



Journal of Teaching Language Skills
(JTLS)

39(2), Summer 2020, pp. 89-119

DOI: 10.22099/jtls.2020.37659.2850

Online ISSN: 2717-1604

Print ISSN: 2008-8191

Research Paper

The Role of Post-listening Vocabulary-focused Activities and Multimedia Presentation on Receptive and Productive Vocabulary Development

Salva Kazemipour Khabbazi

Ali Akbar Ansarin**

*

*Department of English Language and Literature, Faculty of Persian
Literature and Foreign Languages, University of Tabriz, Tabriz, Iran*

Abstract

Designing appropriate materials and activities to enhance vocabulary learning is one of the primary goals of language courses. Among the claims about efficient pedagogical tasks is the Involvement Load Hypothesis (Laufer & Hulstijn, 2001) according to which vocabulary development is contingent on the amount of cognitive process a task involves. Building on the previous research on this hypothesis, this study seeks to explore the extent to which three post-listening tasks of different involvement loads along with single-versus dual-mode glosses affect receptive and productive vocabulary learning of advanced English as a foreign language (EFL) learners. The participants ($n = 204$) listened to an expository text while twenty lexical items were presented through text-only or text-plus-picture glosses. Subsequently, they were assigned to one of the task type groups (sentence writing, gap-fill, and paragraph writing). Paribakht and Wesche's (1997) Vocabulary Knowledge Scale was employed as the pre and post-tests. The results pointed to the efficacy of dual-mode over single-mode glosses concerning both receptive and productive vocabulary learning. The sentence writing task produced the highest scores in the vocabulary tests, whereas the gap-fill and paragraph writing tasks did not significantly differ. Hence, the results only partially confirmed the Involvement Load Hypothesis. The current study corroborates other studies verifying the modality effect on vocabulary learning and suggests that task type is a significant factor in EFL vocabulary learning.

Received: 07/07/2020 Accepted: 18/10/2020

* PhD Candidate, Email: Kazemipour_s@yahoo.com, Corresponding Author

** Professor, Email: Ansalak@gmail.com

Keywords: Single-mode glosses, Dual-mode glosses, Task involvement load, Receptive vocabulary development, Productive vocabulary development

While improving second language (L2) vocabulary is one of the main goals of language teachers and learners, developing activities to achieve this goal is not always well-defined due to a large number of factors affecting it (Schmitt, 2008). Although listening and reading have always been regarded as principal sources of meaning-oriented input for EFL (English as a foreign language) vocabulary acquisition (Vidal, 2011), there is a growing emphasis on the significance of explicit and elaborate word-focused activities (Laufer, 2005; Schmitt, 2008). Among the studies supporting such activities, Laufer and Hulstijn (2001) put forward the Involvement Load Hypothesis (ILH), according to which tasks that bring about higher cognitive loads improve vocabulary acquisition to a large extent. The constituents of need, search, and evaluation in each task determine how effective a task is in vocabulary development. Previous studies on the effect of task involvement load, however, have only investigated the hypothesis concerning single tasks or in post-reading activities (Kim, 2011; Laufer, 2003; Yang, Shintani, Li, & Zhang, 2017).

Likewise, cognitive approaches to language learning have also emphasized the effect of multimedia learning on vocabulary development. According to Cognitive Theory of Multimedia Learning (CTML) proposed by Mayer (2005, 2009) and Cognitive Load Theory (CLT) (Sweller, 2010; Sweller, Ayres, & Kalyuga, 2011), written information together with visual aid will possibly overload the working memory by competing for its visual resources. Consequently, presenting materials in an audiovisual manner is proposed to impose equal loads on two of the working memory channels; that is, visual and auditory (Inan, et al., 2013). One of the means to realize this aim

is the use of glosses to represent new lexical items in the course of reading and listening activities (Nation, 2013). The impact of type (picture, video, or text) and mode (single, dual or multimodal) of glosses along with the system of instruction (system- or self-paced) are open to dispute (Akbulut, 2007; Boers, Warren, He, & Deconinck, 2017; Mohsen & Balakumar, 2011). Nevertheless, the modality effect and the use of glosses have been explored on text comprehension or vocabulary learning through reading, with very few studies on lexical retention through listening (Çakmak & Erçetin, 2017). Hence, there is an apparent need for further research on the use of glosses in a listening text to enhance vocabulary acquisition. Moreover, research on task-induced involvement load has just examined the effect of tasks on reading comprehension, or on vocabulary learning through a reading text, and not through a listening text. Thus, this paper first aimed to look more thoroughly into the modality effect and the use of different gloss modes to improve both receptive and productive vocabulary development. Second, it sought to inspect if tasks of different involvement loads mediate lexical retention through listening to an audio text and in post-listening activities.

Multimedia Glossing and Vocabulary Learning

One particular concern of the present study was to explore the effect of two gloss modes on receptive and productive vocabulary learning of EFL learners. According to CLT, selection, organization, and integration of visual and auditory information occur through working memory (WM). This information is further conveyed to the long-term memory (Sweller et al., 2011). Contrary to the long-term memory, the capacity of WM is very limited (Sweller, 2005; 2010). Different subcomponents of WM are responsible for different modes of input. Therefore, textual material accompanied by visual aid supposedly overloads WM by merely engaging its visual resources

(Mayer, 2005, 2009). On the other hand, audiovisual materials, which include a listening text plus visual presentations, occupies both visual and auditory limited channels. Audiovisual materials are believed to decrease the load on WM (Mayer, 2009; 2014). Research has pointed to the advantages of multimedia instruction (Austin, 2009; Brünken, Plass, & Leutner, 2004; Brunyé, Taylor, Rapp, & Sapiro, 2006; Gyselinc, Jamet, & Dubois, 2008; Kozan, Erçetin, & Richardson, 2015). Moreover, the provision of glosses along with the texts to enhance text comprehension and vocabulary learning has been proposed (Nation, 2013). Glosses are translations or meanings of unknown words in a text. They can take the form of texts, images, or videos. Facilitating text comprehension, presenting correct meanings to prevent incorrect guesses, less time-consumption compared to dictionary use, and fostering vocabulary learning are among multimedia glosses' advantages (Lenders, 2008; Nation, 2013). Several studies have pointed to the effectiveness of glosses in comprehension and vocabulary retention (He, 2019; Jung, 2016; Rassaei, 2017; Webb, 2010). Nonetheless, there is disagreement on the efficiency of the gloss type; that is text, image, or video; single, dual or multimodal modes; and whether system- or self-paced pace instructions are more effective (Acha, 2009; Al-Seghayer, 2001; Boers, Warren, He, et al., 2017; Chun & Plass, 1996a, 1996b; Yeh & Wang, 2003).

Furthermore, most of the experiments have focused on the effect of audiovisual aid on reading comprehension, or on vocabulary learning through reading. Very few have examined the effect through a listening text (Çakmak & Erçetin, 2017). Taken together, there is an apparent disparity observed in the findings of the investigations plus the gap in the literature for the modality effect in listening and vocabulary acquisition through listening to an auditory text. Therefore, one of the objectives of the present research was to look into

the effect of single- and dual-mode multimedia glosses on both receptive and productive vocabulary development of proficient EFL learners.

Involvement Load Hypothesis and Post-listening Word-focused Activities

The Involvement Load Hypothesis (ILH), put forward by Laufer and Hulstijn (2001), maintains that retention of unknown words is contingent upon three factors: need, search, and evaluation, which might be present or absent in a task. These motivational-cognitive constructs determine the involvement load of the tasks and the degree to which the tasks will lead to vocabulary retention. It is assumed that the higher the involvement load of a task, the more effectual it would be. Need is the motivational dimension of ILH. It is the drive to adhere to the requirements of the task, either internally or externally imposed. It is moderate with an index of 1 when it is externally induced, for example by the teacher. If the need is intrinsically induced by the learner, it is strong with an index of 2. The two cognitive dimensions of the ILH; that is search and evaluation, primarily involve noticing and allocating attention to the forms and meanings of vocabulary items. When learners strive to match the meaning and form of an unknown word with the help of external sources such as the teacher or a dictionary, search is present (index of 1). It has an index of 0 when it is absent. With two degrees of prominence, evaluation can be moderate (index = 1) or strong (index = 2). When comparing some words or the meanings of one word is necessitated, it is moderate. When it requires learners to assess the appropriateness of using a word in a new context, evaluation is strong. ILH has a quantifiable nature that allows researchers and teachers to measure the effectiveness of different tasks in fostering vocabulary development. The sum of the indices of these three constructs would be the involvement load of a task ranging from 0 to 5. For example, according to Laufer and Hulstijn (2001), a composition writing task, in which the concepts

are chosen by the teacher and the learners must look up the form in the target language, the involvement load index is 4 (moderate need, search, and strong evaluation). However, the same writing task can be done with learners selecting the concepts and finding their target equivalent. In this case, the involvement load index is 5 (strong need, search, and strong evaluation). It is presumed that higher involvement loads would lead to better learning. Nevertheless, evaluation has been stated to be a stronger determiner of task efficacy (Kim, 2011; Yang et al., 2017).

In their study, Laufer and Hulstijn (2001) could provide partial and full support to ILH. Through three vocabulary tasks of varying involvement loads, they examined their effect on incidental receptive vocabulary development. The reading comprehension task (index = 1, the need is moderate, no search or evaluation) involved reading the text with the glossed words. The gap-fill task was a text with deleted target words for the participants to fill the blanks (index = 2, moderate need and evaluation, but search is absent). The last task (index = 3, as the need is moderate, search is absent, and evaluation is strong) asked the participants to write a composition using the selected words. The experiment in one of the settings fully supported the ILH with the composition group outperforming the gap-fill group along with the gap-fill group outperforming the reading group. In the experiment in the second context of the research, the ILH was only partially confirmed. The composition task was found to be surpassingly more functional than the other tasks, but the difference in the gap-fill and reading group performance was insignificant. Laufer and Hulstijn (2001) concluded that “evaluation” might be more effective than the other two constructs.

Only few of the follow-up studies provided full support for the hypothesis (Min, 2008), with most of them only partially supporting the ILH with mixed results (Bao, 2015, Kim, 2011; Yang et al., 2017). They have indicated that

there exist other variables that affect the results. For example, Kim (2011) conducted a replicate of the study by Laufer and Hulstijn (2001), comparing two proficiency levels and controlling time on task. As the immediate and delayed posttests respectively provided partial and full support for the ILH, the findings were comparable to those of Laufer and Hulstijn (2001). Bao (2015) also investigated the hypothesis in a sentence-context design and by output tasks. In this study, it was observed that the tasks with the same involvement load do not necessarily lead to similar vocabulary knowledge. The tasks with higher involvement loads may not lead to better learning either. Other reasons besides the involvement load index of the tasks should be explored. All the studies on ILH, however, shared reading as the means to present the new vocabulary items. No study, as far as we are aware, has explored the effect of post-listening vocabulary tasks with varying involvement loads on both receptive and productive vocabulary development of proficient EFL students. Therefore, in addition to exploring the differential effect of single- versus dual-mode glosses, the second objective of the current study was to look into how word-focused tasks of different involvement loads can mediate receptive and productive vocabulary development of words presented through two gloss modes in a listening context. To address the mentioned gaps, this research sought to answer the subsequent questions:

1. Do different modes of input (i.e. text-only versus text-plus-picture glosses) vary in their efficacy in receptive and productive vocabulary development of EFL learners?
2. Do tasks of different involvement loads (i.e. sentence writing, gap-fill, and paragraph writing) lead to differential gains of receptive and productive word knowledge for EFL learners?

3. Are there any interaction effects of different modes of input (i.e. text-only versus text-plus-picture glosses) and tasks of different involvement loads on receptive and productive vocabulary development of EFL learners?

Method

Participants

The research was conducted with the participants ($n = 204$, 146 females and 58 males; $M = 20.47$ years) of three public universities of Iran. They were enrolled in six intact classes (two classes in each university including juniors and seniors) in English language and literature major in the English Departments. As their scores were beyond 75 percent on the special English test of the National University Entrance Examination, they can be considered high-intermediate or advanced learners (as B2 level or independent users according to CEFR). There were initially 240 participants with some dropping out of the study or being excluded from the analyses as their z-scores were beyond the normal range.

Research Design

A pretest, treatment, and posttest design was followed in the current study. There was the vocabulary pretest in the first session and the treatment together with the immediate vocabulary posttest after two weeks in the second session. The study was conducted as an independent experiment and the treatment was unrelated to the main courses of the participants. The vocabulary tests measured both receptive and productive EFL vocabulary knowledge (dependent variables) of the participants. In the treatment session, the participants listened to a listening text throughout which twenty words were presented. The vocabulary items were presented using single-mode (text-only) or dual-mode (text-plus-picture) gloss types. Immediately after

listening to the text, the participants answered five comprehension questions, which was subsequently followed by the tasks. The participants in each class were randomly divided into three groups and completed one of the three vocabulary tasks; that is Paragraph writing, Gap-fill, and Sentence writing. Then, they were asked to complete the vocabulary posttest. Regarding the input modality, the study had a within-subjects design, while apropos of task type, it had a between-subjects design. In none of the stages of the research, were the participants allowed to discuss the vocabulary items or use dictionaries or any other devices.

Materials and Instruments

The expository listening text used in the study was a lecture about opossum and its defensive mechanisms. As the participants were advanced learners of EFL, the text was selected from the Third Edition of Longman Preparation Course for the TOEFL iBT Test (Phillips, 2015, p.172 [*mp3 039-040*]). Initially, twenty-five vocabulary items were chosen, five of which were omitted from the list after a pilot test with similar participants. Ten of the twenty items presented through text-only gloss mode (i.e. the item accompanied by its Persian translation) were: bluff (v.), claw (n.), crawl (v.), foul (adj.), hiss (v.), intimidation (n.), project (onto) (v.), reiterate (v.), snarl (v.), stiffen up (v.). The other ten vocabulary items; that is cling (v.), emit (v.), marsupial (n.), opossum (n.), outrun (v.), play possum (phr.), pouch (n.), racket (n.), stench (n.), venom (n.), were presented through text-plus-picture gloss mode (i.e. the word with Persian translation together with a picture representing the meaning).

The video including the listening text with corresponding slides of translations or translations and pictures was presented via the video projector in the classes. The slides appeared on the screen instantaneously after the

intended words were pronounced by the lecturer. As the aim was to measure the amount of vocabulary acquisition in a multimedia environment, system-paced instruction was adopted. It means the participants could not control the speed of presentation, nor could they pause the listening text or check the glosses again. After piloting the whole procedure with similar participants, the researchers decided that each slide remain on the screen for 4 seconds.

Post-listening activities. Following the listening text, the participants were randomly divided into three groups of Paragraph writing ($n = 69$), Gap-fill ($n = 68$), and Sentence writing ($n = 67$). Although the allotted time for all tasks was 30 minutes, the participants were asked to write down the amount of time they spent on the task on the papers. It was anticipated that the different nature of the tasks together with varying capacities of the participants would lead to differential allocated time, which might have affected the results.

The paragraph writing task required the participants to summarize the listening text in no more than 100 words. The involvement load index of this task was 0 (no need, no search, and no evaluation), as the participants were not asked to use the selected vocabulary, nor was the focus of the task on them. Subsequent text analysis revealed that the chosen words were not used in the participants' writings. The gap-fill task required the participants to read a text of 318 words about opossum and to fill in the blanks choosing among the words which included twenty words with four distractors. There is no search with a moderate need (as the completion of the task was required by the researcher rather than by the learners to learn the meanings) and moderate evaluation (the task required comparisons among the meanings of the words to fill in the gaps), so the involvement load index was 2. The participants of the Sentence writing group needed to provide an authentic sentence for each word. Each sentence had to include more than seven words. They were provided with the L1 translation of the words to refer to during the task if

needed. The involvement load index of this task is 3, as there is no search, moderate need (likewise, the task is required by someone other than the learners and their needs) and strong evaluation (the task included assessment of the words in a new context and using the words in original sentences) (Kim, 2011).

Vocabulary test. To measure the vocabulary knowledge, Paribakht and Weshche's (1997) Vocabulary Knowledge Scale (VKS) was employed as the pretest and posttest. Previously used in studies on vocabulary knowledge (File & Adams, 2010; Kim, 2011; Yang et al., 2017), this scale has a five-point scale of vocabulary knowledge which enables not only the assessment of the participants' performance but also their own judgment of the amount of knowledge they possess regarding the target words. It fits the purpose of this study as it could evaluate both receptive and productive vocabulary knowledge. In addition, even minor changes in the participants' vocabulary development could be tracked through the pre and post-tests.

The pre and post-tests included 24 items, 20 target items selected from the listening text as well as 4 distractors. The vocabulary items appeared in a random order in the pre and post-tests besides the sentence writing and gap-fill tasks to avoid rote learning and carry-over effect from previous sessions. Each participant was given two scores according to their performance on the VKS tests. One of the scores was based on their performance on the receptive categories; that is, categories 1-4 of the test. The other score was given to their performance on the productive category of the test; that is, category 5. The first and second categories of receptive knowledge were 1: "I do not remember having seen this word before" and 2: "I have already seen this word, but I don't remember the meaning", that only required a tick mark in the box without provision of any answers. The third and fourth categories, 3: "I have already seen this word, and I suppose the meaning is", 4: "I know this word,

it means”, required the provision of a synonym or the Persian translation of the words. To demonstrate their productive knowledge of the words, the participants had to use the words in a grammatically and semantically sentence (category 5: “I am able to use this word in a sentence”).

Scoring of the VKS tests was based on Paribakht and Wesche’s (1997) scoring system in which self-report of the amount of knowledge was transferred into a word knowledge of the item. In this system, ticking the first box implies no knowledge of the word; therefore, no score was given. A score of 1 was awarded to the second category which stated that the word sounds familiar. Scores of 2 and 3 were given to the third and fourth categories respectively. However, if the provided synonym or translation was not correct, a score of 2 was given despite putting a checkmark in these categories. Concerning the productive category; that is, the fifth category, if the sentence was meaningful and semantically appropriate, a score of 4 and if the sentence was also grammatical, a score of 5 was awarded.

All the instructions were given in L1 to ensure comprehension of the procedure. Moreover, as the aim was to investigate intentional vocabulary learning through two gloss modes and post-listening activities, the participants were informed of the forthcoming posttest.

Data Analysis

The participants’ gain scores on the pre and post VKS tests (i.e. posttest - pretest) were calculated for both receptive and productive knowledge in text-only and text-plus-picture gloss modes. Therefore, there were four scores for each participant: receptive and productive knowledge in single gloss mode, as well as receptive and productive knowledge in dual gloss mode. As the input mode was the within-subjects independent variable, while task type was the between-subjects independent variable, we opted for a mixed design. In order

to answer the research questions, a Factorial 2 (two gloss types) \times 3 (three task types) Mixed Analysis of Variance was conducted. To control its effect on the participants' scores, time on task was inserted into the model as the covariate. All the analyses were executed using SPSS version 23.

Results

Descriptive results of the gain scores of the participants in the VSK pre and posttests are represented in Table 1.

Table 1.

Descriptive Statistics for gain Scores on VKS Tests

	Receptive Knowledge		Productive Knowledge	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Single gloss mode	11.92	5.246	22.08	11.767
Dual gloss mode	15.50	5.091	28.29	11.443

n: 204

Aimed to determine the effect of gloss mode and task type on receptive and productive vocabulary development of EFL learners, a mixed factorial ANOVA was performed which gave us the main effect of the two factors together with their interaction and the amount of variance explained by the covariate. The assumptions of the test (homogeneity of regression, Box's equality of covariance, homogeneity of variance, and linearity of the relation between the covariate and the dependent variable) were all checked ($p > 0.05$).

Table 2.

Main and Interaction Effects of Gloss Mode, Task Type, and Time on Task on Receptive Vocabulary Development

	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>	<i>Partial Eta Squared</i>
Gloss Mode	166.298	15.524	0.0001	0.072
Task Type	420.184	11.391	0.0001	0.102
TimeOnTask	2.253	0.061	0.805	0.001
GlossMode*TaskType	17.273	1.612	0.202	0.016
GlossMode*TimeOnTask	2.947	0.275	0.601	0.001

Table 2 and 3 summarize the results for the main and interaction effect of gloss mode, task type, and time on task on receptive and productive vocabulary development.

Single- versus Dual-Mode Glosses and Vocabulary Development

The first question concerned whether there was a statistically significant mean difference between the gain receptive and productive scores for words under text-only and text-plus-picture gloss modes. The results indicate that there was a significant variation in the effect of two gloss modes on both receptive and productive vocabulary learning. For receptive vocabulary development of the participants, a main effect of input mode [$F(1, 200) = 15.524, p = 0.0001, \text{partial } \eta^2 = 0.072$] was found. According to the results, there was a significant difference between the receptive gain scores for words under text-only ($M = 11.92, SD = 5.24$), and under text-plus-picture glosses ($M = 15.50, SD = 5.09$). With a medium effect size (partial $\eta^2 = 0.072$), text-plus-picture glosses led to increased receptive vocabulary development.

Table 3.

Main and Interaction Effects of Gloss Mode, Task Type, and Time on Task on Productive Vocabulary Development

	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>	<i>Partial Eta Squared</i>
Gloss Mode	206.141	4.028	0.046	0.020
Task Type	3196.577	20.127	0.0001	0.168
TimeOnTask	45.736	0.288	0.529	0.001
GlossMode*TaskType	212.866	4.160	0.017	0.040
GlossMode*TimeOnTask	29.335	0.573	0.450	0.003

Similarly, productive gain scores of vocabulary in text-only ($M = 22.08$, $SD = 11.76$), and text-plus-picture glosses ($M = 28.29$, $SD = 11.44$), were found to be significantly different. The main effect of input mode on productive vocabulary development was also significant [$F(1, 200) = 4.028$, $p = 0.046$, partial $\eta^2 = 0.020$]. Therefore, the null hypothesis, suggesting that single and dual gloss modes do not vary in their efficacy on receptive and productive vocabulary development, was rejected. The results supported the modality effect by demonstrating multimedia gloss mode effect on both receptive and productive vocabulary development, with dual-mode glosses (i.e. text-plus-picture) leading to more vocabulary learning.

Post-Listening Word-Focused Activities and Vocabulary Development

To answer the second research question of whether tasks of different involvement loads lead to differential gains in receptive and productive vocabulary knowledge, the main effect of task type was looked into. Table 4 presents the descriptive statistics for the performance of the three task type groups on VKS tests.

Table 4.

Descriptive Statistics for the Performance on VKS by Different Task Type Groups

	<i>N</i>	Receptive Single-mode		Receptive Dual-mode		Productive Single-mode		Productive Dual-mode	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Paragraph	69	10.16	3.79	14.26	4.81	16.25	9.59	23.09	10.51
Gap-fill	68	10.69	5.25	14.87	5.31	19.24	10.55	27.54	11.38
Sentence	67	14.97	5.24	17.43	4.62	30.97	9.68	34.42	9.49

As presented in Table 2, the results indicate a significant difference between paragraph writing, gap-fill, and sentence writing conditions. There was a significant main effect of task type on the receptive vocabulary development of the participants [$F(2, 200) = 11.391, p = 0.0001, \text{partial } \eta^2 = 0.102$]. The same held true for productive vocabulary knowledge. The main effect of task type on productive vocabulary knowledge of the participants was statistically significant. The F ratio of productive vocabulary knowledge was $F(2, 200) = 20.127, p = 0.0001, \text{partial } \eta^2 = 0.168$. As the partial eta squared signifies, the effect size of the difference between groups for productive word knowledge is slightly higher than that of receptive word knowledge. A Tukey HSD post hoc test demonstrated that under receptive vocabulary development for both gloss modes, the mean score of the Sentence writing group was significantly higher than the other two groups. This is while the difference between the Gap-fill and the Paragraph writing rendered insignificant. The test also demonstrated that under productive vocabulary development for single-mode glossed words, while the sentence writing task produced the best scores, the other two groups' performance did not differ. Concerning productive vocabulary development under dual-mode glosses, the performance of all three task type groups varied. The Sentence writing group significantly outperformed the other two, with very high and high effect sizes,

and the Gap-fill group performed better than the Paragraph writing group, with a medium effect size. Table 5 summarizes the results of Tukey HSD along with the Cohen's d (Cohen, 1988).

Table 5.

Levels of Significance and Effect Sizes between Three Task Type Groups

	Receptive Single mode		Receptive Dual mode		Productive Single mode		Productive Dual mode	
	<i>Sig.</i>	<i>d</i>	<i>Sig.</i>	<i>d</i>	<i>Sig.</i>	<i>d</i>	<i>Sig.</i>	<i>d</i>
Sentence- paragraph writing	.000	1.05	.001	.67	.000	.54	.000	1.31
Sentence writing- Gap-fill	.000	.81	.008	.51	.000	.25	.001	.65
Paragraph writing- Gap-fill	.794		.751		.187		.001	.42

As participants spent unequal amount of time on the tasks, it was necessary to explore the effect of time on task as well. As Table 2 illustrated, insignificant main effect of time on task on receptive [$F(1, 200) = 0.061, p = 0.805$, partial $\eta^2 = 0.0001$] and on productive vocabulary development [$F(1, 200) = 0.288, p = 0.592$, partial $\eta^2 = 0.0001$] was found. Nor did time on task interact with gloss type (under single gloss mode: [$F(1, 200) = 0.275, p = 0.601$, partial $\eta^2 = 0.0001$], under dual gloss mode: [$F(1, 200) = 0.573, p = 0.450$, partial $\eta^2 = 0.0003$]). The results showed that time on task as a covariate explained an insignificant amount of variance in the dependent measure; that is, receptive and productive vocabulary development. Thus, it can be concluded that task type effect was not mediated by the allocated time on tasks.

On the whole, the results partially supported the ILH. While the sentence writing task led to better results compared to the other two task types, the Gap-fill and Paragraph writing groups did not perform significantly different, except for the productive word knowledge for text-plus-picture glosses. This is despite the differences in the involvement load indices of gap-fill and paragraph writing tasks.

The Interaction between Gloss Mode and Post-Listening Activities

It was observed that gloss mode and task type have independently played a role in vocabulary development (Table 2). The third question concerned the interaction between the gloss mode and tasks of varying involvement loads. As demonstrated in Table 2, the interaction of gloss mode and task type on receptive vocabulary development was insignificant ($F(2, 200) = 17.273, p = 0.202$, partial $\eta^2 = 0.016$) failing to prove that the effect of gloss mode was different among the three groups of participants completing different tasks.

On the other hand, regarding productive vocabulary development, it was observed that the interaction of gloss mode and task type was statistically significant, as presented in table 3 ($F(2,200) = 4.160, p = 0.017$, partial $\eta^2 = 0.040$). This interaction, though not great, implies that the difference in the vocabulary development rate for words glossed textually or visually was not consistent across the three groups of participants completing tasks with various involvement loads.

Discussion

This study aimed at looking into the differences in the acquisition rate of receptive versus productive vocabulary knowledge triggered by two different gloss modes (textual and textual-plus-pictorial). In addition, it attempted to investigate the degree to which task type and involvement load of the task

affect vocabulary learning. Previous studies on the effect of multimedia glosses and tasks with various involvement indices mainly investigated word development in terms of meaning recognition or form recall. The present study extended earlier research by focusing on both receptive and productive vocabulary development. While form recall in other research (Eckerth & Tavakoli, 2010; Hazrat, 2015; Webb & Kagimoto, 2009) implied the provision of L2 form of the word, productive vocabulary knowledge in the present study entailed the ability to utilize the target word in a sentence which is both meaningful and grammatical. It is worth mentioning that the participants completed a pretest to ensure taking into account the amount of their familiarity or unfamiliarity with the target words plus ruling out any pre-existing differences among the subjects. Furthermore, prior to the treatment, the participants were informed of the vocabulary posttest, so any changes in their performances were considered as a result of intentional learning.

In both receptive and productive vocabulary development, significant differences between single and dual gloss modes were yielded, which are indicative of multimedia advantage (Mayer & Moreno, 2003; Mayer, 2009). In tune with several other studies, dual-mode glosses proved to be superior to single-mode ones (Akbulut, 2007; Kozan et al. 2015; Lin & Tseng, 2012; Sadeghi, Khezrlou, & Modirkhameneh, 2017). One of the contributing factors which brings about the modality effect is system-paced instruction. In self-paced instruction, the participants are able to check the unknown words as many times as they need or scroll back and forth through the words as much as they need. In such instruction, the participants control the reading/listening speed. In system-paced instruction, however, the participants have no control over the pace of the presentation and the speed is decided by the researcher/instructors. Most of the studies have favored system-pacing over self-pacing instruction (Ginns, 2005; Schmidt-Weigand, Kohnert, &

Glowalla, 2010; Tabbers, 2002), with some finding no modality effect under self-pacing (e.g., Wouters, Paas, & van Merriënboer, 2009), or reverse modality effect (i.e. when the learners benefit more from textual materials than auditory ones) when the instruction was self-paced (Tabber, Martens, & van Merriënboer, 2004). These findings demonstrate the stronger effect of system-paced materials over self-paced ones.

The results are in line with some other studies which found the modality effect to be influential on the acquisition rate (e.g. Brünken et al. 2004; Jones, 2007; Jones & Plass, 2002; Kozan, et al., 2015; Ramezani & Faez, 2019; Türk & Erçetin, 2012). Our results, however, run counter to the results obtained in a number of studies that did not reveal any differences between different types of gloss modes (Cottom, 2010; Çakmak & Erçetin, 2017; Tebbers et al. 2004). Besides the instruction pace, the proficiency level of the participants has been mentioned as another factor that has caused inconsistencies in the results. According to Schmitt (2010), a threshold level of vocabulary knowledge exists before which learners cannot benefit from exposure to new words through glosses. A meta-analysis on the effect of multimedia glosses versus no glosses on reading comprehension and vocabulary learning indicated that beginner learners do not benefit from glosses as much as advanced learners (Abraham, 2008). While Cottom (2010) and Çakmak & Erçetin (2017), who found no difference between single- and dual-mode glosses, had low proficient participants, the participants in the present study were proficient learners who benefited from dual-mode glosses more. The results also support CLT, according to which audiovisual presentations are superior and lead to more learning (Sweller et al., 2011). Furthermore, more noticing occurs through dual-mode glosses because they attract learners' attention, therefore improving vocabulary learning (Boers, Warren, Grimshaw, & Siyanova-Chanturia, 2017; Nation, 2001).

As an attempt to provide empirical evidence for the ILH (Laufer & Hulstijn, 2001), the present study also examined how tasks with different involvement loads affect vocabulary development. Our overall results provided partial support to the hypothesis in that the participants completing the sentence writing task which induced the highest load indicated superior performance over the others, corroborating previous research on the subject (Kim, 2011; Min, 2008). The performance of the Paragraph writing and Gap-fill groups, however, did not vary. The results of mixed factorial ANOVA suggested a significant main effect of task type on both receptive and productive vocabulary learning under both gloss modes. The Tukey HSD test yielded clearer evidence for the distinction among the three tasks. The Sentence writing group did surpass the Gap-fill and Paragraph writing groups in receptive vocabulary development for verbally and visually glossed words. In both of the input modes, the difference between the Sentence writing and Gap-fill groups was less than that of the Sentence writing and Paragraph writing, implying a greater effect of the gap-fill task compared to the paragraph writing task.

It is interesting that for words under text-plus-picture glosses, the effect sizes of the differences between three task type groups are lower than those of verbally glossed. In other words, the effect size of the difference between the Sentence writing group with the Gap-fill and Paragraph writing groups in text-only mode was $d = .81$, and $d = 1.05$. On the other hand, Cohen's d for the sentence writing versus gap-fill tasks was $d = .51$, and for the sentence writing versus paragraph writing was $d = .67$ in text-plus-picture mode. It was observed that dual-mode glosses decreased the size of the differences between the three task type groups. This point signifies an interaction, though not so great, between task type and input mode. Comparably, Boers, Warren, He, et al. (2017) stated that adding visuals facilitates meaning development, the

aspect assessed in the receptive part of the VKS tests. In the same way, our results confirmed Jones and Plass's (2002) assertion that pictures foster meaning recall in the posttests. In spite of this observed point, mixed factorial ANOVA, with time on task as a covariate, did not demonstrate a significant interaction between gloss mode and task type. Therefore, the effect of gloss mode was consistent through all three task groups. The findings, nevertheless, run counter to the findings of Yang et al. (2017) who detected significant preeminence of both the Sentence writing and Gap-fill groups in enhancing vocabulary development.

With regard to productive vocabulary development, the results are rather different. For words under single-mode glosses, despite a statistically insignificant difference between the Gap-fill and Paragraph writing groups, the former outperformed the latter. Here again, the Sentence writing group's scores were the highest. The fact that the sentence writing and gap-fill tasks, albeit to a lower extent, provided the opportunity to use and learn the new words in a new context is one reason for this superiority over paragraph writing task. The Paragraph writing group, on the other hand, did not use, neither encountered them in a new context, thus the weaker performance.

Nonetheless, dual-mode glosses brought about the most changes in the results of productive vocabulary learning. The results suggested a significant interaction of gloss mode and task type for productive vocabulary knowledge. Although the performance of the Sentence writing group remained superior ($M = 34.42$), the differential effect of the gap-fill and paragraph writing tasks rendered significant. With a larger mean ($M = 27.54$), the Gap-fill group outscored the Paragraph writing group ($M = 23.09$). It was only this aspect of word knowledge that lent full support to the ILH. Seemingly, as shown in Table 3, the interaction of the gloss mode and task type is at its highest for productive vocabulary learning in text-plus-picture mode. The underlying

reason might be that as a result of dual-mode glosses, the Sentence writing and Gap-fill groups' mastery of the word meanings were stronger, which was reinforced by the application of the words in a different context (Nation, 2001). Apparently, the effect of the tasks was strengthened through the dual input mode by constructing a deeper vocabulary knowledge, in Schmitt's words (2010, p. 15). Furthermore, as it is stated elsewhere (Eckerth & Tavakoli, 2012), the higher demands of the sentence writing task led to more learning than the gap-filling task, which in turn benefited the learners more than the paragraph writing task. The results are additionally in line with previous research (Huang, Willson, & Eslami, 2012; Yang et al., 2017) which pointed to the better outcomes of production tasks (i.e. sentence writing task in the present study) which impose more cognitive load and foster more vocabulary learning.

Neither the main effect of time on task nor its interaction with gloss mode was significant. The results imply that dual-mode glosses and productive tasks with strong evaluation (i.e. sentence writing task) would most likely lead to more retention no matter how much time it is spent on completing the tasks. Therefore, the findings of this study support previous research (Kim, 2011; Laufer & Hulstijn, 2001), suggesting that the variation in the amount of vocabulary development is an outcome of the nature of the task rather than the amount of time participants are engaged in the activities.

Conclusion

The present study shows how single- and dual-mode glosses together with tasks of varying involvement loads affect receptive and productive vocabulary development of proficient EFL learners. The findings implied that the modality effect may hold true for both receptive and productive vocabulary learning in system-paced instruction. Text-plus-picture glosses

enhanced vocabulary acquisition more than text-only glosses. Furthermore, the overall results pointed to the positive effect of word-focused post-listening activities. The ILH was partially supported, as the sentence writing task with the highest involvement load produced the best performance in both receptive and productive knowledge. The gap-fill and the paragraph writing tasks, despite different involvement loads, did not bring about different performances in the VKS tests. Accordingly, the results verify that the “evaluation” component of the ILH is the most significant factor in vocabulary learning (Kim, 2011). Furthermore, the interaction of modality effect and task type was only observed for productive vocabulary development. Moreover, time on task was not a significant factor affecting the amount of knowledge gained through the tasks. It is apparent that by presenting the materials through dual-mode glosses plus providing post-listening activities, we would enhance the receptive and productive knowledge of the learners. The post-listening activities which require production seem to benefit learners better as these activities engage learners in absorbing different aspects of the new words (Yang et al. 2017).

To sum, the present study aimed to fill the gaps in the literature on the relationships between the modality effect, task-induced involvement load, and vocabulary learning. It constitutes a significant complement to the previous research in that it brought forward the effect of post-listening word-focused activities and their interaction with the modality effect on vocabulary acquisition of proficient EFL learners. The pedagogical implication of the current study is that EFL teachers and material designers should bear in mind that the materials are best presented with accompanying visual aids and the activities implemented to enhance vocabulary learning should involve tasks that induce higher loads and require the production of the new vocabulary items in new contexts.

References

- Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, 21(3), 199–226. <http://dx.doi.org/10.1080/09588220802090246>
- Acha, J. (2009). The effectiveness of multimedia programs in children's vocabulary learning. *British Journal of Educational Technology*, 40, 23–31. <https://doi.org/10.1111/j.1467-8535.2007.00800.x>
- Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. *Instructional Science*, 35, 499–517. <http://doi:10.1007/s11251-007-9016-7>
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202–232. <https://doi.org/10.1251/25117>
- Austin, K. A. (2009). Multimedia learning: cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules. *Computers & Education*, 53, 1339–1354.
- Bao, G. (2015). Task type effects on English as a Foreign Language learners' acquisition of receptive and productive vocabulary knowledge. *System*, 53, 84–95. <http://dx.doi.org/10.1016/j.system.2015.07.006>
- Boers, F., Warren, P., He, L., & Deconinck, J. (2017). Does adding pictures to glosses enhance vocabulary uptake from reading? *System*, 66, 113–129. <http://dx.doi.org/10.1080/09588221.2017.1356335>
- Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning*, 30(7), 709–725. <http://doi.org/10.1080/09588221.2017.1356335>

- Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: auditory load and modality effects. *Instructional Science*, 32, 115–132. <http://doi:10.1023/B:TRUC.0000021812.96911.c5>
- Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology*, 20, 917–940. <http://doi.org/10.1002/acp.1236>
- Çakmak, F., & Erçetin, G. (2017). Effects of gloss type on text recall and incidental vocabulary learning in mobile-assisted L2 listening. *ReCALL*, 30(1), 24–47. DOI: <http://doi.org/10.1017/S0958344017000155>
- Chun, D. M., & Plass, J. L. (1996a). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183–198. <https://doi.org/10.1111/j.1540-4781.1996.tb01159.x>
- Chun, D. M., & Plass, J. L. (1996b). Facilitating reading comprehension with multimedia. *System*, 24(4), 503–519. [https://doi.org/10.1016/S0346-251X\(96\)00038-3](https://doi.org/10.1016/S0346-251X(96)00038-3)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Cottam, M. E. (2011). *The effects of visual and textual annotations on Spanish listening comprehension, vocabulary acquisition and cognitive load* (Unpublished doctoral dissertation). Arizona, USA: Arizona State University. Retrieved from <http://www.learntechlib.org/p/118131>
- Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research*, 16, 227–252. <http://doi:10.1177/1362168811431377>

- File, A., & Adams, R. (2010). Should vocabulary instruction be integrated or isolated? *TESOL Quarterly*, *44*, 222–249. <http://doi:10.5054/tq.2010.219943>
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*, 313–331. <http://doi:10.1016/j.learninstruc.2005.07.001>
- Gyselinck, V., Jamet, E., & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology*, *22*, 353–374. <http://doi.org/10.1002/acp.1411>
- Hazrat, M. (2015). The roles of working memory components in vocabulary knowledge. *International Journal of English Language and Literature Studies*, *4*(2), 58–67. <http://doi:10.18488/journal.23/2015.4.2/23.2.58.67>
- He, X. (2019). The effect of glossing on incidental vocabulary acquisition in the learning of Chinese as a second language. *Global Chinese*, *5*(1), 25–43. <https://doi.org/10.1515/glochi-2019-0002>
- Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *The Modern Language Journal*, *96*, 544–557. <https://doi.org/10.1111/j.1540-4781.2012.01394.x>
- Inan, F. A., Crooks, S. M., Cheon, J., Ari, F., Flores, R., Kurucay, M., & Paniukov, D. (2013). The reverse modality effect: Examining student learning from interactive computer-based instruction. *British Journal of Educational Technology*, *46*(1), 123–130. <http://doi:10.1111/bjet.12129>
- Jones, L. C. (2007). Effects of collaboration and multimedia annotations on vocabulary learning and listening comprehension. *CALICO Journal*, *24*(1), 33–58. <https://doi.org/10.1558/cj.v24i1.33-58>
- Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The*

- Modern Language Journal*, 86(4), 546–561.
<https://doi.org/10.1111/1540-4781.00160>
- Jung, J. (2016). Effects of glosses on learning of L2 grammar and vocabulary. *Language Teaching Research*, 20(1), 92–112. doi: 10.1177/1362168815571151
- Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 61, 100–140. <http://doi.org/10.1111/j.1467-9922.2011.00644.x>
- Kozan, K., Erçetin, G., & Richardson, J. C. (2015). Input modality and working memory: Effects on second language text comprehension in a multimedia learning environment. *System*, 55, 63–73. <http://dx.doi.org/10.1016/j.system.2015.09.001>
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review*, 59 (4), 567–587. <http://doi:10.3138/cmlr.59.4.567>
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5, 223–250. <https://doi.org/10.1075/eurosla.5.11lau>
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26. <http://doi:10.1093/applin/22.1.1>
- Lenders, O. (2008). Electronic glossing - is it worth the effort? *Computer Assisted Language Learning*, 21(5), 457–481. <http://doi:10.1080/09588220802447933>
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.
- Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52. http://doi:10.1207/S15326985EP3801_6
- Min, H. T. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning* 58(1), 73–115. [http://doi: 10.1111/j.1467-9922.2007.00435.x](http://doi:10.1111/j.1467-9922.2007.00435.x)
- Mohsen, M. A., & Balakumar, M. (2011). A review of multimedia glosses and their effects on L2 vocabulary acquisition in CALL literature. *ReCALL*, 23(2), 135–159. [http://doi: 10.1017/S095834401100005X](http://doi:10.1017/S095834401100005X)
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Harlow: Longman.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.
- Paribakht, S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 147–200). Norwood, NJ: Ablex.
- Phillips, D. (2015). *Longman preparation course for the TOEFL iBT test*. White Plains, NY: Pearson Education.
- Ramezanali, N., & Faez, F. (2019). Vocabulary learning and retention through multimedia glossing. *Language Learning & Technology*, 23(2), 105–124. <https://doi.org/10125/44685>

- Rassaei, E. (2017). Computer-mediated textual and audio glosses, perceptual style, and L2 vocabulary learning. *Language Teaching Research*, 22(6), 657–675. <https://doi.org/10.1177/1362168817690183>
- Sadeghi, K., Khezrlou, S., & Modirhameneh, S. (2017). Calling Iranian learners of L2 English: Effect of gloss type on lexical retention and reading performance under different learning conditions. *Journal of Research in Reading*, 40(S1), S66–S86. <http://doi.org/10.1111/1467-9817.12088>
- Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction*, 20(2), 100–110. <http://doi.org/10.1016/j.learninstruc.2009.02.011>
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12, 329–363. <http://doi.org/10.1177/1362168808089921>
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York: Palgrave Macmillan.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review*, 22, 123–138. <https://doi.org/10.1007/s10648-010-9128-5>.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Tabbers, H. K. (2002). *The modality of texts in multimedia instructions: Refining the design guidelines*. (Unpublished doctoral dissertation).

- Heerlen: Open University of the Netherlands. Retrieved from <http://hdl.handle.net/1820/1025>.
- Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, *74*, 71–81. <https://doi.org/10.1348/000709904322848824>
- Türk, E., & Erçetin, G. (2012). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning*, *27*, 1–25. <http://doi:10.1080/09588221.2012.692384>
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. <http://doi:10.1111/j.1467-9922.2010.00593.x>
- Webb, S. (2010). Using glossaries to increase the lexical coverage of television programs. *Reading in a Foreign Language*, *22*(1), 201–221.
- Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, *43*(1), 55–77.
- Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2009). Observational learning from animated models: Effects of modality and reflection on transfer. *Contemporary Educational Psychology*, *34*, 1–8. <http://doi:10.1016/j.cedpsych.2008.03.001>
- Yang, Y., Shintani, N., Li, S., & Zhang, Y. (2017). The effectiveness of post-reading word-focused activities and their associations with working memory. *System*, *70*, 38–49. <http://doi:10.1016/j.system.2017.09.012>
- Yeh, Y., & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal*, *21*(1), 131–144. <https://www.learntechlib.org/p/97205/>