Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 20 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i20.50247

PAPER

Beware of What Lurks in the Margin: Glosses, Mobile Digital Games, and Incidental Vocabulary Learning

Amin Rasti-Behbahani, Mohammed Abdulgalil Abugohar(✉), Nagwa Mohammed Khallaf

Centre for Preparatory Studies, Dhofar University, Salalah, Oman

mabugohar@du.edu.om

ABSTRACT

The study was set to investigate the effectiveness of meaning-given (meaning-G) and multiple-choice (choice-G) glosses in enhancing incidental vocabulary learning by word learning activities developed with a tablet digital game (DGBVL). Fifty-four Persian speakers (13–14 years old) were divided into meaning, choice, and control (CG) groups randomly after their vocabulary knowledge sizes were measured (2100–2400). After administering two pre-tests, measuring receptive and productive knowledge of 20 concrete nouns, participants played a commercial adventure game. They read a modified game guide and completed a chapter of the game in pairs. Meaning-G had a single Persian definition, choice-G had three, and CG had none on the page margins. Three weeks later, the participants took post-tests, receptive and productive, without forewarning. Moreover, qualitative data were collected. Results showed that 1) DGBVL activities could enhance incidental learning of the target words regardless of the gloss types; 2) contrary to our expectations, meaning-G enhanced target word productive knowledge acquisition much more effectively than choice-G; and 3) glosses led to different vocabulary learning strategies. This study extends the literature by advising prospective teachers to employ meaning-G to enhance vocabulary acquisition through DGBVL activities effectively.

KEYWORDS

meaning-given glosses, multiple-choice glosses, digital game-based learning, vocabulary acquisition, incidental learning, mobile video games

1 INTRODUCTION

It has been widely discussed that the knowledge that is gained through incidental vocabulary learning [IVL] is crucial for lexical development [35] because, in addition to form-meaning, other knowledge such as register, collocation, etc. are enhanced by IVL [35]. There are two widely used definitions of IVL, and both apply to this study. In applied linguistics, IVL refers to the acquisition of some words by doing a language activity that is not designed for learning words [35], and in psychology, "the absence

Rasti-Behbahani, A., Abugohar, M.A., Khallaf, N.M. (2024). Beware of What Lurks in the Margin: Glosses, Mobile Digital Games, and Incidental Vocabulary Learning. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(20), pp. 169–185. https://doi.org/10.3991/ijim.v18i20.50247

Article submitted 2024-05-25. Revision uploaded 2024-07-17. Final acceptance 2024-07-17.

of test announcement" distinguishes incidental and intentional learning [2]. Despite the importance of IVL, many studies show that IVL usually does not lead to large vocabulary gains [35]. Therefore, different techniques were introduced [27], especially with the introduction of handheld devices such as tablets into the classrooms, to accelerate the rate and quality of IVL gains. Glossing [17] and interactive content such as tablet digital games are two techniques that are the focus of this study.

Glosses or annotations are written on the text margins for further clarification and better comprehension of texts [2]. It has been discussed that glosses can support IVL for significant reasons: firstly, they are easily accessible and prevent wrong guesses [11]; secondly, glosses can consolidate form-meaning connections because they keep the reading process uninterrupted [26]; thirdly, glosses can enhance word retention by prompting more lexical [13] and multiple encounters with the target words [34].

In addition to glosses, the significant effect of interactive content such as digital games on IVL has been widely reported in previous studies [1], [4], [5], [6], [23], [24]. Most of these studies discuss the superiority of commercial digital game-based vocabulary learning (DGBVL) activities [24] to traditional paper-based activities due to digital games' inherent elements such as "interactivity, rules, goals, challenge, risk, fantasy, curiosity, control, rich images, animations, videos, visuals, audios, and interactive dialogues" [20].

Hence, to enhance the IVL gains, understanding how to implement glosses in interactive vocabulary learning activities such as DGBVL activities properly is essential. This study, then, is an attempt to investigate the effect of glosses on incidental vocabulary learning in tablet DGBVL activities, which has rarely been studied.

1.1 Glosses and incidental vocabulary learning

Glosses can be categorized based on their type (in-text, margin, etc.), language (L1 or L2), modality (e.g., text + video), and cognitive stimulating methods (e.g., multiple-choice glosses) [2]. According to recent meta-analysis studies, glossed reading supports IVL effectively by strengthening the form-meaning link and enhancing word use knowledge [2], [36]. Despite their advantages, an examination of their types or conditions is recommended [37].

In paper-based settings, for example, [12] studied the effects of meaning-given glosses (meaning-G), in which a translation of target words is presented on the margin of a text; multiple-choice glosses (choice-G), in which two or more definitions are provided on the margin of a text and L2 learners must choose one; and no glosses on IVL. By experimenting on 65 Dutch learners, he found that glosses were effective for IVL, although the results showed a small effect size. He also found that meaning-G could enhance the form-meaning connection in an immediate post-test more effectively. [13] studied the effect of glosses on 85 undergraduate-level advanced Spanish learners' vocabulary recall and learning. After assigning them to meaning-G L2 glosses, meaning-G L1 glosses, and no glosses conditions, immediate form-recall and translation tests were administered. Glosses helped with more recall and outperformance in the translation test. [11] tested advanced French learners' IVL by providing L2 meaning-G and found that in addition to dictionary use, glosses could support IVL and eliminate the possibility of wrong guesses in reading with no glosses. [34] studied the effects of meaning-G, choice-G, appositives (providing an in-text definition for a target word), and translation on 231 Japanese students. After administering unannounced both immediate and delayed post-tests, he found that both

L1 meaning-G and choice-G could enhance IVL similarly, and their impacts were greater than appositives and the translation task. [27] investigated the effects of L1 choice-G, L2 text reconstruction, and a combination of both on 76 advanced German learners' IVL gains. The results of an immediate and a five-week delayed posttest showed that the choice-G condition led to deeper gains in both receptive and productive vocabulary knowledge, whereas its effect diminished over time while the effect of the combination condition lasted longer. Lastly, [33] compared the effects of L2 textual, pictorial, and L2 textual + pictorial glosses on 56 high-intermediate English learners' incidental acquisition of six pseudowords. They found that pictorial glosses could promote word meaning retention the best and enhanced the previous studies that indicated gloss type can be a determining factor in successful incidental vocabulary learning.

Since computers were introduced into classrooms, computerized or multimedia glosses such as video, audio, pictorial, and hyperlinked ones have also become popular. The motivations for using computerized glosses were usually generated from theories such as dual coding theory [18], the theory of multimedia learning [16], and cognitive load theory [21]. The dual coding theory discusses that learning from two channels, e.g., visual and aural, leads to enhanced learning [18]. Moreover, the cognitive load theory posits that human cognitive capabilities and processing capacity are limited; therefore, learning activities must be tailored for a better effect [21]. Also, the theory of multimedia learning offers a framework to optimize cognitive load and audio-visual inputs [16]. Relying on those theories, the literature on glossing and IVL expanded. For instance, [37] recruited 195 Japanese university students and examined whether L1 only, L1 + picture, L2 only, and L2 + picture glosses can enhance the IVL gain of 14 English words differently in a multimedia context. He found that glosses could enhance immediate recall and recognition, and the difference between L1 and L2 glosses was not significant. Also, he found that pictorial cues were more effective than glossed-only text and explained the results by relying on dual coding theory. Finally, [22] found that multimodal glosses (text + audio or video) can support IVL more effectively.

Why glosses can be effective was also studied. For instance, [26] studied the effect of glosses through think-aloud protocols. They found that texts + gloss can support IVL because glosses encourage finding meaning and can enhance form-meaning links firmly. Moreover, regarding the importance of vocabulary learning strategies [10], [25] studied the effects of the quality and quantity of lexical processing strategies that are triggered by meaning-G and choice-G on IVL. She found that the integration of mental sources such as meta-cognition and semantic elaboration and the frequent attempts to find and evaluate the word meanings are the main reasons for the establishment and strength of the form-meaning connection. She discussed that, because choice-G triggers such conditions but meaning-G provokes only meta-cognitive resources and linear text processing, choice-G can lead to stronger form-meaning connections, and consequently more effective incidental vocabulary learning.

As is evident, the literature on glossing and IVL is vast in both paper- and computer-based contexts. Overall, it is noticeable that regardless of type, L1 glosses lead to superior gains; the effect of choice-G was much greater than other gloss types—mainly due to the effect of inferencing [9]; meaning-G, as well as hyperlinked and interlinear glosses, had similar effects; and different modes of glossing had similar effects on IVL [2], [36]. Hence, this study focuses specifically on the effects of L1 meaning-G and L1 choice-G on IVL in an interactive computer-based context—digital games—to extend the current literature on L1 glossing.

1.2 Digital games and incidental vocabulary learning

[24] defines DGBVL as the educational use of interactive software developed for entertainment purposes, or video ordigital games, to learn vocabulary; therefore, acquiring vocabulary items incidentally from a digital game that is played for fun rather than learning new words can be considered DGB-IVL. It has been shown that language learners' IVL gains can be enhanced greatly by such interactive DGBVL activities. For instance, [24] found that the productive recognition of form-meaning can be enhanced by adventure DGBVL activities. [4] also found that adventure DGBVL activities that include word-focused exercises can enhance IVL and retention more effectively than those without word-focused exercises, which enhances the findings by [32], who found that anticipating learners' behaviors and providing timely interventions can enhance learning. Overall, according to recent meta-analysis studies, DGBVL activities are significantly effective and preferable to many other IVL activity types [6], [14].

Furthermore, reasons for the positive effects of DGBVL activities on IVL have been discussed. [6] found that a digital game's inherent element, i.e., "challenge," plays a significant role in supporting IVL. [19] found that game-based lectures can change learners' attitudes positively towards and motivate them to learn. Moreover, [23] discussed that owing to the essential psychological factors for vocabulary acquisition, such as motivation [28], authenticity, repetition, instantiation, dual encoding, interactivity, and feedback, DGBVL activities can offer rich opportunities for enhancing incidental vocabulary learning.

For IVL gains to increase significantly, modification of digital games through the addition of glosses is critical [4], [32]; however, in DGBVL literature this issue has rarely been addressed. To the best of our knowledge, only two studies were found that have addressed this issue. [1] investigated the effect of multimedia glosses (L1 text + picture, L2 text + picture, and picture only) on the immediate and delayed incidental acquisition of 12 concrete nouns by 162 Arabic speakers. He found that although all three gloss types could enhance vocabulary acquisition, L1 + picture glosses were more effective. [4] studied the effects of L1 glosses on 21 eighteen-year-old Chinese speakers. They modified a digital game on-screen texts by adding Chinese translations (L1 meaning-G) to 28 target words in parentheses next to each word. They found that the addition of L1 glosses could enhance IVL in DGBVL activities.

Since meaning-G, as well as hyperlinked and interlinear glosses, can have similar effects [2], [36], we can deduce that [1] and [4] studied only multimedia and L1 glosses that are almost similar in their quality of contribution to supporting IVL. Most importantly, the majority of studies in glossing literature did not compare the acquisition of receptive and productive word knowledge with each other. Hence, investigating the effect of choice G on both receptive and productive knowledge seems necessary. Therefore, the answers to the following questions are sought in this study:

1. Do DGBVL activities affect IVL?
2. To what extent does providing meaning-G and choice-G L1 glosses make a significant difference in IVL through a DGBVL activity?

2 METHODOLOGY

This study was quasi-experimental. Also, based on the definitions of IVL discussed above, the nature of vocabulary learning was incidental because participants played the game with no intention of learning target words, and the post-tests were administered without forewarning.

2.1 Participants

Based on the convenience-sampling method, participants of this study were selected among junior high school students who studied English four hours per week in evening classes for one year. Selected participants were 54 randomly selected Persian speakers, comprising 23 males and 31 females whose age range was between 13 and 14 years old. They were categorized as lower-intermediate based on their school's in-house proficiency test. However, the *vocabulary size test (VST) (14,000 Version, receptive)* was also administered and revealed that participants' breadth of English word knowledge was from 2100 to 3400 words (M = 2727.77, SD = 379.28). Therefore, based on the common european framework of reference (CEFR), their proficiency in English was between A2 and B1. They were accustomed to playing digital games in their free time and had no familiarity with the selected game in this study.

2.2 Materials

The digital game. We used *Haunted Hotel: Death Sentence Collector's Edition* for this study, which was an adventure game in which gamers help a detective unravel the secret of his missing friend by solving object-based puzzles (see Figure 1). Boosting intrinsic motivation [20], and enhancing target language learning through interesting and engaging puzzles and dialogues [4] were two main reasons why an adventure game was selected.

Fig. 1. Item-based puzzle

Game guide. After downloading a game guide, or a guide text that gives gamers hints to solve puzzles in the game, from the game's developer website, we modified the first chapter for this study. Firstly, picture guides were deleted; next, numbers were assigned to the sentences. Moreover, since glosses on the page margins are preferable by many learners [13], [36], we placed the target word Persian definitions on the page margins.

Providing the game guide could reduce time spent on the task and hamper the negative effects of extensive interactivity [24]. Participants were supposed to read the game guide and follow the steps precisely to complete the game successfully.

Target words. The game guide text frequency profile was checked first by Lextutor. Next, twenty low- to infrequent concrete nouns of inanimate objects such as *skull* (K4) and *shack* (K7) were selected. The results in Table 7 confirm that all selected target words were new for the participants. When selecting the target words, we followed four steps: First, we made sure that participants needed to know

them to solve game puzzles. Secondly, they had not to be repeated more than two times in the game guide (Nation, 2022). Then, we selected only nouns because of the ease of acquisition and reducing the burden of word difficulty in the learning process [29]. Finally, words lower than the K4 frequency band were not considered (refer to Table 1).

Table 1. Target words

Target Words	Freq.	K		Target Words	Freq.	K
Glove	1	4		Embers	1	9
Bracelet	1	7		Glue	2	4
Closet	1	6		Lens	2	4
Debris	1	5		Reel	2	5
Drape	1	5		Sack	2	4
Tag	1	4		Rust	2	5
Latch	1	6		Shovel	2	6
Skull	1	4		Pouch	2	7
Kettle	1	6		Medallion	2	9
Shack	1	7		Dossier	2	11

Task design. After modifying the game guide, three different versions were developed. In the first version (refer to Table 2), a single definition in Persian was added in front of each target word on the right page margins (meaning-G group).

Table 2. Meaning-G game guide text

8. Now, check the skull; pick up the key.	جمجمه

Note: جمجمه means skull.

In the second version (refer to Table 3), three definitions in Persian were added in front of each target word (choice-G group). The frequency of repeated definitions was controlled not to exceed twice [9].

Table 3. Choice-G game guide text

	جمجمه
8. Now, check the skull; pick up the key.	فیلم
	النگو

Notes: جمجمه means skull; فیلم means reel; and النگو means bracelet.

The third version included no definitions (control group) (refer to Table 4).

Table 4. Control game guide text

8. Now, check the skull; pick up the key.	

The main instruction for all three versions was similar: *read the game guide and follow the steps carefully while playing the game.*

2.3 Instruments

Vocabulary size test. A paper-based version of receptive VST (rasch reliability indices > 0.96) was administered to check participants' homogeneity. It included 140 multiple-choice questions, and each question had four options to choose from. With no time limit, participants were asked to answer all 140 questions.

Achievement tests. A receptive meaning recognition test checked participants' knowledge of meaning by selecting the proper Persian definitions (refer to Table 5). Also, a productive form recall test checked participants' knowledge of the form by asking them to fill in the blanks. They had to recall the proper English forms (refer to Table 6). Cues and sentences were provided for them. The total score of the test was 20, and the weight of each correct answer was one. With no time limit, participants took both tests once as a pre-test and another time as a delayed post-test (it is discussed later). The pre-test helped us to understand if the target words were new to the participants. It is worth noting that to avoid the effect of tests on each other, the productive test was administered first. In the examples below, in the receptive test, the correct answer is **a**, and in the productive test, **skull** is the correct answer.

Table 5. Receptive test: Question 12

12. don't touch **embers**! You can burn yourself.			
a) زغال داغ	b) کتری	c) گاز پیک نیک	d) کبریت

Note: زغال داغ is Embers; کتری is kettle; گاز پیک نیک is burner; and کبریت is matches.

Table 6. Productive test: Question 13

13. In a car accident, her **sk**____ cracked badly and she fell into a coma.

Four professors of applied linguistics checked the face validity of the tests. Later, we pilot-tested them. In so doing, we randomly administered the tests to nine English learners who were not participants in this study. Internal consistency of the tests gained through Cronbach's alpha [31] was satisfactory (receptive test = 0.76; productive test = 0.72).

Think-aloud protocols. Qualitative data was collected from eight participants only by concurrent think-aloud method [8] for a better understanding of the effect of glosses. The instruction was carefully designed so that the possible effects of the instruction could be hampered [8]. Then, the instruction was: '*When playing, try to speak out your thoughts. You can also discuss with your partner about your thoughts*'. Participants were randomly checked through surveillance cameras because the validity of the data could be impaired by the presence of researchers in the same space [8]. Think-aloud participants verbalized their thoughts in Persian.

Interview. An "exit interview" is recommended [3] for enriching the elicited concurrent think-aloud data because individual differences and the methods used can affect the quality and quantity of verbalized data. Hence, an exit interview was given.

The exit interview was semi-structured and contained two types of questions. The first group of questions was to know how participants learned or understood the target words. The second group of questions explored how multimedia learning factors affected learning. Each participant was interviewed individually in Persian.

2.4 Procedure

Firstly, 71 English learners volunteered to participate. After their families signed the consent letters, volunteers sat for the paper-based VST and the pre-tests (see Figure 2). Next, because participants who had a vocabulary size of less than 2000 word families would not understand the game guide, we selected only those who scored higher (N = 54). Moreover, pre-test results showed that the selected words were new for all participants (refer to Table 7). After that, we randomly gave them meaning-G (N = 18), choice-G (N = 18), or no gloss (N = 18) activities. Later, for collecting think-aloud data, we randomly selected two pairs from both meaning-G and choice-G groups (N = 8). The participants completed the tasks in pairs.

In the second stage of the study, two days before their main task, the think-aloud pairs were instructed on how to: use a tablet to play the game cooperatively, become familiar with being recorded, and think aloud in a 30-minute warm-up session [8]. Two days after the warm-up session, we gave all participants an orientation session on how to complete their DGBVL activities on their tablets cooperatively. In three different rooms, all participants were given their specific game guides and were asked to complete the first chapter by reading and following the instructions in the game guide on their tablets. The think-aloud pairs' actions and voices were recorded with a camcorder. Due to unskippable video cutscenes in the game, participants spent nearly 70 minutes completing their activities.

In the third stage, firstly, the think-aloud participants were interviewed individually for about 20 minutes each immediately right after their activity completion; however, we did not forewarn them about the exit interview because their cognitive processes during activity completion could have been affected [3]. Non-think-aloud participants left the rooms immediately after their activity completion. Finally, we administered the delayed post-tests three weeks later to check all participants' retention because [37] found that word knowledge gains through L1 glosses would vanish more quickly and a stable vocabulary gain lasts at least three weeks [24]. It is worth noting that the participants had no access to the game or materials after their playing session.

2.5 Data analysis

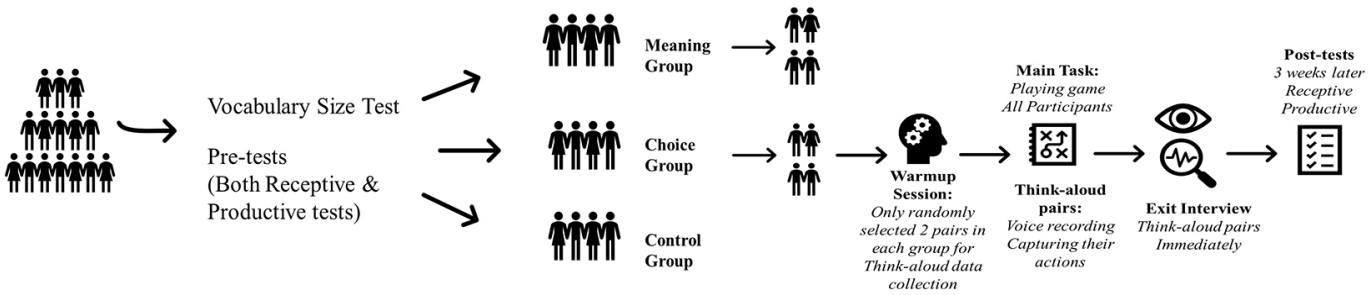

Fig. 2. The process of data collection

To answer both the first and second research questions, we collected quantitative data and used *SPSS* version 23 and *Jamovi* (2024) for data analysis. The quantitative data were analyzed by formulas such as *two related-samples wilcoxon signed rank Tests* to test differences within groups and the *Kruskal-Wallis test* to test the differences between groups. The rank-biserial correlation and eta-squared formulas [31] were used to calculate the effect sizes for the former tests, respectively.

Qualitative data were also collected to understand how each gloss could affect IVL gains and thereby answer the second research question elaboratively. We followed the inductive content analysis conventions because knowledge of the effect of glossing on IVL gains in DGBVL activities is fragmented [7]. To content-analyze the data manually, we followed these steps: 1) Our unit of analysis was "sentence"; 2) data was transcribed first and it was read and reread several times for familiarization with the data; 3) while reading data, recurring patterns and themes were coded; 4) after extracting the emerged patterns and their codes, 5) the pertinent patterns were grouped; 6) we identified and tagged the most recurring pattern in each group; 7) later, the patterns were labeled and abstracted; and 8) finally, we categorized the findings to explain the qualitative results. An applied linguistics professor supervised the whole process.

3 RESULTS

3.1 Quantitative results

According to the Shapiro-Wilk test, the distribution of test scores was not normal. Therefore, the Kruskal-Wallis test result was run on the pre-tests (refer to Table 7) and showed that there was no significant difference ($p > 0.05$) between participants, meaning that the selected target words were new for all participants.

Table 7. Comparison of pretests

Tests	Groups	Mean Rank	N	χ^2	df	p
Receptive (Pre-Test)	Meaning G	26.39	18	0.14	2	0.92
	Choice G	28.22				
	Control G	27.89				
Productive (Pre-Test)	Meaning G	26.33	18	0.15	2	0.92
	Choice G	28.08				
	Control G	28.08				

RQ1: To what extent do DGBVL activities make significant differences in IVL?

Table 8. The Results of the Wilcoxon signed-rank test

		N	Mean Rank	Sum of Ranks	Z	p	Rank-Biserial Correlation
Receptive	Negative Ranks	0	.00	.00	–6.16	.00	0.96
	Positive Ranks	50	25.50	1275.00			
	Ties	4					
	Total	54					
Productive	Negative Ranks	0	.00	.00	–6.04	.00	0.97
	Positive Ranks	48	24.50	1176.00			
	Ties	6					
	Total	54					

The results in Table 8 indicate that although there were differences in gloss types, the DGBVL activities enhanced receptive and productive knowledge gains with large effect sizes. However, the DGBVL activities were not effective on the acquisition of receptive knowledge for four participants, and productive knowledge for six participants.

RQ2: To what extent do providing meaning-G and choice-G L1 glosses make a significant difference in IVL through a DGBVL activity?

Table 9. Results of the Kruskal-Wallis test with dwass-steel-critchlow-fligner (DSCF) PostHoc test

Tests	Groups	Mean Rank	N	χ^2	df	p	Pairwise Comparison	DSCF p	Eta-Squared (η^2)
Receptive Gain Scores	Meaning G	31.89	18	7.2	2	.02	MeaningG – ChoiceG	0.96	0.06
	Choice G	31.17					MeaningG – Control	0.06	
	Control G	19.44					ChoiceG – Control	0.04	
Productive Gain Scores	Meaning G	39.33	18	21.7	2	.00	MeaningG – ChoiceG	0.04	0.34
	Choice G	28.06					MeaningG – Control	0.00	
	Control G	15.11					ChoiceG – Control	0.01	

Note: Gain score = pre-test score – post-test score.

As shown in Table 9, at $p \leq 0.05$, there was a significant difference between the meaning-G, choice-G, and control groups in their overall performance. There was also a large effect size for productive knowledge. However, the results of the DSCF posthoc test show no significant difference between meaning-G and choice-G in the receptive test, but it shows a weakly significant difference in the productive test. Also, the DSCF posthoc test [30] shows that meaning-G and the control group had no significant difference in receptive tests. Overall, based on Table 9, the glossed DGBVL activities could enhance productive knowledge better than receptive. Also, meaning-G outperformed choice-G in the productive test.

3.2 Qualitative results

For further investigation of the effect of glosses (RQ2), think-aloud data were collected. The content analysis of the qualitative data revealed eight patterns that were later grouped into two larger categories, namely, *general strategies* and *exclusive strategies*. The former ones were used by *all* participants, while the latter ones were used by participants in each group to cope with the target words. It facilitated data interpretation.

General strategies. The emerged general strategies were later grouped into five categories and labeled as follows:

Information search. The game guide and the game were the main sources of information in which participants searched for information about the target words. These were coded as *information searches*.

A: We must go out first. | B: No! Let's read the game-guide again and decide what to do. | A & B: {game-game information search}.

For example, here, both participants look through the game-guide text for their next action.

Negotiation. Having found the correct information, participants negotiated and implemented a strategy for solving a problem. These instances were coded as negotiation.

A: Tap on the latch, find ... | B: What is a latch? Isn't it {a Persian word for the oil}? | A&B: {they search for in-guide information}. | B: Yeah, the latch is {Persian word for the latch}

For instance, the target word *latch* and its meaning are being negotiated here.

Turn-taking. After participants had negotiated, they tested their ideas by taking turns to either control the game by holding the tablet and tapping or reading the game guide in turns. We coded these instances as turn-taking.

A: Let me read the rest and decide. / B: {B taps to check his own idea} | A: Tap on the main door {game-guide text} | A: Please give me the tablet. Read!

For example, here, A takes a turn to read and B taps on the screen.

Trial-and-error. Participants' final strategy was trial-and-error when none of their attempts succeeded. They tapped randomly on the objects on the screen until they found the right object by chance. These were tagged with trial-and-error.

A: {reads from the game guide} Tap on the hammer | B: Is hammer {Persian definition}? | A: I'm not sure, but it must be here. Just tap around to find it.

In this excerpt, participants try to find a *hammer* by trial-and-error strategy.

In the exit interview, using the trial-and-error strategy was confirmed by some of the participants.

Researcher: Did any trick help you with unknown words during the game?

A: When we couldn't find them, we tapped on everything until we found it. | B: When we ran out of ideas, we kept tapping on every corner of the game or a scene.

Review. Finally, participants were regularly reorganizing their thoughts by reviewing their previously completed plans or actions. These were labeled as reviews. In the following example, the places and number of *nails* were reviewed.

A: Look, this is Isaac's reel in the inventory; we need those two parts for the camera. |B: We already picked those two parts. | A: What is a plank? | B: {Persian definition}. We already picked that up too. So, we can go out of the room now.

Exclusive strategies. Due to the differences in the gloss types, meaning-G and choice-G groups employed group-specific strategies. By using these strategies, participants solved the problems in which knowledge of the target words was necessary.

Meaning-G group. Encountering the target words, participants with meaning-G used a strategy that we named later the *enhancement* strategy because the target words or their Persian definitions were read out loud by them repeatedly. For example, participant A repeatedly read out loud the Persian definition of the word *debris,* and participant B repeatedly voiced it while he was searching for it on the screen.

A: Find and tap on debris. {Persian definition of matches is enunciated loudly} | B: {Persian definition of matches is being repeated while he is searching}

This was mentioned in the exit interview too.

Researcher: What have you learned the most from the words? | A: Their pronunciations and definitions. Moreover, I have some pictures of them in my mind. | B: Both pronunciations and definitions.

Choice-G group. The participants in the Choice-G group employed two strategies that we later called *inferencing from contexts* [26] and *hypothesizing*. To infer meanings of target words by inferring from context strategy, participants searched the game and the game guide for any contextual clues that helped them guess the meaning. For example, participants A and B guessed the meaning of the target word *magnifier* by inferencing from the contextual information:

A: Tap on the pouch in your bag. Then, tap on the magnifier. | B: Pouch means {Persian definition}. Try to find a {Persian definition of pouch} somewhere.

Also, in the exit interview, this was mentioned.

Researcher: Did any trick help you with unknown words during the game?

A: Clues and wordssuch as, for example, red or yellow were helpful to guess what an unknown word means because we could have focused on red things only, for instance. | B: The words that surrounded the new words helped me a lot.

The choice-G group participants also implemented a hypothesizing strategy. Because there were three definitions for each target word, they hypothesized their future actions based on each definition. Moreover, based on the available contextual clues and guided by their inferencing strategies, they reshaped their hypotheses. For example, for the word *sign*, participants A and B had three definitions and, thus, developed three hypotheses that determined their future actions.

A: Check the door sign; then, pick up the nails and plank. It says... | B: {Persian definitions for ember, sign, and shack}. Tap on {Persian definition for shack} to check. | A: {Taps} | B: Now we know that it is not {Persian definition for shack}. Check with {Persian definition for embers} now.

Table 10. Table of frequencies for the strategies

	Meaning-G		Choice-G	
Information Search	156	35.2%	106	18.6%
Negotiation	81	18.3%	150	26.3%
Turn-taking	38	8.6%	61	10.7%
Trial-and-Error	40	9%	49	8.6%
Review	56	12.7%	18	3.1%
Enhancing	71	16%	–	–
Inferencing	–	–	62	10.9%
Hypothesizing	–	–	124	21.8%
Total	**442**		**570**	

4 DISCUSSION

The main goal of this study was to investigate the effect of DGBVL activities, modified by meaning-given and multiple-choice glosses, on IVL. For the first question, the results supported the previous studies and revealed that DGBVL activities, regardless of their modifications, could enhance IVL and increase receptive and productive knowledge gains [4], [5], [15], [24]. This can be explained in light of qualitative findings. Evidently, the game and the game-guide could offer a rich context that facilitates guessing, comprehension, and form-meaning links. In other words, inherent elements of digital games such as graphics, audio, animations, etc. in addition to the game-guide text could have enriched the context [4], [19], [20]. Hence, factors that can strengthen form-meaning links were enhanced and resulted in significant IVL gains. Referring to Paivio's dual coding theory, we can explain further that IVL gains were high because the digital game and the game-guide could deliver multimodal inputs through different channels, such as visual, aural, and textual,

which could lead to the generation of meaningful links between the target words' form and meaning.

Another important finding is that DGBVL activities with meaning-G and choice-G could enhance IVL gains and quality more effectively than the non-gloss activity (refer to Table 9). This finding enhances the previous findings [2], [33], [36]. One reason is that glosses can make inputs more cognitively manageable for learners [2]. Think-aloud data showed that due to glosses, participants employed different strategies to deal with both activity completion and target words. In other words, glosses assisted them with input management through the provocation of necessary processes for a significant IVL. Therefore, participants with glosses could benefit from the DGBVL activities significantly more than no gloss participants. Moreover, because glosses can draw learners' attention to both the form and meaning of the target words and, hence, the generation of strong form-meaning connections [26], DGBVL activities with glosses could be more successful in IVL enhancement. Another explanation is that because participants needed to go back and forth between the glosses, game-guide text, and the game, they had a chance to process the target words' form and meaning more deeply than the control group [13]. Then, participants had the chance to increase the quality and quantity of their target word acquisition owing to factors such as frequency of exposure, saliency, attention [26], [34], and proper strategy use [10].

Surprisingly, the results, however, showed that multiple-choice glosses did not enhance incidental productive knowledge gain as well as meaning-given glosses (refer to Table 9). Although this result seems to agree with the finding by [12], it is contradictory to the majority of previous studies [2], [26], [36]. The unexpected effect of multiple-choice glosses can be explained by comparing the categories found through the content analysis of think-aloud data [7]. Because two sets of strategies, i.e., general and exclusive, were found, the comparison was done from two different perspectives.

Firstly, concerning general strategies, Table 10 shows that in the *information search* strategy, there is a significant difference between groups. Participants in the meaning-G group searched for information (156) more frequently than those in the choice-G group (106) did. In other words, meaning-G participants were exposed to the text and the target word forms more frequently. The recorded video clips showed that meaning-G referred to the game guide text as their primary source for seeking information more frequently than other groups. Thus, more exposure to the target words increased the chance of form acquisition for meaning-G participants. This is in line with studies that emphasize the role of frequency of exposure [9], [17], [35].

The above discussion is also enhanced by the frequencies of the *review* strategy. Meaning-G participants reviewed (56) their previous actions significantly more often than choice-G participants did (18), which indicates even more exposure to forms. Such analyses enhance acquisition because they encourage language learners to revise and develop their linguistic knowledge. Thus, meaning-G had a superior performance due to the frequency of exposure.

The *negotiation* strategy can be another contributing factor to differences in performance between the groups. Table 10 shows that there is a significant difference in employing *negotiation* between choice-G (150) and meaning-G (81) which can explain choice-G performances. Choice-G participants might have had too many options to follow because of their higher number of negotiations [25]. In other words, they needed to work with too many ideas and hypotheses, which might have caused a cognitive overload. Subsequently, their memory did not have enough room

for processing the target word forms effectively. This could have also forced them to rely on hypothesizing (124) more than on the other general strategies. The frequency of the *trial-and-error* (49) strategy confirms this reasoning because it was less than expected for the choice-G. In other words, the *trial-and-error* strategy was not used more often because choice-G participants' minds were flooded with hypotheses and options. Moreover, the higher frequency of *negotiation* strategy might have caused the choice-G group's lower number of exposures to the target word forms because they were so busy negotiating and hypothesizing that they might have ignored noticing the English forms of the target words in the text [27].

Secondly, comparing the exclusive strategies can also help with the explanation of the results. [10] discussed that context can affect the learner's choice of strategy. Therefore, the choice-G group might have selected strategies that were inappropriate for the interactive multimedia context and led to their poor productive knowledge gain.

The inverse effect of DGBVL context on the learning outcome is worth discussing. [16] emphasized the role of cognitive overload in learning from multimedia materials. According to [21], cognitive overload can hamper learning. Hence, choice-G participants might have experienced cognitive overload due to the DGBVL activity and context. But how? According to [16], the *split-attention* effect, or the inappropriate distribution of learners' attention between sources of information, can be the reason. In other words, the choice-G participants needed to focus on both the game guide and playing the game. However, because multiple-choice glosses needed attention too, they needed to devote more attentional sources to process information, which might have led to their cognitive overload and insufficient target word form processing.

However, for the meaning-G group, the split-attention effect was probably prevented by the *signaling* technique [16], or cues–in this case, marginal glosses—that helped with processing the material effectively.

4.1 Conclusion

This study has limitations. Firstly, we only used concrete nouns, and this must be considered in future studies. Moreover, the outcomes of this study are limited to commercial adventure (hidden object) games only. Hence, in future studies, the genre of digital games should be considered as a factor. Finally, in this study, only one component of word knowledge [17], i.e., form-meaning connection, was studied.The main goal of the current study was to explore the effectiveness of two crucial gloss types, namely, meaning-given and multiple-choice, on IVL through DGBVL activities. The outcomes of this study, firstly, showed that DGBVL activities can boost IVL without regard for the inclusion or type of glosses. The second major finding was that adding glosses to DGBVL activities can enhance IVL more effectively. The last crucial finding was that multiple-choice glosses did not support IVL productive knowledge gain as effectively as meaning-G did. Based on these findings, two general conclusions can be drawn. Firstly, using DGBVL activities and glosses is encouraged because DGBVL activities in any form or design can support IVL. Secondly, as [12] warned, glosses, however, must be implemented in the DGBVL activities carefully. If a DGBVL activity is considered for out-of-school or self-learning, teachers are advised to implement meaning-given glosses and avoid multiple-choice glosses because they can lead to cognitive overload and weak form-meaning links.

5 REFERENCES

[1] E. A. Alghamdi, "Effects of using multimodal glosses in video games to enhance incidental vocabulary learning and retention," in *Handbook of Research on Integrating Technology Into Contemporary Language Learning and Teaching*, B. Zou and M. Thomas, Eds., 2018, pp. 427–448. https://doi.org/10.4018/978-1-5225-5140-9.ch021

[2] F. Boers, "Glossing and vocabulary learning," *Language Teaching*, vol. 55, no. 1, pp. 1–23, 2022. https://doi.org/10.1017/S0261444821000252

[3] E. Charters, "The use of think-aloud methods in qualitative research: An introduction to think-aloud methods," *Brock Education: A Journal of Educational Research and Practice*, vol. 12, no. 2, pp. 68–82, 2003. https://doi.org/10.26522/brocked.v12i2.38

[4] H. J. H. Chen, H. L. Hsu, and Z. H. Chen, "A study on the effect of adding L1 glosses in the subtitle of an adventure game for vocabulary learning," *Interactive Learning Environments*, pp. 1–17, 2020. https://doi.org/10.1080/10494820.2020.1863233

[5] H.-J. H. Chen, H.-L. Hsu, Z.-H. Chen, and A. G. Todd, "Investigating the impact of integrating vocabulary exercises into an adventure video game on second vocabulary learning," *Journal of Educational Computing Research*, vol. 59, no. 2, pp. 318–341, 2020. https://doi.org/10.1177/0735633120963750

[6] M. H. Chen, W. T. Tseng, and T. Y. Hsiao, "The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis," *British Journal of Educational Technology*, vol. 49, no. 1, pp. 69–77, 2018. https://doi.org/10.1111/bjet.12526

[7] S. Elo and H. Kyngäs, "The qualitative content analysis process," *Journal of Advanced Nursing*, vol. 62, no. 1, pp. 107–115, 2008. https://doi.org/10.1111/j.1365-2648.2007.04569.x

[8] K. A. Ericsson and H. A. Simon, "Verbal reports as data," *Psychological Review*, vol. 87, no. 3, pp. 215–251, 1980. https://doi.org/10.1037/0033-295X.87.3.215

[9] C. A. Fraser, "Lexical processing strategy use and vocabulary learning through reading," *Studies in Second Language Acquisition*, vol. 21, no. 2, pp. 225–241, 1999. https://doi.org/10.1017/S0272263199002041

[10] Y. Gu and R. K. Johnson, "Vocabulary learning strategies and language learning outcomes," *Language Learning*, vol. 46, no. 4, pp. 643–679, 1996. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x

[11] J. H. Hulstijn, M. Hollander, and T. Greidanus, "Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words," *Modern Language Journal*, vol. 80, no. 3, pp. 327–339, 1996. https://doi.org/10.1111/j.1540-4781.1996.tb01614.x

[12] J. H. Hulstijn, "Retention of inferred and given word meanings: Experiments in incidental vocabulary learning," in *Vocabulary and Applied Linguistics*, P. J. L. Arnaud and H. Béjoint, Eds., Palgrave Macmillan, London, pp. 113–125, 1992. https://doi.org/10.1007/978-1-349-12396-4_11

[13] G. M. Jacobs, P. Dufon, and F. C. Hong, "L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge," *Journal of Research in Reading*, vol. 17, no. 1, pp. 19–28, 1994. https://doi.org/10.1111/j.1467-9817.1994.tb00049.x

[14] S. Kaeophanuek and K. Chaisriya, "A conceptual framework for an adaptive tutorial system with gamification to enhance digital literacy," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 19, pp. 44–60, 2022. https://doi.org/10.3991/ijim.v16i19.32997

[15] V. Maloku, M. Ebner, and M. Ebner, "A mobile application for school children controlled by external bluetooth devices," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 12, no. 5, pp. 81–96, 2018. https://doi.org/10.3991/ijim.v12i5.8961

[16] R. E. Mayer and R. Moreno, "Nine ways to reduce cognitive load in multimedia learning," *Educational Psychologist*, vol. 38, no. 1, pp. 43–52, 2003. https://doi.org/10.1207/S15326985EP3801_6

[17] I. S. P. Nation, *Learning Vocabulary in Another Language*, Cambridge: Cambridge University Press, 2001. https://doi.org/10.1017/CBO9781139524759

[18] A. Paivio, "Dual coding theory: Retrospect and current status," *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, vol. 45, no. 3, pp. 255–287, 1991. https://doi.org/10.1037/h0084295

[19] S. Papadakis and M. Kalogiannakis, "Evaluating the effectiveness of a game-based learning approach in modifying students' behavioural outcomes and competence, in an introductory programming course. A case study in Greece," *International Journal of Teaching and Case Studies*, vol. 10, no. 3, pp. 235–250, 2019. https://doi.org/10.1504/IJTCS.2019.10024369

[20] M. Pivec, O. Dziabenko, and I. Schinnerl, "Aspects of game-based learning," in *3rd International Conference on Knowledge Management*, Graz, Austria, 2003, pp. 216–225.

[21] J. L. Plass, R. Moreno, and R. Brünken, Eds., *Cognitive Load Theory*, New York: Cambridge University Press, 2010. https://doi.org/10.1017/CBO9780511844744

[22] N. Ramezanali and F. Faez, "Vocabulary learning and retention through multimedia glossing," *Language Learning & Technology*, vol. 23, no. 2, pp. 105–124, 2019. https://doi.org/10125/44685

[23] A. Rasti-Behbahani, "Why digital games can be advantageous in vocabulary learning," *Theory and Practice in Language Studies*, vol. 11, no. 2, 2021. https://doi.org/10.17507/tpls.1102.01

[24] Rasti-Behbahani and M. Shahbazi, "Investigating the effectiveness of a digital game-based task on the acquisition of word knowledge," *Computer Assisted Language Learning*, vol. 35, no. 8, pp. 1920–1945, 2020. https://doi.org/10.1080/09588221.2020.1846567

[25] S. Rott, "Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened," *Reading in a Foreign Language*, vol. 17, no. 2, pp. 95–124, 2005.

[26] S. Rott and J. Williams, "Making form-meaning connections while reading: A qualitative analysis of word processing," *Reading in a Foreign Language*, vol. 15, no. 1, pp. 45–75, 2003.

[27] S. Rott, J. Williams, and R. Cameron, "The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention," *Language Teaching Research*, vol. 6, no. 3, pp. 183–222, 2002. https://doi.org/10.1191/1362168802lr108oa

[28] D. Safitri *et al.*, "Improvement of student learning motivation through word-wall-based digital game media," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 6, pp. 188–205, 2022. https://doi.org/10.3991/ijim.v16i06.25729

[29] A. J. Sökmen, "Current trends in teaching second language vocabulary," in *Vocabulary: Description, Acquisition, and Pedagogy*, N. Schmitt and M. McCarthy, Eds., Cambridge: Cambridge University Press, 1997, pp. 237–257.

[30] "Open statistical software for the desktop and cloud," Jamovi. Retrieved from https://www.jamovi.org, 2024.

[31] M. Tomczak and E. Tomczak, "The need to report effect size estimates revisited. An overview of some recommended measures of effect size," *Trends in Sport Sciences*, vol. 1, no. 21, pp. 19–25, 2014.

[32] N. Vidakis *et al.*, "Generating education in-game data: The case of an ancient theatre serious game," in *Proceedings of the 11th International Conference on Computer Supported Education: CSEDU,* vol. 1, 2019, pp. 36–43. https://doi.org/10.5220/0007810800360043

[33] P. Warren, F. Boers, G. Grimshaw, and A. Siyanova-Chanturia, "The effect of gloss type on learners' intake of new words during reading: evidence from eye-tracking," *Studies in Second Language Acquisition*, vol. 40, no. 4, pp. 883–906, 2018. https://doi.org/10.1017/S0272263118000177

[34] Y. Watanabe, "Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary," *Studies in Second Language Acquisition*, vol. 19, no. 3, pp. 287–327, 1997. https://doi.org/10.1017/S027226319700301X

[35] S. Webb, "Incidental vocabulary learning," in *The Routledge Handbook of Vocabulary Studies*, S. Webb, Ed., New York: Routledge, pp. 225–239, 2020. https://doi.org/10.4324/9780429291586-15

[36] A. Yanagisawa, S. Webb, and T. Uchihara, "How do different forms of glossing contribute to L2 vocabulary learning from reading?: A meta-regression analysis," *Studies in Second Language Acquisition*, vol. 42, no. 2, pp. 411–438, 2020. https://doi.org/10.1017/S0272263119000688

[37] M. Yoshii, "L1 and L2 glosses: Their effects on incidental vocabulary learning," *Language Learning & Technology*, vol. 10, no. 3, pp. 85–101, 2006.

6 AUTHORS

Amin Rasti-Behbahani is an English lecturer at Dhofar University, Centre for Preparatory Studies. His research interests are digital-game-based learning, and vocabulary acquisition.

Mohammed Abdulgalil Abugohar is an English lecturer at Dhofar University, Centre for Preparatory Studies. His research interests are language technologies, assessment, and pedagogy (E-mail: mabugohar@du.edu.om).

Nagwa Mohammed Khallaf is an English lecturer at Dhofar University, Centre for Preparatory Studies. Her research interests are pedagogy, curriculum & instruction, and language technology.