ISELT 109
Interdisciplinary Studies in English Language Teaching

Pages 109 to 125

Glossed Full Captions Versus Glossed Keyword Captions: Which One Is More Effective for Incidental Vocabulary Learning from Viewing?

Maryam Ehsani [1]*, Hossein Karami [1]
[1] University of Tehran, Tehran, Iran

10.22080/iselt.2026.30913.1155

Received
December 26, 2025
Accepted
April 28, 2026
Available online
May 3, 2026

Keywords:
Audiovisual Input, Caption, Incidental Vocabulary Learning, Gloss, Viewing.

Abstract

To date, only one study (Teng, 2022b), which was conducted with Chinese primary school students, has compared the effectiveness of glossed full captions with glossed keyword captions for incidental vocabulary learning through video viewing, and how these two captioning formats operate with other learner populations, for example adult EFL learners, remains unexplored. Recognizing the scarcity of research on this topic, the current investigation sought to compare the effects of these two glossed captions on incidental vocabulary learning from viewing. To this end, 54 adult EFL learners from three intact university classes were allocated to one of the following conditions: 1) a test-only control condition (N = 18), 2) viewing with glossed full captions (N = 18), and 3) viewing with glossed keyword captions (N = 18). Learning of the 20 target items was assessed at the levels of form and meaning recall. Mixed-design ANOVAs and a MANOVA were used to analyze the data. Regarding form recall, the pretest-immediate posttest results showed that glossed keyword-captioned viewing produced significantly larger vocabulary gains than the other conditions. As for meaning recall, the glossed fully-captioned viewing condition resulted in higher gains, although it failed to reach statistical significance. These findings suggest that glossed full and glossed keyword captions may support distinct dimensions of vocabulary knowledge differentially. Pedagogically, the results highlight the importance of aligning caption design with specific language learning goals.

1. INTRODUCTION

Vocabulary knowledge is central to core language competencies, namely, listening, speaking, reading, and writing (Kaivanpanah & Parvin, 2019; Wang, 2020). In addition to the large number of words that learners need to acquire (Nation, 2006), each lexical item involves multiple dimensions of knowledge (e.g., form, meaning, and use), which makes vocabulary learning a particularly demanding task. Given the limited classroom time, explicit teaching of all the vocabulary items that one needs to know to communicate effectively in an L2 appears to be a

* **Corresponding Author:** Maryam Ehsani, Ph.D. Candidate, Department of English Language and Literature, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran, **Email**: maryam.ehsani1996@ut.ac.ir

cumbersome task. For that reason, the use of different sources of input as support and supplement for explicit L2 instruction has been widely encouraged.

Although the findings of prior studies reveal that incidental vocabulary acquisition can happen through various input source, (e.g., Jin & Webb, 2020; Peters, 2019; Vu & Peters, 2022), research indicates that vocabulary learning rates from exposure to input-only are quite low (e.g., Peters & Webb, 2018). This has made vocabulary acquisition scholars propose a number of techniques to enhance the vocabulary gains. One such technique is providing glosses (i.e., L1 translations or L2 definitions for unfamiliar lexical items) for vocabulary items in reading (e.g., Jung, 2022; Yanagisawa et al., 2020) as well as in captioned viewing (e.g., Hsieh, 2020; Teng, 2023).

Glossing has been extensively studied in the realm of incidental vocabulary learning from reading and a number of meta-analyses have been published on this topic (e.g., Kim et al., 2024; Yanagisawa et al., 2020; Zhang & Ma, 2024). However, during the past decade, the increasing use of audio-visual materials has led some researchers to examine how glossing can be effectively utilized in captioned videos to enhance vocabulary gains. Nevertheless, as acknowledged by vocabulary scholars (e.g., Montero Perez, 2022), empirical work addressing glossing within video-based input remains limited. To the best of the researchers' knowledge, although some studies compared either glossed keyword captions (e.g., Montero Perez et al., 2018) or glossed full captions (e.g., Hsieh, 2020; Teng, 2023) with different captioning types which did not provide direct access to the meanings of the target vocabulary, to date, only Teng's (2022b) study has compared the two types of glossed captioning conditions directly. This comparison seems more logical given the fact that both captioning types clarify the meanings of target vocabulary and there is no risk of incorrect inferences.

Primary school students participated in Teng's (2022b) study and the comparative effects of glossed full captions and glossed keyword captions on other learner groups, for example, university students, remain unstudied. Additionally, that study assessed form and meaning recognition and recall knowledge was not tapped into. Further research on form and meaning recall acquisition is required because developing productive knowledge (recall) is more demanding than developing receptive knowledge (recognition) (Pellicer-Sánchez, 2017; Pellicer-Sánchez et al., 2022), and this aspect of lexical knowledge is what L2 learners need to be equipped with to comprehend L2 input and communicate effectively. The present study made a comparison between vocabulary form and meaning recall gains resulting from three conditions, namely, test-only control, glossed fully-captioned viewing, and glossed keyword captioned viewing.

2. LITERATURE REVIEW

Incidental Vocabulary Learning through Viewing

Incidental vocabulary learning refers to the acquisition of lexical items that occurs as a secondary outcome of engaging in meaning-focused activities, such as video viewing for comprehension (Webb, 2020). Recognizing the popularity of audio-visual materials among language learners (Lindgren & Muñoz, 2013; Peters, 2018), researchers have conducted a large number of studies during the past decade to probe the possible effects of these input types on L2 learners' language development. A central and extensively researched facet of language within this field of inquiry is vocabulary knowledge. The effect of different video genres on L2 vocabulary development has been investigated during the recent years. For example, watching YouTube vlogs (e.g., Arndt & Woore, 2018), documentaries (e.g., Feng & Webb, 2020), cartoons (Dashtestani & Hojatpanah,

2023), and movies (e.g., Majuddin et al., 2021) has been detected to contribute to L2 vocabulary development.

Studies on incidental vocabulary learning from video viewing reveal that vocabulary gains are rather small (Peters & Webb, 2018; Puimège & Peters, 2019, 2020). For instance, as for single words, participants in Peters and Webb's (2018) study learned approximately four words out of 64 target items after viewing a sixty-minute television program- equivalent to relative gains of 8% on the meaning-recall measure and nearly 14% on the meaning-recognition measure. Although these gains are small, they can be considered encouraging (Peters & Webb, 2018), given that vocabulary acquisition is an incremental process (Schmitt, 1998).

Increasing the Chances of Incidental Vocabulary Learning through Viewing

Researchers have endeavored to find ways to accelerate the rate of incidental vocabulary acquisition across different input modalities. One of these methods is providing glosses for unknown lexical items. Glosses are added to texts to clarify the meanings of unfamiliar words (Boers et al., 2017). The contribution of various gloss types to incidental vocabulary acquisition from reading remains an unresolved question. To clarify the picture, several meta-analytic studies have been undertaken recently. However, paradoxically, their findings do not always concur. Analyzing 42 primary studies, Yanagisawa et al. (2020) established that reading glossed texts led to significantly greater lexical development than reading non-glossed texts, and Zhang and Ma (2024) reported a medium effect for textual glossing on vocabulary learning. In another study, Mohsen et al. (2024) discovered that although glossing had a generally large effect on L2 vocabulary acquisition, its effect size ranged from small to large, contingent upon the learners' proficiency level.

With respect to the language of glosses in reading, findings generally support the use of L1 glosses (Kim et al., 2024; Yanagisawa et al., 2020); nevertheless, Kim et al. (2024) reported that L2 proficiency might play a moderating role as L1 glossing proved particularly beneficial for beginner-level learners when compared to those at intermediate proficiency and beyond. Yet the meta-analysis of 20 primary studies conducted by Zhang and Ma (2024) failed to find any definite effect for the language of glosses. Accordingly, Zhang and Ma argued that what is of importance is the clarity of textual glosses, not their language (L1, L2). Moreover, it has been found that glossing affects different aspects of word knowledge differently (Jung, 2022). Jung detected that glossing affected meaning recognition positively but form recognition negatively.

Captions are the on-screen text that matches the spoken language (Sydorenko, 2010; Vanderplank, 2016a, 2016b). There is empirical evidence demonstrating that captioned viewing can facilitate the incidental acquisition of vocabulary (Ahrabi Fakhr et al., 2021; Kurokawa et al., 2024; Teng, 2022a, 2024). To illustrate, Teng (2019) examined the influence of three L2 captioning conditions on the incidental uptake of 15 single words among 257 primary school ESL students from six classes. According to the results, participants exposed to full captions significantly outperformed those in the keyword-caption and no-caption conditions.

During the last decade, several studies have explored the combined effect of glossing and captioned viewing (in either keyword or fully captioned videos) on vocabulary acquisition. Some studies have investigated only glossed full captions to examine their standalone effect. For example, Fievez et al. (2023) carried out a study investigating whether out-of-class viewing of an entire season of a French television series (a total of 307 minutes) accompanied by glossed captions facilitated incidental vocabulary learning. To this end, they allocated low- to high-intermediate French learners to either a test-only control group (N = 37) or a treatment group (N = 65). Incidental learning of 78 single-word targets was assessed using three instruments: a written form-

recall test, and aural and written meaning-recall tests. The findings showed that learners recalled almost 35% of the target meanings and 28% of the target forms.

Several studies have compared glossed full captions with either no captions or non-glossed full captions to better clarify the possible added benefit of glossing. Hsu (2018) compared three captioning conditions (i.e., uncaptioned, captioned, and glossed captioned) in relation to incidental acquisition of business-related vocabulary while viewing three 15-minute business videos. The findings demonstrated that glossed captioned viewing led to significantly higher scores than the other two conditions on both immediate and delayed posttests. Moreover, participants expressed favorable perceptions regarding the usefulness of glossed captions for incidental vocabulary development.

In another study, Hsieh (2020) investigated how five captioning conditions influenced vocabulary development among low-intermediate EFL learners. The captioning conditions included: 1) no captions (NC), 2) full captions with the audio removed (FCNA), 3) full captions (FC), 4) full captions with the target word underlined (FCHTW), and 5) full captions with both an underlined target word and an L1 gloss (FCL1). The findings indicated a differential impact based on test type. For form recognition, there was no significant advantage for glossing or highlighting over standard full captions (FC, FCHTW, and FCL1 performed equally). However, for tests of meaning recognition and recall, a clear hierarchy emerged. The FCL1 condition (glossed captions) was superior to all other groups. Furthermore, the standard FC condition outperformed the NC, FCNA, and FCHTW conditions, which did not differ from each other. This pattern suggests a potential attentional trade-off in the FCHTW condition. While aiding form recognition, this condition appeared to inhibit the development of meaning knowledge. This suggests that visual enhancement alone (underlining) directs attention to orthographic form at the cost of semantic processing, whereas the addition of an L1 gloss provides the necessary semantic support for deeper lexical integration.

Using a mixed-methods approach, Wu and Yang (2022) investigated the effect of glossing and textual enhancement of single words in a captioned video on incidental lexical gains among low-intermediate Chinese EFL learners. Participants, in their four intact class groups, were randomly allocated to one of these conditions: 1) full captions with highlighted target words and L1 glosses (FCL1), 2) bilingual full captions, 3) bilingual full captions with highlighted target words, and 4) no captions (NC). Immediate vocabulary posttests were given to the learners after viewing the video twice. No significant differences in form-recognition performance emerged among the four groups. Regarding meaning recall, the FCL1 group significantly outscored the other groups. With respect to meaning recognition, the three captioned groups significantly outperformed the NC group.

Other research studies have compared glossed full captions or glossed keyword captions with non-glossed full captions and non-glossed keyword captions to examine their relative effects. In a study with a between-subjects design, Teng (2023) explored incidental vocabulary acquisition from three captioning conditions: 1) full captions + L1 glosses, 2) full captions, and 3) keyword captions. He randomly assigned 125 intermediate-level primary school ESL learners to three experimental groups and one control group that watched uncaptioned videos. Analyses of the posttests demonstrated that the three captioned conditions significantly outperformed the control group. Furthermore, the results demonstrated that viewing with glossed captions produced the greatest form recognition and meaning recall gains. The pretest score, which was a covariate, was shown to be significantly associated with the immediate scores in both meaning recognition and meaning recall. However, it was significantly related to delayed scores only at the form recognition level.

Extending this line of research, Montero Perez et al. (2018) aimed to examine the effects of the following captioning types on L2 vocabulary learning: 1) no captions, 2) full captions, 3) keyword captions, and 4) glossed keyword captions. Intermediate to high intermediate L2 French learners were assigned to the experimental conditions. Target items were 12 single words as well as 6 multiword items (MWIs), the majority of which occurred only once. The findings revealed that the glossed keyword captions group achieved significantly higher meaning-recall scores than the other groups. The remaining differences in meaning-recall scores were not statistically significant. In terms of form recognition, the glossed keyword captions group significantly outperformed the no captions and the full captions group. However, it did not differ from the keyword captions group significantly.

Finally, to date, only one study has directly compared glossed full captions with glossed keyword captions to determine their relative effectiveness. In that study, Teng (2022b) compared four captioning conditions, namely, L1-glossed full captions, L1-glossed keyword captions, full captions, and keyword captions, in terms of their effects on ESL primary school students' incidental vocabulary development. Learning outcomes were assessed using tests that targeted form and meaning recognition, and the ability to use the words. According to the results, glossed full captions brought about the greatest vocabulary gains, irrespective of the word knowledge aspect. Although reaching valuable findings, primary school students participated in that study and research needs to be carried out to compare the effects of glossed full captions and glossed keyword captions on other learner groups. Furthermore, that study did not investigate recall knowledge. Research indicates that building productive (recall) vocabulary knowledge is more demanding for learners than building receptive (recognition) knowledge (Pellicer-Sánchez, 2017; Pellicer-Sánchez et al., 2022) and this dimension of lexical knowledge is crucial for enabling L2 learners to comprehend L2 materials and to communicate effectively in an L2. The current study compared adult EFL learners' vocabulary learning at form- and meaning-recall levels resulting from three conditions, namely, test-only control, glossed fully-captioned viewing, and glossed keyword captioned viewing.

Apart from empirical considerations, there are two theories that are strongly associated with using multimodal input in language learning and teaching, namely, Dual Coding Theory (DCT) and Multimedia Learning. In his DCT, Paivio (1990) assumes that memory and cognition operate through two distinct symbolic systems, one handling verbal information and the other nonverbal information. Borrowing tenets from the DCT, Mayer (2002, 2021) introduced the "cognitive theory of multimedia learning" to explain how learning from words and pictures takes place. According to this theory, presentation of "words and pictures rather than words alone" can facilitate learning. Although not originally designed for language learning, dual-coding and multimedia learning theories suggest that L2 learners can improve their lexical knowledge through multimodal input like captioned viewing with glossed vocabulary items.

A core assumption of the "cognitive theory of multimedia learning" is the limited capacity assumption, which posits that individuals can process only a finite amount of information in each cognitive channel at one time (Baddeley, 2010). Accordingly, one can postulate that if one adds glossed full captions to videos, a cognitive overload may occur and learners may fail to notice and learn the glossed items, especially when the primary goal is content comprehension. However, addition of glossed keyword captions may prevent this cognitive overload from happening because less written text is presented on the screen. Empirical research is required to clarify which of these two scenarios would result in cognitive overload and which would lead to successful vocabulary learning. This is what the present study aimed to explore. This study attempted to address the following questions:

1. To what extent do adult EFL learners demonstrate gains in form recall and meaning recall of target vocabulary items under each condition (i.e., test-only control, glossed full captions, and glossed keyword captions)?

2. Are there statistically significant differences among the three conditions in terms of relative gains in form recall and meaning recall?

3. METHODOLOGY AND DESIGN

This study included three conditions: 1. no viewing (test-only control), 2. viewing with glossed full captions and 3. viewing with glossed keyword captions. Full captions were a verbatim transcription of the dialogues and keyword captions involved those words and phrases that were essential for the learners to comprehend the content of the videos (Teng, 2019, 2023). Target items were embedded among the selected keywords, enabling comparison across experimental conditions. The keywords were identified by three English lecturers. Figure 1 illustrates the glossed full-caption condition, while Figure 2 provides a screenshot of the glossed keyword-caption condition.

Figure 1: A Screenshot of Glossed Full Captions

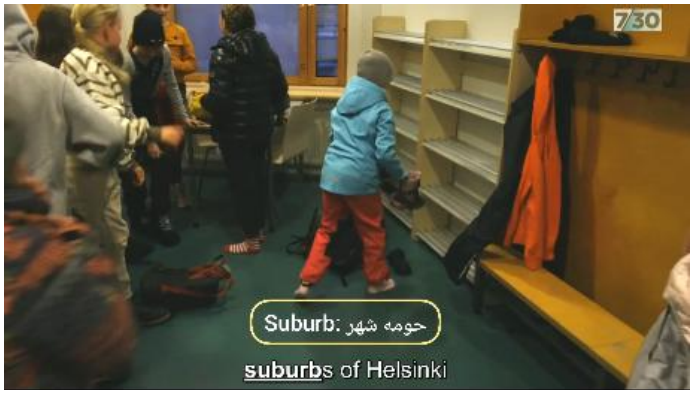

Figure 2: A Screenshot of Glossed Keyword Captions

Participants

Initially, 105 first-year undergraduate students from a public university in southern Iran participated in the experiment. However, the results of 51 students were removed from the main data because of two reasons: 1. Some of them were absent in one of the testing sessions, 2. Others did not know a minimum of 2,000 most frequent English words, according to the results of Webb et al.'s (2017) Updated Vocabulary Levels Test (UVLT). Fifty-four learners (39 males, 15 females) from three intact classes whose ages ranged from 18 to 25 (Mean = 18.90, *SD* = 1.49) participated in the main experiment. Participants were from three intact English for General Purposes (EGP) classes. Although the course content was General English in all cases, in accordance with the university's administrative arrangements, each intact class comprised students from a different engineering field, namely, computer engineering (N = 18), chemical engineering (N = 18), and electronic engineering (N = 18). Among the three intact classes, one was randomly assigned to the test-only group, while the remaining two classes were allocated to the experimental conditions.

Videos

Video selection for this study was guided by several criteria. First, the content needed to be engaging to hold participants' attention. Second, the video had to contain multiple vocabulary items with which the participants were unlikely to be familiar. Third, it required a reasonable difficulty level based on factors such as lexical coverage, grammar, and dialogue speed (with lexical coverage assessed using RANGE software). A pilot study with 23 students who were similar to the participants of the main study in terms of language proficiency and background and comments from the English teachers of the same intact classes confirmed the videos' appropriateness.

Considering the above-mentioned criteria, two YouTube videos were selected for this study. The first video, entitled 'Why Finland's schools outperform most others across the developed world", was a news video from Australian ABC news YouTube channel. It was 6 minutes and 48 seconds long, with a total of 957 running words. Analysis of the script using RANGE (Nation & Heatley, 2002) indicated that the most frequent 1,000-, 2,000-, and 3,000-word families accounted for 87.70%, 94.89%, and 97.21% of the video's vocabulary, respectively. The second video, entitled 'Top 10 technologies to learn in 2024', was 7 minutes and 53 seconds long and included 1,130 running words. Running the script of the video in RANGE showed that the most frequent 1,000-, 2,000-, and 3,000-word families resulted in 69.32%, 84.51%, and 92.76% coverage of the video's vocabulary, respectively. The participants did not report any major difficulties in understanding the videos. This assured us that the videos were not too demanding for the target students. Moreover, findings from the pilot study indicated that generally, the learners found the videos interesting and appropriate in terms of difficulty.

Target Items

As a result of comments from three of the students' EGP lecturers who were invited to read the transcripts of the two videos, a set of single words and multiword items (MWIs) was identified as potentially unfamiliar to the students. A total of 36 single words and MWIs (18 for each video) were pilot tested with 23 students who were of comparable language proficiency and background to the main study's participants. Nine single words that were unfamiliar to 80% or more of the pilot-study participants were chosen as target items for video 1 which was about Finland's educational system. Except for one item, all single words appeared only once in the video. Regarding the second video, which was about technology, 11 items, including 6 single words and 5 collocations, that were unknown to 80% or more of the pilot-study participants were chosen as

target items. This resulted in a total of 20 target items (15 single words, 5 collocations) for the two videos.

Instruments

Two paper-and-pencil tests with different order of items were given to participants as the pretests and immediate posttests. The first test measured participants' form recall and the second test assessed their meaning recall of the target vocabulary. In the first test, a Persian meaning was provided for each single word or collocation and students were required to write the equivalent English form. In the second test, participants were instructed to either translate each target item into Persian or provide an English definition. The tests comprised 20 target items (15 single words and 5 collocations) along with 16 distractors. The distractors consisted of high-frequency items known by 80% or more of the pilot study participants, and they were incorporated to motivate the participants to make a serious attempt to complete the tests. Participants received examples as well as explicit instructions in both English and Persian to guarantee that they understood how to complete the tests properly.

Procedure

This study was carried out in three successive weeks within normal class time. In Week 1, informed consent was obtained in Persian, and the three groups completed the paper-and-pencil UVLT to evaluate their pre-existing vocabulary knowledge. A one-way ANOVA showed no significant variation in the groups' UVLT mean scores ($p > 0.05$), confirming equivalent baseline vocabulary levels. Given that a substantial body of evidence has demonstrated a robust relationship between vocabulary size and general L2 proficiency (Miralpeix & Muñoz, 2018; Stæhr, 2008; Uchihara & Clenton, 2020), the equivalent UVLT scores across the three groups provide reasonable evidence that they did not differ systematically in general English proficiency. After a week, the participants were given the form and meaning recall pretests. To reduce potential test-induced effects, the form-recall test was administered first. Once all answer sheets related to this test had been collected, the meaning-recall test was administered. In Week 3, the experimental groups watched the videos in their allocated conditions. All classrooms had a multimedia system, which consisted of a computer, a projector, and high-quality speakers. To encourage participants to pay attention to video content rather than intentionally noticing unknown items, they were informed that they would be asked questions about the content of the videos afterward. Immediately after viewing, the experimental groups completed the immediate posttests and answered an open-ended question about what they learned in terms of content from each video. Each posttest lasted approximately 13 minutes. The control group did not watch any videos but took the pretests and posttests. In line with previous research (Puimège & Peters, 2019; Vu & Peters, 2022), to prevent any test order and practice effects, the items on the posttest were presented in a different sequence than on the pretest. Participants were instructed to attend to the video content and to refrain from using mobile phones, dictionaries, or taking notes during the experiment. The first researcher monitored participant compliance with the instructions. The actual aim of the research was disclosed to participants after the posttests.

Scoring

Participants' responses were scored using a dichotomous scale (1 for correct, 0 for incorrect). A lenient scoring system was employed, in which minor errors received a full mark (e.g., using a plural instead of a singular form). All pretests and posttests were initially scored by the first

researcher, followed by a random secondary scoring of 50% of the tests by a Persian-speaking TEFL PhD holder. Regarding form recall, the inter-rater reliability estimates were .98, and .96 for the pretest and immediate posttest, respectively. As for meaning recall, the reliability estimates were .99, and .98 for the pretest and immediate posttest, respectively. Any discrepancies were resolved through discussion.

Data Analysis

A target item was coded as learned when it was unknown on the pretest but correctly produced on the posttest (an absolute gain). Two 2 × 3 mixed within-between subjects ANOVAs were used to answer RQ1 (within-subjects IV: time; between-subjects IV: group).

According to Rodgers and Webb (2020), absolute gains may not accurately represent vocabulary development. This is because participants who score higher on the pretest have less room for improvement than those who score lower. Therefore, following previous studies (Fievez et al., 2020; Rodgers & Webb, 2020), to take into account the differences between participants' pretest scores, relative gains were also calculated, using the following formula:

[absolute gains/ (number of target items - number of known words)] × 100.

A MANOVA was conducted to answer RQ2 (DVs: relative form-recall and meaning-recall gains; IV: group). All analyses were conducted in SPSS version 26.0.

4. RESULTS

The open-ended question "what did you learn regarding the content?" produced references to the central topics of the videos as well as some details, indicating that the participants experienced no major difficulty comprehending the videos' main ideas.

Concerning vocabulary, first, two one-way ANOVAs were performed on form-recall and meaning-recall pretest scores, respectively. Regarding form recall, Group 1 had the best performance, followed by Group 3, and Group 2. Levene's test confirmed homogeneity of variance for form-recall pretest scores across the three groups ($p = 0.139$, $p > 0.05$). The ANOVA result showed that the form-recall pretest scores did not differ from each other significantly, F (2, 51) = 0.833, $p = 0.441$, $p > 0.05$. With respect to meaning recall, Group 2 gained the highest score, followed by Group 3 and Group 1. The assumption of equal variances was met for the meaning-recall pretest, as evidenced by a non-significant Levene's test ($p = 0.279$, $p > 0.05$). The ANOVA result showed that the meaning-recall pretest scores did not differ from each other significantly, F (2, 51) = 0.638, $p = 0.533$, $p > 0.05$.

Within-Group Differences (RQ1)

Two 2 × 3 mixed within-between subjects ANOVAs were conducted to assess improvement in form and meaning recall from pretest to immediate posttest across groups. Time (pretest, immediate posttest) was used as the within-subjects factor and Group was the between-subjects factor. For both form and meaning recall, because Box's M test revealed a violation of the homogeneity of inter-correlations assumption ($p = 0.000$, $p < 0.05$ [in both cases]), instead of Wilks' lambda, Pillai's trace was used to interpret main effects in the F-statistic test (Cramer & Howitt, 2004, as cited in Teng, 2022a).

Regarding form recall, the Multivariate Tests showed a significant Time × Group interaction, F (2, 51) = 11.57, $p = 0.000$, $p < 0.05$, partial $\eta^2 = 0.312$ (a large effect size according to Cohen's (1988) benchmarks), suggesting that the effect of group varied over the two forms recall testing phases. Three paired-samples t-tests were conducted to further investigate the

interaction effect (with a Bonferroni-adjusted alpha of 0.017). The analysis revealed that all three groups had experienced significant form-recall improvements (Group 1: $p = 0.006$, Group 2: $p = 0.001$, Group 3: $p = 0.000$, $p < 0.017$ in all three cases).

Concerning meaning recall, the Multivariate Tests indicated no significant Time × Group interaction, F (2, 51) = 2.90, $p = 0.064$, $p > 0.05$. The within-subjects factor "Time" yielded a significant effect, F (2, 51) = 35.69, $p = 0.000$, $p < 0.05$, partial $\eta^2 = 0.412$ (large effect). This result indicates a significant increase in meaning-recall performance. The between-subjects variable "Group" was not significant, F (2, 51) = 2.79, $p = 0.071$, $p > 0.05$. This result demonstrates that all groups improved significantly to a similar degree from the pretest to the immediate posttest. Table 1 displays the descriptive results of pretests and posttests for each group.

Table 1: Form- and Meaning-Recall Pretest and Posttest Scores per Group

	Pretest/Form		Posttest/Form		Pretest/Meaning		Posttest/Meaning	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Group 1	0.72	1.49	1.50	2.23	0.72	1.45	2.50	3.11
Group 2	0.22	0.55	1.06	0.80	1.28	1.23	5.78	5.60
Group 3	0.44	1.25	3.17	2.18	1.11	1.81	3.50	2.73
Total	0.46	1.16	1.91	2.04	1.04	1.50	3.93	4.18

Note. **Max. = 20**

Between-Group Differences (RQ2)

Table 2 shows the descriptive statistics for absolute and relative gains with respect to form and meaning. Regarding form recall, participants who viewed the videos with glossed keyword captions significantly outperformed both the control group and those who viewed glossed full captions. However, with regard to recalling the meaning, participants who watched glossed fully-captioned videos obtained higher gains than the control group and those who watched the videos with glossed keyword captions. The descriptive statistics showed that for form recall, participants who viewed the videos with glossed keyword captions achieved a relative score of 13.85% but the control group and the group who watched glossed fully-captioned videos achieved relative gains of 4.31% and 4.17%, respectively. The descriptive statistics demonstrated that concerning relative gains in recalling the meanings of the target items, those who watched glossed fully-captioned videos had a better performance ($M = 24.22$) compared to the control group ($M = 9.77$) and those who watched the videos with glossed keyword-captions ($M = 12.72$).

Table 2: Descriptive Statistics for Absolute and Relative Gains in Terms of Form and Meaning

		Absolute/Form	Relative/Form	Absolute/Meaning	Relative/Meaning
Group	Group 1	.78	4.31	1.78	9.77
	Group 2	.83	4.17	4.50	24.22
	Group 3	2.72	13.85	2.39	12.72

To compare the three groups' performance on form and meaning recall, a MANOVA was conducted using relative gains in form and meaning recall as dependent variables and group membership as the independent variable. As stated in the "Data analysis" section, the reason for preferring relative gains over actual scores was that participants who score higher on the pretest have less room for improvement than those who score lower, and therefore, actual scores may not

provide an accurate representation of lexical development (Rodgers & Webb, 2020). Box's M test revealed a significance value below .05, indicating a violation of the homogeneity of variance assumption for MANOVA. Consequently, to interpret the main effects, Pillai's trace was employed as the F-statistic instead of Wilks' lambda. The MANOVA results can be seen in Table 3.

Table 3: Multivariate Tests for MANOVA

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Group	Pillai's Trace	.413	6.641	4.000	102.000	.000	.207

The multivariate tests revealed that group affected relative gains significantly, V = 0.413, F = 6.641, p = 0.000, p < .05, partial η2 = 0.207 (large effect). Univariate tests followed the multivariate tests, separated by test type (see Table 4). A significant group effect was detected for form recall (F = 10.438, p = .000, p < .05, partial η^2 = .290, indicating a large effect), whereas no significant group effect was observed for meaning recall (F = 2.793, p = .071, p > .05). Table 5 presents Tukey post hoc analysis for relative form-recall gains. The results revealed that Group 3 (glossed keyword-captioned viewing) achieved significantly higher scores than both Group 1 (the control group) and Group 2 (glossed fully-captioned viewing) (p = 0.001, p <.05 in both cases). However, no significant disparities emerged between Group 1 and Group 2 (p = 1, p > 0.05).

Table 4: Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Group	Relative gain/ Form	1109.835	2	554.918	10.438	.000	.290
	Relative gain/Meaning	2099.260	2	1049.630	2.793	.071	.099
Error	Relative gain/ Form	2711.358	51	53.164			
	Relative gain/ Meaning	19167.834	51	375.840			
Total	Relative gain/ Form	6813.499	54				
	Relative gain/Meaning	34358.403	54				
Corrected Total	Relative gain/ Form	3821.193	53				
	Relative gain/Meaning	21267.094	53				

Table 5: Pairwise Comparisons of the Relative Form-Recall Gains

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Relative gain/ Form	Group 1	Group 2	.144	2.430	1.000	-5.873	6.160
		Group 3	-9.544*	2.430	.001	-15.561	-3.528
	Group 2	Group 1	-.144	2.430	1.000	-6.160	5.873
		Group 3	-9.688*	2.430	.001	-15.705	-3.671
	Group 3	Group 1	9.544*	2.430	.001	3.528	15.561
		Group 2	9.688*	2.430	.001	3.671	15.705

5. DISCUSSION

This study aimed to compare the effects of viewing videos with glossed full captions and glossed keyword captions on adult EFL learners' incidental vocabulary learning. Regarding RQ1, the findings confirmed that vocabulary learning took place under all three conditions: both experimental treatments as well as the control condition, at both form- and meaning-recall levels. Significant pretest-posttest improvement in the test-only control group has been documented in previous studies as well. For example, the test-only control group in Fievez et al.'s (2023) study made significant pretest-immediate posttest improvements in terms of form- and meaning-recall of target single words. Furthermore, in the study conducted by Rodgers and Webb (2020) about single-word learning from uncaptioned viewing, the test-only control group made significant lexical gains from the pretest to the posttest. A similar pretest-immediate posttest increase was observed in Webb et al.'s (2013) research on incidental acquisition of collocations from reading-while-listening. In their research, the test-only control group, whose members did not engage with the graded reader through either reading or listening, still demonstrated significant improvement from the form-recognition pretest to the posttest, suggesting a test-induced learning effect.

The descriptive statistics showed that in terms of relative gains for recalling forms, the glossed keyword-captioned viewing group achieved the highest gains, followed by the test-only control group and the glossed fully-captioned viewing group, which performed at a very similar level. In answer to the first part of RQ2 about form-recall, Tukey post hoc analyses on immediate form-recall scores revealed that Group 3 (glossed keyword-captioned viewing) achieved significantly better results than both Group 1 (test-only control) and Group 2 (glossed fully-captioned viewing). However, Group 1 and Group 2 did not differ significantly. The better performance of the glossed keyword-captioned viewing compared to the glossed fully-captioned viewing group in terms of form was in contrast with Teng (2022b), in which glossed full captions resulted in superior immediate form-recognition gains to glossed keyword captions. But it appears that this finding aligns with the limited capacity assumption of the "cognitive theory of multimedia learning" (Baddeley, 2010; Mayer, 2002, 2021). In fact, one explanation for the significantly higher form-recall scores in the glossed keyword-captioned viewing condition compared to the glossed fully-captioned viewing condition is that by adding glossed full captions to videos a cognitive overload occurred in learners' minds and thus they failed to notice and learn the forms of the glossed items. However, addition of glossed keyword captions prevented this cognitive overload from happening to some degree, because less written text was presented on the screen.

The descriptive statistics showed that in terms of relative gains for recalling meanings, the glossed fully-captioned viewing group achieved the highest gains, followed by the glossed keyword-captioned viewing group, while the test-only control group showed the lowest gains. As explained in the "Results" section, the relative gains showed that on average, the control group acquired less vocabulary knowledge across the two tests. Additionally, in both tests, the score of the control group increased from the pretest to the posttest. This result suggested that the learning gains might not be merely due to the interventions but could also stem from being exposed to the target items during the pretests. This phenomenon, called the testing effect, is frequently reported in vocabulary acquisition research (Jin & Webb, 2020; Webb et al., 2013), and it is the reason why inclusion of a control group is essential to verify that additional gains in experimental groups are the result of the treatments, not tests. Nevertheless, the finding that, on average, the control group achieved lower immediate posttest scores suggests that at least some learning occurred due to the experimental treatments.

Concerning the second part of RQ2, Tukey post hoc analyses on meaning-recall relative gains revealed no significant between-group disparities. However, the superior performance of the glossed fully-captioned viewing group compared to the glossed keyword-captioned viewing was in line with Teng (2022b), in which glossed full captions led to better immediate meaning-recognition gains than glossed keyword captions. Additionally, the effectiveness of full captions rather than keyword captions for meaning recall was consistent with Teng's (2019) study, in which fully-captioned videos led to more incidental meaning-recall learning than keyword-captioned videos. The finding that the glossed fully-captioned viewing condition led to a lower form recall score but a higher meaning recall score was congruent with Jung's (2022) reading study, which found that glossing enhanced recognition of meaning but hindered recognition of form.

Taken together, these results suggest that glossed full captions and glossed keyword captions support different aspects of vocabulary knowledge. Specifically, glossed keyword captions appear more effective for enhancing recall of lexical forms, whereas glossed full captions may better support the recall of lexical meanings. From a theoretical perspective, the present findings offer valuable implications for both "Dual Coding Theory" (Paivio, 1990) and the "cognitive theory of multimedia learning" (Mayer, 2002, 2021). The observed benefits of the two glossed caption types seem to support the core assumption of these frameworks: that multimodal input integrating verbal and non-verbal stimuli can facilitate learning more effectively than unimodal input. However, the differing effects of full and keyword captions suggest that multimodality does not intrinsically guarantee optimal learning. Instead, the specific design of multimodal materials (e.g., the choice between full or keyword captions) must take cognitive load into account. The finding that full captions may hinder form recall but promote meaning recall shows that multimodal input can generate unequal benefits across different vocabulary dimensions, a result that underscores the importance of designing theoretically informed captioning interventions.

Finally, by addressing a previously underexplored comparison, namely, glossed full captions versus glossed keyword captions, the present study contributes to filling two key empirical gaps. First, it extends prior research that has focused predominantly on children (e.g., Teng, 2022b) by examining adult EFL learners. Second, it adds evidence regarding recall-based knowledge, a dimension often overlooked even though it is deemed essential for enabling L2 learners to comprehend and communicate effectively. Together, these contributions help move the field toward a more comprehensive understanding of how various glossed caption types influence different vocabulary dimensions.

6. CONCLUSION

Like any research, the present investigation has several limitations. First, participants consisted of Iranian university students from a single university. To be able to generalize the results, future studies involving learners from other academic levels and cultural backgrounds are required. Second, glosses were provided in students' L1. Future research can explore the influence of L2 glosses and other glossing formats (e.g., multimedia, multiple-choice, etc.). Third, future studies should include delayed posttests to assess retention. Finally, longitudinal studies can be carried out to yield more illuminating insights into the topic in question.

As mentioned earlier, a substantial body of research has examined how glossing influences word learning during reading (e.g., Jung, 2022; Yanagisawa et al., 2020). However, research investigating the effectiveness of glossing and different gloss types for vocabulary acquisition through captioned video viewing remains limited. The present study explored this issue and was among the first attempts to directly compare two previously less researched caption conditions: glossed keywords versus glossed full captions. Acknowledging the limitations, the present study showed that glossed keyword-captioned viewing can lead to significant improvements in terms of form-recall, and glossed fully-captioned viewing can be utilized to develop meaning-recall.

Taking a pedagogical perspective, the findings of this study offer insights that may assist L2 learners, teachers, and materials developers in selecting optimal captioning conditions to enhance incidental vocabulary acquisition through video viewing. This study showed that captioned viewing with glosses can lead to vocabulary learning. Specifically, glossed keyword-captioned viewing can be used to improve form recall and the glossed fully-captioned viewing can be utilized to develop meaning recall. With recent technological advancements, L2 teachers and materials developers can add captions and glosses to videos to help learners acquire unfamiliar vocabulary items.

References

Ahrabi Fakhr, M., Borzabadi Farahani, D., & Khomeijani Farahani, A. A. (2021). Incidental vocabulary learning and retention from audiovisual input and factors affecting them. *English Teaching & Learning*, *45*(2), 167–188. https://doi.org/10.1007/s42321-020-00066-y

Arndt, H., & Woore, R. (2018). Vocabulary learning from watching YouTube videos and reading blog posts. *Language Learning & Technology*, *22*(3), 124–142.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), 136–140. https://doi.org/10.1016/j.cub.2009.12.014

Boers, F., Demecheleer, M., He, L., Deconinck, J., Stengers, H., & Eyckmans, J. (2017). Typographic enhancement of multiword units in second language text. *International Journal of Applied Linguistics*, *27*(2), 448–469. https://doi.org/10.1111/ijal.12141

Cohen, J.W. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Dashtestani, S. R & Hojatpanah, S. (2023). The effect of interactive videos on Iranian EFL learners' lexical knowledge in the context of secondary schools. *Journal of Foreign Language Research*, 13 (1), 157-175. http://doi.org/ 10.22059/jflr.2023.350567.993

Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening, and viewing: Which mode of input is most effective? *Studies in Second Language Acquisition*, *42*(3), 499–523. https://doi.org/10.1017/S0272263119000494

Fievez, I., Montero Perez, M., Cornillie, F., & Desmet, P. (2020). Vocabulary learning through viewing captioned or subtitled videos and the role of learner- and word-related factors. *CALICO Journal*, *37*(3), 233–253. https://doi.org/10.1558/cj.39370

Fievez, I., Montero Perez, M., Cornillie, F., & Desmet, P. (2023). Promoting incidental vocabulary learning through watching a French Netflix series with glossed captions. *Computer Assisted Language Learning*, *36*(1–2), 26–51. https://doi.org/10.1080/09588221.2021.1899244

Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning*, *33*(5–6), 567–589. https://doi.org/10.1080/09588221.2019.1577898

Hsu, H.-T. (2018). Incidental professional vocabulary acquisition of EFL business learners: Effect of captioned video with glosses as a multimedia annotation. *The JALT CALL Journal*, *14*(2), 119–142. https://doi.org/10.29140/jaltcall.v14n2.227

Jin, Z., & Webb, S. (2020). Incidental vocabulary learning through listening to teacher talk. *The Modern Language Journal*, *104*(3), 550–566. https://doi.org/10.1111/modl.12661

Jung, J. (2022). The impact of glossing and reading activity manipulation on learning of L2 lexico-grammatical and lexical items. *Language Teaching Research*, *26*(4), 777–798. https://doi.org/10.1177/1362168820911198

Kaivanpanah, Sh., Parvin, M. (2019). The contribution of breadth and depth of vocabulary knowledge to EFL learners' summary writing success. *Journal of Foreign Language Research*, 9 (2), 423-444. https://doi.org/10.22059/jflr.2019.260808.527

Kim, H. S., Lee, J. H., & Lee, H. (2024). The relative effects of L1 and L2 glosses on L2 learning: A meta-analysis. *Language Teaching Research*, *28*(1), 7–28. https://doi.org/10.1177/1362168820981394

Kurokawa, S., Hein, A. M., & Uchihara, T. (2025). Incidental vocabulary acquisition through captioned viewing: A meta-analysis. *Language Learning*, *75*(4), 939–987. https://doi.org/10.1111/lang.12697

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1), 105–129. https://doi.org/10.1080/14790718.2012.679275

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental acquisition of multiword expressions through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition*, *43*(5), 985–1008. https://doi.org/10.1017/S0272263121000036

Mayer, R. E. (2002). Multimedia learning. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 41, pp. 85–139). Academic Press. https://doi.org/10.1016/S0079-7421(02)80005-6

Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition*, *10*(2), 229–240. https://doi.org/10.1016/j.jarmac.2021.03.007

Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, *56*(1), 1–24. https://doi.org/10.1515/iral-2017-0016

Mohsen, M. A., Mahdi, H. S., & Alkhammash, R. (2024). Multimedia glosses and their impact on second language vocabulary acquisition: Insights from a meta-analysis and document co-citation analysis. *Innovation in Language Learning and Teaching*, *18*(2), 109–124. https://doi.org/10.1080/17501229.2023.2236084

Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, *55*(2), 163–192. https://doi.org/10.1017/S0261444821000501

Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, *31*(1–2), 1–26. https://doi.org/10.1080/09588221.2017.1375960

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). *Range: A program for the analysis of vocabulary in texts* [Computer software]. Victoria University of Wellington. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-analysis-programs.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research*, *21*(3), 381–402. https://doi.org/10.1177/1362168815618428

Pellicer-Sánchez, A., Siyanova-Chanturia, A., & Parente, F. (2022). The effect of frequency of exposure on the processing and learning of collocations: A comparison of first and second language readers' eye movements. *Applied Psycholinguistics*, *43*(3), 727–756. https://doi.org/10.1017/S014271642200011X

Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, *169*(1), 142–168. https://doi.org/10.1075/itl.00010.pet

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly*, *53*(4), 1008–1032. https://doi.org/10.1002/tesq.531

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing l2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/S0272263117000407

Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audiovisual input: An exploratory study into incidental learning of single words and formulaic sequences. *The Language Learning Journal*, *47*(4), 424–438. https://doi.org/10.1080/09571736.2019.1638630

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing l2 television and factors that affect learning. *Studies in Second Language Acquisition*, *42*(3), 525–549. https://doi.org/10.1017/S027226311900055X

Rodgers, M. P. H., & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics*, *171*(2), 191–220. https://doi.org/10.1075/itl.18034.rod

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, *48*(2), 281–317. https://doi.org/10.1111/1467-9922.00042

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139–152. https://doi.org/10.1080/09571730802389975

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, *14*(2), 50–73.

Teng, M. F. (2019). Incidental vocabulary learning for primary school students: The effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher*, *46*(1), 113–136. https://doi.org/10.1007/s13384-018-0279-6

Teng, M. F. (2022a). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System*, *105*, 102736. https://doi.org/10.1016/j.system.2022.102736

Teng, M. F. (2022b). Vocabulary learning through videos: Captions, advance-organizer strategy, and their combination. *Computer Assisted Language Learning*, *35*(3), 518–550. https://doi.org/10.1080/09588221.2020.1720253

Teng, M. F. (2025). Effectiveness of captioned videos for incidental vocabulary learning and retention: The role of working memory. *Computer Assisted Language Learning*, *38*(1–2), 206–234. https://doi.org/10.1080/09588221.2023.2173613

Teng, M. F. (2024). Working memory and prior vocabulary knowledge in incidental vocabulary learning from listening, reading, reading-while-listening, and viewing captioned videos. *System*, *124*, 103381. https://doi.org/10.1016/j.system.2024.103381

Uchihara, T., & Clenton, J. (2020). Investigating the role of vocabulary size in second language speaking ability. *Language Teaching Research*, *24*(4), 540–556. https://doi.org/10.1177/1362168818799371

Vanderplank, R. (2016a). 'Effects of' and 'effects with' captions: How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching*, *49*(2), 235–250. https://doi.org/10.1017/S0261444813000207

Vanderplank, R. (2016b). The state of the art II: Selected research on other issues in watching captioned TV, films and video. In *Captioned media in foreign language learning and teaching* (pp. 105–148). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-50045-8_5

Vu, D. V., & Peters, E. (2022). Learning vocabulary from reading-only, reading-while-listening, and reading with textual input enhancement: Insights from Vietnamese EFL learners. *RELC Journal*, *53*(1), 85–100. https://doi.org/10.1177/0033688220911485

Wang, Q. (2020). Incidental L2 vocabulary acquisition and reading: Concerns, progresses and future directions. *Chinese Journal of Applied Linguistics*, *43*(4), 469–488. https://doi.org/10.1515/CJAL-2020-0031

Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Routledge. https://doi.org/10.4324/9780429291586-15

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*(1), 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL: International Journal of Applied Linguistics, 168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Wu, H., & Yang, X. (2022). Effectiveness of textually-enhanced captions on Chinese high-school EFL learners' incidental vocabulary learning. *Porta Linguarum*, *38*, 209–228. https://doi.org/10.30827/portalin.vi38.23511

Yanagisawa, A., Webb, S., & Uchihara, T. (2020). How do different forms of glossing contribute to l2 vocabulary learning from reading?: A meta-regression analysis. *Studies in Second Language Acquisition*, *42*(2), 411–438. https://doi.org/10.1017/S0272263119000688

Zhang, C., & Ma, R. (2024). The effect of textual glosses on L2 vocabulary acquisition: A meta-analysis. *Language Teaching Research, 28*(3), 967–986. https://doi.org/10.1177/13621688211011511