

Research paper

The Effect of On-Screen Text through Imagery on Vocabulary Development and Retention

Yasaman Seif ¹, & Saeid Najafi Sarem ^{2*}

¹Department of English, Ha. C., Islamic Azad University, Hamedan, Iran.

²Department of English, Ha. C., Islamic Azad University, Hamedan, Iran.

Seif, Y., & Najafi Sarem, S. (2024). The effect of on-screen text through imagery on vocabulary development and retention. *Journal of new advances in English Language Teaching and Applied Linguistics*, 6(2), 127-148
Doi: 10.22034/Jeltal.2024.6.2.6

*corresponding author: saeidnajafi@iau.ac.ir

Abstract

Recent research has highlighted the potential of multimodal approaches in enhancing vocabulary acquisition among language learners. Studies have shown that combining visual and textual elements can significantly improve vocabulary retention and recall. In the same direction, this study investigated the effect of on-screen text through imagery on vocabulary development and retention among Iranian EFL learners. Fifty intermediate-level female students aged 15-19 were selected using a two-stage sampling method and randomly assigned to experimental and control groups. The Solution Placement Test was used to ensure homogeneity, while researcher-developed vocabulary pre- and post-tests assessed learning outcomes. Over ten 10-minute sessions, the experimental group viewed videos with L2 subtitles and highlighted words with imagery support, while the control group watched the same videos with only L1 subtitles. Data were analyzed using Analysis of Covariance (ANCOVA). Results revealed significant positive effects of the treatment on both vocabulary development and retention. The experimental group significantly outperformed the control group in both immediate and delayed posttests, demonstrating the effectiveness of combining on-screen text with imagery for enhancing vocabulary acquisition and long-term retention in EFL contexts. These findings have important implications for curriculum design and teaching practices in language classrooms, suggesting the integration of multimodal approaches to support vocabulary learning

Keywords: EFL Learners, Imagery, Multimodal learning, On-screen Text, Vocabulary Development, Vocabulary Retention

Received: 2024-05-22 **Revised** 2024-08-22 **Accepted** 2024-09-28

Introduction

Vocabulary acquisition is a cornerstone of second language learning, as it plays a critical role in enabling learners to communicate effectively and understand written and spoken language. Recent studies have highlighted that various teaching methods can substantially improve vocabulary acquisition and retention. Among these, the use of audiovisual materials that integrate on-screen text with visual imagery has attracted attention for its potential to support language learning.

In recent years, the contribution of audiovisual input to language learning has been widely explored. Research suggests combining auditory and visual stimuli enhances cognitive processing, thereby improving vocabulary learning outcomes (Peters & Webb, 2018; Winke, Gass, & Sydorenko, 2010). The dual coding theory, proposed by Paivio (1986), provides a theoretical framework for this approach, asserting that information presented in both verbal and visual formats is more likely to be effectively encoded and retained in memory. This theory supports the integration of on-screen text and imagery in educational settings.

Additionally, studies have demonstrated that learners often benefit from watching videos with on-screen text, such as captions or subtitles, which serve as contextual aids to comprehension and vocabulary learning (Montero Perez et al., 2014). For instance, Rodgers (2018) found a positive relationship between imagery and vocabulary acquisition when paired with auditory input. However, despite these promising findings, questions remain about how different forms of on-screen text interact with visual imagery to influence learning outcomes.

While the role of audiovisual input in vocabulary acquisition has been extensively studied, limited attention has been given to the combined effect of on-screen text and imagery on vocabulary retention. Most existing research has primarily focused on subtitles or captions without examining how these features contribute to vocabulary development when integrated with visual stimuli (Teng, 2021). This gap in the literature underscores the importance of further research better to understand these interactions and their implications for TEFL practices.

This study seeks to address this gap by examining how on-screen text and imagery affect vocabulary development and retention among English as foreign language learners. It also aims to provide practical insights into effective teaching strategies that can enhance vocabulary acquisition in TEFL contexts.

Theoretical framework

Dual coding theory (DCT)

Dual Coding Theory (DCT), introduced by Allan Paivio in the 1970s, proposes that humans process and store information through two distinct channels: verbal and visual. According to this theory, combining verbal and visual inputs enhances learning by offering multiple pathways for encoding and retrieval. Paivio (1986) emphasizes that “when information is transmitted through verbal and nonverbal channels, it is represented more fully, leading to stronger comprehension and greater recall” (p. 254). This dual-channel approach is particularly advantageous in educational contexts, where learners benefit from the simultaneous activation of both types of information to improve vocabulary acquisition and retention.

Numerous studies have validated the effectiveness of DCT in vocabulary instruction. For example, Yanasugondha (2017) found that learners exposed to L2 (English) and L1 (Thai)

translations alongside pictorial aids achieved significantly higher vocabulary recognition and retention compared to those relying solely on verbal or visual cues. These results align with Read's (2004) assertion that "systematic learning of individual words can provide a good foundation of vocabulary development" (p. 41). Such findings underscore the necessity of employing strategies that integrate verbal and visual coding systems to optimize language learning outcomes.

More recent studies have further supported and expanded upon DCT. In a quasi-experimental study, Wooten and Cuevas (2024) demonstrated that students exposed to dual coding strategies in a 5th-grade social studies classroom significantly outperformed their peers in vocabulary retention and comprehension. Similarly, Chen et al. (2022) highlighted the role of computer-aided dual coding in English vocabulary instruction, showing that integrating verbal and visual elements through digital tools enhances vocabulary acquisition and retention. These findings reaffirm the relevance of DCT in modern educational contexts.

DCT also facilitates the development of connections between visual and verbal mental representations, aiding comprehension and fostering meaning construction. As Mohamed (2021) points out, "by simultaneously activating verbal and visual coding systems, learners can enhance their ability to create meaningful associations with new vocabulary" (p. 371). This integration is particularly valuable for language learners, who often struggle with retaining new vocabulary due to limited opportunities for practice outside the classroom.

Beyond its theoretical contributions, DCT has practical applications in contemporary education, especially with the growing prevalence of digital learning tools. Incorporating multimedia elements such as animations, images, and on-screen text aligns with the principles of DCT, creating rich and engaging learning experiences. For instance, research by Mayer (2001) highlights that learners exposed to audiovisual materials tend to outperform their peers in

vocabulary assessments, suggesting that DCT-based instructional strategies improve immediate learning outcomes and enhance long-term retention.

Visual imagery in learning

Visual imagery is a cognitive strategy that enhances learning by enabling individuals to form mental representations of information. This approach not only improves memory retention but also aids comprehension and problem-solving. Recent research underscores the importance of visual imagery in educational contexts, particularly in language learning. Zeman et al. (2020) explain that “visual imagery acts as a bridge between perception and cognition, allowing learners to conceptualize ideas and deepen their understanding” (p. 5). This capacity to create mental images fosters more profound learning experiences, especially when acquiring new vocabulary.

The significance of visual imagery in learning is rooted in several theoretical models. For instance, the Psycho Neuromuscular Theory suggests that mental imagery activates neural pathways similar to those used in actual physical practice, thereby improving performance and retention. Al-Jarf (2021) emphasizes that "visualizing vocabulary items within meaningful contexts engages cognitive mechanisms that promote deeper encoding and recall" (p. 87). This theoretical foundation indicates that visualizing new words and their meanings can effectively simulate real-life learning experiences, leading to enhanced recall. Dance et al. (2023) also note that “the neural circuits triggered during imagery tasks closely mirror those activated during direct experiences, strengthening memory formation” (p. 12).

Empirical studies further validate the role of visual imagery in vocabulary acquisition. For example, Milton et al. (2021) reported that learners who employed visual imagery techniques while studying new words achieved higher retention rates compared to those using only verbal methods.

Their findings reveal that “integrating visual elements into vocabulary instruction not only supports memorization but also helps learners establish meaningful connections between words and their definitions” (p. 8). These results highlight the value of incorporating visual strategies into language teaching methodologies.

Additionally, research has demonstrated the broader cognitive benefits of visual imagery, particularly its impact on working memory and information processing. Learners who develop strong visual imagery skills often exhibit enhanced performance in tasks involving memory recall and recognition. As highlighted by Richardson (2020), “Visual imagery facilitates deeper cognitive processing and fosters robust connections between memory and learning” (p. 142). Strengthening these skills among language learners can significantly contribute to their vocabulary development.

The role of on-screen text in language learning

On-screen text is a vital element in language learning, particularly when combined with audiovisual content. As part of multimedia resources, subtitles and captions serve multiple functions that enhance both comprehension and vocabulary acquisition. Liu (2024) highlights that “on-screen texts play a critical role in improving comprehension and reducing cognitive demands, thereby enriching the learning experience in second language acquisition” (p. 5). This highlights the need for presenting on-screen text effectively in educational contexts.

A key advantage of on-screen text lies in its ability to provide contextual clues that support learners in understanding spoken language. Subtitles, for example, allow learners to connect spoken words with their written forms, reinforcing vocabulary retention and comprehension. Research by Montero Perez et al. (2022) emphasizes that “on-screen text not only aids in word

recognition but also improves listening comprehension by serving as a visual aid” (p. 10). This process aligns with the principles of Dual Coding Theory, which proposes that combining verbal and visual elements enhances learning outcomes.

Various formats of on-screen text cater to distinct educational purposes. Keyword captions, for instance, emphasize essential vocabulary, while glossed captions offer explanations or definitions alongside target words. Empirical research supports the effectiveness of these tailored approaches. Wang et al. (2023) note that “keyword-highlighted subtitles significantly boost vocabulary acquisition among language learners” (p. 15), underscoring the value of designing on-screen text to meet specific learner needs.

The cognitive load theory explains how on-screen text influences learners’ mental processing. Properly designed on-screen text can alleviate cognitive demands by reducing redundancy and ensuring efficient information delivery. This is particularly crucial in language learning, where learners must simultaneously process auditory and visual inputs. Liu (2024) asserts that “well-crafted on-screen texts minimize extraneous cognitive load, enabling learners to focus on critical content” (p. 6).

Previous studies

Recent research has examined the effects of on-screen text and visual imagery on vocabulary acquisition, offering valuable insights into their role in language learning. Ahrabi Fakhri et al. (2021) studied the influence of visual imagery on incidental vocabulary acquisition through audiovisual materials. Their findings revealed that participants exposed to vivid visual imagery exhibited significantly better vocabulary retention than those provided with minimal imagery. The authors concluded that “vivid images accompanying new vocabulary items enhance semantic

processing and support memory retention” (p. 182). This study underscores the importance of visual context in facilitating language learning.

Similarly, Milton et al. (2021) explored the impact of various types of on-screen text, such as subtitles and captions, on vocabulary learning. The study found that learners who used captions achieved significantly greater vocabulary gains than those without access to on-screen text. The researchers noted, “Captions offer a dual coding experience that strengthens word recognition and comprehension, resulting in improved retention” (p. 8). This finding highlights the critical role of integrating effective on-screen text into language instruction to enhance learning outcomes.

Peters et al. (2016) further investigated the role of visual imagery in audiovisual input on vocabulary acquisition. Their study demonstrated that words accompanied by strong visual imagery were learned more effectively than those without such support. They stated, “Visual aids in audiovisual materials promote understanding and establish memorable associations that aid vocabulary retention” (p. 396). This emphasizes educators' need to incorporate visually rich content alongside verbal instruction to improve vocabulary acquisition.

Statement of the Problem

Although substantial evidence highlights the positive effects of on-screen text and visual imagery on vocabulary acquisition, limited research exists on their combined effects and specific applications in language learning contexts. While prior studies have shown the individual benefits of these elements, little is known about how their interaction comprehensively influences vocabulary retention and development. This study addresses these gaps by exploring the combined impact of on-screen text and visual imagery on vocabulary acquisition among English language learners.

This research holds significance for instructional practices within TEFL, as it aims to provide empirical evidence for effective multimedia strategies. By optimizing the use of audiovisual materials, educators can develop pedagogical approaches tailored to the diverse needs of learners. The study focuses on how different types of on-screen text can enhance vocabulary learning and retention when paired with visual imagery. Consequently, this research is guided by two primary questions and hypotheses given below:

Research questions

1. Does on-screen text through imagery have any positive effect on vocabulary development?
2. Does on-screen text through imagery have any positive effect on vocabulary retention?

Research hypotheses

1. On-screen text through imagery does not have any positive effect on vocabulary development.
2. Does on-screen text through imagery does not have any positive effect on vocabulary retention.

Methodology

Participants

The study targeted 95 female intermediate Iranian EFL learners aged 15 to 19, enrolled in an English language institute in Hamedan, Iran. A two-stage sampling method was employed. Initially, 80 participants were conveniently selected. The Solution Placement Test (Edwards, 2007) was administered to ensure homogeneity in their English language proficiency. Based on the test results, 50 participants with scores ranging from 25 to 40—indicative of intermediate-level proficiency according to Edwards (2007)—were selected as the final sample. The selected

participants were randomly assigned to two equal groups: experimental and control groups, each comprising 25 learners. This division ensured comparability between groups prior to the intervention.

Instrumentation

Solution placement test

The Solution Placement Test (Edwards, 2007) consisted of three sections. The first section assessed grammar and vocabulary through 50 multiple-choice items, while the second section evaluated reading comprehension skills with 10 multiple-choice questions. The third section was an optional writing task designed to measure written proficiency. The grammar, vocabulary, and reading sections were administered in a single 45-minute session, adhering to the standard procedure outlined for the test.

Vocabulary pretest

A researcher-developed vocabulary pretest consisting of 30 multiple-choice items was created based on excerpts from Earth Planet. To ensure the test's validity and reliability, it was piloted with 20 students resembling the study's target participants. The pretest's reliability, calculated using the Brown Prophecy Formula, was 0.84, indicating an acceptable level of internal consistency.

Vocabulary posttest

The vocabulary posttest was designed parallel to the pretest, comprising 30 multiple-choice items based on the same source (Earth Planet). It underwent a similar piloting process, with a

reliability coefficient 0.93, as determined by the Brown Prophecy Formula. This reliability level is considered robust, as per Larson-Hall (2008).

Materials

The materials utilized in this study included a series of video excerpts from the documentary Earth Planet, featuring subtitles in both the learners' first language (L1) and the target language (L2). These videos were carefully selected to align with the participants' proficiency level and the study's instructional objectives, ensuring their relevance for vocabulary development and retention. The use of dual subtitles aimed to enhance comprehension and facilitate the association between L1 and L2 vocabulary, making the materials an integral part of the experimental design.

Procedure

The initial population consisted of 95 female intermediate Iranian EFL learners aged 15 to 19, enrolled in an English language institute in Hamedan, Iran. Using a two-stage sampling method, 80 participants were first conveniently selected. The Solution Placement Test (Edwards, 2007) was administered to these participants to ensure homogeneity in English proficiency. Based on their scores, 50 students whose proficiency level fell within the intermediate range (scores between 25 and 40) were identified and selected as the final sample. These participants were then randomly assigned into two equal groups: experimental (N = 25) and control (N = 25).

The study began with the administration of a vocabulary pretest for both groups. This pretest, developed specifically for the study, served as a baseline to evaluate participants' initial vocabulary knowledge.

The treatment phase consisted of 10 sessions, each lasting approximately 10 minutes. During these sessions, the experimental group was exposed to videos from the documentary Earth Planet, featuring L2 subtitles and 10 highlighted words per session. Each highlighted word was supported by imagery, creating a one-to-one correspondence between the word and its visual representation on the screen. This design aimed to enhance vocabulary acquisition through visual and contextual support. In contrast, the control group viewed the same videos but with L1 subtitles and without highlighted words or imagery. Immediately after the tenth session, a vocabulary posttest was administered to both groups to measure their short-term vocabulary acquisition. The posttest was re-administered two weeks later to assess long-term retention. This delayed posttest provided insights into the effectiveness of the treatment in supporting vocabulary retention over time.

Design of the study

This study employed a quasi-experimental design using a pretest-posttest control group model. This approach was selected due to practical limitations in the educational setting, where random assignment of participants to groups was not feasible. The participants were already organized into intact classes, making true randomization impossible. Quasi-experimental designs are commonly used in such contexts because they allow for the comparison of groups while maintaining ecological validity. By using a control group and controlling for pretest scores through ANCOVA, the study aims to provide reliable evidence of the treatment's effects despite the absence of full randomization.

EG	T1	X	T2
CG	T2	-----	T2

Data Analysis

To analyze the collected data, Analysis of Covariance (ANCOVA) was employed. ANCOVA was chosen to determine whether significant differences existed between the experimental and control groups on the dependent variables while controlling for pretest scores. This statistical method adjusts for any initial differences between the groups, ensuring a more accurate evaluation of the treatment effect. In this study, ANCOVA was run twice via SPSS software to assess the dependent variables: vocabulary development (immediate posttest scores) and vocabulary retention (delayed posttest scores). This approach ensured that the effect of the treatment was isolated and accurately measured. The statistical results are presented and discussed in subsequent sections.

Table 1
Descriptive Statistics of the Groups

group		pretest	posttest
experimental group	Mean	10.76	14.60
	N	25	25
	Std. Deviation	2.107	2.062
	Kurtosis	.670	.213
	Skewness	.693	-.372
control group	Mean	11.16	11.84
	N	25	25
	Std. Deviation	2.154	2.734
	Kurtosis	.088	-.270
	Skewness	.210	.448
Total	Mean	10.96	13.22
	N	50	50
	Std. Deviation	2.119	2.772
	Kurtosis	.115	-.698
	Skewness	.430	-.190

Table 1 provides the descriptive statistics for the experimental and control groups on the pretest and posttest, highlighting changes in mean scores and standard deviations. For the experimental group, the mean score on the pretest was 10.76 (SD = 2.107), which increased to 14.60 (SD = 2.062) on the posttest, reflecting notable improvement. In contrast, the control group showed a

pretest mean score of 11.16 (SD = 2.154), with a modest increase to 11.84 (SD = 2.734) on the posttest. Additionally, the skewness and kurtosis values for both groups on the pretest and posttest fall within the acceptable range of -2 to +2, indicating that the distribution of the data adheres to the assumptions of normality, as outlined by Tabachnick and Fidell (2007). These results suggest that the statistical analyses are based on appropriately distributed data, lending validity to subsequent findings.

To further validate the data distribution, the normality of the experimental group's pretest and posttest scores was examined using the Kolmogorov-Smirnov test, as summarized in Table 2.

Table 2
The Normality of the Experimental Group's Scores on Pretest and Posttest

		One-Sample Kolmogorov-Smirnov Test^c	
		pretest	posttest
N		25	25
Normal Parameters ^{a,b}	Mean	10.76	14.60
	Std. Deviation	2.107	2.062
Most Extreme Differences	Absolute	.175	.177
	Positive	.175	.129
	Negative	-.119	-.177
Kolmogorov-Smirnov Z		.873	.885
Asymp. Sig. (2-tailed)		.431	.414
a. Test distribution is Normal.			
b. Calculated from data.			
c. group = experimental group			

The results of the Kolmogorov-Smirnov test, as shown in Table 2, indicate that the experimental group's scores on both the pretest and posttest are normally distributed, with p-values greater than 0.05 ($p > 0.05$). This finding confirms that the data meet the assumptions required for subsequent parametric analyses. Similarly, the normality of the control group's pretest and posttest scores was assessed using the Kolmogorov-Smirnov test, with the results presented in Table 3.

Table 3
The Normality of the Control Group's Scores no Pretest and Posttest

		pretest	posttest
One-Sample Kolmogorov-Smirnov Test^c			
N		25	25
Normal Parameters ^{a,b}	Mean	11.16	11.84
	Std. Deviation	2.154	2.734
Most Extreme Differences	Absolute	.148	.157
	Positive	.148	.157
	Negative	-.132	-.076
Kolmogorov-Smirnov Z		.741	.783
Asymp. Sig. (2-tailed)		.642	.572
a. Test distribution is Normal.			
b. Calculated from data.			
c. group = control group			

The results of the Kolmogorov-Smirnov test in Table 3 indicate that the control group's scores on both the pretest and posttest follow a normal distribution, as the p-values exceed 0.05 ($p > 0.05$).

To ensure the assumptions for conducting ANCOVA are met, the homogeneity of regression slopes was examined, and the results are illustrated in Figure 1.

Figure 1
The Homogeneity of the Slope of Regression Lines

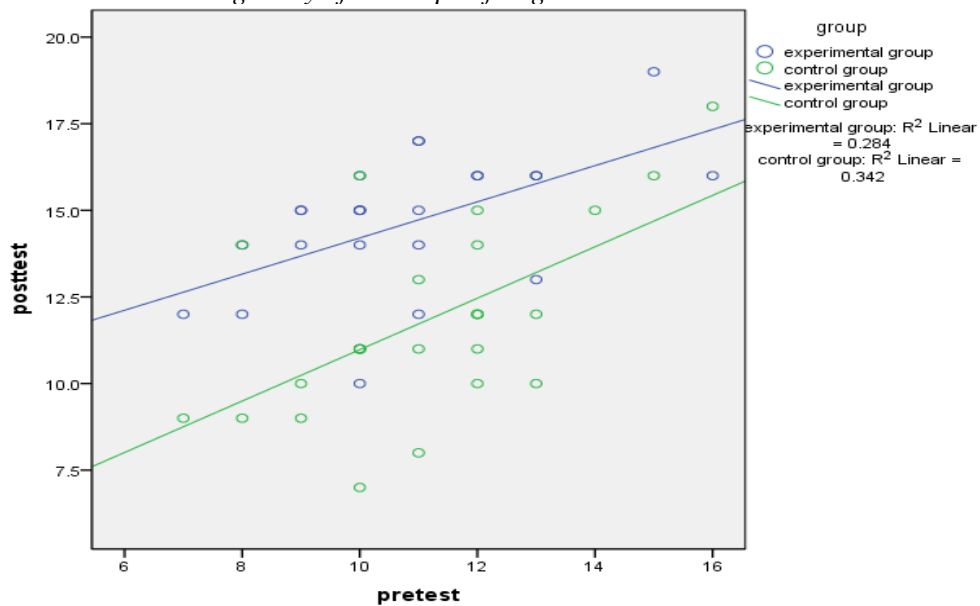


Figure 1 demonstrates a linear relationship between the pretest and posttest scores for both groups. The parallel nature of the regression lines, with no intersection, suggests that the regression slopes are homogeneous across the groups. This observation is further supported by the inferential statistics provided in Table 4, under the Tests of Between-Subjects Effects.

Table 4
Tests of Between-Subjects Effects Showing the Homogeneity of Slope of Regression Lines

Tests of Between-Subjects Effects					
Dependent Variable:posttest					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	185.557 ^a	3	61.852	14.895	.000
Intercept	68.865	1	68.865	16.583	.000
group	12.893	1	12.893	3.105	.085
pretest	86.982	1	86.982	20.946	.000
group * pretest	2.643	1	2.643	.637	.429
Error	191.023	46	4.153		
Total	915.000	50			
Corrected Total	376.580	49			

a. R Squared = .493 (Adjusted R Squared = .460)

Table 4 confirms that the regression slopes are homogeneous across all groups, as evidenced by the non-significant result [$F(1, 46) = 0.637, p > 0.05$].

To further ensure the validity of the ANCOVA analysis, Levene's test was conducted to verify the assumption of homogeneity of variances across groups. Table 5 presents the results of this test.

Table 5.
Levene's Test of Equality of Error Variances^a

Levene's Test of Equality of Error Variances^a					
Dependent Variable:posttest					
F	df1	df2	Sig.		
.716	1	48	.402		

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + pretest + group

The results presented in Table 5 indicate that the assumption of homogeneity of variances is met, as evidenced by the non-significant Levene's test [$F(1,48) = 0.716, p > 0.05$].

To assess the impact of the treatment on vocabulary development, an Analysis of Covariance (ANCOVA) was conducted. Table 6 presents the results of the between-subjects effects, which provide insight into the significance of the treatment's effect while controlling for pre-existing differences.

Table 6
T B-S E Showing the Effect of the Treatment on Vocabulary Development

Tests of Between-Subjects Effects						
Dependent Variable: posttest	Type III	df	Mean	F	Sig.	Partial
Source	Sum of		Square			Eta Squared
	Squares					
Corrected Model	182.914 ^a	2	91.457	22.195	.000	.486
Intercept	68.766	1	68.766	16.688	.000	.262
pretest	87.694	1	87.694	21.282	.000	.312
group	112.501	1	112.501	27.302	.000	.367
Error	193.666	47	4.121			
Total	9115.000	50				
Corrected Total	376.580	49				

a. R Squared = .486 (Adjusted R Squared = .464)

The results displayed in Table 6 reveal a statistically significant main effect of the treatment (on-screen text through imagery) on vocabulary development [$F(1,47) = 27.302, p < 0.001$, partial $\eta^2 = 0.367$]. This finding indicates that the intervention had a substantial impact on participants' vocabulary development. Regarding the second dependent variable, vocabulary retention, an additional ANCOVA was performed after confirming that all necessary assumptions were met. The results of this analysis are presented in Table 7.

Table 7
T B-S E Showing the Effect of the Treatment on Vocabulary Retention

Tests of Between-Subjects Effects						
Dependent Variable: posttest	Type III	df	Mean	F	Sig.	Partial
Source	Sum of		Square			Eta Squared
	Squares					
Corrected Model	185.096 ^a	2	92.548	22.785	.000	.492
Intercept	78.607	1	78.607	19.353	.000	.292
pretest	81.416	1	81.416	20.044	.000	.299
group	120.921	1	120.921	29.770	.000	.388
Error	190.904	47	4.062			
Total	9354.000	50				
Corrected Total	376.000	49				

a. R Squared = .492 (Adjusted R Squared = .471)

The results presented in Table 7 demonstrate a statistically significant main effect of the treatment (on-screen text through imagery) on vocabulary retention [$F(1,47) = 29.770, p < 0.001$, partial $\eta^2 = 0.388$]. This finding indicates that the intervention had a substantial and significant impact on participants' ability to retain vocabulary over time.

Discussion

This study investigated the effect of on-screen text through imagery on vocabulary development and retention among Iranian EFL learners. The results provide compelling evidence for the effectiveness of this approach in enhancing both vocabulary acquisition and long-term retention.

The ANCOVA results revealed a significant main effect of the treatment on vocabulary development [$F(1,47) = 27.302, p < 0.001$, partial $\eta^2 = 0.367$] and vocabulary retention [$F(1,47) =$

29.770, $p < 0.001$, partial $\eta^2 = 0.388$]. These findings support the rejection of both null hypotheses, confirming that on-screen text combined with imagery significantly enhances vocabulary development and retention among EFL learners.

These results align with previous studies demonstrating the efficacy of multimodal approaches in language learning. For instance, Mayer and Moreno (1998) found that combining visual and verbal information enhances learning outcomes by reducing cognitive load. Similarly, Montero Perez et al. (2022) reported that on-screen text, such as subtitles, significantly improves comprehension and vocabulary acquisition by providing visual references for spoken words.

The significant effects observed in both vocabulary development and retention can be explained through the lens of Dual Coding Theory (DCT) proposed by Paivio (1986). The combination of verbal (on-screen text) and visual (imagery) information likely facilitated deeper encoding and stronger memory traces, as posited by DCT. This dual-channel processing appears to have created more robust mental representations of the vocabulary items, leading to improved acquisition and retention.

Moreover, our findings extend the understanding provided by Peters et al. (2016), who found that words presented with strong visual imagery were learned more effectively and retained better than those without such support. The present study demonstrates the sustained impact of combined on-screen text and imagery on vocabulary retention over time, suggesting that this approach not only aids initial acquisition but also supports long-term memory consolidation.

The results also corroborate the findings of Liu (2024), who emphasized the importance of on-screen texts in promoting comprehension and reducing cognitive load in second language acquisition. The combination of on-screen text with imagery in our study likely optimized the cognitive processing of new vocabulary items, leading to enhanced learning outcomes.

Conclusion

This study provides compelling evidence for the effectiveness of combining on-screen text with imagery in promoting vocabulary acquisition and retention among Iranian EFL learners. The findings highlight the value of multimodal instructional approaches, particularly in the context of vocabulary learning.

The results can be interpreted through the lens of Dual Coding Theory and cognitive load theory, which suggest that integrating verbal and visual inputs enhances cognitive efficiency and memory retention. This is consistent with current insights into how the brain processes and stores information, especially in language-learning environments.

These findings carry important implications for EFL teaching and curriculum design. Educators and material developers are encouraged to incorporate on-screen text accompanied by meaningful imagery into vocabulary instruction to optimize learning outcomes. Such strategies may prove particularly advantageous in digital and multimedia-based language learning settings.

However, the study has certain limitations. The exclusive focus on Iranian EFL learners may restrict the applicability of its findings to other cultural and linguistic contexts. Moreover, the long-term impacts of this method beyond the study's retention test period remain unexplored and warrant further research.

Future studies could explore various aspects, such as comparing the effects of static versus animated imagery on vocabulary learning, evaluating learner-generated imagery versus instructor-provided visuals, examining the effectiveness of this approach across diverse proficiency levels and age groups, and conducting longitudinal research to assess sustained vocabulary retention.

References

- Al-Jarf, R. (2021). The impact of mental imagery on vocabulary retention among EFL learners. *Journal of Language Education Research*, 18(3), 85–97.
- Chen, X., Li, H., & Zhao, Y. (2022). Application of computer-aided dual coding in English vocabulary teaching. *Advances in Cognitive Psychology*, 18(1), 1–12. <https://doi.org/10.1155/2022/5951844>
- Dance, R., Stevenson, L., & Fisher, G. (2023). Neural dynamics of mental imagery in language learning. *Applied Psycholinguistics*, 44(1), 11–25.
- Edwards, C. (2007). Solutions placement test. Oxford: Oxford University Press.
- Larson-Hall, J. (2008). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.
- Liu, F. (2024). A Review of On-Screen Texts in Multimedia and Audiovisual Inputs for Second Language Learning. *Applied Linguistics*, 45(1), 5-20.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90(2), 312-320.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.
- Milton, J., Richards, P., & Walker, T. (2021). Visual strategies for vocabulary acquisition: A longitudinal study. *Language Learning*, 71(2), 7–15. <https://doi.org/10.1111/lang.12345>
- Milton, F., Zeman, A., & Dance, C. (2021). Visual imagery techniques enhance vocabulary acquisition: A study on learning outcomes. *Educational Psychology Review*, 33(2), 1-20.
- Mohamed, M. (2021). Dual coding and its role in vocabulary retention. *Journal of Language and Literacy Studies*, 12(4), 365–380.
- Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). The effect of on-screen text in audiovisual input for L2 vocabulary learning: A review. *TESOL Quarterly*, 53(4), 1008-1032.
- Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2022). On-screen texts in audiovisual input for L2 vocabulary learning: A review of empirical studies. *Language Teaching*, 55(2), 163-192.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Peters, E., Webb, S., & Montero Perez, M. (2016). The influence of visual imagery on incidental vocabulary learning from audiovisual input: Evidence from a documentary study. *World Journal of English Language*, 12(1), 395-404.
- Peters, E., & Webb, S. (2018). Vocabulary acquisition from audiovisual input: The role of subtitles and captions. *Language Learning*, 68(3), 648-680.
- Read, J. (2004). Research in vocabulary acquisition. In N. Schmitt (Ed.), *Researching vocabulary: A vocabulary research manual* (pp. 35–50). Oxford: Oxford University Press.
- Richardson, J. T. E. (2020). *Mental imagery and memory: Applications for education and language learning*. New York: Routledge.
- Rodgers, M. P. (2018). The effect of imagery on word learning: Evidence from audiovisual input. *Applied Linguistics*, 39(5), 753-779.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education.
- Teng, L. S. (2021). Incidental vocabulary acquisition from captioned videos: A review of empirical studies. *Language Teaching Research*, 25(1), 45-65.

- Wang, Y., Chen, H., & Zhang, L. (2023). The impact of keyword captions on vocabulary acquisition: An experimental study. *Educational Technology Research and Development*, 71(1), 15-32.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The role of captions in second language listening comprehension: A review of the literature. *Language Learning*, 60(3), 665-695.
- Wooten, J., & Cuevas, J. (2024). The impact of dual coding strategies on domain-specific vocabulary and comprehension. *International Journal of Social Studies Education*, 19(2), 143–162.
- Yanasugondha, M. (2017). The impact of dual coding on vocabulary retention in Thai EFL learners. *Journal of English Language Teaching*, 11(3), 56–67.
- Zeman, A., Dewar, M., & Della Sala, S. (2020). Visual imagery ability and its implications for cognitive processes: A review of recent findings. *Cognitive Science*, 44(1), e12880.