THE IMPACT OF ASYNCHRONOUS COMPUTER-MEDIATED INSTRUCTION (CAI) ON EFL LEARNERS' VOCABULARY UPTAKE ACROSS DIFFERENT PROFICIENCY LEVELS

by **Zahra Fakher Ajabshir**
University of Bonab, Iran
fakherzahra @ yahoo.com
and **Karim Sadeghi**
Urmia University, Iran
k.sadeghi @ urmia.ac.ir

Abstract

This study investigated the effect of computer-assisted instruction (CAI) on adult second language (L2) learners' vocabulary recognition and production across high and low proficiency levels. Seventy-four participants were assigned to experimental (CAI) and control groups. All participants in the CAI group were categorized into high and low proficiency levels, based on their L2 vocabulary knowledge. The treatment lasted for one semester, 16 sessions, during which 16 passages were covered. While the CAI group individually worked on the passages uploaded on the CALL software, enhanced with hypertext annotations, the control group read the same passages through traditional teacher-directed instruction. The comparison of vocabulary recognition and production pretest and post-test scores revealed the significant effect of CAI on L2 learners' vocabulary uptake in the immediate and delayed post-tests. While no significant differences were found between the high and low proficiency participants in terms of their improvement from production pretest to the post-test, lower-level participants revealed more vocabulary recognition gains. The findings have pedagogical implications for L2 teachers, practitioners, and courseware designers in that they can rely on CALL software as a viable scaffolding tool for L2 vocabulary growth.

Keywords: asynchronous CALL; computer-assisted instruction (CAI); hypertext annotations; vocabulary uptake

1. Introduction

Almost all second language (L2) learners and teachers are well aware of the fact that learning an L2 involves acquiring a large number of words. Experts and researchers in second and foreign language acquisition have increasingly emphasized the importance of investing in vocabulary learning. According to Gardner (2011), L2 learners' most important goal is to work

for communicating effectively in the target language. Without a good command of L2 lexical knowledge they may face communication breakdown. Indeed, vocabulary acts as a bridge between other language competences which are essential for learners' effective communication (Schmitt, 2010). Nevertheless, L2 language learners have difficulty with vocabulary learning. Only a few learners come close to the threshold level in terms of the breadth and depth of vocabulary knowledge (Qin, 2012), thus, it is important to systematically design the learning and teaching of a large number of new words.

Despite the importance of vocabulary acquisition, it seems that the traditional approaches to vocabulary learning have been ineffective, and teachers and scholars must look for alternative approaches to make vocabulary learning an enjoyable and autonomous practice. In recent years, advances in computer technology and the individuals' easy access to the Internet have opened new paths to instruction in a variety of scientific fields. Accordingly, language learning has been informed by the introduction of new instructional technologies. In the domain of technology-enhanced language learning (TELL), computer-assisted language learning (CALL) and computer-assisted instruction (CAI) have been increasingly applied by language teachers and practitioners to numerous aspects of L2 instruction in synchronous real-time and/or asynchronous delayed-time modes (O'Neil, Fisher, & Newbold, 2004).

Coincidental with the application of CALL technologies in L2 classes, a number of CALL-oriented studies have been conducted by L2 teachers and researchers; nevertheless, the literature in the field on different L2 aspects (e.g., Kılıçkaya, 2015; Li, 2018; Mohamadi, 2018) in general and L2 vocabulary (e.g., Khezerlou, Ellis & Sadeghi, 2017; Tsai, in press; Yun, 2011) in particular is yet inconclusive, and more studies are needed to arrive at a more robust picture of the role technology plays in vocabulary instruction. The need is even more evident in the context of Iran as a developing country with weak online infrastructures, which do not easily allow the implementation of CALL. Moreover, most of the existing studies have addressed a single level of proficiency (e.g., Tsai, in press; Wang, 2014), and few studies (Chen, Chen & Yang, 2019; Gorjian, Moosavinia, Ebrahimi Kavari, Asgari & Hydarei, 2011) have looked at the differential impacts of CALL on participants at more or less proficiency levels.

Given the shortcomings of previous research and the call for more studies on the effects of CALL adoption especially in rarely touched contexts, this study aims to investigate the effect of CALL on the acquisition of L2 academic vocabulary among Iranian university students. Moreover, it explores whether the effect of CALL differs as far as the proficiency

level of the students is concerned. The following research questions were specifically addressed:

1. Is there any significant effect for asynchronous CAI on Iranian L2 learners' vocabulary recognition?
2. Is there any significant effect for asynchronous CAI on Iranian L2 learners' vocabulary production?
3. Does the effect of asynchronous CAI, if any, differ across learners at different proficiency levels?
4. Is the effect of asynchronous CAI, if any, retained over a long time?

2. Literature review

2.1. Background

With the advances made in information technology and the growing use of the Internet, computer technology has permeated educational contexts. Ever since, in the field of EFL, many teachers, educators, and practitioners have adopted CALL-afforded technological platforms as an alternative to or a complement for their conventional instructional approaches. Meanwhile, in vocabulary acquisition, the affordances available in CALL (in terms of numerous contacts with the lexical items and the provision of valuable information regarding vocabulary use, spelling, pronunciation, and collocational patterns in multimedia environments) have received lots of attention.

Adopting technology for learning/teaching purposes is supported theoretically and empirically. Theoretically, Paivio's (1991) dual coding theory and the generative theory of multimedia learning (Mayer, 1997) lend support to the use of multi-modal technological interfaces. These scholars argue that different modes of presentation (verbal, pictorial, and textual) may collaboratively provide more favorable conditions for the acquisition of instructional objectives. While employing different modes simultaneously, the burden placed on the working memory will be reduced, and it can process the information in a less demanding way. Apart from the theoretical support, a wealth of empirical studies (e.g., Chen et al., 2019; Eftekhari & Sotoudehnama, 2018; Li, 2018; Mohamadi, 2018) also offer evidence on the advantages associated with computer technology.

This study does not involve many multimedia modes, since we limited our experimentation to the textual annotation. Accordingly, our study borrows its theoretical foundation from Vygotsky's (1978) sociocultural theory, and in particular the notion of

scaffolding. According to Frawley and Lantolf (1985), the learning process may be regulated by the individual's interaction with others (other-regulation) or with the tools or mediational means (object-regulation). Based on this view, a number of tools act as a buffer between the learner and the social environment and mediate the relationship between the learner and the social world (Lantolf, 2000). From among a variety of tools, computer-mediated platforms can be regarded as scaffolding tools or mediational means for promoting learners' L2 knowledge in general and vocabulary knowledge in particular. Drawing on this theoretical foundation, the purpose of this study was to investigate the effect of asynchronous CAI on Iranian university students' L2 vocabulary uptake.

Two important aspects of vocabulary include the recognition and production of lexical items. According to Harley (2008), "recognizing a word occurs when we uniquely access its representation in the mental lexicon" (p. 207). In production, on the other hand, we go from three phases of conceptualizing the message, formulating it into a linguistic form, and executing it by phonetic planning and articulation. While recognition involves the activation of existing memory traces, production demands searching within the mental representations of the already acquired knowledge (Cariana & Lee, 2001). Given the various underlying processes involved in recognition and production, different factors might affect each of them, amongst them, according to Lee and Pulido (2016), the proficiency level of the individuals. A further purpose of this study was thus to explore whether the effect of asynchronous CAI on L2 vocabulary recognition and production is significantly different for learners at different proficiency levels.

2.2. CALL and vocabulary acquisition

In L2 vocabulary acquisition, computer-mediated annotations or glosses can be employed to clarify the meaning of unknown words. They have the potentiality of assisting learners in an adaptive, autonomous, and individualized context. These annotations might be L1 translations, L2 synonyms, definitions, exemplifications, visuals, or a combination of them. There is ample evidence that the use of hypertext glosses affects the reading skill and vocabulary gains in a variety of ways. According to Abuseileek (2008), the incorporation of CALL *per se* does not explain the overall vocabulary acquisition, but it does so via increasing the retention time and decreasing the vocabulary look-up time. Some other studies (e.g., Su, Li, Liang, & Tsai, in press; Wang, 2016) also attributed the beneficial effects of CALL to learners' positive attitudes, perceptions, and motivation towards reading enhanced by hypertexts.

In pedagogical contexts informed by CALL technology, learners may be engaged in an online synchronous (e.g., video conferencing and chatting) or offline asynchronous (e.g., email and blog) CALL. As stated by Abrams (2003), synchronous and asynchronous CALL are similar and different in a number of ways. Both of them offer affordances in terms of more opportunities for language use, increased amounts of input and output, and more interaction and negotiation. Synchronous mode is simultaneous and requires immediate response and feedback, not allowing for external support. Asynchronous forum, on the other hand, is not subject to time constraints, and learners are able to learn the language at their own pace without being interrupted by the factors inherent in traditional face-to-face modes. Due to the affordance provided in asynchronous mode in terms of reflection on one's ideas, it results in the production of more sophisticated lexicon and syntactically more complex language (Zapata & Sagarra, 2007). According to Fitzpatrick and Donnelly (2010), decisions on whether to adopt a synchronous/asynchronous approach are contingent upon a number of factors including individual dimensions, preferences, aims, purposes, and institutional and pedagogical objectives.

The interface between synchronous/asynchronous CALL and L2 vocabulary acquisition has been examined in a number of studies, providing evidence on the preference of CALL over traditional approaches. In a meta-analysis, Chiu (2013) found an overall average effect of CALL on L2 vocabulary development. She enumerated four important moderators of vocabulary learning in CALL: treatment duration, participants' educational level, game-based learning, and the instruction of the teacher. Learners who received CALL treatment over a short period of time (about a month) benefited from this type of instruction more than those who were exposed to similar instruction in the long run. Moreover, CALL proved more effective for students at high educational levels (e.g., university level) compared with elementary levels. Instruction via CALL without the games appeared to be better than game-based instruction. Finally, autonomous student-centered learning led to better outcomes than teacher-directed instruction. Similar observations in term of the advantages of CALL were also reported by Wang (2016), Mirzaei, Rahimi Domakani and Rahimi (2016), and Tsai (in press).

A number of CALL-focused studies have found that learners' vocabulary retention may vary as a function of their proficiency level, amongst other factors. The corresponding vocabulary growth was found to be dissimilar for low and high proficiency learners. Some studies offered evidence on better vocabulary gains in advanced learners (e.g., Abraham, 2008; Gorgian et al., 2011). Abraham (2008), in synthesizing the findings of previous studies on the impact of glosses on reading comprehension and incidental vocabulary acquisition, found a

small effect size for beginners compared with intermediate and advanced learners. She, however, argued that since the number of studies was limited for each of the instructional levels, these conclusions were tentative. In a further study, Gorjian and his colleagues (2011) found that low achievers benefited from CALL in vocabulary retention (as shown by their immediate post-test scores), but high achievers demonstrated gains in both vocabulary retention and recall (as shown by their delayed post-test scores).

The better gains of the advanced learners were not reported in all studies, with some research documenting better scores of low proficiency learners (e.g., Chen et al., 2019; Li, 2010; Yun, 2011). Li (2010) explored the short-term and long-term effects of using computer-mediated dictionaries on Chinese English as a Second Language (ESL) students' retention of vocabulary items across different proficiency levels. During the treatment, the participants were required to read stories in two conditions: with and without the support of monolingual print dictionaries and/or bilingual electronic dictionaries. This was followed by some vocabulary tests based on the reading texts. The results showed that low-ability learners outperformed their high-level counterparts. Similar findings were reported by Yun (2011), who conducted a meta-analysis to synthesize the findings of some previous studies that compared the effect of computer-mediated glosses on L2 reading and vocabulary retention among learners exposed to these glosses versus those who used traditional techniques. He found the positive impact of computer-mediated glosses on these measures. In particular, it appeared that, in comparison with intermediate and higher-level learners, lower-level learners were more likely to get advantage from multiple multimedia glosses. This concurs with the findings of Chen (2019), who developed a corpus-based paraphrasing system, assisting learners to expand the knowledge of form, meaning, and the use of lexical items and found the better improvement of the weaker students.

Considering the mixed findings in terms of the benefits of CALL to learners at high/low proficiency levels, this domain opts for more studies to arrive at robust findings and generalizable results. Moreover, given the multi-dimensionality of the construct of vocabulary knowledge (Zhong, 2014), vocabulary-focused research should be designed in a way to account for different aspects of word knowledge, including recognition and production. Accordingly, this study aims at examining the effects of asynchronous CALL on Iranian EFL learners' acquisition of vocabulary knowledge across high and low proficiency levels. It also explores whether the vocabulary gains, if any, are maintained over a long time.

3. The present study

3.1. Design

This study adopted a quasi-experimental design with two intact classes to investigate the impact of asynchronous computer-assisted instruction (CAI) on L2 learners' vocabulary uptake across different proficiency levels. The instruction type was the independent variable and learners' comprehension and production of vocabulary represented the dependent variables. The proficiency level of the learners served as a moderating variable.

3.2. Participants

The participants of this study were selected from two intact classes in a national University in East Azarbaijan Province, Iran. They were 79 (45 male and 34 female) freshman students in the age range of 18 to 27 (M = 23.6). The first language of the participants was either Farsi or Azari Turkish, and they had an average of 6.5 years of formal pre-university English learning. Based on their scores in the quick placement test (QPT), the participants were at high-intermediate (N = 37) and low-intermediate (N = 42) proficiency levels. They enrolled in a general English course which is an obligatory course for all university students. There was a 7% subject attrition. Since some students (N = 3) did not take the post-test or were not present in some treatment sessions (N = 2), they were excluded from the final analysis. So, from the original pool of 79 participants, the data from 74 (N = 74) participants including 41 males and 33 females were analyzed.

The classes were assigned to CAI (N = 43; 24 males and 19 females) and control groups (N = 31; 17 males and 14 females). Moreover, based on their scores in the QPT, the participants in the CAI group were assigned to the high (N = 23) and low proficiency (N = 20) levels.

3.3. Instrumentation

Quick placement test (QPT): QPT is a standardized test with established reliability and validity developed by Oxford University Press and University of Cambridge Local Examinations Syndicate. It includes two parts, 60 items, with the second part including more difficult items. For this study, the first part of the test was used. It included 40 multiple-choice items, 25 items for vocabulary and 15 cloze items. The test took about 45 minutes to complete. The internal consistency of the test was also acceptable as indicated by a Cronbach's alpha coefficient of .77.

Vocabulary pre/post-test: Two isomorphic researcher-made versions of a vocabulary test were administered as the pretest and the post-test. The vocabulary items included in each test were meticulously chosen from the students' course book to represent all chapters covered. Each test included 40 multiple-choice and 40 completion items, which were used to assess the participants' recognition and production of L2 vocabulary prior to and following the treatment. In each of the tests, the items were equally distributed in terms of the word classes (parts of speech) including five items for each of the noun, adjective, verb, and adverbs. Moreover, to control the effect of word frequency level, an attempt was made to choose the lexical items in the test stem and response options (in the case of the recognition test) from among the 4,000 and 5,000 word frequency levels. In scoring the recognition test, each correct answer was given 0.05 point. There was no penalty for wrong answers. The criteria used for scoring the production tests were based on the lenient/strict evaluation proposed by Yoshi and Flaitz (2002). Accordingly, partially correct answers were given 0.25 points while the fully correct answers were allocated 0.5 point. Otherwise, no point was assigned to the responses. The maximum score for each of the recognition and production tests was 20.

The original version of the pretest was piloted with 20 students similar to the target population. Following the pilot test and having consulted with two specialists in the field, some items were removed or replaced. The reliabilities of the tests were also measured by Cronbach's alpha coefficient and found to be acceptable yielding .78 and .73 coefficients for the pretest and the post-test, respectively. The subjective judgments of two experts in TEFL were used to check the content validity. Moreover, the inter-rater reliability measures of the production test scores were verified by Spearman-Brown Formula with coefficients of .79 and .83 for the production pretest and post-test, respectively.

Questionnaire: A background questionnaire was administered to elicit the participants' prior knowledge and experience in using computer technology. The 20 items in the questionnaire were adapted from Warner (2004), and in addition to demographic questions (name, age, gender, native language), it included three parts:

(a) questions on the participants' amount of access to the computer/Internet [how many hours a day they used computers at campus, in dorm/home, at work (if they had a job), and other (specify)],

(b) the purposes for which they used the computer (word processing, email, World Wide Web, chat and online discussions, and gaming)

(c) their assessment of the extent to which they used the technology for a variety of purposes (for career, communication with people, learning about people and

cultures, overcoming weaknesses, getting a sense of belonging to a community, and enhancing the creativity).

The participants provided their responses to part (b) and (c) on a 5-point Likert scale. The participants' responses to the questionnaire items were used to assign them to control and experimental groups, with the students with more prior knowledge and further experience of the computer technology being placed into the experimental group. This was done to ensure that the experimental group's computer literacy (as a construct irrelevant variance) or a lack thereof did not affect their post-test performance.

Textbook: To pass the general English course, all students needed to cover eight chapters (16 passages) of *Active Skills for Reading 2* (Anderson, 2008). It remains one of the major sources for General English courses in most of the universities in Iran. The book has been authored with an intermediate-level audience in mind who aim at increasing their general and academic English knowledge and are preparing for standardized tests. It includes a variety of passages with interesting and engaging topics (e.g., young athletes, human body, leisure time, and music). Each unit contains some brainstorming questions followed by two passages. After each passage, there are several activities including reading comprehension exercise, activities designed to promote learners' critical thinking skills, vocabulary matching exercises, and completion type exercises. For this study, the focus was mainly on the activities that aimed at promoting L2 vocabulary knowledge.

Vocabulary building software: The vocabulary building software employed in this study was *Learning with Texts*, version 14. It is a tool to support the general and academic language learning through reading, listening, and testing the words in the context. It has a lot of user-friendly resources; however, not all of these facilities were used in this study. One of the most important and useful features of this software – which is hardly present in similar types of vocabulary software – is its potential for users to upload their content. The main feature of the software used in this study was the glossed dictionary which enabled the learners to look up the words' L1 translation, synonyms, and parts of speech. Also, specific modules have been incorporated into this software to provide the words' pronunciation.

3.4. Procedure

Prior to the treatment and based on the participants' responses to the background questionnaire, the participants with more prior experience with CALL applications were assigned to the experimental (CAI) group while those with lesser experience sat in the control group. Moreover, based on their QPT scores, all participants in the CAI group, in an uninformed way,

were assigned to high and low proficiency levels. In the introductory session, the CAI group was familiarized with the type of instruction and materials used and received technical training on how to use the vocabulary software.

The CAI group received the treatment in the language laboratory equipped with PCs with the vocabulary building software installed, headphones, and a good Internet connection. Before each session, the reading passage that would be covered in that session was uploaded by the teacher (one of the researchers). At the very beginning, the learners listened to the passage through headphones. Following this, they read the passage on their own. While encountering a new word, they clicked on it. A window then opened displaying the information needed to clarify the meaning of that word like synonyms, antonyms, definitions, and sample sentences including the word. Halfway through the task, the teacher walked around and assisted the students if they encountered any problems.

While the CAI group received the treatment in the language laboratory, the control group attended the sessions in a normal classroom with no computing facilities. The participants in the control group covered the same passages individually with no access to the CALL software. Prior to reading each paragraph, the meanings of the unknown words were clarified by the teacher through verbal cues including synonyms, definitions, and examples.

After reading each passage, sample textbook activities with a major focus on vocabulary development were done by both groups. While each participant accomplished the activities individually, the teacher monitored their performance and offered them feedback and assistance. Finally, the answers were checked and shared with the whole class.

The treatment lasted three months, 16 sessions, once a week and twice every other week for 90 minutes. Following the treatment, the post-test and four weeks later, the delayed post-test including the vocabulary recognition and production tests were administered to gauge the participants' recognition, production, and retention of vocabulary items after the treatment.

3.5. Data collection and analysis

The data for this study were collected using three instruments: QPT, questionnaire, and vocabulary pre/post-test. QPT was administered before the treatment to check the participants' general English proficiency level. The questionnaire was completed by the students to assign them to control and experimental groups. Vocabulary pretest and post-test, including a similar number of multiple-choice and completion items for the recognition and production of vocabulary, were administered to gauge the participants' knowledge prior to and following the treatment.

The data collected were analyzed using SPSS version 22. A series of inferential and parametric statistics were run. After ensuring the normal distribution of the data, a series of tests were run to answer the questions. The level of significance was set at 0.05. As for research questions 1 and 2, two independent samples *t*-tests were conducted to compare CAI and control groups' post-test scores in both measures of recognition and production of vocabulary items. Moreover, two paired samples *t*-tests were run to make sure that the CAI group improved from pretest to post-test in recognition and production of vocabulary. As for the third research question, a series of ANOVA tests were run to compare high, low, and control groups in vocabulary recognition and production at pretest and post-test stages. Using the Scheffe test, post-hoc paired comparisons were also run to locate where the difference between the groups lies. To answer research question 4, two paired samples *t*-tests on recognition and production post-test and delayed post-test scores of the CAI group were conducted to check whether the effect of the instruction was durable over time.

The following section presents the results, which are organized around the three main areas of focus in this study: (a) the effect of CAI on vocabulary uptake, (b) the effect of CAI across high and low proficiency levels, and (c) the long-term effect of CAI on vocabulary uptake**.** Finally, the results obtained from the questionnaire were discussed.

4. Results

4.1. The effect of CAI on vocabulary uptake

The first and second research questions addressed the effect of CAI on vocabulary recognition and production. Table 1 shows the descriptive statistics for CAI and control groups in the pretest and the post-test. The comparison of means shows that the CAI group outperformed in vocabulary recognition post-test (M = 13.43; SD = 2.53) compared with the pretest (M = 9.69; SD = 3.24). Similarly, the mean increased from production pretest (M = 10.03; SD = 3.18) to post-test (M = 13.03; SD = 3.01).

Table 1. Descriptive statistics for the pretest and the post-test scores

	N	Mean	SD	Min.	Max.
Control, Recognition Post-test	31	10.27	2.01	8.00	13.00
Control, Production Pretest	31	9.98	1.23	8.50	13.50
CAI, Recognition Pretest	43	9.69	3.24	5.50	14.0
Control, Production Post-test	31	10.23	4.09	7.5	15.00
CAI, Recognition Post-test	43	13.43	2.53	12.0	19.0
CAI, Production Pretest	43	10.03	3.18	7.50	18.0
CAI, Production Post-test	43	13.03	3.01	9.00	16.5
Control, Recognition Pretest	31	9.61	1.22	8.50	14.00

Independent samples *t*-tests were conducted to compare the means of the CAI and control groups in vocabulary recognition and production post-tests (Table 2).

Table 2. Independent samples *t*-test for vocabulary recognition and production post-test scores of CAI and control groups

				Paired Differences 95% Confidence Interval				
	Mean	**SD**	**SEM**	**Lower**	**Upper**	***t***	***df***	**Sig.(two-tailed)**
CAI, Control Rec.	3.16*	0.35	0.29	2.56	5.98	9.74	43	0.00
CAI, Control Pro.	2.90*	0.47	1.3	.020	3.07	10.23	39	0.00

Note. The mean difference is significant at the 0.05 level.
Rec = Recognition, Pro = Production

The results of independent samples *t*-tests, as illustrated in Table 2, show a significant difference between the post-test scores of control and CAI groups in vocabulary recognition ($t = 9.74$, $p < 0.05$) and production ($t = 10.23$, $p < 0.05$). To assess the CAI group's improvement from pretest to post-test (within-group comparison), paired samples *t*-tests were conducted (Table 3).

Table 3. Paired samples t-test of CAI group's scores for the vocabulary recognition and production tests

					Paired Differences 95% Confidence Interval			
	Mean	**SD**	**SEM**	**Lower**	**Upper**	***t***	***df***	**Sig.(two-tailed)**
Rec. pretest-post-test	3.74*	.23	.43	-2.38	3.98	10.26	21	0.00
Pro. pretest-post-test	3.10*	.45	1.2	-3.02	4.64	12.76	21	0.00

Note. Rec = Recognition, Pro = Production

The results of paired samples *t*-tests (Table 3) comparing the effect of CAI on vocabulary uptake show that significant differences exist between the participants' mean scores in recognition ($t = 10.26$; $p < 0.05$) and production post-test ($t = 12.76$; $p < 0.05$) compared with the pretest. Thus, in response to the first and second research questions, it can be concluded that CAI had a statistically significant effect on vocabulary uptake of Iranian L2 learners.

4.2. The effect of CAI across high and low proficiency levels

The focus of the third research question was the effect of CAI on vocabulary uptake across the participants at high and low proficiency levels. Table 4 shows the descriptive statistics related to the pretest and post-test scores of control, high, and low groups. As shown in the table, there are differences in the vocabulary recognition and production mean scores of the high group (Recognition M = 14.32; SD = 3.07; Production M = 13.49; SD = 1.02) and low group (Recognition M = 12.03; SD = 1.09; Production M = 12.58; SD = 2.73).

Table 4. Results of the pretest and post-test scores of control, high, and low groups

	N	M	SD	Min.	Max.
High Group's Recognition Pretest	23	11.24	4.03	10.0	15.00
High Group's Recognition Post-test	23	14.32	3.07	12.0	19.00
High Group's Production Pretest	23	10.63	3.21	7.50	19.00
High Group's Production Post-test	23	13.49	1.02	11.5	16.50
Low Group's Recognition Pretest	20	8.12	2.81	5.00	11.00
Low Group's Recognition Post-test	20	12.03	1.09	9.50	15.00
Low Group's Production Pretest	20	9.54	4.31	7.00	12.50
Low Group's Production Post-test	20	12.58	2.73	9.00	14.50

To compare the significance of differences between the mean scores across the high, low, and control groups, two one-way ANOVA tests of between-subjects effects were run to compare the recognition and production scores of the three groups (Table 5).

Table 5. ANOVA tests comparing the improvement of high, low, and control groups from vocabulary recognition and production pretest to post-test

	Sum of squares	*df*	Mean Squares	*F*	Sig.
Rec. Between Groups	326.74	2	163.37	36.12	0.00
Rec. Within Groups	215.02	47	5.63		
Total	541.76				
Pro. Between Groups	298.20	2	203.25	29.86	0.00
Prod. Within Groups	167.32	47	12.02		
Total	465.52				

Note. Rec = Recognition, Pro = Production

As Table 5 shows, there is a significant difference between the three groups in terms of vocabulary recognition ($F = 36.12$, $p < 0.05$) and production scores ($F = 29.86$, $p < 0.05$).

To determine the location of the difference, post hoc pairwise comparisons were run. Tables 6 and 7 present the results of paired comparisons.

Table 6. Post hoc pairwise comparisons of vocabulary recognition gains

		Mean Difference	SEM	95% Confidence Interval		Sig.
				Lower	Upper	
High Group	**Control**	4.05*	0.21	1.73	4.43	0.01
	Low	2.29*	0.72	0.11	5.01	0.00
Low Group	**Control**	1.76*	0.54	1.24	3.81	0.00
	High	-2.29*	0.72	-5.01	-0.11	0.00
Control Group	**High**	-4.05*	0.21	-4.43	-1.73	0.01
	Low	-1.76*	0.54	-3.81	-1.24	0.00

Note. $P < 0.05$

As shown in Table 6, the vocabulary recognition post-test scores are significantly different in three groups. The high group outperformed control (mean difference = 4.05; SD = 0.21) and low groups (mean difference = 2.29; SD = 0.72). Likewise, a significant difference was found between control and low groups (mean difference = 1.76; SD = 0.54). A further point is that although both of the experimental groups appeared to perform significantly better in vocabulary recognition post-test compared with the pretest, as observable in Table 4, the recognition gains were higher in the case of the low group (pretest-post-test mean difference = 3.91) compared with the high group (pretest-post-test mean difference = 3.08). It can be concluded that asynchronous CAI had an effect on adult L2 learners' vocabulary recognition, and the effect was higher for low proficiency learners.

Table 7. Post hoc pairwise comparisons of vocabulary production gains

		Mean Difference	SEM	95% Confidence Interval		Sig.
				Lower	Upper	
High Group	**Control**	3.26*	0.61	.0.71	5.21	.001
	Low	0.91	0.73	-5.01	3.53	.067
Low Group	**Control**	2.35*	0.46	1.28	3.84	.000
	High	-0.91	0.73	-3.53	5.01	.067
Control Group	**High**	-3.26*	0.61	-5.21	-0.71	.001
	Low	-2.35*	0.46	-3.83	1.28	.000

As presented in Table 7, the high and low groups performed significantly better than the control group on vocabulary production post-test (high and control group's mean difference = 3.26, $p < 0.05$; low and control group's mean difference = 2.35; $p < 0.05$). However, no significant difference was found between the high and low group in the production of L2 vocabulary when exposed to asynchronous CAI (high and low group's mean difference = 0.91, $p > 0.05$). Thus, there was an interaction between the proficiency level and the type of vocabulary tests. In response to the third research question, it can be concluded that while both high and low groups outperformed their pretests in vocabulary recognition and production post-tests, the low group revealed higher recognition gains than the high group.

4.3. The long-term effect of CAI on vocabulary uptake

To address research question 4, which addressed the long-term effect of CAI on vocabulary uptake, the post-test and delayed post-test scores of the CAI group were compared. Table 8 shows the descriptive statistics.

Table 8. Descriptive statistics for the post-test and delayed post-test scores of the CAI group

	N	M	SD	Min.	Max.
Recognition Post-test	43	13.43	2.53	12.0	19.0
Recognition Delayed Post-test	43	13.29	3.21	11.0	17.0
Production Post-test	43	13.03	3.01	9.00	16.5
Production Delayed Post-test	43	12.97	3.45	7.00	18.5

To compare the significance of differences between the post-test and delayed post-test mean scores, two paired samples *t*-tests were run.

Table 9. Paired samples *t*-test for the post-test and delayed post-test scores of the CAI group

	Paired Differences 95% Confidence Interval							
	Mean	**SD**	**SEM**	**Lower**	**Upper**	***t***	***df***	**Sig.(two -tailed)**
Rec. Post-Delayed	0.14	.23	.63	0.07	1.23	1.47	47	.15
Pro. Post-Delayed	0.06	1.45	1.42	-1.08	2.69	9.53	38	.07

Note. Rec = Recognition, Pro = Production

Table 9 shows no significant difference ($p = .15$) between the vocabulary recognition scores in the post-test and the delayed post-test (M = 0.14, SD = 0.23) with t (47) = 1.47, $p > 0.05$. Similarly, vocabulary production scores revealed no significance difference between the

post-test and delayed post-test scores (M = 0.06, SD = 1.45) with t (38) = 9.53, $p > 0.05$. Thus, in response to research question 4, it can be concluded that the effects of CAI on vocabulary uptake were retained over four weeks.

4.4. Questionnaire results

The results of the first part of the questionnaire suggested that while almost all learners had access to computers and the Internet, the place and the amount of time they spent on computers per day were different across participants. The majority of the students reported that they had access to computers at campus (4.2), in dorm (3.6), at home (4.3), and in dorm computer center (2.02). Few of them (1.6%) reported computer access at work, and only 1.21 percent used their friends' computers.

Table 10. Participants' responses to the questionnaire (part 1)

The place where you access the computer	**Hours per day**
At campus computer	4.2
In dorm room	3.6
In dorm computer center	2.02
At the place where you live (if not a dorm)	4.3
At work	1.6
From a friends computer Other (please specify)	0.21
Other (please specify)	0.03

Concerning the second part of the questionnaire (the purposes for which the computer/Internet was used), surfing the World Wide Web received the highest score (79**%**), followed by word processing (53%) and email (48%). Using the technology for online chatting and discussion in groups did not receive as much ratings (21% and 17%, respectively). Eighteen percent of the participants reported that they used the computer for gaming purposes.

Table 11. Participants' responses to the questionnaire (part 2)

	often	**sometimes**	**rarely**	**never**
Word processing	53%	15%	29%	3%
E-mail	48%	21%	24%	7%
World Wide Web	79%	13%	5%	3%
Online chatting and discussion	21%	15%	53%	11%
Gaming	18%	22%	29%	31%

The last part of the questionnaire addressed the participants' ratings of the extent to which they used the computer/Internet for a variety of purposes. Most of them (84%) believed that computers are useful for their future careers. Three-quarters of them (75%) reported that technology assisted them to overcome the weaknesses and obstacles. Seventy-one percent and 64 percent used it for communication with other people and for learning about other people and cultures, respectively. Some students (29%) found computers less threatening than face-to-face communication. About half of the students (52%) felt a sense of belonging to a community while using the technology, and a little more students (60%) perceived technology as a means for fostering creativity.

Table 12. Participants' responses to the questionnaire (part 3)

Please rate each of the following questions	**strongly agree**	**agree**	**no opinion**	**disagree**	**strongly disagree**
Learning how to use computers is important for my career.	79%	5%	3%	6%	7%
I enjoy using computer to communicate with people around the world.	43%	28%	14%	5%	10%
Using the Internet is a good way to learn about different people and cultures.	46%	18%	21%	13%	2%
Computers help people overcome weakness and powerlessness.	57%	18%	13%	5%	7%
I am less afraid to contact people by e-mail than in person.	17%	12%	23%	35%	13%
Using e-mail and the Internet makes me feel part of a community	34%	18%	17%	23%	8%
Working with computer makes me more creative.	48%	12%	25%	43%	12%

5. Discussion

This study aimed to investigate the effect of asynchronous CAI on vocabulary uptake among adult L2 learners and to explore whether the vocabulary gains differed across learners at high and low proficiency levels. It was found that CAI had a positive effect on vocabulary recognition and production among adult L2 learners; however, an interaction was found between the proficiency level and the type of vocabulary tests. The differences in pretest and post-test means reveal that while both high and low proficiency groups showed almost similar degrees of production gains, low proficiency participants revealed higher recognition gains. Finally, the effect of CAI was found to be durable over four weeks.

The outperformance of the experimental group compared with the control group documents the beneficial effect of CAI on L2 vocabulary development, which has been corroborated in some studies (e.g., Su et al., 2019; Tsai, in press; Wang, 2014). These studies argued that online vocabulary tools provide opportunities for vocabulary practice and enrichment, an affordance non-existent in traditional approaches to vocabulary instruction. Boers, Warren, Grimshaw, and Slyanova-Chanturia (2017) argued that using different forms of glosses afforded by online tools brings about learners' mental engagement with the target word and hence promotes the acquisition of different aspects of the word. This finding is also consistent with the SCT in that the technological tools provide affordances for learning and regulate the learning process, providing mediated assistance to learners and help them move from object-regulation towards autonomous functioning or self-regulation.

The higher recognition gains observed in the lower level participants concur with the findings of some studies (e.g., Chen, et al., 2019; Li, 2010; Yun, 2011) which claimed that the online vocabulary enhancement tools including the computer-mediated dictionaries and particularly bilingual dictionaries conform to low-ability learners' learning styles and preferences. Zapata and Sagarra (2007) argued that while processing the unknown words, low-ability learners experience greater difficulty and high cognitive load and are likely to avoid allocating much time and mental operations to process these words. The provision of computer-mediated aids helps "enhancing cognitive resources and lead[s] learners to engage in deeper processing when needed" (p. 168).

The improvement of the weak learners, however, contradicts the results of some studies (e.g., Abraham, 2008; Gorjian et al., 2011) that reported higher performance gains by advanced learners and attributed this to the dual code theory (Paivio, 1991). According to this theory, two mental systems or codes, including verbal and non-verbal, account for the knowledge of language and knowledge of the world. When applied to L2 vocabulary acquisition, by using

multiple (auditory and visual) glosses of retrieving new vocabulary, knowledge of the lexical item is established as a result of the simultaneous engagement of auditory and visual memories. However, not a variety of hypertext glosses were used in this study, with the major gloss being the textual one. Although the pronunciation module has been incorporated in the program as well, it assisted the participants' mastery of phonological form rather than meaning and did not contribute to learners' test performances.

Vocabulary recognition gains differed across varied proficiency levels, however, a similar trend was not observable for production gains. While both high and low groups outperformed their vocabulary production pretest scores, not a significant difference was found between the two groups in their vocabulary production gains. Moreover, as revealed by their delayed post-test scores, both groups were found to retain the vocabulary gains after four weeks. It can be concluded that both groups not only used the textual glosses to make sense of the general meaning of the text, but also retained the lexical associations and cues for future use. This corroborates the findings of some previous studies (e.g., Li, 2010; Rimrott, 2010) that documented the long-lasting effect of CALL on vocabulary acquisition.

6. Conclusion, limitations and suggestions for future studies

The limitations of this study should be acknowledged. As a novel experience, exposure to CAI in the beginning sessions entailed some degree of resistance on the part of the students with a lesser degree of ambiguity tolerance. This was, however, alleviated gradually, and despite an awkward commencement, the students were comfortable with the technology in the later sessions. A further limitation relates to employing a small population and a single type of gloss, i.e., textual. Future research may address the implementation of CALL with a bigger sample size, a variety of annotation types, using more rigid measures, and over a prolonged period of time to provide a detailed account of how the incorporation of CALL technology alone or as an extracurricular program affects the development of different aspects of L2 in general and L2 vocabulary skill in particular.

This study contributes to CALL research by providing evidence on the affordances offered by CAI in vocabulary recognition and production among L2 learners at high and low proficiency levels in both the short- and the long run. Some pedagogical implications may be drawn. Teachers, L2 practitioners, and material designers are suggested to incorporate the technology in the design of the curricula as an aiding tool in conventional face-to-face instructional contexts. The overall better improvement of the low-level group suggests that CAI may be potentially more effective for weaker students, provided that the affordances of this

technology are appropriately tailored to the proficiency level of the students (Çakmak & ErÇetin, 2018). Various features of the CAI interface may be customized to cater for the needs of a variety of students with different educational goals and objectives.

While implementing CAI, it should be noted that technology does not necessarily guarantee success. The educational system's online infrastructures, the stakeholders' computer literacy, the properties of the program, and learners' attributes including their learning styles and preferences (verbalizer or visualizer) are the factors in need of consideration in adopting/adapting the CAI. Moreover, it should be born in mind that, as suggested by Kowie and Sakui (2013), employing computer technology in language learning should not be a replacement for the whole learning/teaching practices and processes, but rather it can be a complement facilitating these practices.

References

Abraham, B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199-226.

Abrams, Z. I. (2003). The effect of synchronous and asynchronous CMC on oral performance in German. *The Modern Language Journal, 87*(2), 157-167.

AbuSeileek, A. F. (2008). Hypermedia annotation presentation: Learners' preferences and effect on EFL reading comprehension and vocabulary acquisition. *CALICO Journal*, *25*(2), 260-275.

Anderson, N. (2008). *Active Skills for Reading: Book 4* (2nd ed.). Boston: Thomson, Heinle.

Boers, F., Warren, P., Grimshaw, G., & Slyanova-Chanturia, A. (2017). On the benefits of multimedia annotations for vocabulary uptake from reading. *Computer Assisted Language Learning*, *30*(7), 709-725.

Cariana, R. B., & Lee, D. (2001). The effects of recognition and recall study tasks with feedback in a computer-based vocabulary lesson. *Educational Technology Research & Development, 49*(3), 23-36.

Çakmak, F., & ErÇetin, G. (2018). Effects of gloss type on text recall and incidental vocabulary learning in mobile-assisted L2 listening. *ReCALL*, *30*(1), 1-24.

Chiu, Y. (2013). Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology*, *44*(2), E52-E56.

Eftekhari, M., & Sotoudehnama, E. (2018). Effectiveness of computer-assisted argument mapping for comprehension, recall, and retention, *ReCALL*, *30*(3), 337-354.

Fitzpatrick, N., & Donnelly, R. (2010). *Do You See What I Mean? Computer-Mediated Discourse Analysis.* Dublin: IGI Global.

Frawley, W., & Lantolf, J. P. (1985). Second language discourse: A Vygotskyan perspective. *Applied Linguistics, 6*, 19+44.

Gardner, H. (2011). *Frames of Mind: The Theory of Multiple Intelligences.* New York: Basic Books.

Gorjian, B., Moosavinia, S. R., Ebrahimi Kavari, K., Asgari, P., & Hydarei, A. (2011). The impact of asynchronous computer-assisted language learning approaches on English as a foreign language high and low achievers' vocabulary retention and recall. *Computer Assisted Language Learning, 24*(5), 383+391.

Harley, T. A. (2008). *The Psychology of Language* (Third Edition). New York: Psychology Press.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use and the reoccurrence of unknown words. *The Modern Language Journal, 80*(3), 110-132.

Khezerlou, S., Ellis, R., & Sadeghi, K. (2017). Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *System, 65*, 104-116.

Kılıçkaya, F. (2015) Computer-based grammar instruction in an EFL context: Improving the effectiveness of teaching adverbial clauses. *Computer Assisted Language Learning*, *28*(4), 325-340.

Kowie, N., & Sakui, K. (2013). It's never too late: an overview of e-learning. *ELT Journal, 67*(4), 459-467.

Li, J. (2010). Learning vocabulary via computer-assisted scaffolding for text processing. *Computer Assisted Language Learning*, *23*(3), 253-275.

Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning*, *31*(8), 882-904.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

Mirzaei, A., Rahimi Domakani, M., & Rahimi, S. (2016). Computerized lexis-based instruction in EFL classrooms: Using multi-purpose LexisBOARD to teach L2 vocabulary. *ReCALL, 28*(1), 22-43.

Mohamadi, Z. (2018). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation, 59*, 29-40.

O'Neil, C. A., Fisher, C. A., & Newbold, S. K. (2004). *Developing an Online Course: Best Practices for Nurse Educators.* New York: Springer.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287.

Qin, C. (2012). Didactic considerations of vocabulary breadth and depth in EFL/ESL contexts-a literature review. Retrieved 18 January 2017 from https://gupea.ub.gu.se/bitstream/2077/38777/1/gupea_2077_38777_1.pdf

Rimrott, A. (2010). *Computer-Assisted Vocabulary Learning: Multimedia Annotations, Word Concordances, and Individualized Instruction*. (Doctoral dissertation, Simon Fraser University, Burnaby, British Columbia, Canada). Retrieved 20 February 2017 from http://summit.sfu.ca/item/11876

Rogers, E. M. (1995) *Diffusion of Innovations*. (4th ed.). New York: Free Press.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. London: Palgrave Macmillan.

Su, Y., Li, Y., Liang, J. C., & Tsai, C. C. (2019). Moving literature circles into wiki-based environment: the role of online self-regulation in EFL learners' attitude toward collaborative learning. *Computer Assisted Language Learning*, *32*(5-6), 556-586.

Tsai, K. J. (in press). Corpora and dictionaries as learning aids: Inductive versus deductive approaches to constructing vocabulary knowledge. *Computer Assisted Language Learning*. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/09588221.2018.1527366?journalCode=ncal20

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Cambridge, Massachusetts: Harvard University Press.

Wang, A. (2014). Managing student participation: Teacher strategies in a virtual EFL course. *JALT CALL Journal*, *10*(2), 105-125.

Wang, Y. H. (2016). Promoting contextual vocabulary learning through an adaptive computer-assisted EFL reading system. *Journal of Computer Assisted Learning*, *32*(4), 291-303.

Warner, C. N. (2004). It's just a game, right? Types of play in foreign language CMC. *Language Learning & Technology*, *8*(2), 69-87.

Yoshi, M., & Flaitz, J. (2002) Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO Journal*, *27*(3), 517-528.

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning*, *24*(1), 39-58.

Zapata, G. C., & Sagarra, N. (2007). CALL on hold: The delayed benefits of an online workbook on L2 vocabulary learning. *Computer Assisted Language Learning*, *20*(2), 153-171.

Zhong, H. (2014). *The Interface between Receptive and Productive Vocabulary Knowledge: Vocabulary Knowledge as a Multi-Aspect Construct*. (Doctoral Dissertation, University of Sydney, Sydney, Australia). Retrieved 21 July 2019 from https://ses.library.usyd.edu.au/handle/2123/12472