



Readers, Players, and Watchers: Short and Long-term Vocabulary Retention through Digital Video Games

Mohsen Ebrahimzadeh (Corresponding author)

Department of Foreign Languages and Linguistics, Shiraz University, Fars, Iran

E-mail: s.mebrahimzade@rose.shirazu.ac.ir

Sepideh Alavi

Department of Foreign Languages and Linguistics, Shiraz University, Fars, Iran

E-mail: salavi@rose.shirazu.ac.ir

Received: 05-11-2016

Accepted: 11-01-2017

Advance Access Published: April 2017

Published: 01-07-2017

doi:10.7575/aiac.ijalel.v.6n.4p.52

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.6n.4p.52>

Abstract

The study investigated short- and long-term vocabulary retention through a digital video game. Participants were 241 male students (age 12–18) randomly assigned to three treatments, namely, Readers who learned vocabulary through intensive reading, Players who learned vocabulary through playing a digital video game, and Watchers who learned by watching two classmates play the digital video game. Twenty one words (simple and compound nouns) were pretested and then followed by the interventions which lasted for five weeks. The same pretest was posttested twice: As a one month and a three months delayed posttest after the study to measure short- and long-term vocabulary retention respectively. Results indicated that the Players and Watchers outperformed the Readers in terms of both short- and long-term vocabulary retention. The Readers were the only group to forget a significant number of words on the three-months delayed posttest. Only the Players' treatment was equally effective for learning both simple and compound nouns. The Players showed the highest mean score on both posttests and the least drop from the short- to long-term posttest. It is concluded that digital video games can help students retain vocabulary over a longer time period.

Keywords: digital video games, game-based language learning, vocabulary retention, vocabulary learning, reading-based vocabulary learning

1. Introduction

1.1 Language learning through Digital Video Games (DVGs)

During the past few decades, researchers have shown a growing interest in the prospect of DVGs used for language learning (Kettemann, 1995; Sylvén & Sundqvist, 2012) especially with young learners since DVGs are quite popular among them (Prensky, 2007; Sylvén & Sundqvist, 2012). Game-based learning is defined as 'any initiative that combines or mixes video games and education' (Tsai & Fan, 2013, p. 115). Reinhardt and Sykes (2012) distinguished between game-based and game-enhanced language learning. They called the use of an educational DVG (i.e., a DVG specifically developed for language learning purposes) a game-based approach. Game-enhanced learning, on the other hand, refers to the use of commercial-off-the-shelf DVGs for language learning purposes.

Reinhardt and Sykes (2012) hold that commercial DVGs have advantages that educational games lack. For example, massively multiplayer online DVGs offer contact with a huge number of native speakers. Also, being commercial, they would probably be seen as genuine activities by the learner. Lastly, these DVGs are cultural products and thus incorporate narratives that promote not only language learning but also critical cultural awareness. Thus, it becomes important to investigate the use of commercial DVGs as supplemental activities, since a uniquely game-based approach might convey a less important role for the teacher while the multiplicity of instruments should in fact help the practice become more effective (Reinhardt & Sykes, 2012).

Interacting with DVGs, gamers go through many action-reflection loops (Buckley & Anderson, 2006) where learners construct meaning through activities they conceive meaningful and important by going through a series of trial and errors. To show how games work, Ang and Zaphiris (2008) compared them to other instructional media/instruments. Based on them, texts only provide clues as to what a concept or object is but pictures depict it at a certain point in time. Videos offer more information on how the subject changes through time. A toy car, however, provides still more information by modelling a real car's behavior (e.g., how it moves, gas, brakes).

DVGs are a kind of multimedia that refers to presentation of spoken or written words accompanied by static or dynamic images. Multimedia learning is defined as 'people build[ing] mental representations from words ... and pictures' (Mayer, 2014, p. 3). Thus, multimedia learning involves learners constructing knowledge from this representation. Multimedia instruction, therefore, would be the design of multimedia learning environments that promote this knowledge construction. For example, researchers have found that learners do better on retention tests if words and pictures are presented together (Mayer & Anderson, 1991, 1992).

According to Gee (2003) there are 36 learning principles at work in good games such as self-knowledge, amplification of input, achievement, and practice principle, to name a few. The practice principle, for example, states that ‘learners get lots and lots of practice in a context where the practice is not boring (i.e., in a virtual world that is compelling to learners on their own terms and where the learners experience ongoing success). They spend lots of time on task’ (Gee, 2003, p. 208).

1.2 Vocabulary retention and digital video games

Knowledge of a word over a period of time or retention is of pivotal importance (Min, 2008). According to Nation (2001) successful vocabulary retention depends on three processes, namely, noticing, retrieval, and creative use. Noticing is a private experience necessary for converting input into intake. Retrieving a word from memory for successive use can result in better retention. Creative use refers to the utilization of learned vocabulary in novel contexts/ways such as through retelling tasks.

DVGs are media with inherent vocabulary learning potentials. Gamers often have to interact with items in order to make progress. In other words, they constantly reuse and analyze these items to discover their characteristics and abilities. As such, gamers would have repeated interactions with these items in new contexts. These items have textual (names and explanations) as well as visual representation (thumbnails, color codes, and visual effects). Therefore, multimedia presentation may help the learners strengthen what they code (Hulstijn, 2001) into their memory regarding the meaning of a specific item.

The impact of a commercial DVG on vocabulary retention was investigated by Hitosugi, Schmidt, and Hayashi (2014). Participants of the study were 20 EFL Japanese students with average-high and intermediate-low proficiency in English. The study included two separate treatments (five sessions). Task sheets with embedded new vocabulary and the DVG as a supplementary activity were used in the first treatment. This treatment involved no immediate posttests and teaching of the new vocabulary. In the second treatment, target vocabulary was explicitly introduced, immediate posttests were administered, and the DVG scores comprised part of the course grade. Additionally, 25 textbook words were included in this treatment. Results showed that after five weeks, DVG-based vocabulary retention was the same as the immediate posttest. A significant number of textbook vocabulary items, however, were forgotten after five weeks. The authors concluded that DVGs promote deep learning.

Smith et al. (2013) investigated vocabulary retention through DVGs embedded in eBooks compared to hardcopy booklets. Participants of the study were 57 male and female EFL undergraduate, intermediate, Chinese students. Data was collected through a vocabulary pre- and posttest. The participants underwent both interventions in a counterbalanced manner. The game-mediated intervention was found to yield better results.

1.3 Reading comprehension modes and vocabulary learning

Vocabulary learning through readings involves direct (e.g., glosses, mnemonic devices, lexicography) and indirect (e.g., inferring meaning from context, using background knowledge, making predictions) techniques which are employed in two reading comprehension modes, namely, intensive and extensive reading. Intensive reading promotes direct teaching of vocabulary, usually through a passage preceded by pre-reading and followed by post-reading activities. Extensive reading involves reading lots of books without the aim of doing post-reading activities. Extensive reading texts should be strictly controlled for the amount of unfamiliar vocabulary items to enable the reader to guess the meanings.

Both modes have attracted criticism regarding the amount of vocabulary that can be acquired through each. Nation and Waring (1997), for example, reasoned that intensive reading fails to fully prepare ESL/EFL learners. On the other hand, the right amount of new vocabulary and number of repetitions for example, are issues debated on regarding extensive reading. To compare, some empirical evidence indicates that the extensive mode is relatively ineffective when measured against the intensive mode. Min (2008), for example, looked into this issue by examining vocabulary retention among EFL, secondary school, Chinese students. Participants were divided into two groups with one working on readings with vocabulary activities while the other included selected texts plus thematically related materials without vocabulary activities. The study introduced the same target vocabulary items to both groups in five weeks. Through a posttest, the results showed the treatment with vocabulary-focused activities to be more effective for retention.

1.4 Focus and questions of the study

Game-mediated learning proponents have argued that DVGs are a suitable medium regarding educational purposes (e.g., Gee, 2003; Gibson, Aldrich, & Prensky, 2006) which has raised the interest of researchers in the area of second language acquisition as well (Reinhardt & Sykes, 2012). However, few empirical studies have been conducted that relate gaming experience to second language development (Cornillie, Thorne, & Desmet, 2012) and the major part of existing literature consists of studies conducted outside the classroom (Thomas, 2012). Thus, the present study evaluated short- and long-term vocabulary retention through DVGs in high schools to address this gap. In a larger scale, whether the growing DVG industry could be used in service of language learning is the idea focused in this paper.

By short-term retention we mean the amount of vocabulary content retained over the course of a month, and by long-term retention, we refer to vocabulary retention after three months. Thus, we administered the same vocabulary posttest twice: Once after one and again after three months’ delay. Since Iranian high schools mostly lack sufficient equipment, we investigated vocabulary retention through DVGs in two approaches. The first approach, Players, provided all students with computers to play the game. In the second approach, Watchers, two students played and the rest watched. The following questions were investigated:

1. How do DVGs affect high school EFL students’ short- and long-term vocabulary retention?
2. How does playing individually affect short- and long-term vocabulary retention as compared to watching others play a DVG?

2. Method

2.1 Participants

Chosen through cluster sampling, participants were 241 Iranian, male, junior and senior high school students (age 12–18) from one junior ($N = 61$) and two senior ($N = 153$) high schools in Shiraz. The students were not familiar with game-enhanced language learning. Some of them attended private language institutes but the majority studied English only in school. Their textbooks did not include educational games or multimedia content except for an audio CD. Senior high school classes were randomly assigned to one of the three treatments (Readers: $N = 75$, Players: $N = 65$, Watchers: $N = 74$) but junior high school classes were limited to either Players' or Watchers' treatment because they did not meet the language background requirements explained by the Cambridge English Language Assessment (2013) syllabus used as the base of the Readers' treatment. The syllabus offers a comprehensive approach to testing English proficiency of learners in high schools and reaches up to the Common European Framework of Reference A2 level. The Headway placement test published by Oxford University Press in 2012 classified the majority of the participants (87.9 %) as A1. Twenty five students were removed from final analyses, because they had either (a) played the game at home; (b) cheated in the exams; or (c) missed more than one session.

Prior to the study, permissions to conduct the investigation were acquired from Ministry of Education and high school officials. Also, this study was approved by the research ethics committee of Shiraz University. Students were informed of the scientific nature of this investigation and participation was voluntary. They did not know about its purpose, however, to avoid sensitization of any kind toward the target vocabulary. Those who chose not to participate, received a worksheet based on their course book to practice.

2.2 Materials and instruments

All materials/instruments were piloted four months before the study for observation and revision purposes. Participants of the pilot study were similar in English proficiency, age, and gender to those of this study.

2.2.1 Target vocabulary items

Twenty one simple and compound nouns (Appendix 1) were extracted from a map called Defense of the Ancients released by IceFrog (2015) for the commercial DVG Warcraft III: The Frozen Throne (Blizzard, 2003). Word selection was based on several criteria; for example, some words were required for the gameplay. Others were items that needed to be purchased using gold pieces (the DVG's currency); we had to choose items that required a manageable amount of gold pieces considering class time. Additionally, items with thumbnails more expressive of their meaning were prioritized. Moreover, since these items affected the avatars, their effects had to be immediately observable in order for the students to be able to make educated guesses on their Persian equivalent. Word length, thus, was affected by these criteria.

The DVG dictated the presentation order of the items. Each word was repeated four times during its respective session for the Readers (oral repetitions excluded) since it was tried to avoid raising student sensitivity. These words were mostly unknown to the participants because they (a) were not among high-frequency words and (b) did not exist in students' English textbooks. Pretesting the word list also supported this assumption (Table 1).

2.2.2 Worksheets and reading passages

We developed five reading passages (Appendix 2), each including 600–650 words, through which the vocabulary items were introduced to Readers. The DVG's plot served as the base of these passages. Like Players and Watchers, Readers had 50 minutes to go over a worksheet including one reading passage preceded and followed by several activities. Since students were mostly at the A1 level, the readings were developed based on the A2 level of Cambridge English Language Assessment (2013) syllabus to have the text flow one step above the learners' proficiency conforming to Krashen's (1982) $i+1$ comprehensible input hypothesis.

All readings were written based on the extensive reading criterion stating that above 95 % of the text should be familiar to the reader (Hsueh-Chao & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). To achieve this, a word-list was developed based on the headwords introduced by the Cambridge English Language Assessment (2013) syllabus (the Flyers' stage). This word-list was programmed to work with the Range (Nation, 2002). Each reading was analyzed by this software to ensure good statistics regarding the amount of new words, number of times they were repeated, and text length.

Worksheets started with pre-reading activities such as brainstorming and prior knowledge activation. They ended in post-reading activities including guessing from context, fill in the blanks, matching, multiple-choice questions, comprehension check, summary writing, and a word puzzle. Such activities were included based on the idea that readings combined with enhancement activities result in better vocabulary learning (Min, 2008; Paribakht & Wesche, 1997). Repetition and translation activities were carried out orally to encourage student participation. The readings were meaning-focused input activities (Nation, 2009) in that they represented a story unfolding session by session. The meaning of new words were glossed in English in the right margin as simplified definitions to represent an element of language-focused learning (Nation, 2009). They were also orally explained in Persian, the students' mother tongue. Fill in the blanks and inferential activities were added to promote meaning-focused output (Nation, 2009). Also, each worksheet included images extracted from the game that helped understand the target vocabulary either by depicting it or through visualization of the concept of that item.

2.2.3 The vocabulary pretest and short-/long-term retention posttests

This test included 21 multiple-choice (four alternatives) vocabulary items (Appendix 3). It was administered to all students before the beginning of the course. Pretest results showed that students were not familiar with the target words (Table 1). This test was administered twice after the study as a one- and three-months delayed posttest referred to as short- and long-term posttests respectively. It took the students less than 10 minutes to return the papers.

2.2.4 The digital video game

The study used a Real-Time Strategy video game called *Warcraft III: The Frozen Throne* (Blizzard, 2003). In this genre, gamers position and maneuver units and structures under their control to secure areas of the map and/or destroy their opponents' assets (Rollings & Adams, 2003). Choosing this commercial DVG was based on the learning opportunities it offered, suitability, GameFlow criteria (Sweetser & Wyeth, 2005), and technical implementation issues (hardware, software, and gameplay-training requirements).

The learning opportunities offered in this DVG can be based on noticing, repetition, and generative use (Nation, 2001) while also adhering to Nunan's (1999) presentation, practice, production model. Through an overhead projector, first, steps for obtaining certain items were shown to Players/Watchers (noticing and presentation). Next, using their backpacks, the students carried their items around and examined/used them whenever they wished (repetition and practice). The DVG, finally, allowed manipulation of owned items to create other items (generative use and production). Other than player-game interaction, the game offered collaborative and competitive player-player interaction due to its multiplayer nature. The DVG also provided feedback which is in line with educational implications of the socio-cultural theory (Vygotsky, 1978) and GameFlow criteria (Sweetser & Wyeth, 2005).

Regarding suitability, according to the Entertainment Software Rating Board and Pan European Game Information, the DVG is suitable for users 12 years old and above. Additionally, ign.com and gamefaqs.com gave it very high popularity scores (9 out of 10 and 88 out of 100 respectively) suggesting its attractiveness.

2.2.5 Game pictures, cinematics, and cutscenes

Relevant images from the DVG were used to highlight the target items and steps to create them for both Players and Watchers. Each session started with a video for the Readers to help them visualize the specific part of the story they were going to read about. The Players/Watchers also watched the same videos to gain a sense of purpose and awareness regarding what was going on in the game or whereabouts of the characters and their motives. All videos had hardcoded Persian subtitles to help student comprehension.

2.3 Procedure

The Headway placement test and the vocabulary pretest were administered two weeks before the study. The experiment continued for five sessions—one session a week lasting 50 minutes. Each session, 3–6 vocabulary items—out of the total 21—were introduced through all interventions (the same items were introduced to all groups as they proceeded). The students were not aware of the two posttests. Each session, students of all classes, except for the Watchers, were randomly divided into several groups. We tried to include five students in each group, a number dictated by the DVG, but sometimes due to different student numbers there were groups of four. The Watchers were always divided into two groups and one student from each group was randomly selected to play. The following subsections present in depth explanations of each group.

2.3.1 The Readers' treatment

The Readers started with watching a video from the game that depicted the part of the story they were going to read about. Their worksheets started with pre-reading activities (brainstorming and activating prior knowledge). Next, each passage was intensively read and translated into Persian by the researcher. To help structure their knowledge and deepen their cognitive processing (Fincher-Kiefer, Post, Greene, & Voss, 1988), students were asked to retell parts of the story in their own words. This also allowed for the production of comprehensible output which is necessary for successful language learning (Swain, 1985; Swain & Lapkin, 1995). Students were allowed to use both English and Persian because of their low proficiency. At the end of each reading, post-reading activities were completed in the class. These activities required them to go through the reading again and again therefore increasing the chance of encountering the target vocabulary items. All this was done in groups. Group members were randomly selected each session to avoid ordering effects. As the students went through their worksheets, they could ask for guidance from the researcher as well.

2.3.2 The Players' and Watchers' treatment

These two groups too watched the same videos like the Readers at the outset of each session. They were trained on how to interact with the DVG beforehand. When starting a match, two groups of five (or less) joined a local server and began playing. Similar to the Readers, each session lasted about 50 minutes. Instructions on finding/creating the items were presented through an overhead projector. For example, an image of the DVG screen was captured; then, numbered highlights were added to it in order to show the steps for creating/buying an item. These guides were shown through the projector on-demand to avoid confusion and anxiety. The vocabulary items were considered as a means to an end—winning the game—not the end itself. However, students did require to pay attention to them as part of the gameplay.

Students kept playing until they had gathered enough gold pieces for buying an item. Killing enemy champions and minions or destroying enemy structures, for example, gave the player some gold pieces. When a student had gathered enough gold pieces, he would go back to base to purchase the items. After purchasing an item, it was stored in the avatar's backpack which was accessible anytime. Through the backpack, students could also combine some items to create a stronger item. While seeing its thumbnail in the backpack, a student could hover over an item to see a textual description of that item. Furthermore, they could rebuy or simply drop the items and pick them up again to see which attributes (strength, agility, damage, armor, intelligence, hitpoints, and mana) of the avatar were affected by it. Also, they could examine the color codes used in the game (e.g., green represented hitpoints and blue had to do with mana). Thus, students saw each vocabulary item through textual (names and descriptions) and visual means (thumbnails, visual effects, and color codes)—otherwise known as multimedia presentation (Clark & Mayer, 2011). Through all these and the occasional leading questions asked by the researcher, students decided on a Persian equivalent for each item. Put briefly, students saw the item description, thumbnail, and effects; then, they decided on its meaning. As noted earlier, the

vocabulary items were presented, practiced, and produced (Nunan, 1999) or in other words noticed, repeated and generated (Nation, 2001). The way Arnseth (2006) put it, through this DVG gamers learn the vocabulary to play rather than playing the game to learn the vocabulary.

To win, team members had to improve their avatar's attributes such as strength, agility, damage, armor, intelligence, hitpoints, and mana, all made possible by purchasing the items. Prompts covered information on the required amount of gold, units, environment, item combinations, abilities, and resources, all contributing to a constant flow of student-software interaction. Student-student interaction mainly involved devising plans of action and items to purchase. These interactions were done in Persian, but items names were mentioned in English.

Watchers' treatment was the same as the Players' with the exception that in this group only two classmates played and others watched. Computer-controlled teammates and opponents were included to energize the competition and avoid boredom. Each player had half the class as his team who supported him by providing hints and encouragement.

2.3.3 Short- and long-term vocabulary retention posttests

One month after the study, students sat the short-term vocabulary retention posttest. Three months after the study, the same test was administered as the long-term vocabulary retention posttest. The students were not informed about these tests to ensure their lack of prior preparation.

2.4 Data analysis

Data analysis involved several ANOVAs and t-tests run through SPSS (v. 23) to examine the differences in vocabulary retention scores. The alpha level was set at .01 to reduce the effects of cooperative learning (teamwork) and in part observe the ANOVA assumption of having independent observations (Stevens, 2009).

3. Results and discussion

3.1 Results

Table 1 shows the descriptive statistics of the pretest and posttests (short- and long-term). A one-way between-groups ANOVA was run on the vocabulary pretest scores to examine homogeneity which showed no difference in students' knowledge of the target vocabulary items ($p = .058$).

		N	Min	Max	Mean	Std. Deviation
Readers	Pretest	70	0	15	6.27	3.98
	Short-term	70	3	18	10.63	3.17
	Long-term	68	0	19	9.47	3.65
Players	Pretest	64	0	20	5.44	4.65
	Short-term	62	2	21	14.85	4.22
	Long-term	56	4	21	14.34	4.46
Watchers	Pretest	71	0	13	4.61	3.71
	Short-term	73	6	21	14.07	3.90
	Long-term	73	0	21	13.25	4.02

The first question focused on the effect of DVGs on short- and long-term vocabulary retention. A mixed between-within subjects ANOVA was run to answer this question. Table 2 shows that there was a statistically significant main effect (time: $p = .000$) with a large effect size (Partial eta squared = .750). Thus, short- and long-term retention scores have had significantly higher means than the pretest. However, there has been an interaction effect (time*group: $p = .000$) which means that one variable was affected by another. This indicates that the main effects may not be appropriately generalizable.

Table 2. Multivariate Tests^a for the main effect and interaction effects

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial eta squared
time	Wilks' Lambda	.250	268.125 ^b	2.000	179.000	.000	.750
time * Group	Wilks' Lambda	.721	15.937 ^b	4.000	358.000	.000	.151

a. Design: Intercept + Grp

Within Subjects Design: time

b. Exact statistic

After examining the within subjects' effects (time), we go over the between subjects' effects (interventions: Readers, Players, and Watchers). Table 3 shows that there has been a statistically significant difference ($p = .000$) between the three groups with a medium-large effect size (Partial eta squared = .119).

Table 3. Comparison of the three groups

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial eta squared
Intercept	57849.468	1	57849.468	1981.942	.000	.917
Group	711.786	2	355.893	12.193	.000	.119
Error	5253.889	180	29.188			

Observing a significant difference between the treatments, Post-hoc Tukey HSD tests (Table 4) were run to find out which groups differed. As shown in Table 4, the Readers differed significantly from both Players ($p = .000$) and Watchers ($p = .003$). Players and Watchers did not show a significant difference ($p = .195$). This answers the second question which asked if Players and Watchers differed. In sum, both Players and Watchers outperformed the Readers but they did not differ from one another.

Table 4. Examination of the differences between the groups

(I) Group Name	(J) Group Name	Mean Difference (I-J)	Std. Error	Sig.
Readers	Players	-2.8316*	.58799	.000
	Watchers	-1.8468*	.54877	.003
Players	Readers	2.8316*	.58799	.000
	Watchers	.9847	.56795	.195
Watchers	Readers	1.8468*	.54877	.003
	Players	-.9847	.56795	.195

Based on observed means.

The error term is Mean Square (Error) = 9.729.

*. The mean difference is significant at the .01 level.

Based on the data, vocabulary attrition was examined as well. Thus, three paired-samples t-tests were run to compare the short- and long-term vocabulary retention scores of the three groups (Table 5). Although all groups forgot some vocabulary during the interval (positive t values), only the Readers forgot a significant amount ($M_1 = 10.80$, $SD_1 = 3.15$; $M_2 = 9.31$, $SD_2 = 3.68$) with a large effect size (Eta squared = .22). Players ($M_1 = 15.17$, $SD_1 = 3.90$; $M_2 = 14.37$, $SD_2 = 4.54$) and Watchers ($M_1 = 14.04$, $SD_1 = 3.92$; $M_2 = 13.39$, $SD_2 = 3.86$) did not show significant decreases over the interval. In sum, vocabulary attrition over a period of two months was significant only among the Readers.

Table 5. Comparison of short- and long term retention scores

	Paired Differences			t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean			
Readers	1.48438	2.83959	.35495	4.182	63	.000
Players	.79630	3.61015	.49128	1.621	53	.111
Watchers	.65278	2.90778	.34269	1.905	71	.061

Significance level = .01

Lastly, the study included simple ($N = 10$) and compound ($N = 11$) nouns as target vocabulary items. We compared the simple vocabulary items against the compound ones to examine the effectiveness of each treatment. Table 6 shows the results of three paired samples t tests run on the short-term vocabulary posttest scores separately for each treatment (results of the long-term test are not presented since they were similar). As shown, there was a significant difference between learning simple ($M = 1.60$, $SD = .20$) and compound nouns ($M = 1.41$, $SD = .17$) through the Readers' treatment with a large effect size (Eta squared = .49). Also, a significant difference between learning simple ($M = 1.72$, $SD = .19$) and compound ($M = 1.62$, $SD = .23$) nouns was found through the Watchers' treatment with a large effect size (Eta squared = .22). The Players' treatment, however, showed no significant difference between learning simple ($M = 1.74$, $SD = .19$) and compound ($M = 1.70$, $SD = .23$) nouns.

Table 6. Comparison of learning the simple and compound target vocabulary items through different treatments

	Mean difference	Std. Deviation	t	df	Sig.
Readers	.20	.20	7.885	66	.000
Players	.04	.19	1.824	60	.073
Watchers	.10	.20	4.457	71	.000

3.2 Discussion

The Players and Watchers performed significantly better than the Readers on the short- and long-term vocabulary retention posttests. This finding agrees with Hitosugi et al.'s (2014) conclusion in that DVG use resulted in higher retention. The finding also agrees with Smith et al. (2013) in that DVG use showed better outcomes regarding vocabulary learning. The need to go beyond short-term learning has been implied in the literature (Folse, 2006; Waring & Takaki, 2003). The retention scores suggest that, on average, about 15 words were retained by Players, 14 by Watchers, and 10 by Readers after a month's delay. After three months' delay, Players and Watchers forgot less than a word (mean difference: .80 and .65 respectively) which was not significant. Readers, however, forgot 1.48 words which was statistically significant. Thus, the Readers forgot about twice as much as the Players and Watchers.

Agreeing with Meyers (2010) and Smith et al. (2013), DVG use seems to have brought about generative learning advantages defined as generation of associations between new vocabulary and long-term memory. This could have well helped the better long-term retention of words among the Players and Watchers. In other words, as explained by Groot (2000) and Ma and Kelly (2006), inclusion of multiple modalities seem to help create stronger connections in the mind therefore strengthening the learning of new words. Thus, the results support the idea explained by Clark and Mayer (2011), Mayer (2002), and Thorne and Fischer (2012) that students retain vocabulary better through multimedia (which already includes words) than just words. A study by Cobb and Horst (2011) found that the participants needed to interact with a commercial game called Word Coach for two months to show significant gains in vocabulary knowledge. The present study, however, found such gains in a shorter period. Still, a direct comparison would be risky since the designs are different as well as participants' age.

The multiplicity of different sources enabled students to seek information in different places based on their needs and interests (Reinhardt & Sykes, 2012) which may have helped long-term memory activation. Based on the involvement load hypothesis (Laufer & Hulstijn, 2001) in terms of need, search, and evaluation, vocabulary retention is stronger through practices that induce higher levels of involvement load. The Readers needed the target vocabulary items to answer the worksheet questions so there was a moderate need. These items and their target meanings, however, were glossed in the margins so there was no involvement load in terms of the search factor. A moderate load in terms of evaluation existed based on the activity types. The game groups, however, experienced a stronger involvement load since the need factor (to understand the game better and accomplish more) was self-imposed. Moreover, the students had to infer the meanings based on several in- and out-of-game clues. It could be assumed, thus, that the search factor was moderate. Lastly, the evaluation factor seemed to be strong since the students had to weigh different suggested Persian equivalents for each item against the clues and choose the best. Thus, higher involvement seems to have helped achieve better retention, supporting Laufer and Hulstijn's (2001) conception. This finding also agrees with Hulstijn (1992) in that when word meanings are inferred, they would be retained better compared to the time when meanings are provided directly. It also supports the findings of a study by Newton (1995) in that negotiation for meaning of a word improves its retention.

The results also showed a significant difference in the retention of simple and compound nouns through the Readers' and the Watchers' treatments in favor of simple nouns. The Players' treatment, however, indicated no such difference. While vocabulary retention through the Players' and the Watchers' treatments was found superior to the Readers' treatment, the Players' treatment seems most effective. In sum, the Players treatment was equally effective for learning simple and compound nouns. The other two treatments, however, worked better with simple nouns.

A remark on playing versus watching a DVG concludes this section. There was no statistically significant difference between Players and Watchers on vocabulary retention. This could be attributed to two main reasons. Firstly, both treatments had discussions on choosing a Persian equivalent in quite a similar quality and used the same instruments. Secondly, while Players' had more hands-on-task experience, Watchers' experienced a more comfortable and relaxed atmosphere, allowing them better focus. These two might have counterbalanced each other out.

4. Summary and conclusion

The study aimed to examine short- and long-term vocabulary retention through a commercial DVG. Participants were divided to three groups, namely, Readers, Players, and Watchers. According to the results, game-learners performed better than the paper-and-pencil learners on both short- and long-term vocabulary retention tests. Considering vocabulary attrition, while the Readers forgot significant amounts of vocabulary knowledge during the interval between the short- and long-term retention tests, the game-learners retained their acquired knowledge. Other than these, it was found that only the Players' treatment was equally effective for learning both simple and compound nouns.

These findings indicate that commercial DVGs can be useful to the language classroom. Although this study focused on a limited number of vocabulary items, the findings through two assessments indicated that DVGs help students retain their vocabulary knowledge for a longer time. This paper demonstrates that implementation of DVGs in classrooms can help bring about significant changes toward more fruitful learning experiences. It should be noted, however, that DVG selection/creation for language learning should be made with special attention to the theoretical knowledge that exists on that component of language. For example, considering vocabulary, there is already an abundant literature explaining how vocabulary learning occurs (e.g., Laufer & Hulstijn, 2001; Nation, 2001). It is important to address these findings when implementing DVGs in classrooms in order to promote language learning. This study used a DVG that, at least partly, employed the conditions and processes leading to successful vocabulary learning which probably have attributed to the results. In sum, it can be concluded that DVGs selected for language learning purposes should be in-line with previous literature on the target content.

Lastly, this investigation faced certain limitations. Firstly, students' computer knowledge and skills in DVGs might have affected the results. Secondly, having independent observations as an ANOVA assumption was not met which

limits generalization of the findings. Thirdly, due to the design, a limited number of vocabulary items could be examined. Fourthly, students' out-of-class interactions may have affected their performance. Fifthly, as the study lasted for five weeks, the interval between items introduced earlier was longer than that for later sessions which may have affected the posttests. Finally, while the Readers saw each word for only one session, the other two groups could revisit the words again since they were always present in the game.

References

- Ang, S. C., & Zaphiris, P. (2008). Computer games and language learning. In T. T. Kidd & H. Song (Eds.), *Handbook of research on instructional systems and technology* (Vol. 1, pp. 449-462). Hershey, PA: IGI Global.
- Arnseth, H. C. (2006). Learning to play or playing to learn - A critical account of the models of communication informing educational research on computer gameplay. *Game Studies: The International Journal of Computer Game Research*, 6(1).
- Blizzard. (2003). *Warcraft III: The Frozen Throne* [Computer software]. USA: Blizzard Entertainment. Retrieved from <http://eu.blizzard.com/en-gb/games/war3/>
- Buckley, K. E., & Anderson, C. A. (2006). A theoretical model of the effects and consequences of playing video games. In P. Vorderer & J. Bryant (Eds.), *Playing video games: Motives, responses, and consequences* (pp. 363-378). Mahwah, NJ: Lawrence Erlbaum.
- Cambridge English Language Assessment. (2013). *Cambridge Young Learners English (YLE)*. Cambridge: Association of Language Testers in Europe.
- Clark, R. C., & Mayer, R. E. (2011). *E-Learning and the science of instruction proven guidelines for consumers and designers of multimedia learning* (3rd ed.). San Francisco, CA: Pfeiffer.
- Cobb, T., & Horst, M. (2011). Does Word Coach coach words? *CALICO Journal*, 28(3), 639-661. doi: 10.11139/cj.28.3.639-661
- Cornillie, F., Thorne, S. L., & Desmet, P. (2012). Digital games for language learning: From hype to insight? [Special issue]. *ReCALL*, 24(3), 243-256. doi: 10.1017/S0958344012000134
- Fincher-Kiefer, R., Post, T. A., Greene, T. R., & Voss, J. F. (1988). On the role of prior knowledge and task demands in the processing of text. *Journal of Memory and Language*, 27(4), 416-428. doi: 10.1016/0749-596X(88)90065-4
- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40(2), 273-293. doi: 10.2307/40264523
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave Macmillan.
- Gibson, D., Aldrich, C., & Prensky, M. (Eds.). (2006). *Games and simulations in online learning: Research and development frameworks*. Hershey, PA: Information Science.
- Groot, P. J. M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology*, 4(1), 60-81.
- Hitosugi, C. I., Schmidt, M., & Hayashi, K. (2014). Digital game-based learning (DGBL) in the L2 classroom: The impact of the UN's off-the-shelf videogame, Food Force, on learner affect and vocabulary retention. *CALICO Journal*, 31(1), 19-39. doi: 10.11139/cj.31.1.19-39
- Hsueh-Chao, M. H., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan.
- Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.
- IceFrog. (2015). *Defense of the Ancients (V. 6.78c AI 1.4e)* [Computer software]. Retrieved from <http://www.playdota.com/>
- Kettemann, B. (1995). How effective is CALL in ELT? *ReCALL*, 7(1), 49-53. doi: 10.1017/S0958344000005103
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26. doi: 10.1093/applin/22.1.1
- Ma, Q., & Kelly, P. (2006). Computer assisted vocabulary learning: Design and evaluation. *Computer Assisted Language Learning*, 19(1), 15-45. doi: 10.1080/09588220600803998
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85-139. doi: 10.1016/S0079-7421(02)80005-6
- Mayer, R. E. (2014). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.
- Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83(4), 484-490. doi: 10.1037/0022-0663.83.4.484
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84(4), 444-452. doi: 10.1037//0022-0663.84.4.444
- Meyers, P. C. (2010). *Incidental foreign language vocabulary learning from generative tasks*. (Unpublished doctoral thesis), Temple University, USA.

- Min, H.-T. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58(1), 73-115. doi: 10.1111/j.1467-9922.2007.00435.x
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2002). Range (V. 1.29) [Computer software]. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. New York, NY: Routledge.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Newton, J. (1995). Task-based interaction and incidental vocabulary learning: A case study. *Second Language Research*, 11(2), 159-176. doi: 10.1177/026765839501100207
- Nunan, D. (1999). *Second language teaching & learning*. Florence, KY: Heinle & Heinle.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). New York, NY: Cambridge University Press.
- Prensky, M. (2007). *Digital game-based learning*. St. Paul, MN: Paragon House.
- Reinhardt, J., & Sykes, J. M. (2012). Conceptualizing digital game-mediated L2 learning and pedagogy: Game-enhanced and game-based research and practice. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 32-49). New York, NY: Palgrave Macmillan.
- Rollings, A., & Adams, E. (2003). *Andrew Rollings and Ernest Adams on game design*. San Francisco, CA: New Riders Publishing.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43. doi: 10.1111/j.1540-4781.2011.01146.x
- Smith, G. G., Li, M., Drobisz, J., Park, H. R., Kim, D., & Smith, S. D. (2013). Play games or study? Computer games in ebooks to learn English vocabulary. *Computers & Education*, 69, 274-286. doi: 10.1016/j.compedu.2013.07.015
- Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York, NY: Routledge.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input and second language acquisition* Rowley, MA: Newbury House.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391. doi: 10.1093/applin/16.3.371
- Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *ACM Computers in Entertainment*, 3(3), 1-24.
- Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners [Special issue]. *ReCALL*, 24(3), 302-321. doi: 10.1017/S095834401200016X
- Thomas, M. (2012). Contextualizing digital game-based language learning: Transformational paradigm shift or business as usual? In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 11-31). Basingstoke, UK: Palgrave Macmillan.
- Thorne, S. L., & Fischer, I. (2012). Online gaming as sociable media. *ALSIC: Apprentissage des Langues et Systèmes d'Information et de Communication*, 15(1), 1-25.
- Tsai, C. W., & Fan, Y. T. (2013). Research trends in game-based learning research in online learning environments: A review of studies published in SSCI-indexed journals from 2003 to 2012. *British Journal of Educational Technology*, 44(5), 115-119. doi: 10.1111/bjet.12031
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130-163.

Appendices

Appendix 1: The target vocabulary list

No	Item	Session	No	Item	Session
1	Damage	1	12	Robe of the magi	3
2	Armor	1	13	Chainmail	3
3	Agility	1	14	Boots of speed	3
4	Healing salve	1	15	Gloves of haste	4
5	Ally	1	16	Broadsword	4
6	Gauntlets of strength	1	17	Quarterstaff	4
7	Ironwood branch	2	18	Claymore	4
8	Status	2	19	Perseverance	5
9	Intelligence	2	20	Power treads	5
10	Mana	2	21	Recipe	5
11	Buckler	3			

Appendix 2: A sample page from the worksheets (Session 4)

Season 2

The Frozen Throne

Lesson 2

Alliance Campaign: To the North

Pre-reading Activities:**Retelling**

Try and remember what happened in the story last session. Present it to the class.

Predicting and Skimming

Look at the title, the pictures, and the phrases below. What do you think happens in the story today?

lord Illidan	you shall be my right hand	to kill Magtheridon	many different soldiers
there are two gates	travel with me to the north	to kill Lich King	to the icy north together

Now, skim the passage and check your predictions.

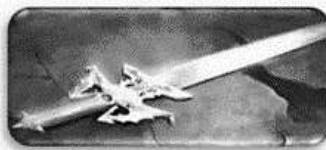


Prince Kael and the Naga finally found Illidan. The first thing Prince Kael asked Illidan was 'can you help us **solve** our magic problem?' 'There is no solution my young Prince; but I can give you a new **source**' answered Illidan. Prince Kael thought for a moment and then said: so be it; from now on, you can think of us as your allies lord Illidan! 'And you shall be my right hand,' Illidan went on 'our first **task** is to kill Magtheridon; he's the lord of this land and we need to take it from him.'

It took them a few days to make their camp. After that, Illidan sent a few **scouts** to Magtheridon's city. When they came back, they reported that a group of undead protect a very strange item; 'it's called the **gloves of haste**' they said. 'We should surely look into that; Kael, that is your first task; bring me the gloves and I shall tell you about your problem' said Illidan. 'Consider it done my lord.' said Kael. The Prince and his men left the camp and found the gloves. Prince Kael quickly returned and gave them to Illidan. He wore them and was surprised; the gloves made him much quicker!



The next morning, Illidan explained to Kael that he met a great **demon** lord named Jaeden; he said: Jaeden promised me strength; and in return, he asked me to kill the Lich King in the north; and he also promised me power beyond **imagination**; now, to kill the Lich King, I need an army; if you travel with me to the north and help me do it, I promise that I'll solve your magic problem. Prince Kael stepped forward and said: I know the legend of the ice **crown** and the frozen throne; my **claymore** is yours to command; we'll go to the icy north together.



Solve: to find the answer to a problem.

Source: a place, person or a thing that you sth from.

Task: a work that you must do.

Scout: a person sent ahead to find information about the enemy.

Gloves of haste: a covering for the hand that makes you act faster.

Demon: very evil.

Imagination: ability to think of new ideas.

Crown: a circle made of gold that kings wear on their head.

Claymore: a large sword with two sharp sides.

Appendix 3: The vocabulary test

نام و نام خانوادگی:		مدرسه:		کلاس:	
معنی لغات و عبارات داده شده در سوالات ۱ تا ۲۱ را از بین گزینه های الف، ب، ج یا د انتخاب کرده و مطابق نمونه دور گزینه مورد نظر خط بکشید. نمونه حل شده معنی کلمه True : الف حقه ب سریع ج  د نور					
۱. معنی لغت Damage : الف شروع	ب آسیب	ج برتاب	د نشان		
۲. معنی لغت Armor : الف زره	ب لباس	ج نقاب	د دستگیر		
۳. معنی لغت Agility : الف صلابت	ب چابکی	ج زورکی	د خشونت		
۴. معنی لغت Intelligence : الف کارایی	ب بینش	ج هوش	د زرنگی		
۵. معنی لغت Ally : الف کوشش	ب عالی	ج متحد	د حمله		
۶. معنی لغت Bucket : الف جامی	ب سیر کوچک	ج بشنه	د دسته بیل		
۷. معنی لغت Status : الف گفته	ب نوع	ج نمر	د وضعیت		
۸. معنی لغت Recipe : الف مطابق	ب آرزوی	ج دستورالعمل	د نیروی		
۹. معنی لغت Chainmail : الف زره زنجیری	ب شمشیر سیاه	ج تیغ سیاه	د سیر قرن		
۱۰. معنی لغت Mann : الف نیروی باستانی	ب جانم	ج نیروی حیات	د فانی		
۱۱. معنی لغت Broadsword : الف شمشیر پهن	ب شمشیر آتشین	ج شمشیر منحنی	د گویار		
۱۲. معنی لغت Quarterstaff : الف تاج امپراتور	ب گسی پول	ج وسایل جزئی	د دعای جنگی		
۱۳. معنی لغت Claymore : الف زره طلائی	ب شمشیر دو دم	ج زره شاهی	د شمشیر هندی		
۱۴. معنی لغت Perseverance : الف شجاعت	ب پرتوکاری	ج استقامت	د جهلشنی		
۱۵. معنی لغت Ironwood Branch : الف چوب آرگیده	ب چوب سیبباز	ج شاخه سرو	د شاخه آهنین		
۱۶. معنی لغت Robe of the Magi : الف ردای حفاظت	ب ردای سرعت	ج ردای زرشن	د ردای پرواز		
۱۷. معنی لغت Gauntlets of Strength : الف زره کوتاه چرمی	ب کوفر قدرت	ج دستکش کوتاه قدرت	د دستکش چرمی		
۱۸. معنی لغت Boots of Speed : الف بوتین آزادی	ب معجون سرعت	ج بوتین سرعت	د معجون آزادی		
۱۹. معنی لغت Healing Salve : الف معجون رهایی	ب دایوی طاقون	ج اکت کش	د برهم شلجاش		
۲۰. معنی لغت Power Treads : الف گامهای پرواز	ب نیروی سه گانه	ج چکمه نبرد	د چکمه رهایی		
۲۱. معنی لغت Gloves of Haste : الف دستکش تریک	ب دستکش بلند شانه	ج دستکش نیرو	د دستکش جامی		