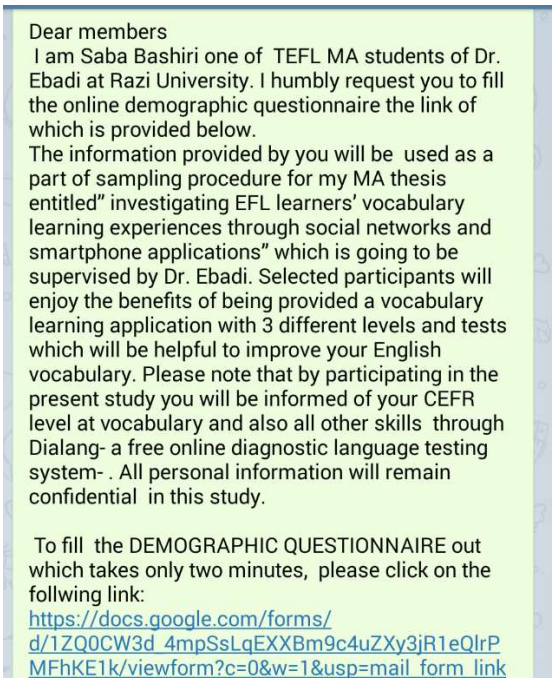
Tabatabaei, O. (2012). The effect of Computer-Assisted Language Learning (CALL) on learning idiomatic expressions: A case of Iranian EFL students. *Journal of Language, Culture, and Translation*, *1*(1), 119-137. Retrieved July 21, 2018 from http://lct.iaush.ac.ir/pdf_5126_c7afc190f26c03c24d1a37360720b95b.html

Vahedi, V. S., Ghonsooly, B., & Pishghadam, R. (2016). Vocabulary glossing: A meta-analysis of the relative effectiveness of different gloss types on L2 vocabulary acquisition. *Teaching English with Technology*, *16*(1), 3-25. Retrieved July 21, 2018 from http://www.tewtjournal.org/?wpdmact=process&did=NDMzLmhvdGxpbms

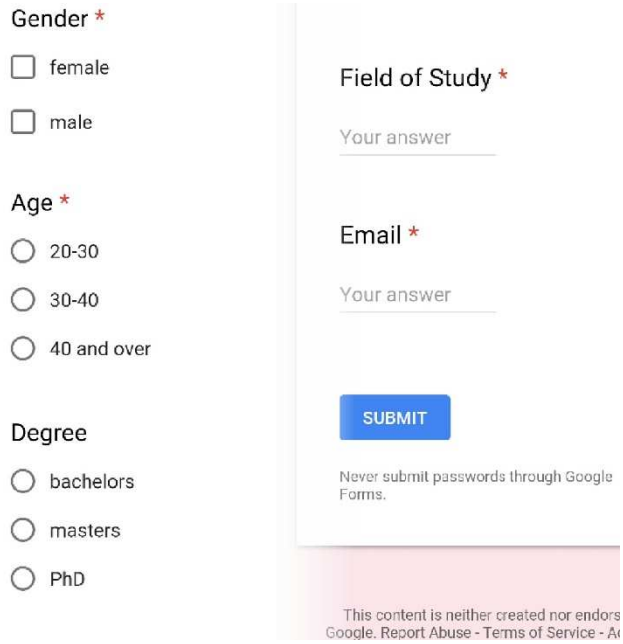
Wang, B. T., Teng, C. W., & Chen, H. T. (2015). Using iPad to facilitate English vocabulary learning. *International Journal of Information and Education Technology*, *5*(2), 100-104. doi: 10.7763/IJIET.2015.V5.48

Appendices

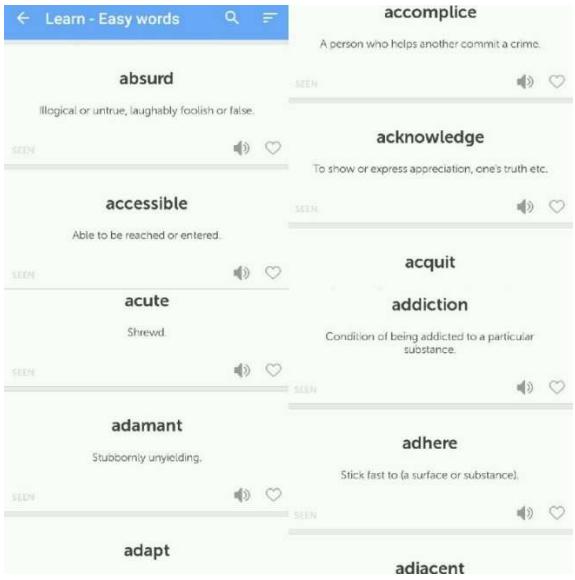
Appendix A: demographic questionnaire

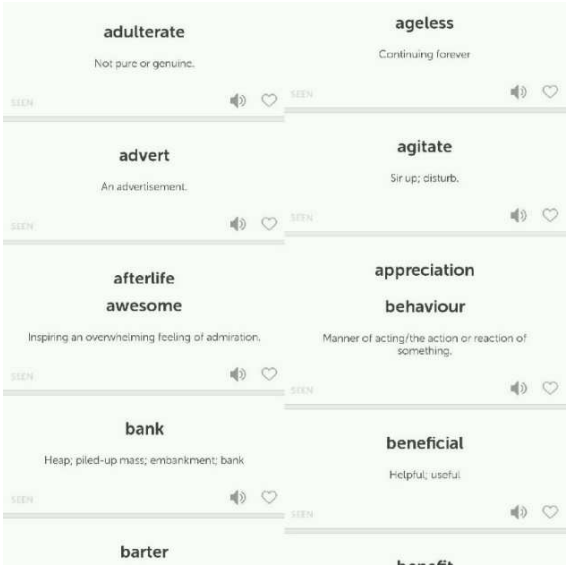

Dear members
I am Saba Bashiri one of TEFL MA students of Dr. Ebadi at Razi University. I humbly request you to fill the online demographic questionnaire the link of which is provided below.
The information provided by you will be used as a part of sampling procedure for my MA thesis entitled" investigating EFL learners' vocabulary learning experiences through social networks and smartphone applications" which is going to be supervised by Dr. Ebadi. Selected participants will enjoy the benefits of being provided a vocabulary learning application with 3 different levels and tests which will be helpful to improve your English vocabulary. Please note that by participating in the present study you will be informed of your CEFR level at vocabulary and also all other skills through Dialang- a free online diagnostic language testing system- . All personal information will remain confidential in this study.

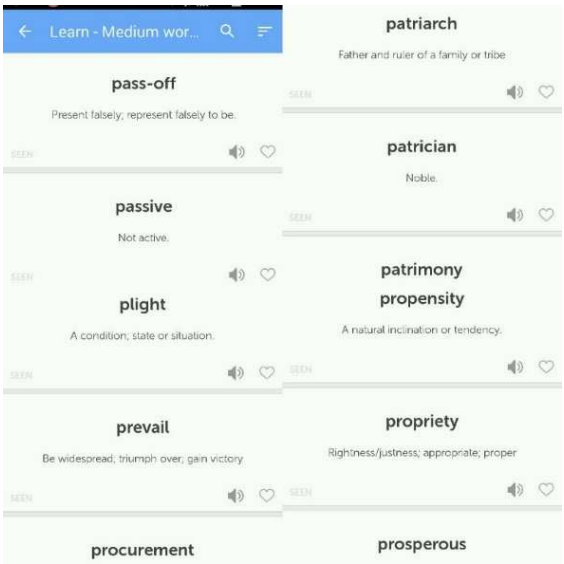
To fill the DEMOGRAPHIC QUESTIONNAIRE out which takes only two minutes, please click on the follwing link:
https://docs.google.com/forms/d/1ZQ0CW3d_4mpSsLqEXXBm9c4uZXy3jR1eQlrPMFhKE1k/viewform?c=0&w=1&usp=mail_form_link

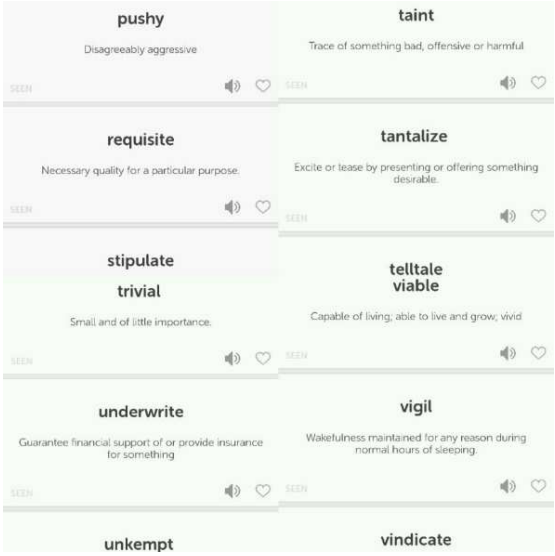

Gender *
female
male
Age *
20-30
30-40
40 and over
Degree
bachelors
masters
PhD
Field of Study *
Your answer
Email *
Your answer
SUBMIT
Never submit passwords through Google Forms.
This content is neither created nor endors
Google. Report Abuse - Terms of Service - A

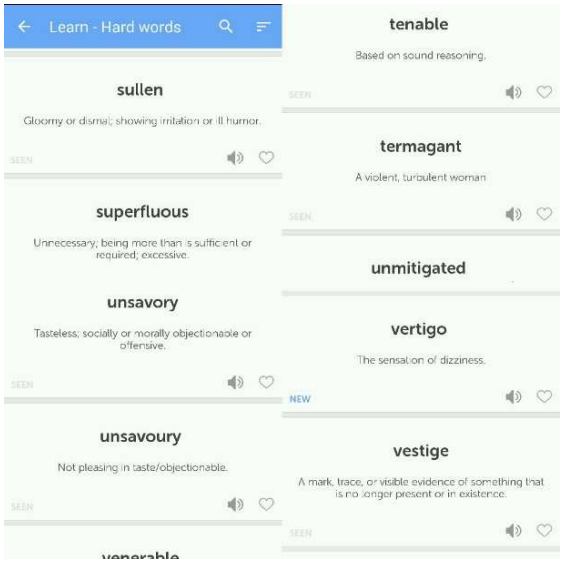
Appendix B: Words studied by participants (2 pictures per level)

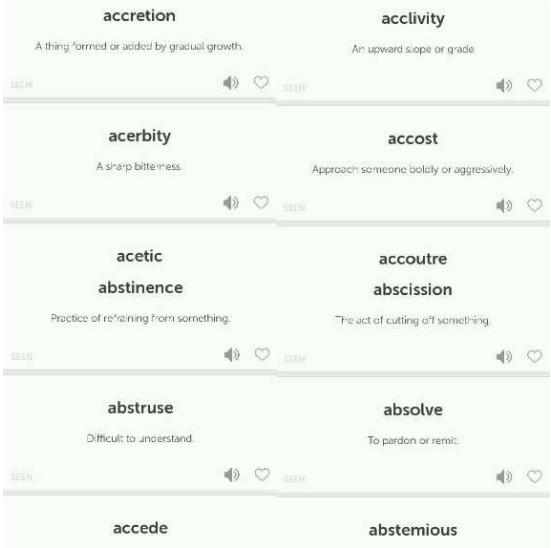












Appendix C: Evaluation questionnaire

Dear Participant

This questionnaire is designed for evaluating vocabulary flashcards 2016 application. Please read the questions carefully and then select the option which is closer to your opinion. The results of filling out this questionnaire will be utilized for conducting MA thesis in TEFL. It is worth mentioning that all personal information will remain confidential. Your precise answers will be a great help in furthering study purposes and improving vocabulary instruction methods.

Thank you very much

Age: gender: educational degree:

Vocabulary proficiency level based on CEFR:

A1 A2 B1 B2 C1 C2

1) I devoted most of my free time during this month on using the app.

1) Yes 2) somewhat 3) not at all

2) Using the app was one the favorite things I could do with my phone.

1) Yes 2) somewhat 3) not at all

3) In case of developing new version of the app I will use it desirably.

1) Yes 2) somewhat 3) not at all

4) I started using the app based on a pre-planned schedule.

1) Yes 2) somewhat 3) not at all

5) I had separate schedules for using each part of the app.

1) Yes 2) somewhat 3) not at all

6) While using the app my main focus was on the form of words.

1) Yes 2) somewhat 3) not at all

7) Using the app challenged my vocabulary ability.

1) Yes 2) somewhat 3) not at all

8) While using the app my main focus was on words meaning and their contextualization.

1) Yes 2) somewhat 3) not at all

9) Words in each level were selected aptly and conveniently.

1) Yes 2) somewhat 3) not at all

10) The quizzes were designed authentically and resembled real life situations.

1) Yes 2) somewhat 3) not at all

11) The app increased my motivation to improve my vocabulary proficiency.

1) Yes 2) somewhat 3) not at all

12) The app made me optimistic about my vocabulary abilities.

1) Yes 2) somewhat 3) not at all

13) I was able to use the app anytime anywhere without any limitations.

1) Yes 2) somewhat 3) not at all

14) I was able to use the app without any specialized skill.

1) Yes 2) somewhat 3) not at all

15) I had enough control over using different parts of the app.

1) Yes 2) somewhat 3) not at all

16) Using the app made me curious to look for similar vocabulary apps.

1) Yes 2) somewhat 3) not at all

17) Provided feedback after quizzes was helpful in reminding me my weak and strong points.

1) Yes 2) somewhat 3) not at all