Effects of Three Types of Glosses on Iraqi EFL Learners' Incidental Vocabulary Learning from Audio-Visual Input

Hind Jasim Hasan
Baquba Technical Institute, Middle Technical University, Ministry of Higher Education and Scientific Research, Iraq

Mehdi Vaez-Dalili
Department of English, Isf.C., Islamic Azad University, Isfahan, Iran

Abstract

This study investigated the effects of three different glossing conditions on Iraqi EFL learners' incidental vocabulary acquisition from captioned videos. It aimed to compare the effectiveness of no gloss (control group), traditional gloss (TG), and multiple-choice gloss (MCG) for vocabulary learning and examine the relationship between prior vocabulary knowledge and learning gain. A quasi-experimental design with pretest-posttest was employed. Ninety intermediate Iraqi university students majoring in ELT participated after being randomly assigned to three groups. Participants completed vocabulary pre-tests assessing prior knowledge levels. They were then randomly assigned to view six animated videos: Two videos without glosses (control), two videos with traditional translations, and two videos with multiple-choice options defining targeted words. Exposure consisted of automatic playback intermittently halting for relevant on-screen annotations. Gains were measured through immediate post-tests without prior notification. Data were analyzed using a one-way ANOVA and a Pearson correlation coefficient. The results of the one-way ANOVA showed significant differences in learning outcomes based on gloss condition. Both gloss groups significantly outperformed the control, evidencing annotation benefits vocabulary uptake beyond mere exposure. Notably, the MCG group demonstrated substantially greater gains than TG, indicating that interactive glossing optimizes learning over simpler formats. Additionally, based on the Pearson correlation coefficient, a strong positive correlation was found between participants' pre-existing vocabulary size and subsequent learning gain.

Keywords: audio-visual input, incidental L2 vocabulary learning, multiple-choice gloss, traditional gloss

Introduction

Vocabulary acquisition is a crucial aspect of second language (L2) learning, as it enables learners to comprehend and communicate effectively (Schmitt, 2000). However, L2 learners often face the challenge of acquiring a large vocabulary, as they need to know at least 3,000 word families for spoken texts and 8,000 word families for written texts to achieve adequate comprehension (Webb & Rodgers, 2009). Two main approaches to vocabulary learning have been identified: intentional learning, where learners deliberately commit words to memory, and incidental learning, where words are acquired by engaging in meaning-focused activities (Hulstijn, 2003; Nation, 2013). Incidental vocabulary learning has gained significant attention, as it aligns with how native speakers acquire their lexicon and allows L2 learners to acquire large quantities of words for communicative competence (Nation, 2013).

While reading has been considered a primary source of incidental vocabulary learning, research has shown that the pick-up rate is relatively low, with learners needing to encounter a new word multiple times to acquire it (Horst et al., 1998). Listening has also been explored as a medium for incidental vocabulary acquisition, but it has been found to be less effective than reading, as more encounters are required for durable effects (van Zeeland & Schmitt, 2013; Vidal, 2011).

Given the limitations of reading and listening, researchers have turned their attention to other sources of incidental vocabulary learning, such as audio-visual (AV) input. AV input, which includes materials like movies, TV shows, and online videos, has several advantages. It is widely available, popular among language learners, and can provide contextual cues through the combination of visual and auditory information (Montero Perez et al., 2025; Rodgers, 2018).

Recent studies have investigated the potential of AV input for incidental vocabulary learning, exploring the role of different types of on-screen text, such as subtitles (L1 text) and captions (L2 text), in facilitating vocabulary acquisition (Montero Perez et al., 2025; Winke et al., 2010). However, the findings have been mixed, with some studies reporting positive effects of on-screen text on vocabulary learning, while others have found limited or no significant differences (Çakmak & Erçetin, 2018).

To further explore the potential of AV input for incidental vocabulary learning, the present study aims to investigate the effects of different types of glosses (i.e., no gloss, traditional gloss, and multiple-choice gloss) on Iraqi EFL learners' incidental vocabulary acquisition. Glosses, which provide brief definitions or translations of unfamiliar words, have been shown to enhance incidental vocabulary learning from written and AV input (Watanabe, 1997; Yeh & Wang, 2003).

The study will examine the differential effects of the three gloss types on learners' acquisition of form, meaning, and productive knowledge of target vocabulary items. The findings of this study will contribute to the growing body of research on incidental vocabulary learning from AV input and provide insights into the effectiveness of different gloss types in facilitating vocabulary acquisition. The results may also have practical implications for designing and implementing AV-based language learning materials and activities.

Review of the Related Literature

The role of multimedia input in second language vocabulary acquisition has garnered significant attention from researchers in recent years, particularly focusing on the effectiveness of various captioning and glossing approaches. As language learning increasingly incorporates digital tools and AV content, understanding how different input modalities affect vocabulary retention has become crucial for both educators and researchers. Studies have investigated various aspects of multimedia-enhanced learning, including the impact of different captioning types, glossing modes, and input modalities on both intentional and incidental vocabulary acquisition. The relevant studies span diverse learning contexts, from traditional classroom settings to mobile-assisted learning environments, and examine multiple factors that affect vocabulary retention, such as learner proficiency, input design, and assessment methods.

Montero Perez et al. (2014) investigated the effects of different types of captioning (no captioning, full captioning, keyword captioning, and full captioning with highlighted keywords) on video comprehension and incidental vocabulary learning. A total of 133 undergraduate students were assigned to the different captioning groups. The results showed that the captioning groups outperformed the control group in form recognition and clip association, but only the keyword captioning and full captioning with highlighted keywords groups outperformed the control group in meaning recognition. Captioning did not affect comprehension or meaning recall. The study suggests that captioning can facilitate students' recognition of unknown words and their initial form-meaning connections.

Hsu (2018) investigated the effect of captioned videos with glosses on EFL business learners' incidental business vocabulary acquisition. The study involved 50 participants who were shown three business-related video clips in three modes: non-captioned, captioned, and captioned with gloss. A pre-test ensured participants lacked familiarity with the target words. After each mode session, participants completed immediate and 3-week-delayed post-tests. The findings revealed that the caption-gloss mode significantly outperformed the other two modes in both immediate and 3-week-delayed post-tests. Retention of the target business words did not decline significantly at the 3-week-delayed post-test. The participants also provided positive feedback regarding the efficacy of the caption-gloss mode for incidental business word acquisition. These results suggest that glosses in captioned videos can improve learners' short- and long-term incidental business vocabulary retention. The study provides valuable insights for educators seeking to enhance EFL learners' vocabulary acquisition through multimedia resources.

Ashcroft et al. (2018) investigated incidental vocabulary learning through watching movies among Japanese university students. The researchers selected the movie "Back to the Future" and designed an experimental study with 187 participants divided into experimental ($n = 143$) and control ($n = 44$) groups. The methodology involved administering identical pre- and post-tests measuring productive vocabulary recall, with the experimental group watching the movie with English captions. The study found a significant mean gain of 1.77 words (4.2%) per student from a list of 42 target words. The researchers attributed the modest vocabulary gains to low word frequency in the movie script and limited attention to vocabulary during viewing. They suggested potential improvements such as multiple viewings, pre-teaching vocabulary, and focused post-viewing activities. The findings indicated that watching movies with captions can result in slight incidental vocabulary learning, highlighting the potential of movies as a supplementary language learning resource.

Çakmak and Erçetin (2018) examined the effects of different types of glosses (textual, pictorial, and textual-plus-pictorial) on text recall and incidental vocabulary learning in a mobile-assisted L2 listening context. Eighty-eight elementary-level students were randomly assigned to the different gloss conditions and a no-gloss control group. The results showed that the gloss conditions outperformed the control group in incidental vocabulary learning, but there were no significant differences among the gloss conditions in text recall or vocabulary learning. The study suggests that glosses can enhance incidental vocabulary learning in mobile-assisted L2 listening, but pictorial glosses may not be as effective as textual glosses due to weak referential connections between verbal and visual information.

Puimège and Peters (2019) investigated the incidental learning of single words and formulaic sequences (FS) through TV viewing among 20 Dutch-speaking EFL learners. The study adopted a pre-test post-test within-subjects design, where participants watched a 30-minute excerpt of the TV show Dragon's Den. Vocabulary learning was measured at three levels: form recall, meaning recall, and form recognition. The results indicated that both single words and FS could be learned incidentally through TV viewing, with learning gains mediated by item-related factors (e.g., concreteness, collocate-node relationship) and learners' prior vocabulary knowledge. Specifically, concrete words and adjective-noun FS were learned more effectively. The study highlighted the potential of TV as a resource for L2 vocabulary acquisition, though it noted limitations such as a small sample size and the need for further research on FS learning.

Kıvrak and Uygun Gökmen (2019) investigated the effects of task-induced involvement load and input modality on incidental vocabulary learning among Turkish EFL learners. The study involved 236 participants at lower- and upper-intermediate proficiency levels, who were assigned to six tasks with varying involvement loads (reading-only, watching-only, gap-filling, and sentence-writing) and input types (written vs. audio-visual). Vocabulary knowledge was assessed through immediate and delayed post-tests measuring productive and receptive word knowledge. The results showed that tasks with higher involvement loads, such as sentence-writing, were more effective for vocabulary retention than gap-filling or reading-only tasks. Additionally, audio-visual input was more beneficial for upper-intermediate learners, particularly in sentence-writing tasks, while written input was more effective for lower-intermediate learners in gap-filling tasks. The findings partially supported the Task-Induced Involvement Load Hypothesis (Laufer & Hulstijn, 2001) and highlighted the role of input modality in vocabulary learning.

A meta-analysis by Ramezanali et al. (2020) examined the effectiveness of adding glossing modes on second language (L2) vocabulary learning. They analyzed 22 studies with 26 independent effect sizes, focusing on single vs. dual and dual vs. triple glossing modes. The researchers found that adding a mode to single textual glossing enhanced vocabulary learning, while adding a mode to dual glossing did not result in significantly better learning. They also found that using more than two gloss modes was not necessary, as it did not always lead to better learning. Moderator variables included learner proficiency, gloss language, text type, and research design. The findings suggest that the effectiveness of additional gloss modes is influenced by a range of variables related to the learner, gloss, text, and research design.

Wang and Lee (2021) investigated the effects of different multimedia glossing modes on EFL learners' vocabulary acquisition and reading comprehension. The study involved 160 Chinese university students who were exposed to one of four glossing conditions: L2 definition only, L2 definition with audio, L2 definition with video, or L2 definition with pictures. Participants read eight hypermediated texts and completed pre- and post-tests for vocabulary and reading comprehension. Results showed that all glossing modes led to vocabulary gains, with L2 definition plus pictures and videos yielding the highest scores. Reading comprehension improved across all conditions, but no significant differences were found between groups. Questionnaire and interview data revealed that students

preferred visual glosses (pictures and videos) over audio or text-only glosses. The study highlighted the effectiveness of multimedia glosses in enhancing vocabulary learning and reading comprehension, particularly during the COVID-19 pandemic.

Zhang and Zou (2022) reviewed 41 articles published in SSCI journals from 2009 to 2020, investigating the effectiveness of various multimedia input modes in promoting second and foreign language (L2) learning. The researchers identified five main types of tools used in the studies: computers, mobile devices, printed materials, audio players, and the integration of PowerPoint slides and stereo speakers. They also categorized the target aspects of L2 knowledge and skills into four main areas: vocabulary knowledge, listening comprehension, reading comprehension, and grammar knowledge. The review revealed that the most effective multimedia input modes varied depending on the target aspect of L2 knowledge or skill. For developing vocabulary and grammar knowledge, the mode with audio-plus-animation-plus-captions/subtitles was found to be most effective. For promoting listening comprehension, audio-plus-animation was the most effective mode, and for facilitating reading comprehension, text-plus-audio was the most effective mode.

Fievez et al. (2023) investigated the impact of extensive exposure to French audio-visual input with glossed captions on incidental vocabulary learning among Dutch-speaking undergraduates. The study involved 102 participants, with 65 in the treatment group watching six episodes of the series Dix pour cent and 37 in a control group. The researchers administered pre-tests and post-tests to measure vocabulary gains, focusing on the influence of learner-related variables (vocabulary size, lookup behavior) and word-related factors (retention interval, frequency of occurrence). Results indicated significant vocabulary gains in the treatment group, with higher vocabulary size correlating with greater learning. The study concluded that watching series with glossed captions can effectively enhance vocabulary acquisition, emphasizing the importance of learner engagement with glossed content.

Teng (2023) investigated the effectiveness of multimedia input on vocabulary learning and retention among Chinese-speaking university students learning English as a foreign language (EFL). The study employed a mixed-method design, with 125 participants randomly assigned to four input conditions: definition-only, definition + word information, definition + word information + audio, and definition + word information + video. Vocabulary knowledge was assessed through pret-ests, immediate post-tests, and delayed tests after two weeks, focusing on receptive and productive vocabulary knowledge. Participants also completed a survey, and five from each group were interviewed. The results showed that the definition + word information + video condition yielded the highest vocabulary learning and retention gains, supported by ANCOVA analyses. Participants perceived the video condition as the most effective, citing its ability to provide contextual clues and cultural knowledge. The findings of the study aligned with Mayer's (2001) cognitive theory of multimedia learning, emphasizing the benefits of dual-channel (visual + verbal) input for vocabulary acquisition.

Kaderoğlu (2024) examined incidental vocabulary learning from audio-visual input among pre-intermediate Turkish learners of English. The study employed a quasi-experimental pre- and post-test design involving 39 university students, who were divided into two groups: one viewing a short video with L2 captions and the other without. Participants' knowledge of 23 target vocabulary items was assessed through form recognition and meaning recall tests before and after the video exposure. The findings indicated significant improvements in both groups, with the captions group outperforming the no-captions group specifically in the meaning recall test after controlling for pre-test scores. This suggests that integrating multimodal input, such as captions, enhances vocabulary learning compared to unimodal input. The study underscored the potential of audio-visual resources for vocabulary acquisition, emphasizing the importance of using captions for better learning outcomes.

Mahdi et al. (2024) conducted a second-order meta-analysis to investigate the effects of multimedia glosses on L2 vocabulary learning. They synthesized findings from seven primary meta-analyses, encompassing 136 studies, to explore the overall effect size and moderating factors. The study found a medium effect size (g = 0.63–0.76) for glosses on vocabulary learning, with beginner-level learners benefiting the most. Recognition tests yielded higher effect sizes than recall tests, and expository texts were more effective than narrative texts. Single-mode glosses outperformed multi-mode glosses, and pop-up/margin glosses were more effective than in-text or bottom glosses. The study highlighted the importance of prior vocabulary knowledge and the role of cognitive load in learning. The authors suggested that future research should explore different gloss types and their impact on reading comprehension.

Zhang and Ma (2024) conducted a meta-analysis to investigate the effects of textual glosses on second language (L2) vocabulary acquisition. The study analyzed 20 empirical studies involving 2,291 participants, focusing on the impact of glossing types (language, location, and mode) and moderator variables (media, academic status, and outcome measures). The results revealed that textual glosses had a medium effect on L2 vocabulary acquisition, with the effect persisting over time. Glossing location and mode significantly influenced vocabulary gains, with in-text glosses and single-mode glosses being more effective. However, glossing language (L1 vs. L2) did not show a significant impact. Outcome measures were the only significant moderator, with productive vocabulary tests showing greater gains than receptive tests. The study concluded that textual glosses are effective for vocabulary learning, particularly when designed with proximity and simplicity in mind.

Montero Perez et al. (2025) explored incidental vocabulary learning through audio-visual (AV) input, focusing on the role of on-screen text and imagery. The authors reviewed 14 studies published between 2010 and 2021, categorizing them based on task conditions that influence incidental learning, such as pre-tests, interim testing, and textual enhancement. They found that AV input, especially with captions, significantly aids vocabulary acquisition, with learners picking up both single words and formulaic expressions. Enhanced captions, such as highlighted words, showed mixed results in boosting learning. Imagery in videos also played a crucial role, as words supported by visuals were more likely to be learned. The authors highlighted methodological considerations, such as test-input modality congruency, and suggested future research should use mixed methods to better understand how learners engage with multimodal input.

In sum, the studies in the literature suggest that multimedia input generally enhances vocabulary acquisition, though its effectiveness varies depending on several key factors. Research consistently demonstrates that captioning and glossing support vocabulary learning, with dual-mode presentations often proving more effective than single-mode approaches. Learner proficiency levels and task-induced involvement significantly influence learning outcomes, with more engaging tasks and appropriate input modalities leading to better retention. The research also highlights the importance of design considerations, such as gloss placement and type, in facilitating vocabulary acquisition. These findings have important implications for educational technology design and pedagogical practices in second language vocabulary instruction. Therefore, the present study aims to investigate the effects of different types of glosses (no gloss, TG, and MCG) on Iraqi EFL learners' incidental vocabulary learning from AV input.

Research Questions

The study will address the following research questions:
1. Is there any significant difference between no gloss (control group), traditional gloss (TG), and multiple-choice gloss (MCG) groups in terms of their effect on Iraqi EFL learners' incidental vocabulary learning through audio-visual (AV) input?
2. Is there a significant relationship between Iraqi EFL learners' prior vocabulary knowledge and their incidental vocabulary learning?

Methodology

Research Design

The study was a quasi-experimental study, which employed a pre-test and post-test design to investigate the effect of different glossing conditions on incidental vocabulary learning through audio-visual input and the relationship between prior vocabulary knowledge and learning. The independent variable was the type of gloss with three levels of (1) traditional gloss (TG), (2) multiple-choice gloss (MCG), and (3) no gloss (NG) conditions. The dependent variable was the incidental L2 vocabulary learning through viewing animated short videos.

Participants

The participants in this research were students majoring in ELT (English Language Teaching) in a state university located in Baquba, Iraq. All participants were of Iraqi nationality, specifically L1 intermediate EFL learners from various intact classes. In order to assess their proficiency level, they took the Oxford Quick Placement Test (OQPT). Initially, the study began with approximately 150 participants. However, some individuals were excluded from the final analysis because they did not complete the assigned tasks or tests. Ultimately, a total of 90 learners were deemed eligible and were included as the final participants of the study. These individuals were evenly and randomly divided into three separate groups: the traditional gloss (TG, *N* = 30), the multiple-choice gloss (MCG, *N* = 30), and the no gloss (NG, *N* = 30) groups.

Data Collection Instruments

The following were used as the instruments in the study: Oxford Quick Placement Test (OQPT), videos, a video platform, glosses, and tests.

Oxford Quick Placement Test

The Oxford Quick Placement Test (OQPT) was used to homogenize participants' language proficiency prior to group assignment. The OQPT contains 60 multiple-choice items that assess grammatical structures and vocabulary knowledge at elementary to advanced levels. The validity and reliability of the test have been established through numerous administrations in previous research (Allen, 2004). It employs a dichotomous scoring where one mark is given for each correct answer, yielding a minimum score of 0 and a maximum of 60.

Based on the OQPT rubric, the learners whose scores are between 30-47 are at the intermediate level. Therefore, after the test was administered to a primary population of 150 students, those who obtained scores in this range were chosen as the final participants, and others were excluded from the study.

Videos

Six concise, informative, and visually engaging videos were carefully selected from the TED-Ed platform, a reputable repository for curated educational videos. The chosen videos included *How sugar affects the brain* (by Nicole Avena), *Where did Earth's water come from?* (by Zachary Metz), *The science of symmetry* (by Colm Kelleher), *Why animals help each other* (by Ashley Ward), *How do ocean currents work?* (by Jennifer Verduin), and *The power of simple words* (by Terin Izil). All selected videos were concise (durations lower than 5 minutes), visually engaging, and covered a variety of topics, making them ideal for an intermediate level of proficiency. Additionally, the videos contained a satisfactory number of unfamiliar words from low-frequency bands, which aligned with our research objectives.

The control group watched *How Sugar Affects the brain* (by Nicole Avena) and *Where did Earth's water come from?* (by Zachary Metz) as the two videos without glosses. The traditional gloss (TG) group watched *The Science of Symmetry* (by Colm Kelleher) and *Why Animals Help Each Other* (by Ashley Ward), and the multiple-choice gloss (MCG) group was exposed to the two videos *How do ocean currents work?* (by Jennifer Verduin) and *The power of simple words* (by Terin Izil), with options defining targeted words.

Video Platform

EdPuzzle, as an open-source and user-friendly platform, was used to incorporate conventional (L1, L2 definition) and interactive explanations (e.g., MCGs) at certain moments within videos. Edpuzzle is an interactive learning platform that allows teachers to create engaging video lessons by adding questions, audio notes, and comments to existing or uploaded videos. It supports videos from sources like YouTube and TED-Ed, and integrates with LMS platforms such as Google Classroom and Canvas. Edpuzzle promotes active learning, personalization, and data-driven insights, making it a valuable tool for enhancing video-based education while saving teachers time and improving student outcomes. The justification for selecting EdPuzzle for this particular investigation lies in its partial availability at no cost, coupled with the effortless ability for the researcher to insert traditional glosses (TGS), multiple-choice glosses (MCGs), as well as True/False question at specific points in the video content, thereby monitoring the accomplishments of learners.

Gloss Types

In the *no gloss (NG) condition*, there was an absence of gloss. The second condition encompassed *transitional glosses (TGs)*. In this particular form of gloss, students were solely exposed to the meaning of the target word in both the first language (L1) and the second language (L2), in conjunction with their respective parts of speech and IPA transcription. A total of 30 target words were selected from the six aforementioned videos. Out of these 30 target words, 10 were nouns, 10 were verbs, and 10 were of the adjective category. The researcher meticulously watched the videos featuring captions and identified the target words, their selection contingent on the researcher's evaluation of the participants' vocabulary size. In order to maximize unfamiliarity amongst the participants, the researcher selected words from lower frequency bands. In the *multiple-choice gloss (MCG) condition*, the participants were presented with the question after the occurrence of the target word within the video. Nevertheless, it is noteworthy to mention that in both conditions, the participants were given the opportunity to observe the excerpt once again, wherein the target word was displayed. All participants in all three conditions were mandated to provide a response of either true or false, thereby ensuring a more acutely discerning and perceptive viewing experience.

Passive Recall Test

A passive recall test was administered as both the pre-test and post-test to assess the participants' incidental vocabulary knowledge. The participants were provided with 30 target words of different parts of speech (i.e., 10 nouns, 10 adjectives, and 10 verbs), with no additional context or cues provided. Their task was to demonstrate their understanding of each word by writing either an L2 synonym, an L2 definition, or an L1 equivalent in the blank spaces provided below each target word. This format was designed to measure their ability to passively recall and write the meanings of the words, offering insights into their incidental vocabulary retention and comprehension before and after the intervention. The participant's passive recall test score ranged between 0 to 30, with 1 point assigned to each correct answer.

Procedure

The procedure began with the recruitment of 150 EFL learners majoring in English Language Teaching (ELT) at a state university in Baquba, Iraq. To ensure homogeneity in language proficiency, all participants took the Oxford Quick Placement Test (OQPT). Based on the OQPT results, 90 participants with scores between 30 and 47, indicating intermediate proficiency, were selected and randomly assigned to one of three groups: TG (*N* = 30), MCG (*N* = 30), or NG (*N* = 30).

The study utilized six TED-Ed videos, each under five minutes in duration, selected for their educational value, visual engagement, and inclusion of low-frequency target words. The videos were divided among the three groups of participants. The glosses were embedded using Edpuzzle, a user-friendly platform that allowed the researcher to insert L1/L2 definitions and multiple-choice questions at specific points in the videos. Edpuzzle also enabled the tracking of participants' interactions with the videos, including their responses to embedded questions.

Before the intervention, all participants completed a pre-test in the form of a passive recall test. This test required participants to provide L2 definitions, L2 synonyms, or L1 equivalents for 30 target words presented in written form. The same test was administered as a post-test after the intervention to measure incidental vocabulary learning. During the intervention, participants watched their assigned videos individually on Edpuzzle. The TG group received traditional glosses, which included L1/L2 definitions and parts of speech. The MCG group encountered multiple-choice questions after the target words appeared, requiring them to select the correct definition. The NG group watched the videos without any glosses, but was still required to answer true/false questions to ensure engagement. All groups were allowed to rewatch video segments containing the target words.

Data collection was conducted over four weeks, with participants completing the tasks in a controlled environment to ensure consistency. The pre-test and post-test responses were scored dichotomously, with one point awarded for each correct answer. The data were analyzed using ANOVA to compare the effects of the glossing conditions on vocabulary learning and the Pearson correlation coefficient to explore the relationship between prior vocabulary knowledge and learning gains.

Results

Descriptive Statistics of Oxford Quick Placement Test (OQPT)

Before conducting the experiment and in order to make sure that the two groups were homogeneous, the results of the performance of the three groups on the Oxford Quick Placement Test (OQPT) were compared (Table 1).

Table 1
Descriptive Statistics for Oxford Quick Placement Test (OQPT)

Group	N	M	SD
NG	30	36.93	3.67
TG	30	37.86	4.53
MCG	30	35.86	3.63

It can be seen that the mean scores of the three groups have seemingly no significant difference. However, to confirm that there is actually no significant difference between the mean scores of the three groups, a one-way ANOVA was conducted. As detailed in Table 2, the analysis revealed no statistically significant difference in the OQPT scores of the NG, TG, and MCG groups, $F(2, 87) = 1.83$, $p = 0.16$.

Table 2
One-way ANOVA Comparing the Mean Scores of NG, TG, and MCG Groups on OQPT

	SS	df	MS	F	p
Between Groups	60.08	2	30.04	1.83	0.16
Within Groups	1424.80	87	16.37		
Total	1484.88	89			

Note. SS = sum of squares; MS = mean square.

Cross-Group Comparison in Terms of Incidental Vocabulary Learning

The first research question examined if there was any significant difference between the no gloss (control group), traditional gloss (TG), and multiple-choice gloss (MCG) groups in terms of their effect on incidental vocabulary learning through audio-visual input. The relevant descriptive statistics are shown in Table 3, showing some differences between the groups (i.e., NG, TG, and MCG).

Table 3
Descriptive Statistics of the Participants' Gains in NG, TG, and MCG Groups

Group	N	Pre-test M	Pre-test SD	Post-test M	Post-test SD	Mean Gain
NG	30	17.83	3.06	18.96	2.27	1.13
TG	30	18.43	4.22	21.13	3.23	2.70
MCG	30	18.26	2.88	25.06	1.76	6.80

In order to make sure about the superficial difference seen in the descriptive statistics, a one-way between-groups analysis of variance was employed to check if the three groups were statistically significant (Table 4). It was found that there was a statistically significant difference at the $p < 0.05$ level in the post-test scores of the three groups: $F(2, 87) = 48.03$, $p = 0.00$.

Table 4
One-way Between-Groups ANOVA Comparing the Means of NG, TG, and MCG Groups

	SS	df	MS	F	p
Between Groups	599.62	2	299.81	48.03	0.00
Within Groups	543.00	87	6.24		
Total	1142.62	89			

Note. SS = sum of squares; MS = mean square.

A significant difference was found in the one-way between-groups ANOVA results since the Sig. value was less than 0.05. Therefore, a post-hoc Scheffe test was run to show exactly where the differences between the three groups occur:

Table 5

Post-Hoc Scheffe Test Indicating Exact Difference between the NG, TG, and MCG Groups

Group		MD	SE	p	95% CI	
					LL	UL
NG	TG	-2.20*	0.64	0.00	-3.81	-0.59
	MCG	-6.23*	0.64	0.00	-7.84	-4.63
TG	NG	2.20*	0.64	0.00	0.59	3.81
	MCG	-4.03*	0.64	0.00	-5.64	-2.43
MCG	NG	6.23*	0.64	0.00	4.63	7.84
	TG	4.03*	0.64	0.00	2.43	5.64

Note. * The mean difference is significant at the 0.05 level.

MD = mean difference; *SE* = standard error; CI = confidence interval; LL = lower limit; UL = upper limit.

Looking down the column labelled Mean Difference, we could see asterisks (*) next to some of the values listed. This means that the two groups being compared are significantly different from one another at the *p* < 0.05 level. The exact significance value is given in the column labelled Sig. Post-hoc comparisons using the Scheffe test indicated that scores of the NG, TG, and MCG groups were statistically significantly different from one another.

Relationship Between Participants' Prior Vocabulary Knowledge and Incidental Vocabulary Learning

The second research question investigated if there was a significant relationship between the participants' prior vocabulary knowledge and incidental vocabulary learning. In order to see the correlation between participants' prior knowledge and their relative gain scores, four Pearson's Correlation Coefficient tests were run using SPSS 26. The results indicated that there were no significant correlations between prior knowledge and incidental vocabulary learning for MCG condition (*r*(30) = 0.20, *p* = 0.34), for TG condition (*r*(30) = 0.29, *p* = 0.10), and for NG condition (*r*(30) = 0.38, *p* = 0.06). However, as shown in Table 6 below, there was a strong, positive correlation between prior knowledge and incidental vocabulary learning when the groups' scores were calculated altogether (*r*(90) = 0.61, *N* = 90, *p* = 0.02). Based on Cohen's (1988) guideline for the interpretation of correlation coefficients, in which correlations above 0.50 are large, it can be seen that the participants' prior knowledge was significantly correlated with their incidental vocabulary learning scores (*r* = 0.61, *p* < 0.05).

Table 6

Pearson Product Correlation Between All Participants' Prior Vocabulary Knowledge and Incidental Vocabulary Learning

		VAR1	VAR2
VAR1	Pearson Correlation	1	0.61*
	Sig. (2-tailed)		0.02
	N	90	90
VAR2	Pearson Correlation	0.61*	1
	Sig. (2-tailed)	0.02	
	N	90	90

Discussion

The present study investigated the effects of different gloss types on Iraqi EFL learners' incidental vocabulary learning from audio-visual input. The findings of the two research questions were as follows: The first conclusion of the study indicates that there was a significant difference among the no gloss (control group), traditional gloss (TG), and multiple-choice gloss (MCG) groups in terms of their effect on incidental vocabulary learning through audio-visual input, and (2) the second conclusion of the study revealed a strong, positive correlation between participants' prior knowledge and incidental vocabulary learning. Both gloss groups significantly outperformed the control, evidencing annotation benefits vocabulary uptake beyond mere exposure. Notably, the MCG group demonstrated substantially greater gains than TG, indicating that interactive glossing optimizes learning over simpler formats.

The significant differences found between the no-gloss, traditional gloss (TG), and multiple-choice gloss (MCG) groups align with previous research demonstrating the effectiveness of glossed input for vocabulary learning. This finding particularly resonates with Wang and Lee's (2021) study, which found that all glossing conditions led to vocabulary gains compared to non-glossed conditions. However, while Wang and Lee found that L2 definition plus pictures and videos yielded the highest scores, the present study suggests that the interactive nature of multiple-choice glosses may provide additional benefits. This difference might be attributed to the enhanced engagement required by MCG, supporting Kıvrak and Uygun Gökmen's (2019) finding that tasks with higher involvement loads lead to better vocabulary retention.

The superior performance of both gloss groups over the control condition supports Hsu's (2018) findings regarding the effectiveness of caption-gloss modes for vocabulary acquisition. However, while Hsu found sustained benefits in delayed post-tests, future research should examine whether the advantages of MCG over TG persist over time, particularly given the different cognitive processes involved in interactive glossing.

Çakmak and Erçetin (2018) similarly found that all gloss conditions outperformed their control group in incidental vocabulary learning. However, while they found no significant differences among different gloss types, the current study revealed that MCG significantly outperformed TG. This divergence might be explained by the different types of glosses examined in that their study compared textual, pictorial, and textual-plus-pictorial glosses, while ours focused on different levels of learner interaction with the glosses.

The particularly strong performance of the MCG group compared to TG suggests that interactive glossing formats may optimize learning outcomes. This finding appears to challenge recent meta-analytic findings by Mahdi et al. (2024), who found that single-mode glosses outperformed multi-mode glosses. However, this apparent contradiction might be explained by considering that MCG, while more interactive, doesn't necessarily constitute a multi-mode gloss in the traditional sense, but rather represents a different approach to learner engagement with single-mode glossing.

The effectiveness of MCG over TG also provides an interesting counterpoint to Zhang and Ma's (2024) meta-analysis, which emphasized the benefits of simplicity in gloss design. While their study found that simpler, in-text glosses were more effective, our

results suggest that adding an interactive element through a multiple-choice format can enhance learning without necessarily increasing cognitive load. This difference might be reconciled by considering that MCG, while more complex than TG in terms of interaction, still maintains clarity and proximity in its presentation.

The strong positive correlation between participants' prior knowledge and incidental vocabulary learning provides an interesting parallel to Fievez et al.'s (2023) findings, where higher vocabulary size correlated with greater learning gains. This consistent pattern across studies suggests that existing vocabulary knowledge serves as a crucial foundation for acquiring new words through glossed audio-visual input. The finding also aligns with Puimège and Peters' (2019) observation that learning gains were mediated by learners' prior vocabulary knowledge, though their study focused specifically on TV viewing without glossing. This finding also aligns with Teng's (2023) emphasis on the importance of learner characteristics in multimedia learning. While Teng focused on the benefits of dual-channel input through definition, word information plus video, the present study suggests that the interaction between prior knowledge and gloss type might be equally important for optimizing vocabulary acquisition.

The positive correlation between prior knowledge and vocabulary learning can be attributed to several factors. Firstly, learners with greater prior knowledge have a more extensive mental schema, which allows them to organize and integrate new vocabulary items more effectively. These learners can relate new words to familiar concepts and contexts, making the learning process more meaningful and memorable. Secondly, learners with higher prior knowledge are more adept at making accurate guesses about the meaning of unfamiliar words encountered in audio-visual input. This guessing ability, coupled with the contextual clues provided by the audio-visual materials, enables learners to infer the meaning of unknown words, leading to incidental vocabulary acquisition.

The findings of the study yield significant theoretical and practical implications, providing support for several fundamental theories in second language acquisition and multimedia learning. The superior performance of both gloss conditions over the control group aligns with cognitive load theory (Sweller et al., 1998), suggesting that glosses reduce extraneous cognitive load during viewing, allowing more resources for meaningful processing. This finding, as discussed earlier, parallels results from studies like Wang and Lee (2021) and Çakmak and Erçetin (2018), who also found benefits of glossed conditions. The particular effectiveness of the MCG condition provides strong support for the Involvement Load Hypothesis (Laufer & Hulstijn, 2001). The interactive nature of multiple-choice glossing appears to optimize the components of need, search, and evaluation, leading to superior retention compared to traditional translation glossing. This finding suggests a reconsideration of simpler gloss formats in favor of more interactive designs that promote deeper cognitive engagement. The strong correlation between prior vocabulary knowledge and learning gains, which aligned with findings from Fievez et al. (2023) and Puimège and Peters (2019), supports constructivist perspectives on vocabulary acquisition. This relationship suggests that new word learning is facilitated by and builds upon existing lexical networks, with more extensive prior knowledge enabling better integration of new vocabulary.

There were certain limitations faced in conducting the present investigation, warranting consideration. The quasi-experimental design poses potential threats to internal validity. The study's focus on receptive knowledge through post-tests, without measuring productive skills or long-term retention, provides a limited view of vocabulary acquisition. This limitation particularly stands out when considering the broader findings in the literature review, where studies like Hsu (2018) incorporated delayed post-tests to assess retention. The cross-sectional nature of the study limits understanding of how and why different gloss types affect learning differently. Additionally, the study's narrow focus on vocabulary acquisition, without examining other outcomes like viewing comprehension or learner perceptions, may overlook important aspects of multimedia learning.

Numerous opportunities remain for additional research to further advance insights into the effects of different glossing techniques on EFL learners' incidental vocabulary acquisition. Longitudinal studies investigating multiple exposures and retention over time could better reflect natural learning processes, addressing a gap noted in the discussion section when comparing with studies like Teng (2023). The research also suggests exploring different types of glosses with varying levels of learner involvement, such as fill-in-the-blank or matching activities. This aligns with the discussion's emphasis on the importance of learner engagement, as highlighted through comparison with Kıvrak and Uygun Gökmen's (2019) findings on task involvement. Investigation of learner variables like proficiency levels, learning styles, and motivation could guide more personalized implementation strategies. Additionally, technological innovations present opportunities for advancing gloss design, including adaptive systems that adjust to learner proficiency and automated vocabulary tracking. This direction connects to the broader findings in Zhang and Zou's (2022) review of multimedia input modes, suggesting the need to optimize different combinations of tools and approaches for various learning contexts.

While the study provides valuable insights into the effectiveness of different glossing approaches for vocabulary acquisition through audio-visual input, it also highlights the need for more comprehensive research incorporating diverse methodological approaches, longitudinal designs, and broader outcome measures. The findings contribute to our understanding of multimedia vocabulary learning while suggesting promising directions for future investigation and practical application in EFL contexts.

Conclusion

The findings contribute to understanding how different glossing approaches affect vocabulary acquisition through audio-visual input. The results suggest that while traditional glossing enhances vocabulary learning compared to no glossing, interactive glossing formats like MCG might provide additional benefits by engaging learners more deeply in meaning construction. This aligns with broader findings in the field about the importance of learner engagement and task involvement in vocabulary acquisition, while also suggesting new directions for gloss design that optimize the balance between interaction and cognitive load.

References

Ashcroft, R. J., Garner, J., & Hadingham, O. (2018). Incidental vocabulary learning through watching movies. *Australian Journal of Applied Linguistics, 1*(3), 135-147. https://dx.doi.org/10.29140/ajal.v1n3.89

Allen, D. (2004). *Oxford placement test*. Oxford University Press.

Çakmak, F., & Erçetin, G. (2018). Effects of gloss type on text recall and incidental vocabulary learning in mobile-assisted L2 listening. *ReCALL, 30*(1), 24-47. https://doi:10.1017/S0958344017000155

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

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

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207-223.

Hsu, H.-T. (2018). Incidental professional vocabulary acquisition of EFL business learners: Effect of captioned video with glosses. *The JALT CALL Journal, 14*(2), 119–142. https://doi.org/10.29140/jaltcall.v14n2.j227

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), The handbook of second language acquisition (pp. 349-381). Blackwell Publishing.

Kaderoğlu, K. (2024). Incidental vocabulary learning from audio-visual input: The case of pre-intermediate Turkish EFL learners. *ELIA, 24*, 177-207. http://dx.doi.org/10.12795/elia.2024.i24.6

Kıvrak, C., & Uygun Gökmen, D. (2019). The effects of task-induced involvement load and input modality on incidental vocabulary learning. *Journal of Language and Linguistic Studies, 15*(4), 1355–1375.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1-26. https://doi.org/10.1093/applin/22.1.1

Mahdi, H. S., Mohsen, M. A., & Almanea, M. (2024). Multimedia glosses and second language vocabulary learning: A second-round meta-analysis. *Acta Psychologica, 248*, 104341. https://doi.org/10.1016/j.actpsy.2024.104341

Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.

Montero Perez, M., Peters, E., Leuven, K. U., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology, 18*(1), 118-141.

Montero Perez, M., Sydorenko, T., & Huntley, L. (2025). Incidental vocabulary learning from watching audio-visual input. In *Researching incidental vocabulary learning in a second language* (pp. 35-50). Rutledge.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audio-visual 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

Ramezanali, N., Uchihara, T., & Faez, F. (2020). Efficacy of multimodal glossing on second language vocabulary learning: A meta-analysis. *TESOL Quarterly, 54*(4), 105-133. https://doi.org/10.1002/tesq.579

Rodgers, M. P. (2018). The images in television programs and the potential for learning unknown words: The relationship between on-screen imagery and vocabulary. *ITL-International Journal of Applied Linguistics*, *169*(1), 191–211. https://doi.org/10.1075/itl.00012.rod

Schmitt, N. (2000). Vocabulary in language teaching. Cambridge University Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285. https://doi.org/10.1207/s15516709cog1202_4

Teng, M. F. (2023). The effectiveness of multimedia input on vocabulary learning and retention. *Innovation in Language Learning and Teaching, 17*(3), 738-754. https://doi.org/10.1080/17501229.2022.2131791

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System, 41*(3), 609-624. https://doi.org/10.1016/j.system.2013.07.012

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219-258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Wang, S., & Lee, C. I. (2021). Multimedia gloss presentation: Learners' preference and the effects on EFL vocabulary learning and reading comprehension. *Frontiers in Psychology, 11,* 602520. https://doi.org/10.3389/fpsyg.2020.602520

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition, 19*(3), 287-307. https://doi.org/10.1017/S027226319700301X

Webb, S., & Rodgers, M. P. (2009). The lexical coverage of movies. *Applied Linguistics, 30*(3), 407-427. https://doi.org/10.1093/applin/amp010

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 66-87.

Yeh, Y., & Wang, C. W. (*2003*). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *Calico Journal, 21*(1), 131-144. https://doi.org/10.1558/cj.v21i1.131-144

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

Zhang, R., & Zou, D. (2022). A state-of-the-art review of the modes and effectiveness of multimedia input for second and foreign language learning. *Computer Assisted Language Learning, 35*(9), 2790–2816. https://doi.org/10.1080/09588221.2021.1896555