

The Impact of Computer-Assisted Language Learning Applications on Incidental Vocabulary Recall and Retention

Fatemeh SHOAEI^{1,*} and Mohammad ALAVI²

¹ Department of English Language and Literature, University of Tehran

² Department of English Language and Literature, University of Tehran

This study reports on an investigation into the effect of Computer Assisted Vocabulary Learning (CAVL) on fostering the incidental vocabulary gain of 62 pre-intermediate teenage learners of English studying in a language institute in Iran. These participants were randomly assigned to CAVL and traditional teacher-led groups. Additionally, the two groups' perceptions towards Computer-Assisted Language Learning (CALL) and Computer-Assisted Learning (CAL) were surveyed and compared. A multimedia application integrating contextual cues, frequency of occurrence, dictionary definitions, and textual and audio annotations, was developed uniquely for the purposes of this experimental research. In order to measure and compare the vocabulary recall and retention of the two groups, a pretest-posttest research design was utilized. The results of the immediate posttest demonstrated that the students in the CAVL group significantly outperformed those of the non-CAVL group, confirming superior recall of the new words. Although the latter groups' scores on the delayed posttest were also higher than the former groups' scores, in comparison to the immediate posttest, a considerable decrease in the CAVL groups' vocabulary retention was observed. Moreover, even though both groups perceived CAL as an effective method of learning, the perceptions of the participants in the CAVL group were more positive toward CALL.

Keywords: Computer Assisted Language Learning (CALL), Computer Assisted Vocabulary Learning (CAVL), incidental vocabulary learning, vocabulary recall, vocabulary retention

1. Introduction

A vast body of research has advocated that second language vocabulary is acquired as a by-product of meaning-focused activities such as extensive reading which do not explicitly target vocabulary learning (Huckin & Coady, 1999). In spite of the apparent vocabulary achievements which surface from such activities, upon closer examination it appears that language learners often fail to notice unfamiliar words because they wrongly assume they know the meaning. Therefore, given the inarguable advantage of incidental learning, vocabulary gain proceeds only incrementally and in small quantities. To this end, the question which arises is how this partial increase in word knowledge can be improved. Hulstijn, Hollander, and Greidanus's (1996) seminal study introduced several factors which favourably affect incidental vocabulary learning, including: (a) the presentation of annotations in the reading passage, (b) the use of a dictionary while occupied in the reading task, and (c) the reoccurrence of new words in the text.

On the other hand, with the evolution of multimedia technology it has become possible to integrate different components constructive to incidental learning into a single multimedia application, thus improving language learning practices. The goal of this study was to examine and provide insight into EFL students' performance and perceptions of vocabulary learning while using computer-based technologies as they are presented with reading texts

containing multimedia annotations, contextual cues, frequency of occurrence, and dictionary definitions. The results of the study have helped reveal aspects of the impact that technology has on foreign language learners.

2. Literature Review

2.1 Importance of Vocabulary Knowledge in Foreign Language Learning

Throughout its varying stages the teaching and learning of vocabulary have been undervalued in the field of second language acquisition (SLA) due to the fact that researchers and teachers have prioritized syntax and phonology as more central to linguistic theory and language pedagogy (Zimmerman, 1997). Still and all, there are many scholars who emphasize the role of vocabulary acquisition in learning a foreign language. Among them, Rivers (1981) stated that without learning an adequate amount of vocabulary the learner was not able to produce the structures necessary for successful communication. Krashen (1989) also expressed that much of the meaning of sentences are carried by the words in those sentences and that is exactly why people travelling abroad prefer carrying dictionaries with them rather than grammar books. Formerly, Saragi, Nation, and Meister (1978) had claimed that reading un-simplified texts required knowledge of at least 3,600 word forms by learners of English; and even then there were still several unknown words on each page, thus making vocabulary acquisition of major importance in learning to read.

2.2 Incidental Vocabulary Learning vs. Intentional Vocabulary Learning

During the last three decades there have been many debates among researchers about effective methodologies, strategies, and techniques for vocabulary learning and teaching. One of the major outlooks put forth by a number of researchers such as Hulstijn (2006), Laufer (2009) and Nation (2001) makes a distinction between incidental and intentional learning of vocabulary. Incidental vocabulary learning has been defined by Huckin and Coady (1999) as learning of unknown words from different contexts as a by-product of meaning-focused communicative activities such as reading, listening, and interaction. On the other hand, Intentional vocabulary learning involves the deliberate retention of words through focal attention placed deliberately on linguistic features.

2.3 Computer Assisted Language Learning

Egbert (2005) defined CALL as learners learning language in any context with the aid of computer technologies. An important emphasis in this definition has language learning as its central focus. Developing a qualified CALL program therefore requires a deep theoretical insight into language teaching methodologies. According to Ma and Kelly (2006), the significance of a CALL program is certified by the methodology behind it, not by the computer technology itself. That is, developers of these programs should take into consideration, the nature of language as well as language learning, throughout the design process.

2.3.1 Computer Assisted Vocabulary Learning

Computer-assisted Vocabulary Learning (CAVL) as a subfield of CALL addresses the importance of vocabulary in second language learning. As Kocaman and Cumaoglu (2014) maintain, CAVL is one of the most widely used applications of CALL. Many researchers

have expressed that computer-mediated instruction has had a substantial influence on the development of language learners' vocabulary skills (e.g. Ahmadian, Amerian, & Goodarzi, 2015; Basoz & Cubukcu, 2014; Chen & Li, 2010; Groot, 2000; Kayaoğlu, Akbaş, & Öztürk, 2011; Maftoon, Hamidi, & Sarem, 2012; Mansouri, 2015; Miles & Kwon, 2008).

In a memory research and vocabulary learning study, Miles and Kwon (2008) divided the participants into five groups: (1) CALL only, (2) CALL and practice, (3) class practice, (4) independent study, (5) control group (no instruction). The main feature of the CALL software was that it was devised to recycle the words through spaced repetition, resulting in the outperformance of the first two groups.

In an article about the influence of virtual learning environments and computer-mediated communication on the improvement of lexical competence, Perez Cañado (2010) outlined the development of superior lexical competence over the traditional system of learning.

In the Iranian context, an experiment carried out by Naraghizadeh and Barimani (2013) compared traditional and computer-assisted teaching approaches. The findings reinforced the fact that there is a link between lexical items and different types of media which fosters foreign language learners' recall and retention of words. There was a significant difference between the experimental and control group on their performance on the pretest and posttest, resulting in improvement in the learners' vocabulary acquisition.

2.3.2 Multimedia Annotations and Vocabulary Learning

Based on an evaluation of CAVL software by Ma and Kelly (2006), programs made up of reading texts with electronic annotations are probably the most popular type of multimedia software, reflecting the prevalent interest in incidental learning. Equipped with an electronic dictionary or glossary, these programs emphasize the comprehension and acquisition of new words as a by-product of the reading process. Hyperlinks to online dictionaries with L2 definitions, L1 translations or graphical images provide the meaning of lexical items as well as their correct usage. As a consequence, by utilizing multimedia effects featuring easy access through a simple click, glosses are made more attractive than traditional paper-based methods.

There is also consensus among scholars and educators alike on the value of designing instructional material using various media. In his study in 2015, Wu concluded that by reading texts on multimedia e-book devices students improves not only students' reading comprehension but also their vocabulary knowledge and their understanding of the associations among words.

The results of Svenconis and Kerst's (2013) experiment offered strong evidence about the effectiveness of the multimedia environment. The experimental group worked with a hypermedia program embedded with the sound factor (i.e., word pronunciation) and achieved the highest test scores among all the participants. The outcome indeed proved that a well-crafted multimedia program promotes effective retention and recall of vocabulary items.

Tang's (2011) paper postulated that to ensure the effectiveness of the multimedia-teaching model towards Chinese students' English learning, language instructors and their students should possess operative ability of the multimedia equipment in the teaching process.

Yoshii (2006) examined the influence of different glosses on incidental vocabulary learning. The research particularly focused on the comparison of L1 and L2 glosses, but failed to find any significant differences between the two types. Although it was not specified whether L1 or L2 glosses are better, the results did confirm that glosses as a whole are

beneficial. The study also contrasted text glosses with and without picture glosses, and the advantage of manifesting pictures was supported.

3. Research Questions

- (1) Is there any statistically significant difference between students' vocabulary learning via computer-assisted instruction in comparison to traditional instruction as reflected in the posttest scores?
- (2) Is there any statistically significant difference between students' vocabulary retention via computer-assisted instruction in comparison to traditional instruction as reflected in the delayed posttest scores?
- (3) What are Iranian EFL learners' perceptions toward the integration of CALL into language teaching?

4. Method

4.1 Participants

This study was conducted with the participation of 62 students learning English at a private language institute. The participants were female teenagers who ranged from 13 to 16 years of age. The students were all at the lower intermediate level of language proficiency and their homogeneity was ensured by way of strict examination procedures, namely the KET proficiency test and interviews, before being enrolled in the research project. Also, the selection and assignment of the subjects to the two different learning conditions was done randomly and based on the institutes' usual procedures.

4.2 Instruments

4.2.1 Pilot Study

A number of measurement tools were used for assessing any changes in the learners' vocabulary gain and retention. But before the start of the actual study, a pilot study was conducted in two stages to validate the measurement tools being used in the main experiment. These measurement tools consisted of a vocabulary knowledge scale (VKS) and a multiple-choice vocabulary test constructed using the VKS.

In the first stage, the VKS was administered to a similar group of students (in terms of age and English proficiency level) in the previous semester (which consisted of 100 participants). The content of this test included a list of 110 main words chosen from the short stories which were to be part of the curriculum in the real experiment. By examining the answers provided by each of the 100 participants, a total of 50 of the more difficult words were chosen. The words were then used to design the 50-item multiple-choice which would make up the measurement tools for the main research procedure.

After the design of the multiple-choice vocabulary test it was presented to 58 students (again with similar characteristics of age and language proficiency) studying at the same language institution for the purpose of evaluating test reliability, item difficulty, and item facility, in addition to estimating the validity of the test. Additionally, a panel of experts in the field of language teaching was consulted. The experts rated each of the questions for clarity and depending on their opinions recommended changes were made.

4.2.2 Pilot Study

Consequently, the instruments for the main study were established based on the results of the pilot study, which consisted of a pretest, immediate posttest, delayed posttest, and a seven-point Likert scale questionnaire.

As for the pedagogical materials, two different modes of instructional content in the form of CD-ROM and hard-copy were integrated into the students' course work.

4.3 Procedure

Four randomized classes with the level of lower intermediate were manipulated for the purpose of this study. Two of the classes were randomly assigned to receive the conventional method of instruction while the other two classes were instructed by means of the computer-mediated method. The classes met two times a week for 90 minutes during an 11 week period consisting of 21 sessions. The first and 19th sessions were allocated, respectively, to the pretest and posttest, while the delayed posttest was held on the last day of the course (or the 21st session). All the participants also answered a questionnaire distributed on session 20 as a means of conveying how they perceived the use of technology in the classroom. Therefore, a total of 18 sessions were left for instructional purposes. Accordingly, 17 reading texts out of the 30 short stories provided in the book were covered in each of the remaining sessions. The teaching content and activities for both groups were the same, the only variation being the medium of instruction. The CAVL group was provided with dictionary definitions and translations in the form of multimedia annotations. The students were asked to read the stories on their own using their personal computers, while at the same time being appointed a headphone for listening to the audio file. Then, they were required to complete the end-of-unit activities and email them for their teacher before the next session of class. The control group, on the other hand, simply followed the reading text in silence as the audio was played for them. They were then given a chance to read the story in pairs and find the definition of unknown words with the aid of either a dictionary or their instructor and peers. At the end of each session the students were asked to complete the activities of that unit as homework which was checked before the start of the new lesson in the following session.

5. Results

5.1 Validation of the Vocabulary Test

Prior to the main study, a pilot study was done in order to validate the test. The test was given to 58 students. The Rasch model was employed for validation. The main purpose of the Rasch analysis was to check the following:

- a. Test dimensionality
- b. Item functioning
- c. Score reliability

The results of a Principal Components Analysis on the measures obtained from the Rasch analysis revealed that the construct underlying the test (presumably, vocabulary knowledge) explained over 22 percent of the variance. The first nuisance factor or secondary dimension explained less than 6 percent of the data. Hence, it may be concluded that only one factor was underlying the test. Also, the inspection of the pattern of loadings on the secondary dimension revealed no meaningful collection of the items around this factor. Therefore, it appears that the test measures only one factor.

The results of the item functioning analysis show if the item functions in the way predicted by the Rasch model. According to Bond and Fox (2007), indices between .70 and 1.30 indicate good fit. Based on their criteria, all items fit the Rasch model.

The reliability of the test turned out to be .87. Considering the small number of participants in the pilot study, it seems that the reliability is high enough.

All in all, it is clear that the test is measuring a single construct, the scores are highly reliable, and the items function as expected.

5.2 Results of the Vocabulary Pretest

In order to make sure that the two groups were at the same level, an independent-samples t-test was run. The results of the t-test shown in Table 1 confirmed that the mean-difference is not significant. Hence, it is concluded that the groups are almost at the same level of knowledge with respect to the vocabulary included in the pretest.

Table 1: Independent-samples t-test for the pretest

	Levene's Test for Equality of Variances		T-Test for Equality of Means		
	F	Sig.	T	df	Sig. (2-tailed)
Equal variances assumed	.554	.459	.189	60	.850
Equal variances not assumed			.189	59.502	.850

5.3 Results of the Immediate and Delayed Vocabulary Posttests

In order to answer the first and second research questions, a mixed between-within subjects ANOVA was run. The first step in inspecting the results of the mixed ANOVA is to check the assumptions. One of the most important assumptions is the sphericity assumption. The assumption “requires that the variance of the population difference scores for any two conditions are the same as the variance of the population difference scores for any other two conditions” (Pallant, 2011, p. 262). The assumption was checked with the help of Mauchly’s Test of Sphericity. The results of the Mauchly’s test revealed that the data violates the assumption.

Whenever the sphericity assumption is violated, it is safer to interpret the multivariate statistics in the SPSS output. This is because the multivariate tests do not require the sphericity assumption. The results of the multivariate tests are displayed in Table 2. Note that only Wilks' Lambda is displayed in the table because other tests had similar results.

Table 2: Multivariate results

	Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Wilks' Lambda	.050	556.278	2.000	59.000	.000	.950
Time * Group	Wilks' Lambda	.124	209.263	2.000	59.000	.000	.876

It appears from the table that main effect for Time is significant which means that the time of the test has had a significant effect on the scores. The interaction effect is also

significant which means that although the Time effect is significant, its effect is not the same across levels of the grouping factor. Specifically, the time effect is not the same for the CAVL and non-CAVL groups. Note also that the effect size estimates for main and interaction terms are both very large. Considering Cohen’s (1988) guidelines for interpreting the effect size estimates, these denote large effect sizes.

The main effect for the grouping factor (i.e., CAVL vs. non-CAVL) also turned out to be significant. The results are displayed in Table 3. Note that the effect size is again large.

Table 3: The Group effect

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	177446.371	1	177446.371	6348.992	.000	.991
Group	4249.038	1	4249.038	152.030	.000	.717
Error	1676.925	60	27.949			

The results are also graphically displayed in Figure 1. It is now more clearly observed that there was an increase in the scores from pretest to posttest in both groups. However, the delayed posttest scores are lower than the posttest scores. The pattern of changes is similar across the CAVL and non-CAVL groups. In addition, it is also observed that the CAVL group outperformed the non-CAVL group on both the posttest and the delayed posttest.

So far, the results of the mixed ANOVA show the general pattern of change in the scores. In order to make more specific comparisons across the means, a series of t-tests were run. It should be noted that when a large number of comparisons are made, the type-one error rate increases. Hence, Bonferroni correction was applied to render the results dependable.

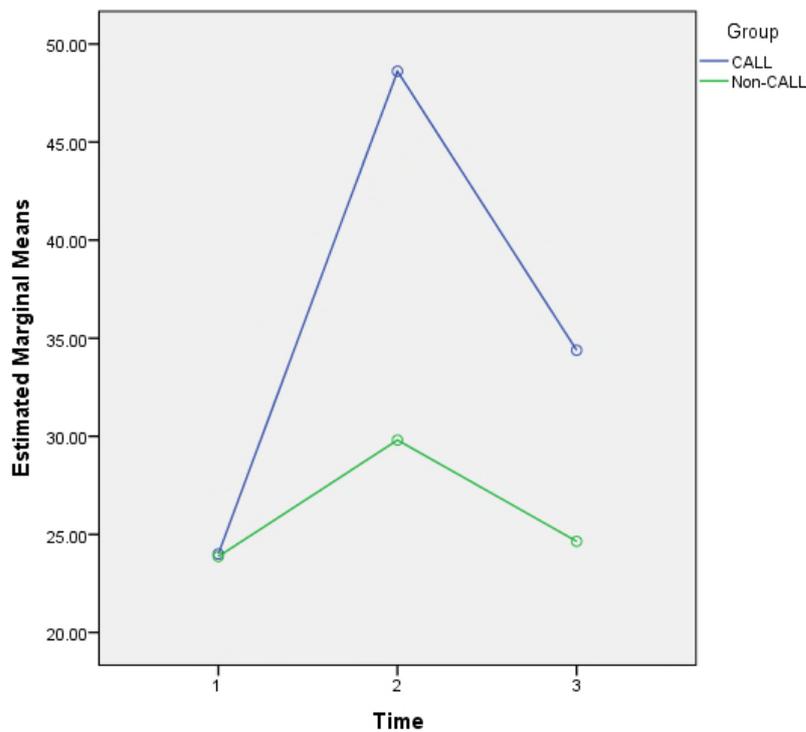


Fig. 1: The Means Across Groups and Time

The results of the t-tests between the pretest and posttest scores are displayed in Table 4. From the first row, it appears that the CAVL group has had a very large increase from the pretest to the posttest. The mean-difference is statistically significant and the effect size is very large. The second row shows that the non-CAVL group has also improved their scores from the pretest to the posttest. However, the increase in the CAVL group scores is more than four times the increase in the scores in the non-CAVL group. In addition, after Bonferroni correction is applied, both differences are still significant.

Table 4: T-test results for the pretest and immediate posttest

Group	Paired Differences			t	df	Sig.(2-tailed)	Eta Squared
	Mean	Standard Deviation	Standard Error Mean				
CAVL	-24.61	3.02	.54	-45.23	30	.000	.98
Non- CAVL	-5.93	4.04	.72	-8.16	30	.000	.68

The comparison between the pretest and delayed posttest scores was done using two paired-samples t-tests. The results appear in Table 5. Remember that both groups (i.e., CAVL and non-CAVL) observed shrinkage in the scores after the posttest. The first row in Table 3 shows that although there was shrinkage in the scores from the posttest to the delayed posttest, the scores that the CAVL group obtained in the delayed posttest is still much higher than the scores they obtained in the pretest. That is, not only did the treatment have a significant effect on vocabulary learning in the CAVL group, it also had a significant effect on vocabulary retention.

On the other hand, the mean-difference between the pretest and delayed posttest scores in the non-CAVL group is not large enough. The mean-difference is only -0.77 (less than a single item). Although the difference is shown to be significant, it is not significant after Bonferroni correction is applied. Also, the Eta Squared shows a moderate effect size. Considering the sample size in the present study, there seems to be little difference between the pretest and delayed posttest scores in the non-CAVL group. Hence, it is observed that although the non-CAVL displayed some learning after the treatment, the vocabulary is not retained for a long time.

Table 5: T-test results for the pretest and delayed posttest

Group	Paired Differences			t	df	Sig.(2-tailed)	Eta Squared
	Mean	Standard Deviation	Standard Error Mean				
CAVL	-10.38	7.95	1.42	-7.268	30	.000	.64
Non- CAVL	-.774	1.99	.35	-2.160	30	.039	.13

5.3 Perception of the Participants toward CAL and CALL

The questionnaire was administered to ascertain whether or not there exists any relationship between the participants' scores and their perceptions toward the use of computers for study purposes. Moreover, the relationship between the CAVL and non-CAVL groups was also measured. In order to see which of the differences are significant, six independent-samples t-tests were run. Again, in order to remove the increased type-one error, Bonferroni correction was applied.

It appears from the Table 6 that the CAVL and non-CAVL groups had statistically significant mean-differences in only two of the components: Effectiveness of CALL vs. non-CALL and Computer Proficiency. Although the mean-difference for “Computer Integration” is also marginally significant, it is no more significant after Bonferroni correction is applied.

Table 6: The results of independent-samples t-tests

	t-test for Equality of Means				
	t	df	Sig.(2-tailed)	Mean Difference	Standard Deviation
Effectiveness of CALL vs. non-CALL	4.493	48.769	.000	6.00	1.335
Teacher influence	1.361	60	.179	1.16	.853
Degree of exhibition to CALL	.510	60	.612	.67	1.327
Surplus Value of CALL	1.095	60	.154	1.25	1.316
Computer Proficiency	4.015	46.306	.000	6.54	1.631
Computer Integration	2.310	60	.024	3.25	1.410

6. Discussion

The purpose of the present study was to investigate the effect of computer-mediated reading of short stories on the incidental learning of vocabulary while at the same time obtaining insight into the perceptions of students towards CALL.

6.1 Effect of Learning Condition on Immediate Vocabulary Recall

The first research question concentrated on the impact of computer-assisted instruction on vocabulary learning. Overall, the results proved that technological support in the foreign language classroom is indeed facilitative toward vocabulary learning, in addition to confirming the previous research findings of CALL researchers such as Chen and Li (2010), Maftoon, Hamidi, and Sarem (2012), and Miles and Kwon (2008).

On the grounds that this study used a software application with different combinations of multimedia annotations arranged as text and audio for student’s vocabulary enhancement, Mayer’s (2005) cognitive theory of multimedia learning can be used to explain this end result. The dual-channel assumption, as one of the components of this theory, asserts that working memory has auditory/verbal and visual/pictorial channels. That is, the visual channel processes information displayed as on-screen text, pictures, or video; while the auditory channel processes sounds (such as narration or non-verbal sounds). Therefore, by making use of both visual and auditory channels, working memory capacity is expanded. Further proof for the effectiveness of multimedia instruction can be put forward through the contiguity and modality principles (Moreno & Mayer, 1999) underlying multimedia learning theory. Mayer and Anderson’s (1992) contiguity principle proposed that the capability of multimedia instruction increases when visual and spoken materials are presented synchronously rather than successively. According to Penney’s (1975) modality principle, auditorily presented information result in higher recall in short-term memory than data displayed visually (Penney, 1989). Considering the theoretical underpinnings set forth by the aforementioned scholars, the multimedia software, which was developed for the purposes of this study, employed a simultaneous combination of on-screen text and narration. As a result, the consumption of both acoustic and visual processing resources brought about effective growth in the learners’ working memory capacity.

6.2 Effect of Learning Condition on Long-term Vocabulary Retention

Another question which prompted the study was whether or not long-term retention of vocabulary through computer-assisted instruction would show superior results in comparison to traditional teacher-led instruction.

All things considered, the fact that the long-term retention rate of these students was half the immediate recall rate cannot be ignored. This leaves us with the question of why students in this study remembered a much smaller proportion of the words they had previously encountered. Before anything else, the spacing effect, a psychological phenomenon dating back to the late 1800s (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006), is set forth. The spacing effect states that learning is enhanced when a given item is studied through spaced presentations instead of massed presentations (Ellis, 1995). In other words, studying a to-be-learned item infrequently over a longer span of time rather than repeatedly studying it in a short time span, yields substantially better learning. In a study comparing spacing and massed conditions, Bahrck and Hall (2005) demonstrated that by creating longer intervals between practice sessions in the latter group, the participants generated superior retention on the final test trial. Other similar studies (such as, Glenberg & Lehmann, 1980; Goossens, Camp, Verkoeijen, Tabbers, & Zwaan, 2012) have found comparable results in favor of the spacing effect as an efficient memory strategy for retaining verbal information. The current study was only designed to look into the effectiveness of computer-assisted vocabulary learning without considering any temporal distribution of practice. Accordingly, by adding a third learning condition to this research design, the effect of spaced repetition of verbal tasks could have also been a factor under close scrutiny which might have led to higher retention rate in this added group of participants.

6.3 Effect of Learning Condition on Students' Perceptions

In general, the perceptions of the students toward computer-assisted learning differed among the CAVL and non-CAVL group. That is the students from the latter class perceived computer integration as being more effective toward their language learning. Also, they perceived themselves as being more proficient in using computers. This could be due to the fact that, as the students engaged in computer-mediated learning throughout the course, they became more motivated and more interested in learning, thus having increased confidence in CAL. It should also be noted that students evolve into autonomous learners in technology-enhanced environments and become more responsible for their own learning process. The teacher's role is more of a facilitator and guide which accommodate students in enhancing the use of software applications in the process of language learning.

In particular, the students in the CAVL group perceived computer-assisted language learning to be more effective than the students in the non-CAVL group. Again, being engaged in the computer-mediated activities had a positive effect on students learning which in turn positively affected their perception toward CALL material. This positive attitude can be attributed to the language learners' gained confidence in their ability to use the multimedia application successfully. Also, younger students usually find computer programs enjoyable and interesting; therefore, the integration of multimedia applications in language classes will definitely increase their motivation and enthusiasm toward learning a foreign language.

7. Conclusion

The results of the research, as measured by immediate and delayed posttests, indicated that the computer-assisted group benefited from the language software. This finding proposes that multimedia-enhanced language learning applications promote incidental vocabulary learning and retention.

When examining the retention of the target words over a two-week period, the results of paired-samples t-tests comparing the posttest and delayed posttest scores indicated that there were significant differences between the two groups with the computer-assisted group showing significant retention rates, while the teacher-led group forgot the learned words soon after.

The students' perceptions towards computer-assisted language learning also correlated with their method of instruction, as the participants in the computer-assisted group perceived computer-assisted language learning to be more effective. Although, the teacher-led group also held a positive attitude toward computer-assisted learning in general, they did not perceive CALL as beneficial toward language gain.

References

- Ahmadian, M., Amerian, M., & Goodarzi, A. (2015). A comparative study of paper-based and computer-based contextualization in vocabulary learning of EFL students. *Advances in Language and Literary Studies*, 6(2), 96-102.
- Bahrick, H. P., & Hall, L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language*, 52(4), 566-577.
- Basoz, T., & Cubukcu, F. (2014). The effectiveness of computer assisted instruction on vocabulary achievement. Retrieved from http://mije.mevlana.edu.tr/archieve/issue_4_1/4_mije_13_77_volume_4_issue_1_page_44_54_PDF.pdf
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological bulletin*, 132(3), 354.
- Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, 18(4), 341-364.
- Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Lawrence Erlbaum Associates.
- Egbert, J. (2005). Conducting Research on CALL. In: Egbert, J. and Petrie, G. M. (Eds.), *CALL research perspectives* (pp. 3-8). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ellis, N. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. *Computer Assisted Language Learning*, 8, 103-128.
- Glenberg, A. M., & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory & Cognition*, 8(6), 528-538.
- Goossens, N. A., Camp, G., Verkoeijen, P. P., Tabbers, H. K., & Zwaan, R. A. (2012). Spreading the words: A spacing effect in vocabulary learning. *Journal of Cognitive Psychology*, 24(8), 965-971.
- Groot, P. J. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology*, 4(1), 60-81.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in second language acquisition*, 21(02), 181-193.
- Hulstijn, J. H. (2006). Psycholinguistic perspectives on second language acquisition. In J. Cummins & C. Davison (Eds.), *The international handbook on English language teaching* (pp. 701-713). New York, NY: Springer Science & Business Media.

- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327-339.
- Kayaoğlu, M. N., Dag Akbaş, R., & Öztürk, Z. (2011). A Small Scale Experimental Study: Using Animations to Learn Vocabulary. *Turkish Online Journal of Educational Technology*, 10(2), 24-30.
- Kocaman, O., & Cumaoglu, G. K. (2014). The Effect of Educational Software (DENIS) and Games on Vocabulary Learning Strategies and Achievement. *Education & Science/Egitim ve Bilim*, 39(176), 305-316.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73, 440-464.
- Laufer, B. (2009). Second language vocabulary acquisition from language input and from form-focused activities. *Language Teaching*, 42(3), 341-354.
- Ma, Q., & Kelly, P. (2006). Computer assisted vocabulary learning: Design and evaluation. *Computer-Assisted Language Learning*, 19(1), 15-45.
- Maftoon, P., Hamidi, H., & Sarem, S. N. (2012). The effects of CALL on vocabulary learning: A case of Iranian intermediate EFL learners. *Broad Research in Artificial Intelligence and Neuroscience*, 3(4), 19-30.
- Mansouri, V. (2015). Vocabulary Instruction: Software Flashcards vs. Word Clouds. *Advances in Language and Literary Studies*, 6(1), 41-45.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge: Cambridge University Press.
- Miles, S., & Kwon, C. (2008). Benefits of using CALL vocabulary programs to provide systematic word recycling. *English Teaching*, 63(1), 199-212.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of educational psychology*, 91(2), 358-368.
- Naraghizadeh, M., & Barimani, S. (2013). The effect of CALL on the vocabulary learning of Iranian EFL learners. *Journal of Academic and Applied Studies*, 3(8), 1-12.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. Retrieved from <http://dx.doi.org/10.1017/CBO9781139524759>
- Pallant, J. (2011). *SPSS Survival Manual 4th edition: A step by step guide to data analysis using SPSS version 18*. Maidenhead, Berkshire: Open University Press.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, 17(4), 398-422.
- Perez Cañado, M. L. (2010). Using virtual learning environments and computer-mediated communication to enhance the lexical competence of pre-service English teachers: a quantitative and qualitative study. *Computer Assisted Language Learning*, 23(2), 129-150.
- Rivers, W. M. (1981). *Teaching foreign language skills*. Chicago: Chicago University Press.
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6(2), 72-78.
- Svenconis, D. J., & Kerst, S. (2013). Investigating the teaching of second-language vocabulary through semantic mapping in a hypertext environment. *CALICO Journal*, 12(2&3), 33-57.
- Tang, X. (2011). Study on the Application of Multimedia in English Language Teaching. In S. Lin & X. Huang (Eds.), *Advances in Computer Science, Environment, Ecoinformatics, and Education* (pp. 293-297). New York, NY: Springer Science & Business Media.

- Wu, T. T. (2015). A learning log analysis of an English-reading e-book system combined with a guidance mechanism. *Interactive Learning Environments*, (ahead-of-print), 1-19.
- Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, 10(3), 85-101.
- Zimmerman, C. (1997). Historical trend in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (5-19). Cambridge: Cambridge University.